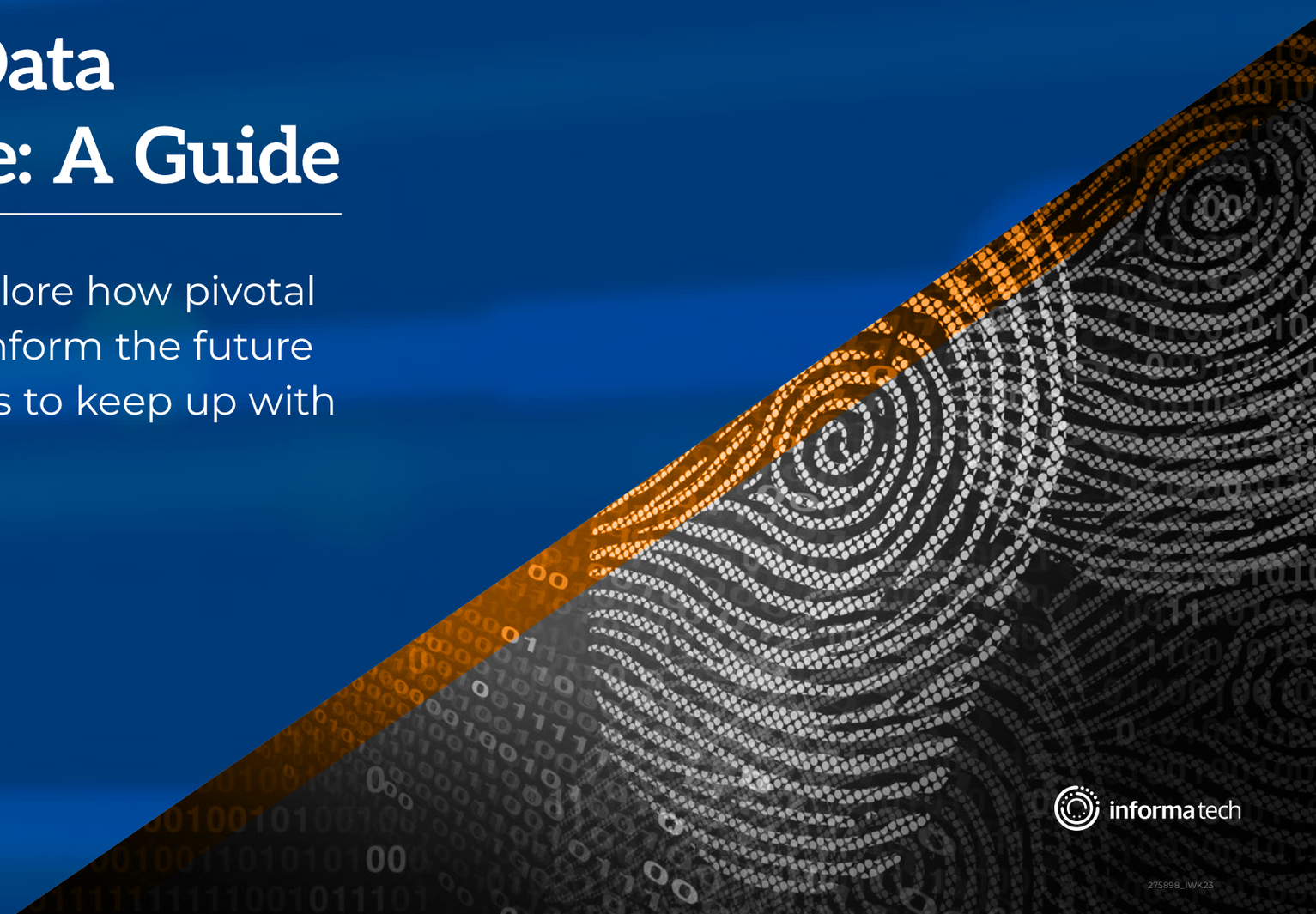
NEXT

Keeping Up With Data Privacy Compliance: A Guide

In this Tech Insight Report, we explore how pivotal moments in data privacy history inform the future of compliance and offer expert tips to keep up with data management best practices.



