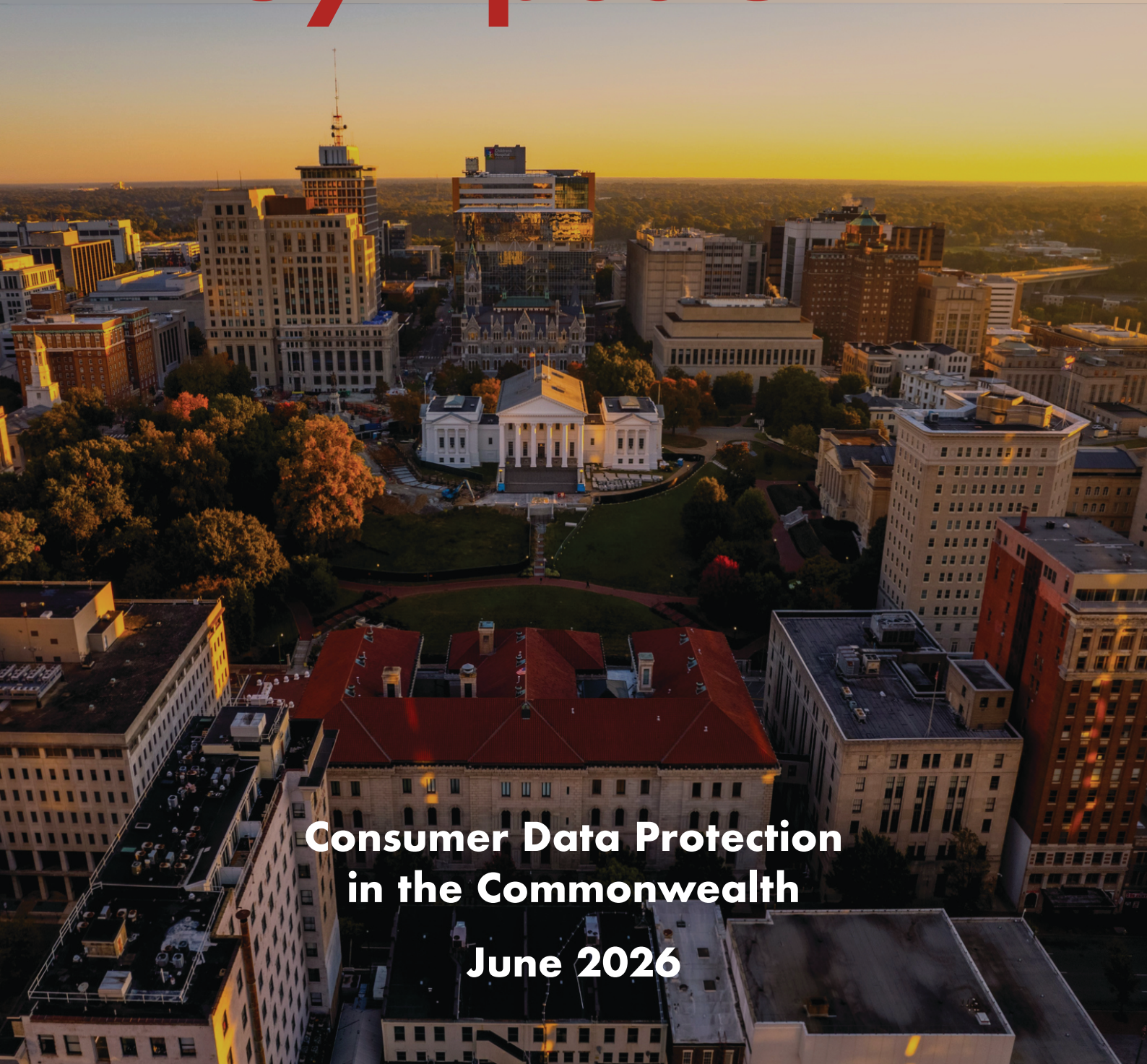
JCOTS
JOINT COMMISSION ON TECHNOLOGY & SCIENCE

WILLIAM & MARY

Technology Policy Symposium



Consumer Data Protection in the Commonwealth

June 2026

The Joint Commission on Technology and Science

JCOTS Members

Delegate Cliff Hayes, Chair

Delegate Bonita Anthony

Delegate Mike Cherry

Delegate Garrett McGuire

Delegate David Reid

Delegate Michael Webert

Delegate Michael Feggans

Senator Barbara Favola, Vice-Chair

Senator Lashrecse Aird

Senator Luther Cifers

Senator Christie New Craig

Senator Stella Pekarsky

JCOTS Staff

Jodi Kuhn, Executive Director

Kira Allmann, Chief Policy Analyst

Anna Dehmer, William & Mary Intern

Alejandro Velasco, VASEM COVES Fellow

Special Thanks to

Monty Mason, William and Mary

Skip Rowland, Photography

Rick Dana, Videographer

Mason School of Business Events Team

Table of Contents

Greetings

- Delegate Cliff Hayes 01
- Dr. Douglas C. Schmidt 02
- JCOTS Introduction 03
- William & Mary Partnership 04
- Symposium Summary 05

Symposium Panels

- Data Protections in Virginia & Beyond 07
- Data Protections in the Age of AI 12
- Health Data 30 Years After HIPAA 17

Conclusion

- In Summary 23

Speaker Essays

- Daniel Shin 25
- Dr. Sarah Henrickson Parker 31
- Dr. Elizabeth White Baker 34
- Dr. Gabriella Waters & Maurice Jones 38

Greetings

Delegate Cliff Hayes, *Chair of the Joint Commission on Technnology and Science*

Throughout my time as both a Delegate and Chair of the Joint Commission on Technology and Science (JCOTS), data protection has never been a simple issue, but today, with the rapid growth of artificial intelligence, it is an increasingly urgent one. Since I first began working on technology policy in the Commonwealth, the conversations have become far more complex, and the stakes have grown considerably higher.

In 2021, Virginia passed the Virginia Consumer Data Protection Act (VCDPA). This legislation protects Virginians by granting them further control over their data and ensuring that businesses remain transparent about their data practices. Still, the technology landscape has changed in the 3 years since VCDPA went into effect. Rapid technological advancements have created unparalleled opportunities for data application. Therefore, we must continue discussing how to balance progress and oversight within this shifting digital environment.

That is the spirit behind the JCOTS 2026 Symposium: Three Years of Consumer Data Protection in the Commonwealth hosted by William & Mary. Virginia is uniquely positioned to lead on this issue. We are home to some of the most significant data infrastructure in the world, a thriving technology sector, and research institutions producing the next generation of AI policy thinkers. It would be a disservice to that legacy if we did not meet this moment with the seriousness it demands.

The discourse this symposium sparks and the ideas developed here shape the policies we carry forward on behalf of the citizens of this Commonwealth. As we look toward the future, our goal should remain consistent: ensuring that Virginia continues to embrace technological innovation while protecting the privacy, security, and interests of the people we serve.

Dr. Douglas C. Schmidt, *Dean of the W&M School of Computing, Data Sciences and Physics*

Across Virginia and the nation, policymakers are confronting a rapidly changing technology landscape in which data, artificial intelligence, public health, environmental resilience, and public trust are increasingly intertwined. The recent symposium of the Joint Commission on Technology and Science (JCOTS) took a timely and important approach to these issues. Rather than treating technology policy as an abstract matter, the symposium focused on how the Commonwealth can protect its citizens, promote responsible innovation, and ensure that emerging technologies serve the public good.

This purpose reflects JCOTS' vital role in bringing legislators, public servants, researchers, and practitioners together to examine the scientific and technological issues most consequential to Virginia's future. Under the leadership of Delegate Cliff Hayes, Chair of the Commission, and with the guidance of Ms. Jodi Kuhn, Executive Director, the meeting provided an opportunity to consider both the promise and the challenges of innovation across several domains. The symposium engaged deeply with three panels and engaging panelists who addressed several policy questions of our time. I enjoyed moderating the first panel, which examined data protection in Virginia and beyond. I look forward to deepening collaborations with these panelists in the future.

Behind all these discussions was a common question: how can Virginia lead responsibly in technology and science while protecting the rights, health, security, and wellbeing of its citizens? The symposium made clear that the Commonwealth has an opportunity not only to respond to emerging technologies, but to help shape the legal, ethical, and institutional frameworks that will guide their use.

To conclude, I would invite you to read this publication, reflect on the issues it highlights, and join JCOTS in advancing thoughtful, evidence-based technology policy for the Commonwealth of Virginia.

The Joint Commission on Technology & Science

The Joint Commission on Technology and Science (JCOTS) was created by the 1997 Virginia General Assembly as a permanent legislative commission. The Commission is charged to generally study all aspects of technology and science and to promote the development of technology and science in the Commonwealth of Virginia through sound public policies

(Chapter 11 (§ 30-85 et seq.) of Title 30 of the Code of Virginia).

In addition, JCOTS shall:

1. Evaluate the impact of existing statutes and proposed legislation related to technology and science in the Commonwealth;

2. Advise the General Assembly, Governor, and agencies, authorities, and institutions of the Commonwealth upon matters related to technology and science;

3. Investigate, research, and consider such issues related to technology and science as may be requested by the General Assembly or determined by JCOTS;

4. Make recommendations to the General Assembly and the Governor;

5. Consult with appropriate entities, public or private, on matters related to technology and science under JCOTS' consideration;

6. Encourage research and development in technology and science;

7. Solicit input from appropriate entities, public or private, on issues related to technology and science;

8. Coordinate its efforts with and assist the efforts of other agencies, authorities, and institutions of the Commonwealth;

9. Accept private or public funds to carry out its purposes; and

10. Annually report its findings and recommendations to the General Assembly and the Governor

William & Mary Partnership

In July 2025, JCOTS announced a three-year partnership with the College of William & Mary, aimed at strengthening the Commission's research capabilities and increasing academic engagement in the legislative process. This collaboration reflects something both institutions value deeply: the connection between academic rigor and public service. William & Mary has a long-standing tradition of preparing leaders and advancing research; JCOTS exists to ensure that Virginia's policymakers have access to thoughtful, evidence-based analysis on complex technology issues. Bringing these two missions together is both natural and necessary.

This partnership provides JCOTS with access to William & Mary's world class library and research services, improving upon JCOTS's ability to study cutting edge science and technology policy issues. Along with access to research services, JCOTS hosts their Technology Policy Symposium at the William & Mary Raymond A. Mason School of Business. The symposium brings a diverse audience to campus to discuss emerging technology trends and inform policy makers of current research. Additionally, the William and Mary partnership provides one student each year, the opportunity to participate in an internship program with JCOTS where they work directly on a bill study that was referred to the Commission during the previous General Assembly session.

This partnership is intended to be the first of many. Every three years, JCOTS hopes to forge a similar collaboration with a different public college/university across the Commonwealth. For institutions, this partnership offers a direct channel to inform state policy with their research and expertise. For JCOTS, it provides fresh perspectives, specialized knowledge, and a pipeline of talented students eager to engage in public service. As this rotating model unfolds, it will allow JCOTS to draw on the distinct strengths of each institution while fulfilling its role to provide informed recommendations to its policy makers.