informa tech

275898_IWK23

Table of Contents

Executive Summary

In spite of their recent popularity, large language models and generative artificial intelligence (AI) have quickly turned into a **data privacy nightmare**. Samsung's early use of OpenAI's ChatGPT, for example, resulted in **leaked internal data**, and Verizon (among others) banned the chatbot over the "risk of losing control of customer information, source code, and more." In this sense, the benefits of data privacy compliance are twofold: reducing (1) potential impact of data breaches and (2) unauthorized access to your company's private intellectual property.

But AI is far from the first compliance challenge in the history of data privacy. Over the past eighteen years, we've seen data breaches, location tracking, and new consumer data protections.

In Keeping Up With Data Privacy Compliance: A Guide, you'll learn how these pivotal moments in data privacy history are informing the future of compliance measures and driving best practices for future-proofing your data management and governance strategy.

Pivotal Moments in Data Privacy History

In the last 18 years, the internet has evolved at warp speed to keep up with busier lives and a craving for mobility, while also trading access for privacy.

By Brandon Taylor

As more and more life changing technology advancements arise, the continued blurring of lines have made American and tech culture synonymous. The cloud, e-commerce, GPS connectivity, remote access, smartphones, and everything in between have changed how we interact with each other and our world for the better (arguably) forever. All companies in every industry collect data in 2022, but how do they "use it"? Consumers expect a product or service in exchange for their personal data, while businesses learned to build trust to exploit opportunity. Surely every business would hold privacy over profit, right? Well, jump in your DeLorean, we're going back to the future.

Historically, finding privacy balance for companies has been key to longevity but not always a priority. Data privacy and protecting consumers is everything, along with even bothering to create (go figure) and even knowing what your policy protects is better. This isn't the most mind-blowing concept, but you'd be surprised at how often companies and their common sense took a back seat to the Almighty Dollar.

ChoicePoint Data Breach (2004-2005)

The first stop in our remade DeLorean is the ChoicePoint Data Breach in early 2005. The data aggregator firm, known for combining information from public and private databases, takes this data and then sells it to private sector firms and government agencies. However, in February 2005, a group of L.A. County fraudsters and their 50 fake businesses were able to dupe ChoicePoint into selling personal consumer information that compromised the lives of 163,000 people. California SB-1386 is a law that requires the disclosure of any data breach to be publicly reported by the company at fault. Security Freeze law discussions and a tarnished image arose along with a $15M penalty for ChoicePoint, which would set a precedent for privacy.

FTC Mails Refund Forms to ChoicePoint Data Breach Victims

Early in 2005, ChoicePoint reported it had handed over consumers' names, addresses, Social Security numbers, and credit reports to fraudsters working out of Los Angeles County.

FTC Sends Message With $15M ChoicePoint Fine: Protect Consumer Data

The $10 million civil penalty portion is the highest fine in FTC history, but paltry compared to the $50 million fine McAfee was ordered to pay the Securities and Exchange Commission earlier this month for allegedly overstating earnings statements.

Law Requires ChoicePoint to Disclose Fraud

An identity-theft ring gained access to 145,000 consumer records held by ChoicePoint, which later notified consumers as required under California law SB-1386.

Smartphones and Location Data (2007-2008)

Just two summers later, a 2007 Tele Atlas survey began to present a new frontier on mobile connectivity. It reported that 84% of consumers wanted GPS capability on their mobile device. The movement for getting a fix on mobile users was fueled by the FCC's 1996 Enhanced 911 Initiative, forcing carriers to provide emergency call location data. Today, location information is key to deliver mobile services, ads, and marketing for many industries. What started out as a sincere way to help civilians in danger, began to morph into pseudo-surveillance, which may provide timely suggestions in certain environments but also intrude on others.