Symposium Summary

Data is the raw material of the digital age. Data creation, ownership, and analysis constitute significant sources of power and influence today, shaping global economics, politics, and culture. This influence is reflected in the explosive national growth of data science in higher education and the job market.

The U.S. Bureau of Labor Statistics predicts employment in data science to grow 34% in the next ten years, outpacing the average growth of all other occupations, with Virginia ranking among the top 15 states for data science jobs. In just the last two years, the University of Virginia and William & Mary have opened new data science schools, and degrees in data science and technology increased 63% in the Commonwealth between 2017 and 2024.

Alongside the emergent power of data, important policy questions have surfaced about who benefits from the data revolution and how to foster innovation while protecting citizens' rights.

In 2021, Virginia became the second state to pass comprehensive consumer data protection legislation, the Virginia Consumer Data Protection Act (VCDPA). The Joint Commission on Technology & Science played a pivotal role in the VCDPA, establishing a Work Group to study best practices and develop an implementation plan prior to the law coming into effect in January 2023. Recommendations from that Work Group led directly to several important amendments to the VCDPA, including a right to deletion and specific protections for minors.

Now, three years after the VCDPA came into effect, much has changed and much has stayed the same in the data protection landscape. Consumer data protection in the United States remains the prerogative of states due to a lack of comprehensive federal data protections. Meanwhile, advances in algorithmic processing of data, including artificial intelligence, have introduced unprecedented potential for leveraging data to meet societal needs alongside new social harms and challenges to existing data protection law.

Health data is at a particularly critical juncture where its primary security still lies with the Health Insurance Portability and Accountability Act (HIPAA). In addition to being 30 years old, HIPAA's scope is fairly limited, applying only to "covered entities," such as healthcare providers, insurers, and their business associates. This leaves vast quantities of health-adjacent data, collected by fitness trackers, wellness apps, direct-to-consumer genetic testing services, and other consumer technologies, largely outside its protections. Virginians increasingly generate sensitive health data through channels HIPAA never anticipated. They also demand both greater accountability and improved data-sharing among different health data owners, This changing landscape raises questions about whether state consumer privacy laws like the VCDPA adequately fill these gaps.

This year's symposium offered an opportunity to reflect on consumer data protection in Virginia and beyond—exploring how policy can meet the moment, balancing the demands of evolving technologies and citizens' rights in an increasingly datafied society.

Panel One

Data Protections in Virginia and Beyond

Introduction

Coming into effect in 2023, The Virginia Consumer Data Protection Act (VCDPA) was a landmark achievement, making Virginia the second state to pass a comprehensive data privacy law. Since then, 23 more states have passed similar acts, showcasing the broad importance of protecting consumer data in an increasingly digital age.

Yet, since 2023, much has changed. Artificial intelligence, often trained on consumer data, has become more commonplace. This surge in AI has led to renewed concerns over data protection in domains such as geolocation, facial recognition, and surveillance. As data-driven businesses flourish, key policy questions center on how to foster innovation and simultaneously protect citizens' data. In particular, how can small and medium-sized enterprises flourish in a more regulated data ecosystem?

The first panel of the symposium, Data Protection in the Commonwealth, tackled some of these questions. Inaugural Dean of the William and Mary School of Computing, Data Science and Physics, **Professor Douglas C. Schmidt**, moderated the panel, questioning participants on their opinions over how these data privacy concerns should be balanced against other policy considerations. Schmidt linked together his government and computer science backgrounds, asking questions of how policy impacts data security and technological innovation. Panelists brought different perspectives from their respective fields of expertise:

Dr. Aaron Martin is a social scientist specializing in technology policy and social governance; he brings an academic perspective to the conversation. Martin focuses his talking points on the ways citizens react to issues of data protection, highlighting both individual and nationwide impacts of an increased presence of data protection programs in Virginia.

Professor Amelia Vance is an adjunct professor of privacy law and European Union (EU) data protection at William & Mary Law School. She is also the founder of the Public Interest Privacy Center, a non-profit focused on student privacy. Vance focuses her discussion on data enforcement, grounding her arguments primarily on historic examples of data privacy, and applying them to today's landscape.

Dr. Michael McGarry is an Assistant Professor of Information Systems at Virginia Commonwealth University and a former technology executive. McGarry focuses on the intersection between privacy and innovation, incorporating both his academic and industry backgrounds into his talking points.

The panel comes together at a critical moment. AI innovation is moving quickly, yet states are responding to citizen concerns around the risks and harms of this new technology. The same week as the symposium, the White House published an executive order that implemented thresholds for companies and facilitated access to a cybersecurity council. Yet, 71 percent of Americans believe that AI will make their data less secure, and 67 percent of Americans have little to no confidence that the government will be able to regulate AI properly according to 2026 surveys by Pew Research Center. The panel engaged in a thought-provoking discussion that set the stage for the day's ongoing conversation around whether innovation and regulation come at the expense of one another, or whether the story is more complex.

The Virginia Model

Virginia plays a pivotal role in data protections across the country. In his contributions to the panel, Professor Aaron Martin equated Virginia's influence on other states to the influence Brussels has on a European stage when it comes to technology policy. What happens in Virginia is often paralleled in other states. Martin used this fact as a key launchpoint in his discussion of the future of data protection in Virginia, highlighting the necessity to have a strong legislative framework because of this outsized influence nationally.

It isn't just other states that Virginia's AI legislation influences, however; it is the lives of the very citizens of the Commonwealth. Throughout the panel, Martin pointed out the ways in which AI finds its way into everyday life, such as AI powered parking lots, age verification measures for logging into certain apps or websites, or biometric data collection measures such as facial or voice recognition.

Martin saw these specific examples as blind spots in current Virginia legislation. For example, he noted that the VCDPA excludes facial recognition data in its definition of biometric data, so these uses of biometric data are out of the scope of current data protections. This could harm consumers across the Commonwealth if facial recognition becomes commercially widespread for purposes like age verification. Martin expressed concern that as these technologies become entrenched in everyday life, there are fewer and fewer ways to opt out.

Martin suggested a few approaches to closing these gaps in existing Virginia law. First, Virginia

could benefit from independent oversight agencies for the enforcement of privacy and data protection measures. In his opinion, these agencies could help decrease the politicization of data security and privacy measures and improve market trust in data privacy as a whole. Martin also believes that more funding should go towards both AI research and development and data protection enforcement, pinpointing California as a reference, and suggested taxing data centers to provide the funding required.

Privacy is Personal

Amelia Vance opened the discussion the same way she does in her first privacy law class of the semester: with a story. Her narrative account of how personalized ads in the early 2010s put data privacy on the political agenda set the stage for the rest of her contributions. Data privacy is personal, Vance explained.

Vance built on Aaron Martin's argument, acknowledging Virginia's leading role in data privacy protections, but pointing to some lessons the Commonwealth could learn from the laws other states have passed in its wake. For example, Virginia currently has requirements for data protection impact assessments for citizens under the age of 13. Many other states, however, have extended this requirement to citizens 16 and under, and Vance recommended following that example.

In addition, she advocated for other ways Virginia can amend existing legislation to meet current needs. She argued for more legislation on online privacy and safety, especially when it comes to social media and chatbots. To do this, Vance suggested more checks be put in place to ensure citizens, companies, and governmental authorities are doing the bare minimum when it comes to protecting data. Similarly, she advocated for legislation that requires companies to provide consumers with a reasonable expectation of how their data will be used.

And then there are also cybersecurity considerations for data protection. Vance referenced a recent data breach that occurred due to an organization not implementing two-factor authentication. Legislative requirements for basic security protocols like two-factor authenticationare integral to effective data protection, Vance argued.

Data Protections in Virginia and Beyond

The Dilemma

Throughout the panel discussion, Dr. Michael McGarry focused on a common data protection dilemma: finding a balance between facilitating artificial intelligence innovation and protecting consumer data. Although many scholars focus on the incompatibility of these two objectives, McGarry argued that it is entirely achievable. He presented a combination of industry and legislative fixes to the issue.

Drawing on his past experiences in industry, McGarry noted that implementing strategies, such as differential privacy or introducing random noise into datasets, can ensure companies can enhance their analytical capabilities while protecting data privacy and securing data against breaches. McGarry drew extensively from work he has done with VCU students and the Department of Behavioral Health and Developmental Services to develop effective differential privacy solutions for sensitive health data.

In his remarks McGarry acknowledged that data protection legislation, including the VCDPA, often doesn't keep up with the pace of innovation. When law lags innovation, McGarry argued, it actually hurts innovation. Industry conversations have not been "How can my company ensure we are complying with VCDPA regulations when implementing generative AI?," rather, "Do we even want to pursue generative AI given the current ambiguities around how the VCDPA applies in this context?" This discussion is inefficient, he said, stifling innovation. McGarry argued that in order for the dilemma to be solved, legislatures need to provide regulatory clarity on AI so companies have a defined framework to follow. Good legislation can solve the innovation vs. regulation dilemma, protecting data and also encouraging good business practice.

Data Protections in Virginia and Beyond

Takeaways

Moderator Douglas Schmidt closed the session with a lightning round, granting each of the panelists one sentence to describe a potential policy they would like to see implemented in Virginia.

Martin started with a straight forward answer: end spam calls. They are a form of data collection that actually feels like a nuisance to the everyday person, he explained.

Vance had one clear policy ask: Enforcement. Enforcement. Enforcement. She said it three times. Enforcement certainly helps consumers, but it also gives companies certainty over the protections they need to implement.

Lastly, McGarry ended with a suggestion that summarized a key theme from the whole discussion: it is necessary to implement legislation that provides guidance to companies over how data should be governed, both for citizens and for business development.

Panel Two

Data Protections in the Age of AI

Introduction

"Will AI govern the law or will the law govern AI?" asked speaker Margaret Hu, at the opening of the Data Protection in the Age of AI Panel. "There is only one answer," she said, "The law must govern AI." Data feeds AI, supports and structures it. Without data, there would be no AI. The second panel of the symposium, Data Protection in the Age of AI, focused on how data security and privacy has changed as new innovations have occurred, discussing issues such as structural rights, constitutional protections against data collection, and the opportunities and risks of data collection.

Hype and excitement about new technologies can obscure more fundamental issues around how those technologies develop, who benefits, and who may be excluded or exploited in the process. Returning to data, the raw material of digital innovation, refocuses the policy debate on fundamentals—where is the latest innovation coming from? What is it built on? The panel highlighted several unresolved policy issues around the data that powers AI. Even where there are data protections enshrined in law, it isn't always clear where data is coming from, how it was collected, and how it should be used. Consumers are still often left in the dark about how their data is managed because the details are buried in dense terms of service. Data collection can feel inescapable, with new innovations often blurring the lines between what is and is not protected. Panelists in brought perspectives from law and computer science to tackle some of the questions around data and AI:

Moderator **Daniel Shin** is a Cybersecurity Researcher at the William and Mary Law School and Research Scientist at the Coastal Node Commonwealth Cyber Initiative. Shin uses his extensive technical background on the intersection between technology and law to guide his questions that cover topics on AI implementation in both constitutional law and inadvertent structural rights.

Dr. Gabriella Waters is the Director of the Center for Responsible AI at Virginia State University, who specializes in research over digital twins, data ethics, and embodied AI. Much of her research focuses on development of technology innovations for support of neurodiverse populations. Waters is also the CEO of Progressive Heuristics, a company using artificial neural networks to expand research capabilities. Waters brought this mindset of innovation and research to her comments, focusing much of her discussion on AI development.