GPS Isn't the Only Tool for Location

Even without a GPS-enabled device, people can still tap into location-based services.

The Promise of Mobile GPS and Location

If 2007 was the year of smartphones, then 2008 promises to be the year of mobile location. Consumers and business users want GPS and other location services on their smartphones. But what does 2008 really hold in store?

Time for Businesses to get Serious About Location Technology

The tech's still in its early days, but tracking a pizza delivery street by street shows what's possible.

The 'Dark Side' of Big Data (2012)

Speaking of predictive analytics, 2012 birthed the concept of "big data." Big data can solve big problems, but big ideas have driven many poor societal decisions over time. Big data has empowered companies to deny services and monetize consumer data. This data aggregation continues to grow today, fueled partly by national security and focused marketing.

Big Data's Dark Side

While big data shows tremendous potential in a variety of industries, such as healthcare, e-commerce and traffic prediction, it has a potential "dark side" as well.

10 Predictions About Big Data

Will big data be a force for good or evil by the end of this decade? See if you agree with expert reactions to new Pew Internet Center research.

Right to Be Forgotten

Following a 2014 ruling by the European Union and the European Court of Justice, the "Right to be Forgotten" provides support to those EU citizens who wish to remove personal data from search engines under EU jurisdiction. But this was easier said than done, and US citizens did not get the same treatment.

Flaws Found in 'Right to be Forgotten' Data Privacy Laws

A study by privacy researchers finds that attempts to obscure online information can be defeated with a bit of effort.

Rethink the Right to be Forgotten

The "right to be forgotten," recognized in Article 17 of the European Union's revision of its 1995 data protection rules, is at once admirable and asinine.

GDPR Emerges (2016-2018)

There are real benefits to be had in understanding and protecting consumer data. The year 2016 brought the European Union's General Data Protection Regulation (GDPR) to the forefront, which forced companies to rethink what practices are done with personal data, and introduced new, real punishments for non-compliance. Amazon, WhatsApp, and Zoom were all penalized in 2021 with a 4% total revenue fine for violations signaling a rebalanced data relationship between people and companies.

5 Common GDPR Misconceptions

A company's effort to comply with GDPR doesn't end on any particular date. The work is ongoing, tied to a recognition of privacy as a fundamental right.

GDPR: A Cost vs. Benefit Analysis

It's a mistake for companies to view compliance with GDPR as just a financial burden. There are real benefits to be had in understanding and protecting customer data.

Equifax Breach (2017)

The Equifax Breach of 2017 taught us the need for companies to test third-party code so that patches can be properly and timely implemented. Equifax was snake bitten when back doors in that sourced code were left open and exploited.

One Point the Equifax Breach Drives Home

Equifax blamed its recent high-profile breach on the Apache Struts Web Framework. As software delivery cycles shrink, developers have to rely on more third-party components, libraries and frameworks. When they do, what are their liabilities and responsibilities?

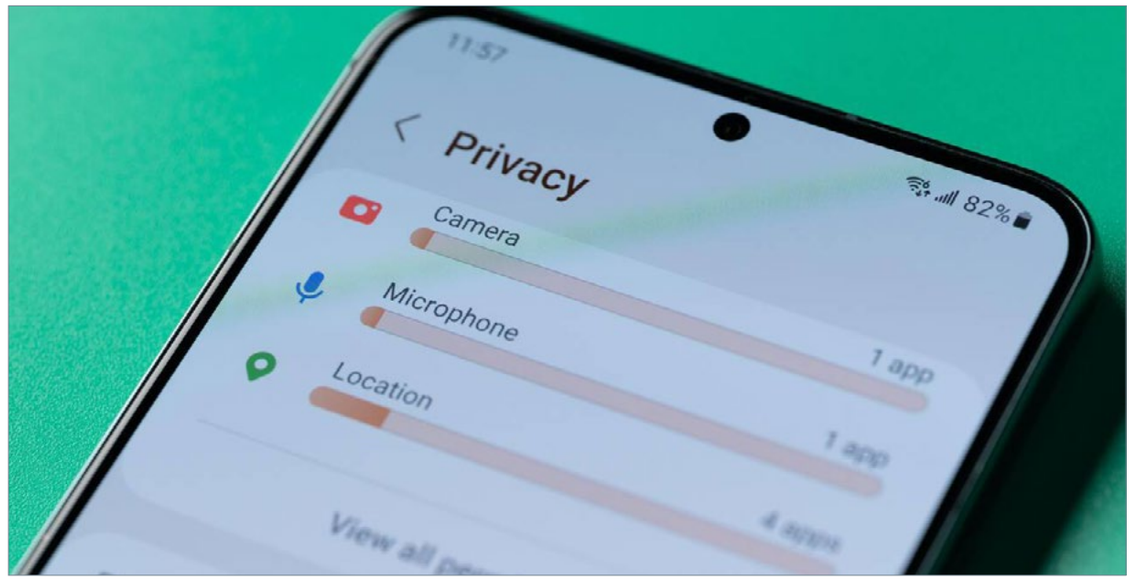
Cambridge Analytica Scandal (2018)

2018's Cambridge Analytica Scandal reiterated the need for CIOs to stop ignoring customer data privacy concerns and how companies can get a handle on their data operations.

Facebook's Cambridge Analytica Trouble Highlights IT Data Privacy Concerns

Facebook has lost billions in market value, been targeted for investigation by 37 states and the FTC, and many users are threatening to delete their accounts. It's past time for enterprise CIOs to stop ignoring customer data privacy concerns.

As we continue to rely on access and convenience, at what point are our Data Privacy issues more than just a personal problem? Great question, stay tuned.



Data Sovereignty, Compliance Shape IT Leadership

A rapidly changing regulatory landscape requires businesses to proactively build data sovereignty considerations into their overall business and risk management strategies.

By Nathan Eddy

Data sovereignty and industry compliance continue to factor in highly to discussions about future organizational IT architectures.

A recent IDC **survey** indicated these two issues will play a central role for IT leaders choosing service providers and making evaluations about their primary datacenter environments.

Meanwhile, the regulatory landscape is changing, and businesses must demonstrate they are meeting their obligations within and across regions despite differing regulations and complexities around where data resides.

For example, data in the cloud could be in a different legal jurisdiction than the business, leading to additional questions on legal obligations.

Businesses are forced to invest more time into compliance considerations, as it's no longer something that can be ignored until it becomes an issue.

The European Union was far ahead in defining what its expectations are through the General Data Protection Regulation (GDPR), however other regions are also introducing their own requirements such as California's Consumer Privacy Act.

Data Sovereignty Competency Matters

"The topic of data sovereignty is more urgent than ever as we try to counter-balance these considerations," explains Jason Conyard, CIO of VMware. "Privacy and privacy-adjacent laws is also an ever-growing topic not only on a national level, but on a consumer level as well."

He points out customers want assurances about their data — how it is used, who it is shared with, and how it is protected.

"If a company can demonstrate competency in meeting its commitments, it builds trust and customer loyalty and ultimately leads to increased profitability," Conyard says.

Spencer Kimball, co-founder and CEO of Cockroach Labs, adds while risk mitigation is the obvious impetus for change, a strategic embrace of the challenge of data sovereignty can pave the way to more frictionless expansion into new markets.

"Very few businesses in today's connected digital economy are not looking towards a future of global expansion," he points out.

He says with the inevitability of new regulations always on the horizon, it's increasingly important to build on infrastructure designed to overcome these challenges.

"The global public cloud is the right substrate, but simply moving workloads built on legacy infrastructure to the cloud isn't enough," Kimball explains. "Instead, architectures must become aware of geographic realities — for example, where must the data be domiciled, and where can it be processed in order to remain compliant.