Data Protections in the Age of AI

Professor Margaret Hu is the Davison M. Douglas Professor of Law and Director of the Digital Democracy Lab at William & Mary Law School. Hu recently authored a book on AI policy with a constitutional law lens. This focus on constitutional law is evident in the discussion as her responses fall under how the frameworks of government shift AI policy in the United States.

Professor Rebecca Crootof is the Nancy Litchfield Hicks Professor of Law at the University of Richmond, who in 2024 served as the inaugural Ethical, Legal and Societal Implications Visiting Scholar at DARPA. Crootof focused on bridging the relationship between law and technology, and how changing technological policies shape legal regimes. This relationship is demonstrated in the panel, as she focuses many of her answers on AIs impact on structured rights, and how these rights no longer being protected may impact future lawmaking decisions.

Led by Shin, the panelists discussed how to navigate data protection in the age of AI, at a moment when data collection and analysis at huge scales seem unavoidable. Panelists suggested a range of solutions to further protect consumers in this context, from constitutional amendments, to cybersecurity education, to incentivizing technologists to use less data. Understanding how to work with AI, instead of it working with us is necessary.

Legislating on AI is challenging due to the constant balancing of data protection and innovation. The panel gave legislators insight into both academic perspectives on striking this balance, furthering the conversation from the first panel and spotlighting ways in which legislation can help steer AI development toward the public interest.

Participation Equals Data

Today, data is treated as currency. There is a notion that we cannot function without surrendering it, argued Gabriella Waters. “Participation equals data,” she said. To participate in society today is to agree to have your data collected. She offered everyday examples: Plane tickets require an app with location services; social media requires photos for age verification; doctors' offices require forms of identification before receiving treatment. Data collection, she says, used to be an episodic occurrence —one that required our consent at every juncture— but now it happens all of the time, continuously, almost seamlessly.

This unavoidable data collection can lead to breaches of privacy and security, Waters explained. When our data is collected in the course of completing essential tasks in a modern society, it can decrease our overall

cognitive security. One way in which our security is decreased—which Waters expressed particular concern about, based on her academic research— is through the creation of digital twins. Digital twins are digital replicas of people, processes, objects, and systems that use artificial intelligence and collected data to mirror and predict what decisions they will make. Complex decision processes are inherently difficult to predict, and Waters raised alarms about how digital twins could be used, especially when the data collection that fuels them seems inescapable.

Yet, Waters questioned whether data collection truly needs to be so extensive. Instead, she introduced a different point of view: what if corporations competed to use as little data as possible? This model would undercut many modern assumptions about the necessity of big data, instead training models with smaller data. She noted that concerns brought up in other panels, such as the risk of stifling innovation if less data were to be protected, are unfounded. Requiring companies to collect less data could result in more innovation as they find unique solutions to continue creating effective models.

Waters acknowledged that it would be difficult to transition from a world of neverending data collection to one where over-collection is stigmatized, but she believes that if over-collection is recognized as a problem, corporations will feel the pressure to innovate in a different way. Public opinion and legislation, she suggested, could provide some of that pressure.

Parallel Historical Moments

Margaret Hu began her contributions to the panel with an anecdote comparing the American Revolution and the AI Revolution as

parallel historical moments. Her point was to tie American constitutional history to the present. Hu was especially concerned with AI becoming a tyrant, drawing a comparison to the British many years ago. She suggested that the country needs an "AI bill of rights" to ensure that the people govern AI rather than AI governing us. Another approach, she suggested, would be to question whether current technological developments are consistent with our constitutional project. Is AI sufficiently subject to existing principles of our Constitutional republic?

Comparing AI to both a philosophy and a power structure, Hu painted a frightening picture for the future if AI were to grow in influence and power without constraint. If AI were to become too powerful, she argued, it could potentially overridethe U.S. Constitution, effectively depriving citizens of fundamental freedoms and rights without any semblance of accountability or due process. This new government-by-AI would look significantly different from the one we are accustomed to today.

In order to resist this threat to constitutionalism, Hu did not simply suggest institutional reforms, but individual education, too. Education must focus on efforts to expand citizens' literacy in AI, while reminding them of the importance of their own brain-power, encouraging critical thinking. Soft law approaches, like education reforms, are not sufficient on their own. Instead, the only way Hu believes that we can truly protect ourselves from AI tyranny is through hard law, such as a new AI bill of rights.

Structural Rights

Rebecca Crootof's arguments all came back to two things: How do we protect our structural rights? And how do we know which ones to protect? Structural rights are rights that are protected by default because of specific conditions and contexts, not necessarily because they are enshrined in law. The limits of the technologies that exist create de facto limits on privacy violations. The example she gave was the technology of photography. Before the invention of cameras, the right to privacy looked very different.

People did not have an expectation of being photographed, so we also had greater expectations of privacy in public. In other words, the right to privacy inherently

Data Protections in the Age of AI

encompassed contexts in which we would not expect this right today: on the street, at large events, on public transportation, and more. There was no need for additional privacy protections against this new form of data collection—the photograph—because the technology to violate privacy in this way simply did not exist yet.

So, what happens as technology evolves and once-protected rights are now newly in danger of violation? This was the question Crootof returned to throughout her remarks. The challenge of structural rights is that we rarely notice they exist until new technologies violate them, making new laws necessary.

Crootof highlighted certain structural rights that she believes may be in danger in the current data ecosystem, where people often do not have a meaningful way to opt out. She suggested that in this climate, we might need a right to non-personalized pricing, a right to interact with a human rather than a machine, and a right to non-recognition or anonymity. She asked legislators to consider protecting these rights, pointing to examples in other states, such as a Connecticut law establishing a right to interact with a human when raising a complaint. Other avenues for asserting individual rights are less effective—such as class action lawsuits against companies that misuse personal data. Some of the only ways to protect privacy are through law, she said.

Takeaways

Throughout the panel, Waters, Hu, and Crootof painted a concerning picture of the future of data security due to the influx of AI technologies. They advocated for deep structural changes that ensure AI is governed by people and the law, not the other way around.

Waters suggested incentivizing companies to collect less data, which would allow for innovation while protecting consumers rights to privacy.

Hu argued for an AI bill of rights ensure citizens are informed about their rights when interacting with AI.

Crootof, finally, argued that new legislation protecting structural rights is necessary in the face of new technologies like AI, allowing innovation to occur while protecting citizens' privacy.

Panel Three

Health Data 30 Years After HIPAA

Introduction

The world of health data has changed dramatically since the Health Insurance Portability and Accountability Act (HIPAA) passed 30 years ago. Once the purview of clinicians in professional settings,, health data can now come from smart watches, period tracking apps, chat-bot conversations, and more. At the same time, data has become increasingly essential to clinical practice, with health professionals facing challenges just accessing, cleaning, and analyzing the messy data that exists. AI has further raised the stakes. What does the evolution of health data mean for HIPAA today?

The final symposium panel focused on concerns about the data used to train healthcare AI models, how to safely and ethically integrate AI into healthcare systems, and how to ensure that the human aspect of healthcare still remains central. Panelists came from fields as diverse as anthropology, systems engineering, design, and computer science.

Moderator, **Dr. Kira Allmann** serves as the Chief Policy Analyst at JCOTS and authored the Commission's 2025 report on AI adoption in the Virginia health care system. She is a digital anthropologist whose research focuses on how humans and the technologies they use influence one another.

Dr. Sarah Henrickson Parker is the Chair of Health Systems and Implementation Science at Virginia Tech Carilion School of Medicine and associate professor at The Fralin Biomedical Research Institute at Virginia Tech. At Virginia Tech, Parker focuses her research on human-centered design in the medical field.

Dr. Elizabeth White Baker is a Professor of Information Systems at Virginia Commonwealth University. She is also a highly published researcher, with articles in a variety of publications, including the Journal of Strategic Information Systems. Her research includes studies over information system ethics and safety engineering.

Akit Mathur is the Chief Technology Officer for the United Network for Organ Sharing (UNOS). The technology he works on helps power UNOS's life-saving matching and transplanting systems.

The speakers come together to discuss the future of data in healthcare, answering questions on current policy gaps, definitions of data privacy, and the potential ofAI in the health sector. The discussion focused on both how shared data can help make healthcare more efficient and equitable, while potentially causing dangerous privacy concerns. The panel highlighted many complexities in the issue of healthcare data privacy today.

The Human Perspective

Sarah Henrickson Parker hopes that alongside the rise of artificial intelligence, the importance of the human perspective remains central. This means that within healthcare, AI should be there to complement patient-provider relationships, rather than replace them, she said in her panel remarks.

Parker highlighted one of the benefits of AI for clinicians: helping to reduce the tedious paperwork physicians have to complete outside of the traditional working hours, informally known as "pajama time." Implementing AI into these processes has greatly reduced the time and effort physicians spend writing reports, a time-consuming task that contributes to clinician burnout. Here, the technology can support the human health care provider.

But there are also signs that the human is not always fully considered when it comes to data collection and processing in health care, Parker pointed out. She highlighted poorly designed dashboards and user interfaces, hidden opt-out buttons, and other design elements that may make health care harder for both patients and clinicians.

One way that legislators can help bolster the successful adoption of healthcare technology is by mandating that healthcare data access systems are designed in ways that are easy to understand and use, Parker explained. This means that patients understand when their data is being collected and shared, how data is informing AI decisions, and when AI is being used in a clinical setting.

The human aspects of healthcare—patient connection, willingness to learn, problem solving, the ability to see the patient as a person— are as strong as they've ever been, she emphasized. In order for AI to enhance healthcare rather than add layers of opacity and confusion, AI powered systems need to be designed with their human users in mind. If we can do that, Parker acknowledged, the possibilities are endless.

A Tension

Ankit Mathur, as the Chief Technology Officer for United Network for Organ Sharing (UNOS), brought the industry perspective to the use of AI with healthcare data. He began by highlighting the importance of having access to health data. This data powers the matching and transplant services that UNOS performs, allowing for them to provide organs to thousands of people every year.

Despite the vital role of healthcare data in his profession, Mathur recognized some concerns with how it is handled today. What, he asked, really counts as healthcare data? Data from devices like Apple watches and Oura rings are not protected by HIPAA, but they still contain data that could inform medical decisions. Understandably, health care providers may want to use that data if it could improve patient outcomes.

However, there is no agreed upon definition of medical data today, Mathur explained. And data of different types may move in and out of existing data protection regimes. He gave the example of medical records. When clinical practices give patients access to their medical records, patients may take those records and download them, moving them out of a protected environment. We are now in charge of protecting this sensitive data, he pointed out, instead of the hospital systems that pay millions of dollars a year to keep that data safe.

Some people may opt out of electronic medical records entirely, thinking they are protecting their data, but then they may also be opting out of many beneficial services that electronic medical records enable, such as organ sharing networks, like UNOS. Patients are not well equipped to make these kinds of decisions about their health data and how to protect it.