This is a problem that extends from the database all the way up to the application logic which processes the data.

A Complex Environment Adds to Challenges

Businesses are running data across multiple third-party datacenters and clouds, which raises questions about where the infrastructure is and how to demonstrate that certifications and obligations are being met.

"It's important that businesses select partners and multi-cloud providers who can certify on their behalf, since the organization is ultimately responsible even if someone else is enabling the transaction," Conyard says.

He points to another interesting factor — that some cloud providers are being barred from, or severely limited from operating in certain jurisdictions, which forces businesses to use more than one provider.

"For example, some cloud providers weren't operating in Russia prior to the invasion of Ukraine, which was exasperated when increased restrictions were put in

place because of the conflict," he says. "This adds another layer of complication to businesses' calculations around service provides."

It's a complicated landscape, which is why businesses must rely on highly competent partners, with the right certifications, who fundamentally understand that data sovereignty is not just a nice to have — it is table stakes.

Kimball agrees careful selection of vendors that provide infrastructure purpose-built to exploit the cloud is a must, but an overreliance on any single cloud service provider (CSP) — especially on CSP-specific infrastructure choices — can lead to unacceptable vendor concentration risk.

"Investing to build a flexible, multi-cloud posture can also be an important prerequisite for expansion, as each cloud vendor has different strengths in presence across different geographies," he explains.

Customer preference for where a service is hosted (the country or region, as well as in which public cloud) can also be a factor, especially where the customer is a business or a government entity.

CIO, Legal, Security Among Key Stakeholders

Kimball explains as demanding compliance requirements continue to evolve, the re-architecture of the tech stack to support the next generations of applications and services has become a strategic priority across the C-suite.

"The time horizon to realize the value of these investments is measured in years, or even decades," he says. "We see this responsibility most commonly falling under the purview of the CIO, with significant execution from chief architects, IT compliance, procurement and legal."

From Conyard's perspective, any large organization should have their privacy team involved in ensuring data compliance, as well as their security, IT and legal teams.

"Many companies are also relying on external counsel to help them navigate the unusual territories," he adds. "While most large companies are familiar with the legal requirements and obligations in the countries they primarily do business in, compliance isn't defined by national borders."

This requires businesses to go to greater lengths to consider relevant jurisdictions and considerations.

They must also know their data — what they have and where they have it — to identify the appropriate requirements.

For example, if businesses have data that includes that of a European Union resident, they have an obligation to fulfill GDPR, no matter the country in which they reside.

"Looking forward, it's crucial that businesses identify their guiding principles," Conyard says. "The choice is doing enough solely to meet legal obligations, or using data compliance as an opportunity to demonstrate to customers and key stakeholders that they take privacy seriously and are a trusted organization in the long term."

The Companies Leading Privacy-Enabling Tech

Selecting a data privacy management vendor requires organizations to carefully consider their specific needs, and the level of help they need to meet compliance requirements.

By Nathan Eddy

The need for data privacy management is being driven by the growing number of regulations — spurred on in a large part by Europe's General Data Protection Regulation (GDPR) legislation — and by the understanding that privacy is a discipline, a posture an organization must take.

Often it is the compliance, the governance, the risk, or the legal department that is tasked with these things, and they start looking for capabilities in the markets for integrated risk management, universal content of preference management, subject rights, request automation, and vendor risk management.

The widespread use of clouds, both public and private, is adding more layers to the issue of data privacy management as organizations turn to data-driven approaches to privacy compliance and governance.

Ways to Approach Privacy

Bart Willemsen, Gartner VP Analyst, who focuses on all privacy-related challenges in an international context, explains that organizations approach their privacy program in a couple of stages.

"First, you must establish it, and that is where the most fundamental of both privacy management and data-centric capabilities are combined," he says.

The typical combination includes mapping of risk, discovery of data, classification, recordkeeping, retention policies, and most importantly, all elements where it touches upon the interaction with the individual data subject.