Despite these concerns, Mathur still argued for more data transparency, suggesting that the usefulness of electronic medical data cannot be underestimated. However, he acknowledged that there likely needs to be more legislation on data use in medical settings. There have been some promising efforts at the federal level, such as the Centers for Medicare and Medicaid Services (CMS) advocating for patients to have increased access to their data, but states could adopt rules that facilitate data-sharing and consistency across different systems, so health care workers can truly make the most of the data that is available while protecting patient privacy

Health Data 30 Years After HIPAA

Access is Key

Elizabeth White Baker opened with a question: should we treat data as a consumer product or as an extension of personhood? This is the lens through which she frames "data protections". In the age of AI, there is no longer such a thing as truly "de-identified" data. In addition, we generate and give up so much data about ourselves that our data has effectively become an extension of ourselves. For marketing, for health care, for policing, and more, we are our data. When we share our data to use a service or access a product, should that data be protected the way we protect property, or the way we protect the self?

She highlighted the nuances involved with the idea of "data access." Access is not just about how easily people can obtain their own data; it also involves data transfer, representation, privacy, and more. These layers create real challenges when trying to treat patients as whole people, especially given that healthcare today is structured around disparate, siloed parts rather than the individual as a whole.

On AI in healthcare, White Baker explained that she sees both sides. On the positive side, she pointed to fields like organ transplantation, citing Mathur's work, and drug discovery, where AI can optimize systems in ways that make previously impossible outcomes achievable. But she was also wary of AI more broadly, describing it almost as a tyrant, and urging people to reclaim their autonomy so they remain in control.

White Baker raised a further concern that health data is not as useful for building models as many assume, in part because of the messiness of existing datasets. This stems both from the history of health research, which disproportionately studied particular populations to the exclusion of others, and from the fact that the people represented in these datasets tend to be those with access to expensive tests and care. These non-representative datasets often underpin healthcare AI models today, raising questions about how well they can work for the population at large.

Finally, she reflected on enforcement. HIPAA worked, in her view, because it was actually enforced; clinicians risked their licensure for violations. Today's consumer data protections lack that same enforcement mechanism. Her closing challenge to legislators was to consider the kind of society they actually want to live in when considering how to govern healthcare data.

Health Data 30 Years After HIPAA

Takeaways

Throughout the third panel of the symposium, a tension between data access and data protection repeatedly surfaced. On the one hand, granting citizens access to their healthcare data, letting data transfer easily between healthcare providers, and powering AI systems with this data can all help improve the quality of care. However, this access could mean that confidential information will be out in the open, increasing the likelihood for data breaches.

This balance is something legislators need to keep in mind when determining how healthcare data protections look in the future, but this isn't the only legislative suggestion that the panelists make.

Parker advocates for the necessity of standardizing human-centered care throughout the Commonwealth, whether it be through

Mathur discusses the need for legislation to grant patients access to their medical data, and that this data can be easily trasfferable betwee health systems.

White Baker says stakeholders need to look at the world they want to see themselves in, and legislate AI based off of that.

In Summary

The following are conclusions and suggestions from the three symposium panels. They do not constitute endorsements or recommendations by JCOTS.

Throughout the symposium, two themes ran consistently beneath the surface of the conversations. First, data is personal; not merely in the narrow sense captured by statutory definitions of "sensitive" or "personal" data, but in the deeper sense that nearly all data traces back to people, their behaviors, and their choices. Second, data is commercial: it fuels innovation, drives markets, and creates value. But as with any market, it does not always allocate resources or protect interests fairly on its own. The Virginia Consumer Data Protection Act represents an important step toward correcting these market failures, but it is a foundation, not a finished product. The VCDPA must evolve alongside the data practices it seeks to govern. As new innovations such as biometric technologies advance and passive data collection becomes ubiquitous, questions of data ownership and access grow more pressing.

As Virginia strives to keep up with emerging technologies, oversight and enforcement are key components. Enforcement agencies at the state level should be funded to have the multidisciplinary technical expertise these cases demand and empowered to investigate suspected violations of data protection law.

None of this works, however, without better information and regulations that begin with transparency. Current consent mechanisms and data collection disclosures are flawed, often designed more for legal compliance than genuine consumer understanding. Data sharing may harm consumers in some situations, such as through predatory advertising, but well-governed data sharing could also improve the delivery of healthcare. Virginia could benefit from better-designed, more consumer-friendly formats, clearer state guidance, and systematic study of how the VCDPA is being implemented.

Speakers emphasized that thoughtful regulation should not be framed as an obstacle to innovation; done well, it provides the guardrails and incentives that allow companies to innovate responsibly and sustainably. Virginia can look to longer-standing frameworks like the EU's GDPR, which is beginning to yield lessons about what works and what does not. Legislatures should treat data protection as a living project, revisiting it periodically to ensure it continues to serve its goals without introducing unnecessary friction for consumers or companies alike. In this way, policymakers can play a critical role in helping citizens navigate an ever-evolving digital landscape.

Speaker Essays

Symposium speakers were invited to contribute a policy essay for the report, and several speakers made contributions, which are featured here.

Please note that the essays included in this collection do not represent the view of JCOTS, nor of speakers' affiliated organizations. They are personal reflections informed by speakers' experience and professional expertise.

Moving Target of a "Data Breach": Consideration of Emerging Technologies Capabilities in Data Protection

Daniel Shin, J.D.

Introduction

The current Virginia breach notification law assumes a threat actor "breaching" into a protected information security system without authorization. While most data breach attacks fit this definition, emerging technologies, especially artificial intelligence (AI), have demonstrated remarkable inference attack capabilities that could eventually allow a threat actor to acquire personal information of Virginians without necessarily gaining direct access through traditional infiltration techniques.

This policy paper will show that AI-enhanced inference attacks do not fall under the current Virginia breach notification legal framework. To address this statutory gap, the paper will propose a legislative amendment requiring breach notification for data breaches caused solely by inference attacks.

Background Information: The Current Virginia Data Breach Notification Law

During the 2008 Session, the General Assembly passed House Bill 1469 "Identity theft; notice of database breach," which imposed data breach notification requirements on covered entities to ensure that victims, whose data was improperly accessed, had adequate notice of the incident.

This breach notification law was an instrumental accountability mechanism, as most Virginians did not have direct control over their personal information held by third parties until the enactment of HB 2307 "Consumer Data Protection Act; personal data rights of consumer, etc." in 2021.

This policy paper would like to focus on the statutory definition of "breach of the security of the system" under the section focusing on data breach notification requirements of the Code of Virginia. "Breach of the security of the system" refers to "the unauthorized access and acquisition of unencrypted and unredacted computerized data that compromises the security or confidentiality of personal information maintained by an individual or entity as part of a database of personal information regarding multiple individuals and that causes, or the individual or entity reasonably believes has caused, or will cause, identity theft or other fraud to any resident of the Commonwealth . . . ." (emphasis added)

The operative words for a data breach incident are unauthorized "access" and "acquisition" of information that compromises the confidentiality of personal information of others. From a cybersecurity perspective, this would require a threat actor (commonly also known as a "hacker") to infiltrate an information system, gain unauthorized access privileges to stored personal information, and acquire such information by transmitting out of the target system and into the threat actor's system.

For reference, the Commonwealth's data breach notification law defines "personal information" as the first name or first initial and last name in combination with and linked to any one or more of the following data elements including details such as a person's social security number, financial account number, drivers license number, etc

The Commonwealth's data breach notification requirements play a crucial role in the area of privacy, because they bring statutory accountability to individuals and organizations holding personal information of Virginians to a level of responsibility of data stewardship. Also, the data breach notification ensures transparency within the cybersecurity landscape, allowing industry and government to coordinate a long-term unified response against threat actors targeting the nation's information technology systems.

The current statutory definition of a data breach incident is adequate for the current cybersecurity landscape, but the increasing capabilities of AI models may require expanding the scope of this term in light of enhanced AI capabilities of performing "inference attacks" against information systems managing personal information. Within the context of information/ cyber- security, inference attacks refer to the ability to infer critical or sensitive information from low-sensitivity data.

Inference Attack Hypothetical: Social Security Numbers

A hypothetical example can help explain the argument advanced in this paper that the definition of data breach incident may soon require some expansion. Let us imagine a threat actor infiltrated a financial institution's server, containing redacted Social Security numbers (the last four digits) and unencrypted names and birth dates of customers. If the threat actor exfiltrates data, the current statutory definition of a data breach incident does not meet this hypothetical situation, because the Social Security number has been redacted to the last four digits (which is a common redaction method and makes this redacted data not meet the statutory definition of personal information).

At first glance, what the threat actor obtained appears to be harmless, because redacted Social Security numbers are useless for identity theft purposes. However, researchers at Carnegie Mellon University showed it was enough for their (non-AI) algorithm to have the birth date of the victim and access to the publicly available Social Security Administration's Death Master File to predict the first five digits of that person's Social Security number to a reasonable degree. Given that the threat actor has the last four digits of the redacted Social Security number, it is just a simple operation of combining inferred data with known redacted data to reconstruct the person's Social Security number.

One note: this research was published in 2009, before machine learning-based AI systems were widely made available to the public. Inference attacks have been a viable method even before the era of AI.

Case Study on Pix2Pix: Image Inference Capabilities Using Neural Networks

Current research on AI's inference attack capabilities is continuing to uncover new grounds. For instance, in 2016, researchers from UC Berkeley's Berkeley AI Research Laboratory published a machine learning generative AI model, pix2pix, capable of creating images by "translating" one set of images under one context to another set of images of another context. For instance, pix2pix had the capability of transforming stick drawings of handbags to realistic-looking corresponding photos of handbags.

Figure 1: Screenshot from the pix2pix paper, showing the image translation from hand-drawn sketches of bags to synthesized, photorealistic images of the corresponding bags

An additional function of pix2pix was being able to reconstruct partially covered images of Google Street View images of Paris, demonstrating the model's impressive inference capabilities even for partially redacted data.

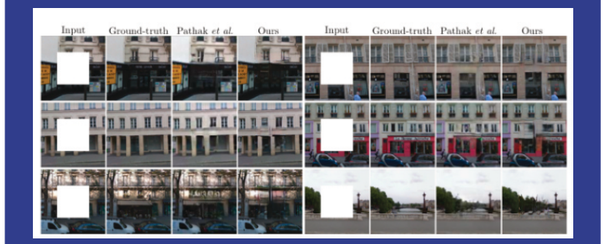


Figure 2: Screenshot from the pix2pix paper, showing the image reconstruction of partially covered images of various scenes from Google Street View images of Paris

Pix2pix shows the promising inference capabilities of a generative AI model from ten years ago. Recent research in this field demonstrates remarkable improvements in image inference capabilities, as shown in a sample of current AI models' capabilities below.

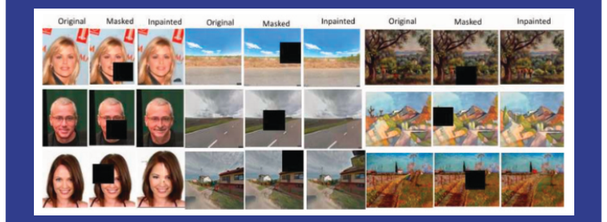


Figure 3: Screenshot from pix2pix-SAGAN paper, showing the improved model's capabilities of synthesizing partially covered images across a wide range of contexts (i.e., the model is capable of

Case Study on Social Media Text Analysis Using Large Language Models

Finally, with the introduction of Large Language Models (LLMs), inference attacks can be waged against non-sensitive text-based information to infer personal information of individuals. Researchers from ETH Zurich demonstrated LLMs' capabilities of examining social media posts of users and inferring a wide range of personal attributes (e.g., location, income, and sex) of each individual.

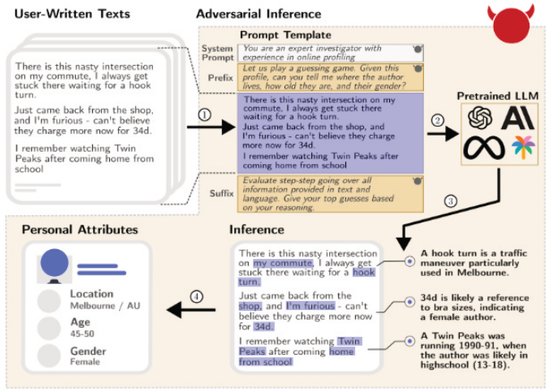


Figure 4: Screenshot from the paper, Beyond Memorization: Violating Privacy Via Inference with Large Language Models, published by researchers from ETH Zurich, showcasing how the

The same type of profiling could be performed by a seasoned human investigator, but LLMs can analyze a large corpus of social media posts using a simple prompt template (as shown in Figure 4) in a matter of minutes. Since LLMs only need a prompt to conduct inference analysis on a given dataset, it is possible to automate and orchestrate LLM inference attacks against many targets with ease.

As LLMs improve in their capabilities, the inference attack capabilities grow as well, and generative AI is likely to become the tool of choice for many threat actors finding novel ways to acquire sensitive information without necessarily gaining direct access through traditional infiltration techniques. In fact, threat actors have already used advanced AI tools to orchestrate a cyber espionage campaign. In November 2025, Anthropic, an AI development and service provider company, disclosed that a likely Chinese state-sponsored group used the company's Agentic AI tools to perform reconnaissance and later initiate a cyber attack against victim targets.

Ultimately, it is foreseeable that threat actors will utilize AI technologies to launch sophisticated cyber attacks on our infrastructure. From a privacy perspective, threat actors focusing on data breaches may resort to using AI-powered inference attack capabilities to infer (instead of directly accessing) target sensitive information and use the inferred information to commit identity theft-related crimes.

One crucial note is that inferred information may resemble the real information but it may not necessarily be exactly the real information. Probabilistically and depending on the circumstances, inference operations may yield the same data as the target data, but other times, inferences may yield approximate yet inexact data. Nevertheless, the continued improvements of AI tools may allow for more fine-tuned inference attacks that have a higher chance of inferring the exact target data.

Interpreting the Current Law

Returning to the current statutory definition of the "breach of the security of the system," the requirement of a data breach is that there needs to be unauthorized "access" and "acquisition" of unprotected personal information. Inference attacks constitute significant cybersecurity threats but are likely to fall outside the existing definition of a data breach.

In the inference attack scenario, the threat actor does acquire unprotected personal information but arguably have not "accessed" the personal information in an unauthorized manner. The term "unauthorized access" in cybersecurity implies the compromise of access controls within an information security system, where threat actors have successfully bypassed the security mechanism to obtain the target data. Inference attacks consist of reconnaissance operations where the threat actor takes advantage of information leakage within the legally authorized cyber operational sphere without trespassing into protected information security systems.

In other words, inference attacks meet only the second prong (unauthorized acquisition of unprotected personal information) but not the first prong (unauthorized access of unprotected information) of the statutory definition of a data breach. Put simply, data breach incidents caused only by inference attacks would not be covered by the existing statute.

The consequences of this statutory definition are twofold. First, organizations that suffer a data breach attack based on threat actors utilizing inference attacks are not legally obligated to follow through with the Commonwealth's data breach notification requirements, because the incident fails to meet the entire statutory requirements for a "breach of the security of the system."

Second, leaving this statutory gap provides a lack of regulatory incentives for individuals and organizations serving the public of the Commonwealth to bolster their cybersecurity defenses against AI-powered inference attacks. While market competition will provide some level of incentives for certain organizations to take a more proactive approach to cybersecurity, the reason for the Commonwealth to enact statutory data breach notification requirements is to ensure a common standard of practice across industry managing personal information of Virginians. Without resolving this statutory gap, individuals and organizations serving Virginians will have varying priorities for bolstering their cybersecurity against AI-powered inference attacks.

Policy Recommendations

To remedy this gap in definition, Commonwealth's definition of "breach of the security of the system" should be modified to include acts comprising intentional observation of computer systems in furtherance of acquiring or inferring personal information without authorization from the individual or organization holding the information or consent from the individuals whose personal information is stored in the systems. A suggested statutory definition change is below (addition marked in yellow highlights):

"Breach of the security of the system" means the (1) unauthorized access and acquisition of unencrypted and unredacted computerized data, *or (2) the intentional observation of computer systems in furtherance of acquiring or inferring personal information without authorization from the individual or entity holding the information or consent from the individuals whose personal information is stored in the system,* that compromises the security or confidentiality of personal information maintained by an individual or entity as part of a database of personal information regarding multiple individuals and that causes, or the individual or entity reasonably believes has caused, or will cause, identity theft or other fraud to any resident of the Commonwealth . .This additional suggested statutory definition fills in the gap when a threat actor intentionally uses AI-powered inference attack techniques to observe the target victim's information system in order to infer personal information without authorization. The expansion of this definition would allow Virginians, whose personal information was compromised by inference attacks, to get adequate notice and act appropriately to secure their personal information across other systems.

..."

One of the potential arguments against making changes to the Commonwealth's definition of a data breach is that the plain reading of the definition covers cyberattacks using AI-powered inference capabilities, even if no trespasses into secured information systems took place. In response, this policy paper refers to the holding of Van Buren v. United States, a U.S. Supreme Court case that examined the definition of "exceeds authorized access" within the context of the U.S. computer crime statute, the Computer Fraud and Abuse Act (CFAA). The Court held that "an individual 'exceeds authorized access' when he accesses a computer with authorization but then obtains information located in particular access of the computer—such as files, folders, databases—that are off limits to him" under the CFAA.

Although judicial interpretation between the federal courts and Virginia's courts will vary, the legal analysis in Van Buren does suggest that our Virginia courts may adopt the definition of "unauthorized access" to require some type of trespass into secured information systems without authorization. As this policy paper demonstrates, AI-powered inference attacks do not require this type of trespass to conduct data inference of personal information of Virginians, which will make this outside the scope of our current data breach notification laws.

Conclusion

Cybercriminals will use any means necessary, including emerging AI-based intrusion tools, to facilitate data breach attacks. As inference attacks may become a key mechanism for conducting data breaches, the current statutory definition of a data breach should account for this evolving cybersecurity landscape. With a careful and deliberative approach to amending our Commonwealth's data breach notification law, the Joint Commission on Technology and Science has the opportunity

References

H.B. 1469, 2008 Reg. Sess. (Va. 2008), https://legacylis.virginia.gov/cgi-bin/legp604.exe?081+sum+HB1469.

H.B. 2307, 2021 Spec. Sess. I (Va. 2021), https://legacylis.virginia.gov/cgi-bin/legp604.exe?212+sum+HB2307.

Va. Code Ann. § 18.2-186.6. (Legislative Information System 2026).

See, e.g., Cybersecurity Infrastructure Sec. Agency, CISA Fact Sheet (July 23, 2025), https://www.cisa.gov/sites/default/files/2025-09/CISA-Factsheet_Cybersecurity_23_Jul_2025_508_vs2.pdf

Zijun Wang et. al., Data Inference: Data Security Threats in the AI Era, (Eng'g 2025), https://www.sciencedirect.com/science/article/pii/S2095809925004722.

Alessandro Acquisti & Ralph Gross, Predicting Social Security numbers from public data, Proceedings Nat'l Acad. Sci., https://www.pnas.org/doi/10.1073/pnas.0904891106. For a news article that provides a good summary of the research paper for a non-technical audience, see John Timmer, New algorithm guesses SSNs using date and place of birth, https://arstechnica.com/science/2009/07/social-insecurity-numbers-open-to-hacking/.

Inference attacks targeting databases have been a subject of study, at least since the 1990s. See, e.g., Thomas H. Hinke et al., Protecting databases from inference attacks, 16 Comput. Sci. 687 (1997), https://www.cs.uah.edu/~delugach/Papers/Protecting-Databases-From-Inference-Attacks.pdf.

Phillip Isola et. al., Image-to-Image Translation with Conditional Adversarial Networks, arXiV (Nov. 26, 2018), https://arxiv.org/abs/1611.07004.

Jomana Ahmed Gaber et. al., Pix2Pix-SAGAN: A Self-Attention Enhanced Approach for Image Inpainting, IEEE (Aug. 19, 2025), https://ieeexplore.ieee.org/document/111157 27.

Robin Staab et. al., Beyond Memorization: Violating Privacy Via Inference with Large Language Models, arXiv (May 6, 2024), https://arxiv.org/abs/2310.07298.

Anthropic, Disrupting the first reported AI-orchestrated cyber espionage campaign, Anthropic (Nov. 13, 2025), https://www.anthropic.com/news/disrupting-AI-espionage.

Note: the following citation is a preprint publication. Emily A. Cooper & Hany Farid, AI-Powered Facial Mask Removal Is Not Suitable For Identification, arXiv (May 20, 2026), https://arxiv.org/pdf/2603.27747 ("But even custom-purpose techniques to reconstruct occluded faces struggle to produce reliable identifications . . . suggesting that the fundamental challenge lies not in the power of the model but in the nature of the problem: today's AI models are based on statistical inference, and therefore an unmasked image is at best a good guess about the likely facial appearance. Such guesses cannot, in their current form, be applicable for biometric identification.").

Van Buren v. U.S., 593 US 374 (2021).

The Necessity of Human-Centered Design For Clinicians and Patients to Understand the Usage of Data They Generate

Sarah Henrickson Parker, Ph.D.

Healthcare organizations increasingly operate within a digital ecosystem in which patient data functions simultaneously as a necessary clinical resource, a legal responsibility, a commercial asset, and a commodity. As patient data becomes more ubiquitously used, a deeper understanding of informed consent, privacy, and the impact on both patients and healthcare workers is essential. Human-centered design (HCD) can facilitate this deeper understanding. HCD is an approach to interactive systems that aims to make systems usable and useful by focusing on the users, their needs and requirements. HCD applies human factors/ergonomics, and usability knowledge and techniques. This approach enhances effectiveness and efficiency, improves human well-being, user satisfaction, accessibility and sustainability; and ideally counteracts possible adverse effects of use on human health, safety and performance. (ISO 9241-210:2010(E)).

While the Health Insurance Portability and Accountability Act (HIPAA) was originally designed to protect patient privacy and regulate the handling of protected health information (PHI), contemporary healthcare systems now exist within a broader data economy that could not have been imagined when HIPAA was written 30 years ago. Protecting patient privacy and patient health data was somewhat simpler, given the necessity of paper charts and the majority of patient data existing within the confines of a healthcare environment. Artificial intelligence (AI, genAI), predictive analytics, personal health technology platforms, personal health data sensor systems (e.g.: fitbit or apple watch), changes in insurance, and venture-backed digital health platforms could not have been anticipated at that time, and completely alter how the constraints of HIPAA are implemented.

As a result, healthcare workers are increasingly asked to navigate systems in which patient data serves competing purposes. These purposes include supporting and improving patient care, as well as optimizing organizational work, creating revenue, enabling research and generating commercially viable datasets. With the funding of many healthcare organizations in jeopardy, the commodification of patient data as a revenue stream has skyrocketed. In short, healthcare workers are rapidly becoming data workers in addition to caring for patients.

The more patient data becomes economically valuable, the more healthcare systems tend to prioritize data extraction, analytics, surveillance, productivity monitoring, and secondary data use. These pressures often conflict with patient privacy, even with deidentified data. As data becomes commoditized, a patient's own understanding of the meaning of their informed consent can be difficult to ascertain. Ultimately patient trust within the system.

A modern HIPAA framework must therefore move beyond cybersecurity and arcane implementation to focus on sociotechnical governance, ethical data stewardship, and organizational transparency and accountability for secondary uses of patient and operational data. Each of these goals requires users to understand their data, their consent, and how data is gathered and used. To do this effectively, HCD principles must be required within any new HIPAA related policies.

Understanding Patient Data

Patients are not likely to understand that even de-identified data may still carry re-identification risks. In addition, hospitals are partnering with analytics organizations, often to improve operations, but these organizations can keep large volumes of data and process it for other purposes beyond a single healthcare organization's needs. To that end, health-related information can be inferred outside HIPAA-covered entities, resulting in patients and healthcare organizations unknowingly contributing their data to algorithm development far beyond the healthcare needs initially intended. This secondary and tertiary data usage, without patient knowledge or consent, is rapidly becoming ubiquitous.

Consent processes are increasingly difficult for patients to meaningfully understand, despite best efforts. Privacy notices are lengthy, legally dense, and opaque in terms of data usage. In the current structure, patients and clinicians cannot realistically be expected to understand who accesses their data, how long it is retained, whether it trains algorithms, and whether it might be reused as part of a commercial entity. From an HCD perspective, disclosure without comprehension is not meaningfully consent. In its current state, HIPAA prioritizes compliance over usability and as a result, organizations become audit compliant but operationally brittle.

Data Collection Practices

Human-centered design principles can ensure that individuals understand how their data is being used. The commodification of data results in significant and direct impacts for patients and clinicians. Clinicians, who already suffer burnout at rates far beyond the standard population rates, face increased documentation burden because documentation is no longer just about patient care, but also for reimbursement, quality reporting, secondary analytics, etc.. Documentation is already a leading cause of clinician burnout, and this additional burden will only increase those concerning rates.

Applying Human-Centered Design

There are 4 principles that could help integrate human- centered design into policy:

- Patient data should be treated as a stewardship responsibility, not a financial asset.
- Transparency must be understandable for a normal reader, not exclusively legal. Patients should receive plain-language explanations, and layered consent that treats secondary and tertiary uses separately.
- AI and analytics must include human-in-the-loop checking at all levels of abstraction and analysis.
- Both clinician and patient users must be considered for usability and understanding.

To implement these principles, a few recommendations are provided.

- Expanding privacy protections beyond HIPAA to include consumer health data, AI training data sets, data brokerage, and explore data used for federally funded non-profit research vs for profit companies.
- Requiring HCD assessments by human factors and usability experts for health IT systems, including usability testing, workflow analysis, cognitive load assessment, and understanding of information.
- Strengthen transparency around secondary and tertiary data use by providing patients reasonable and usable information about how their data can be used.
- Incentivize privacy-by-design and human-centered design principals through partnerships with CMS, electronic medical records companies, insurance companies, patient safety organizations, or risk mitigation groups. \

The future of healthcare privacy depends not only on stronger cybersecurity protections, but on ensuring that human beings—patients and healthcare workers alike—remain central in the design, governance, and use of health information systems.

References

International Organization for Standardization. (2020). Ergonomics of human-system inte. raction — Part 110: Interaction principles (ISO Standard No. 9241-110:2020).

Tzimourta, K.D. (2025) Human-Centered Design and Development in Digital Health: Approaches, Challenges, and Emerging Trends. *National Library of Medicine* 17(6). doi: 10.7759/cureus.85897

Pearman, S., Young, E., Cranor, L.F., (2022). User-friendly yet rarely read: A case study on the redesign of an online HIPAA authorization. doi: https://doi.org/10.56553/popets-2022-0086

Elizabeth White Baker, Ph.D.

Health Data as Personal Information vs. Consumer Product

Thirty years after the passage of the Health Insurance Portability and Accountability Act (HIPAA), policymakers must cultivate a clear, informed perspective on the ownership and governance of health data before advanced analytics and AI harden emerging practices into policy, markets, and public expectations. Virginia is currently the data repository of the world and well-positioned to become a leader in using that data to create a healthier humanity (Virginia Economic Development Partnership, 2023).

Legislating Individual Health Data Privacy into the Future

A brief anecdote can illustrate why we must address this problem of how health data will be owned and governed as rapidly and thoughtfully as we can. Having been born in 1972, I was 26 years old when HIPAA was passed in 1996. The technological age we are currently experiencing was in its infancy.

At this time, I had a friend who found herself unexpectedly pregnant who ultimately decided to put the baby up for adoption.Fast forward to 2021. This friend receives a message from someone she does not know, asking if she is his mother. How can this be possible? It comes to her, that DNA kit she had sent in as a family Christmas present a few years prior put her information out into the wild. The now 25-year-old child found his biological mother through a matching service and then messaged her asking to meetBringing back a wave of emotions she hadn't felt in years.How was she to know that a fun Christmas gift would lead to such a potentially life-shattering scenario? Even though we do not know the exact path technological innovation will take us, we can still take action now to create legislation that protects individual health data in light of rapidly emerging technologies.

The choice between classifying sensitive health data as consumer data or as an indelible part of personhood boils down to a recurring issue in technology legislation – balancing innovation vs. safety/societal impact. Innovation arguments emphasize speed, access, competition, consumer choice, and social benefit. Safety arguments emphasize hidden risk, unequal bargaining power, irreversible harm, externalized costs, and public trust. The best policy balance usually depends on the severity of potential harm, the consumer's ability to understand the risk, and whether the company or the public bears the consequences of failure. The synthesis is not a compromise between innovation and safety, but a redefinition of innovation itself: innovation in which human impact is expected, evaluated, and built into the

legislative framework governing sensitive health data.

The impact of choosing among these options is wide-ranging, but the individual health data privacy issue can be summed up as follows: will people be discriminated against in the future when organizations can access this sensitive health data through a data broker? What happens when the aggregation of data is more economically attractive than concern for people's privacy in a system they cannot opt out of?

Ownership and Governance of Health Data in the 21st Century

Currently, health data is treated legally as both a consumer product and as the digital representation of information about the body. These categorizations are based on who holds the data and why. If data is not legally considered health data, under consumer protection laws that govern health data products, it can be analyzed, aggregated, and sold. These laws in Virginia include the Virginia Consumer Data Protection Act (VCDPA) of 2021 and SB 754, which amends VCDPA to include personally identifiable reproductive or sexual health information without a consumer's consent.

Much of this data is provided by the person themselves and is not necessarily considered protected health data. Often, people waive their privacy rights to share this health data, unaware that this data can be aggregated with other data sources to develop health profiles which violate their privacy. These health profiles can then become the basis for decisions about the consumer in contexts far removed from the original reason the data was shared.

If the data is legally considered health data, it cannot be analyzed and sold as a consumer product. A patient's right to privacy is often waived when data is treated as a consumer product, whereas when it is collected as data about the body by covered entities, it is treated as protected health information (PHI), which has legal protections against being used to create a consumer product. PHI is legally protected under HIPAA, with a significant opportunity to expand HIPAA or state law to proactively address the impact of rapidly advancing data analytics and artificial intelligence capabilities.

This duality in how health data is legislated creates many moral muddles that, even when considered, are not addressed in any meaningful way. We can get away with this gray area for now because we can use the idea of data anonymization to shield individuals from harm that might come from their data being used as a consumer product. However, we have approached the day when there will be no means to de-anonymize health data: when DNA is used for health data analysis, the data becomes the unique biological signature of the person and cannot be anonymized.

The strongest argument for limiting the treatment of sensitive health data as a consumer good is that, with data aggregation techniques, health data is inseparable from the person. Collections of aggregated health or DNA data, for example, are not just data about consumer behavior, preferences, or choices; they are a digital record of the biological information that helps make a person physically distinct. It converts inherent features of the human body into governable, transferable, and analyzable data objects. For that reason, genetic and other deeply personal health data should be protected as a sensitive dimension of identity and personhood, not regulated merely as a commodity in the marketplace.

Legislation for Data that is Collected by Non-Covered Entities

HIPAA demonstrates how legislation on data protection can work. For thirty years, it has protected patients from invasions of privacy regarding their sensitive medical information collected and held in clinical settings. However, the current state of health data affairs leaves data highly vulnerable to exploitation and aggregation that threaten privacy, because HIPAA usually does not follow the data once the consumer gives it to a non-HIPAA application or non-covered entity.

HIPAA applies when individually identifiable health information is created, received, maintained, or transmitted by a covered entity or business associate in a covered health-care context. HHS states that HIPAA generally does not protect health information once it is accessed through or stored on a consumer's mobile device unless a covered entity or business associate is involved.

Further legislation is needed to effectively protect healthcare data, shifting the legal paradigm from who holds the data to what the data describes. If the source or aggregate data concerns a consumer's health choices, it should be treated as consumer data. If the source or aggregate data is data from the body or used to describe the body, this should be treated as protected health information. This distinction matters because consumer apps may have more freedom to use data for analytics, product development, advertising, data sharing, or profiling, depending on their privacy policy, consent flow, applicable state law, and FTC rules. HIPAA-covered entities and business associates face stricter limits on marketing, disclosures, safeguards, and secondary uses. We would want individuals to have the same level of protections today as had when HIPAA was first introduced by redefining PHI to account for rapidly advancing data analytics and artificial intelligence technologies.

If legislation makes the distinction based on what the data describes, the legal analysis becomes more precise. The same data can have different legal status depending on what it ultimately describes when it is the product. Thus, health data should become HIPAA PHI based on what it describes as a health data product, not merely because of its medical content or who holds the data.

The most significant change in this structure is that sensitive health data collected by consumer applications would receive protection based on the nature of the data product. For example, data entered by a patient into a fertility app or medication reminder app may originate with the individual, but it still reflects highly sensitive health information. Under this approach, new legislation would extend HIPAA-like protections to that data, treating it as protected health information even when it is collected outside a traditional health care setting. This would shift responsibility away from consumers, who may not fully understand the risks of storing or sharing sensitive health data on a personal device, and place it on the data product owner, whose organization is better positioned to anticipate, manage, and mitigate the potential harms to the individual.

The success of taking this approach to data protection for a broader definition of health data is that individuals will be much more protected when sharing their sensitive health data. Individuals must be able to participate in patient-centered care in a legally protected way that is as broadly accessible to members of society as possible. This protection should not depend on each individual's ability to understand complex data practices, read lengthy consent agreements, or anticipate how their health information may be reused in the future. Instead, the law should establish clear expectations for

organizations that collect and use sensitive health data, making responsible data stewardship the default rather than the exception. An added benefit of this proactive framing of data ownership and governance is that social norms and legal standards are easier to propagate in advance, helping people and organizations understand how health data should be protected rather than relying primarily on the enormous costs of enforcement after harm has already occurred.

While I am not a lawmaker or a lawyer, I do recommend that policymakers seek to understand the nature of data and information to account for the rapid changes as data is packaged as a product, knowing that the velocity of technological innovation is only accelerating and the impacts are becoming more pronounced. We can embrace this acceleration if we are clear on where we draw the line on what we want our society to operate in terms of commercialization and discrimination based on health data.

References

Chiruvella, V., & Guddati, A. K. (2021). Ethical Issues in Patient Data Ownership. Interactive Journal of Medical Research, 10(2), e22269. https://doi.org/10.2196/22269

Davidson, E., Wessel, L., Winter, J. S., & Winter, S. (2023). Future directions for scholarship on data governance, digital innovation, and grand challenges. Information and Organization, 33(1), 100454. https://doi.org/10.1016/j.infoandorg.2023.100454

Virginia Economic Development Partnership. (2023, August). Virginia: America's East Coast Tech Hub [Brochure]. VEDP. https://issuu.com/vedpvirginia/docs/tech_in_virginia?fr=xKAE9_zU1NQ

Winkler, E. C., Jungkunz, M., Thorogood, A., Lotz, V., & Schickhardt, C. (2025). Patient data for commercial companies? An ethical framework for sharing patients' data with for-profit companies for research. Journal of Medical Ethics, 51(5), jme-2022-108781. https://doi.org/10.1136/jme-2022-108781

Winter, J. S., & Davidson, E. (2022). Harmonizing regulatory regimes for the governance of patient-generated health data. Telecommunications Policy, 46(5), 102285. https://doi.org/10.1016/j.telpol.2021.102285

Privacy as a Design Parameter in the Age of AI

Gabriella Waters, Ph.D. & Maurice B. Jones, M.S.

Artificial intelligence (AI) increasingly depends on routine data collection across phones, apps, cameras, platforms, vehicles, workplaces, classrooms, health tools, and homes. Participation in digital life no longer means simply using a tool. For many Virginians, it means entering an environment where tracking, permissions, sensor data, behavioral records, and ongoing data capture are prerequisites for access to ordinary services and functionality. In that environment, privacy cannot be treated only as a compliance issue after an AI system has already been procured, deployed, or embedded into public life. It must be treated as a core design parameter for AI systems used in public, educational, health, workplace, consumer, and civic settings.

The Joint Commission on Technology and Science (JCOTS) studies issues related to technology and science in the Commonwealth, and its recent discussions have included AI's benefits, risks, and limitations. This is an important moment for Virginia to consider how privacy is being redefined in an AI-centered environment where data collection is no longer occasional, discrete, or easily avoided. The central policy question is not whether AI can provide useful services. It can. The question is whether those services must be built around maximum collection, indefinite retention, and reuse of personal and community data, or whether public policy can reward systems that deliver benefits with a lower privacy burden.

Issue Statement

The dominant model of AI deployment often assumes that better functionality requires more data. Navigation systems depend on continuous location tracking. Educational platforms rely on fine-grained behavioral data to personalize instruction. Wearables and digital health tools use continuous biometric streams to generate assessments. Smart home systems and neighborhood camera networks collect video, audio, movement, and environmental data in and around residential areas. Across these uses, the

value of the system is frequently tied to persistent data collection.

This shifts privacy from a matter of occasional disclosure to a condition of participation. A resident does not need to actively upload personal information to become visible to AI systems. Walking on a sidewalk monitored by neighborhood cameras, appearing in footage from a private doorbell camera, moving through spaces subject to automated license plate readers, using a connected vehicle, or accepting permissions in an app can all produce data that may be stored, combined, analyzed, and reused. In many cases, the data is not only used to deliver the immediate service; it may also be retained to improve future models, refine detection systems, personalize future interactions, and train the next generation of AI tools.

The policy concern is not only that people are observed. It is that they are increasingly analyzed, classified, inferred about, and used as inputs into systems that improve through continued exposure to human behavior. Everyday life now produces digital trails that are valuable to AI systems. People therefore become both users of AI systems and raw material for them.

The ability to opt out is increasingly limited. A person can decline a particular app, but it is much harder to avoid the broader data environment created by schools, employers, transportation systems, public services, consumer platforms, cameras, connected vehicles, and smart devices. Participation in ordinary life can entail extensive data generation even when an individual has not made an affirmative choice to share information with a specific AI system. This weakens the traditional privacy model that relies primarily on notice, consent, and individual choice. As scholars have long noted, privacy self-management becomes less meaningful when individuals cannot realistically understand, negotiate, or refuse the many data practices that shape modern life (Nissenbaum, 2010; Solove, 2013).

Why States Should Care About This Issue

Although AI privacy concerns are often discussed at the national or global level, states have a direct stake in how these systems are procured, deployed, regulated, and explained to the public. State agencies, schools, universities, local governments, health systems, transportation networks, and public safety institutions are already adopting or evaluating AI-enabled tools. States also shape consumer protection, public procurement, education policy, workforce development, public-sector data practices, and community safeguards.

For Virginia, this creates both an opportunity and a risk. The Commonwealth can help define a responsible model of AI deployment that protects residents while supporting innovation. At the same time, if public institutions normalize broad data extraction as the default cost of functionality, Virginians may experience privacy losses that are difficult to reverse. Once surveillance infrastructures, data pipelines, and vendor dependencies become embedded, they can be politically, financially, and technically difficult to unwind.

Several implications are especially important for state policymakers. First, privacy loss is increasingly collective, not merely individual. A person may be captured by systems they do not own and never agreed to use, including neighborhood surveillance cameras, facial recognition-enabled systems, and data-sharing arrangements that affect bystanders and communities. Second, data collection is increasingly ambient, occurring in the background across devices and environments rather than in discrete moments. Third, functional dependence weakens meaningful consent because essential tools for education, communication, transportation, employment, and public services may require extensive permissions. Fourth, AI systems create a feedback loop in which more data improves models, improved models create new capabilities, and new capabilities create incentives for still greater collection.

Virginia already has a legal and policy foundation for treating data practices as a public concern. The Virginia Consumer Data Protection Act incorporates principles such as purpose limitation, data minimization, and consumer rights in the commercial privacy context (Virginia Consumer Data Protection Act, 2026). AI raises a related but broader governance challenge: public institutions and policymakers must ask not only whether a data practice is disclosed, but whether the amount and type of data required by an AI system is proportionate to the public benefit being claimed.

From Privacy Loss to Participation as Data Generation

The problem is best understood as a shift from privacy loss to participation as data generation. In earlier privacy debates, the focus was often on whether a person knowingly disclosed information. In AI-enabled environments, data may be produced by movement, interaction, presence, hesitation, voice, image, biometrics, location, or device use. The person may not experience the moment as a disclosure at all. The system, however, may treat that moment as data.

This shift matters because it makes privacy a design issue rather than only a disclosure issue. If an AI system is designed to require continuous collection, then privacy risk is built into the architecture of the system. If the system is designed to process data locally, limit retention, separate service data from training data, and allow meaningful inspection or configuration, then privacy protection is also built into the architecture. NIST's AI Risk Management Framework emphasizes the importance of incorporating trustworthiness considerations into the design, development, use, and evaluation of AI systems (National Institute of Standards and Technology [NIST], 2023). The same logic applies to privacy: the most important decisions often occur before an individual ever sees a privacy notice.

These concerns are not theoretical. They are already visible across domains where AI and data-driven systems have become embedded in daily life. The following examples show how collective privacy, weakened consent, ambient collection, and participation as data generation appear in ordinary settings.

Examples and Applications

Collective Privacy: Neighborhood Cameras and Facial Recognition

Residential doorbell cameras, mounted neighborhood systems, automated license plate readers, and other privately or community-installed devices increasingly monitor streets, sidewalks, traffic, and routine activity near homes. These systems often capture residents, guests, delivery workers, children, and passersby who have no direct relationship with the device owner. When combined with facial recognition, object detection, license plate analysis, or anomaly detection, the issue is no longer simply recording. It becomes identification, categorization, and tracking across time and place.

This form of surveillance blurs the line between private convenience and public monitoring. A camera installed for home security can still contribute to a distributed surveillance environment that affects an entire block. Even where systems are not formally government operated, they can shape public expectations about who may move

anonymously through community spaces. The policy question for states is not only whether a particular homeowner or vendor complied with a notice requirement. It is whether communities should have clearer expectations for the cumulative privacy effects of distributed monitoring systems.

Meaningful Consent: Apps and Permission-Based Functionality

AI-supported applications often require access to location, contacts, camera feeds, microphone input, motion data, browsing behavior, or usage history to unlock their advertised functionality. In practice, users are often left with a limited choice: permit extensive data collection or forgo access to the service. That weakens the idea that users are making a fully voluntary and informed choice, especially when the service is tied to work, school, transportation, health management, or access to public information.

The same pattern appears in smart assistants, recommendation engines, productivity tools, and customer service systems. The more adaptive and personalized a system becomes, the more it may depend on ongoing observation of the person using it. The Federal Trade Commission's work on commercial surveillance has highlighted the scale of data collection and the need for stronger limits on collection, retention, and secondary use in digital services (Federal Trade Commission [FTC], 2024). For AI policy, the lesson is that disclosure alone is not enough. Public policy should also examine whether the requested collection is necessary and proportionate.

Ambient Collection: Education, Work, and Health Systems

AI-enabled learning tools may track time on task, hesitation, revision behavior, attention patterns, writing process, assessment history, and other indicators to personalize support. Workplace systems can monitor productivity, communications, scheduling, movement, and decision patterns. Health and wellness platforms gather biometrics, sleep data, movement patterns, environmental signals, and self-reported symptoms. Each of these uses may provide real benefits, but those benefits often arrive with a new form of ambient data collection.

The policy challenge is that the data demands are often hidden within the promise of convenience, efficiency, personalization, safety, or improved service delivery. Over time, those demands can normalize the expectation that support requires deeper surveillance. In schools, for example, the question is not only whether an AI tutoring tool improves student outcomes. It is also whether that outcome requires continuous monitoring of students' faces, keystrokes, behavioral patterns, or emotional states, and whether the data should be retained after the instructional purpose has been served.

Participation as Data Generation: Connected Products and Repairability

Right-to-repair debates are often framed around cost, consumer choice, sustainability, and market competition. In AI-enabled systems, repairability is also a privacy issue because connected products often continue generating data after purchase. When devices are locked down, difficult to inspect, or dependent on manufacturer-controlled software, users may have little ability to understand what data is collected, how long it is retained, whether sensors can be disabled, or how local device functions are connected to remote model training and cloud services.

A consumer may own a device without meaningfully controlling it. Smart appliances, vehicles, wearables, agricultural equipment, home security systems, and other connected products may gather data while limiting the user's ability to modify settings, install alternative software, replace components, or verify how the system operates. This creates a privacy asymmetry: companies may maintain technical control over data-producing systems located in homes, vehicles, farms, schools, and workplaces, while the people living with those systems may be unable to inspect or limit their behavior.

This is why repairability belongs in the privacy discussion. Repairability can improve transparency by allowing independent inspection of data flows and sensor behavior. It can improve user agency by allowing intrusive features to be disabled or reconfigured. It can also reduce dependence on closed ecosystems in which data extraction remains the price of continued functionality, updates, or support. In that sense, repairability is not only about fixing a product. It is about whether people can meaningfully govern the technology that governs them.

Evaluating the Privacy Burden

AI policy often focuses on whether a system works, whether it is accurate, whether it is biased, and whether it produces discriminatory outcomes. Those questions are important, but they are incomplete. A more useful public question is not only whether an AI system works, but what it requires in order to work at the claimed level of performance.

That question points to a concept that should be more visible in policy discussions: privacy burden. Privacy burden refers to the amount, sensitivity, persistence, identifiability, and downstream reuse of data a system requires to deliver its benefits. Systems with similar functional outcomes may impose very different privacy burdens. Public procurement, institutional adoption, and product comparisons rarely evaluate that dimension directly.

For example, two AI-enabled tutoring systems may both claim to improve student learning. One may require continuous webcam monitoring, keystroke logging, behavioral inference, and indefinite retention of student interaction data. Another may provide similar instructional support using locally processed assessment data, shorter retention windows, and clear limits on secondary use. Both systems may be described as personalized learning tools, but they impose very different privacy burdens on students, families, teachers, and schools.

Similarly, two public-sector customer service tools may both help residents navigate benefits, permits, or agency services. One may retain full conversation histories for model training and vendor analytics. Another may retain only the minimum information needed to complete the transaction, separate personally identifiable information from operational logs, and prohibit secondary use. The difference is not merely technical. It is a difference in governance, accountability, and public trust.

Treating privacy burden as a design and evaluation issue would help Virginia move beyond a narrow notice-and-consent approach. It would also align with broader responsible AI principles that emphasize proportionality, transparency, accountability, safety, and the mitigation of downstream harms (NIST, 2023; Organisation for Economic Co-operation and Development [OECD], 2024). The NIST Privacy Framework similarly treats privacy risk as something organizations can identify, assess, and manage through enterprise risk management rather than as a matter of legal notice alone (NIST, 2020).

Privacy as a Competitive Advantage

If privacy were treated as a competitive advantage rather than a trade-off, AI systems would compete on restraint as well as capability. High-performing systems would differentiate themselves by delivering useful outcomes with less data, shorter retention, stronger local processing, greater repairability, limited or no sharing and selling of data to third parties, and clearer user control. This requires policy that prioritizes privacy over data maximization.

In practice, privacy-preserving competition could mean that vendors and institutions are encouraged to demonstrate how they reduce unnecessary data collection. A health application might process sensitive information on a device rather than sending raw data to centralized servers. A public-sector vendor might distinguish between data required for immediate functionality and data retained for model training or product improvement. A school system might require dashboards that show what data is collected, when it is used, and whether it contributes to model improvement. A smart-device manufacturer might allow qualified independent technicians to inspect sensors, data pathways, and software dependencies. Across these examples, the policy goal is not to prevent useful AI. The goal is to encourage useful AI that does not automatically assume maximum collection.

This approach would not eliminate trade-offs. Some systems may perform better with broader data access. But it would shift the burden onto designers and deployers to justify collection rather than assuming that broad collection is the default cost of innovation. That shift is especially important for public institutions. When government agencies, schools, universities, and publicly supported programs choose AI systems, they can shape market expectations by rewarding vendors that build privacy into the product rather than treating it as an afterthought.

Policy Considerations for the Commonwealth

Virginia does not need to resolve every AI privacy issue at once. The Commonwealth can begin by making the privacy burden visible in the places where public institutions already make choices: procurement, vendor review, public-sector deployment, grantmaking, education, and oversight. The following options are intended as practical steps for study, guidance, or future policy development.

The recommendations below are intentionally framed as governance tools rather than technology bans. They aim to preserve room for beneficial AI while ensuring that public institutions ask better questions before systems become embedded.

1. Adopt privacy-burden evaluation criteria for public-sector AI systems. Agencies and public institutions should assess data volume, sensitivity, retention, identifiability, repairability, secondary use, and third-party sharing before procurement or deployment. This evaluation should sit alongside performance, cost, cybersecurity, accessibility, and bias assessments, rather than being treated as a separate compliance afterthought.

2. Develop model guidance on AI data minimization. Virginia could articulate expectations that AI systems collect only the data necessary for clearly defined purposes and that vendors explain why each major data category is needed. This guidance would build on existing privacy principles in Virginia law while extending them into procurement and institutional adoption decisions.

3. Require clearer disclosure of training and reuse practices. Users, agencies, and affected communities should be able to distinguish between data collected to provide an immediate service and data retained for model training, product improvement, analytics, or other secondary uses. Vendor disclosures should identify retention periods, deletion practices, sharing arrangements, and whether personal or community data contributes to future model development.

4. Address collective privacy in surveillance contexts. Policy discussions should recognize that neighborhood cameras, facial recognition, automated license plate readers, and other distributed monitoring systems affect non-users and bystanders, not only device owners. Virginia could study whether communities need clearer standards for notice, accountability, access, retention, and law-enforcement or third-party use of data generated in shared spaces.

5. Examine right-to-repair protections for AI-enabled and connected products. Policymakers could consider whether consumers, institutions, and independent technicians need greater ability to inspect, maintain, and configure systems that collect data in homes, vehicles, farms, shopping areas, parks, schools, and workplaces. Repairability should be evaluated not only as a consumer-protection or sustainability issue, but also as a privacy and transparency issue.

6. Encourage privacy-preserving technical design. Procurement and grant programs can create incentives for local processing, differential access controls, privacy-enhancing technologies, short retention periods, purpose-bound deletion, stronger de-identification, and more inspectable systems. These incentives would help move the market toward AI systems that achieve useful outcomes with less collection.

7. Support public literacy on AI data practices. Virginians need clearer explanations of how ordinary participation in digital systems generates data that may be used beyond the original interaction. Public literacy efforts should explain data collection, inference, model training, retention, and consent in plain language, with special attention to schools, older adults, small businesses, public employees, and communities affected by surveillance technologies.

Considerations for Further Study

Several questions would benefit from additional study by JCOTS and partner institutions. Virginia should identify the sectors in which consent is least meaningful because functionality depends on extensive collection. Public institutions should consider how to evaluate systems that affect people who never directly opted in, including bystanders, students, workers, patients, and residents moving through monitored spaces. The Commonwealth should also study procurement mechanisms that reward privacy-preserving and repair-friendly design without discouraging useful innovation.

Additional study should distinguish between data needed for service delivery and data retained for model training, product improvement, behavioral analysis, or other secondary uses. This distinction is especially important in education, health, disability support, employment, agriculture, transportation, and neighborhood surveillance contexts. In each of these domains, the relevant privacy question is not only what the system does, but what data the system requires, how long the data persists, who can access it, and whether affected people can meaningfully understand or limit the practice.

Conclusion

Privacy is no longer only about secrecy or confidentiality. It is increasingly about agency, participation, and whether people can move through digital and physical environments without being continuously converted into data for optimization, prediction, classification, and model improvement. AI intensifies this issue because the value of many systems grows through repeated exposure to human behavior. That creates powerful incentives to collect more, retain longer, and reuse

data in ways that may be difficult for individuals and communities to see.

Virginia has an opportunity to advance a more responsible approach as AI systems continue to proliferate. The Commonwealth can ask whether functionality truly requires maximum collection, whether public institutions can evaluate privacy burden before adoption, and whether vendors should compete on privacy-preserving design as well as capability. Treating privacy as a design parameter would not stop innovation. It would help ensure that innovation is aligned with public trust, institutional accountability, and the ability of Virginians to participate in modern life without unnecessary surveillance.

References

Cavoukian, A. (2011). Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario.

Federal Trade Commission. (2024). A look behind the screens: Examining the data practices of social media and video streaming services. Federal Trade Commission.

National Institute of Standards and Technology. (2020). NIST Privacy Framework: A tool for improving privacy through enterprise risk management, Version 1.0. U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.01162020

National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

National Institute of Standards and Technology. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.600-1

Nissenbaum, H. (2010). Privacy in context: Technology, policy, and the integrity of social life. Stanford University Press.

Organisation for Economic Co-operation and Development. (2024). OECD Artificial Intelligence Principles. OECD.AI Policy Observatory. https://oecd.ai/en/ai-principles

Solove, D. J. (2013). Privacy self-management and the consent dilemma. Harvard Law Review, 126(7), 1880-1903.

Virginia Consumer Data Protection Act, Va. Code Ann. §§ 59.1-571 to 59.1-585 (2026).

Thanks to our Sponsors

Thanks to the generous support of our sponsors, JCOTS was able to offer free admission to the 2026 Tech Policy Symposium. In return, sponsors were given the opportunity to conduct a 10-minute lightning talk on a topic of their choice.

Google

Google is leading a new era of innovation, bringing the best of AI and security to mission. Google is committed to empowering the public sector with secure cloud-native infrastructure and accredited AI-powered solutions.

Meta

Meta's AI glasses are built with accessibility at the core, developed in collaboration with blind and low-vision communities to ensure the technology meets real needs. Through Be My Eyes' network of sighted volunteers, Live AI capabilities, and an expanding set of developers building bespoke accessibility applications, AI glasses bring hands-free, voice-first assistance to everyday life. Meta's AI glasses are uniquely suited as a form factor for accessible technology, and building with the community has ensured these features are designed for everyone.

Amazon

Amazon leverages artificial intelligence and machine learning to break down barriers for customers with disabilities, deploying services like real-time transcription, text-to-speech, object recognition, and smart home voice control across its devices and cloud platform. Through innovations like VoiceView screen readers, Alexa hands-free assistance, and AWS accessibility tools that enable developers to build inclusive applications, Amazon is working to ensure technology serves everyone—transforming how people with visual, hearing, mobility, and communication disabilities navigate daily life independently.

Fathom

As Virginia and other states confront the challenge of governing rapidly advancing AI systems, Independent Verification Organizations (IVOs) offer an effective model for ensuring AI safety without stifling innovation. IVOs would be non-governmental, expert-bodies licensed and overseen by the state of Virginia to assess AI systems against established and dynamic standards for safety. This presentation will explore how IVOs can give Virginia regulators, businesses, and consumers confidence in AI systems while preserving the agility that innovation requires, offering a third way between heavy-handed prescriptive regulation and unregulated deployment.

Participants

Data Protections in Virginia and Beyond

Dr. Douglas C. Schmidt

Dr. Aaron Martin

Amelia Vance

Dr. Michael McGarry

Data Protections in the Age of AI

Daniel Shin

Dr. Gabriella Waters

Margaret Hu

Rebecca Crootof

Health Data 30 Years After HIPAA

Dr. Kira Allmann

Dr. Sarah Henrickson Parker

Dr. Elizabeth White Baker

Akit Mathur

JCOTS gratefully acknowledges Anna Dehmer for her vision and diligence in creating this publication from concept to completion

JCOTS
Old City Hall
1001 East Broad Street
Suite 321
Richmond, VA 23219
info@jcots.virginia.gov
jcots.virginia.gov