"We call that the privacy user experience, which typically starts with transparency — what you put in your notices, your statements, just-in-time messaging, adjusting the customer experience, storyline, the forks in the road that you architect there only to then offer choice," he says. "That's where consent management and preference management comes in."

As data privacy is a discipline that touches on several markets, organizations may have to look at multiple vendors offering their own solutions to different parts of data privacy management issues.

The consent management platform from **Secure Privacy**, for example, can automatically scan the organization's website and create a detailed report of all steps the company must take to make it GDPR or California Consumer Privacy Act (**CCPA**) compliant.

These are just two of the markets for which Secure Privacy has developed data privacy management solutions, in addition to Brazil, Thailand, and Canada.

Available features include the ability to automate cookie consent, visitor preferences, privacy policy & cookie declaration management across international data privacy laws.

Cross-Regulation Requirements

Because of the international nature of contemporary business, Enza Iannopollo, a principal analyst on Forrester's security and risk team privacy management, says organizations should look for providers that offer solutions that satisfy cross-regulation requirements, with automation.

"The privacy market fundamentally is a very broad range of technologies, starting from the regulatory change management, all the way to fundamentally deploying the controls to affect the data," she adds. "Newer providers, those that were created for privacy management specifically, are the ones who have had more success in the market."

Among those vendors is **OneTrust**, which aims to become the "home of the privacy tech ecosystem" offering a range of use cases, including minor and contextual consent, capabilities bolstered by a string of acquisitions, and a large pool of active customers.

The company recently released a Certification Automation product that helps companies attain the new International Organization for Standardization (ISO) 27001:2022 certification.

This certification signals that they have defined and implemented processes across their information security management systems (ISMS) that align with industry best practices.

The solution is designed to help organizations more efficiently scope, assess, and generate evidence to prove compliance across ISO and adjacent security and privacy frameworks, while simplifying preparation for future third-party audits.

Privacy Compliance Management

BigID, meanwhile has increased its scope beyond privacy operationalization for enterprise clients to offer privacy compliance management aimed at small to medium-sized businesses (SMBs).

The company's approach is based off of automated data discovery, helping organizations gain visibility and insight into personal, sensitive, and enterprise data.

One of its more recent products is Hotspot Reporting, which gives organizations the power to visualize and remediate their riskiest data and help prioritize their biggest data vulnerabilities.

Native data deletion capabilities allow organizations to delete personal and sensitive data across their data stores from Snowflake and AWS S3 to mySQL, Google Drive or Teradata.

Securiti offers multi-cloud data protection, governance, and security, underpinned by machine learning capabilities for most of its modules, and boasts partnerships with Workday and Cisco.

The company recently debuted DataControls Cloud, offering a layer of unified data intelligence and controls across all major public clouds, data clouds, SaaS, and private clouds.

The e-discovery and information governance software company Exterro is focused on the legal challenges associated with IT and data, with a platform automating the interconnections of privacy, legal operations, digital investigations, cybersecurity response, compliance, and information governance.

"We see a lot of organizations thinking about privacy as having a basic privacy management software," Iannopollo says. "Whenever I look at a vendor, the very first thing that I ask is always about the possibility to integrate their solutions with the rest of the organization."

Options for Zero Privacy Expertise

The up-and-coming Osano is aimed at organizations that must achieve compliance with privacy regulations but may have zero privacy expertise.

The offerings are not as sophisticated as some of the more established players — only discovery of structured data is supported, for example — but a streamlined approach and support for some third-party risk management could make it an attractive option for those who need help managing a privacy program.

From the perspective of both Iannopollo and Willemsen, the data privacy concern for organizations is complex and multi-faceted, overlapping not only with security but also content management and preference management.

"It's impossible, to be honest, that one single piece of technology is going to provide you with all that you need," Iannopollo adds. "Privacy is how your organization operates with data, and it is everywhere. So, it's very unlikely even thinking that there is a single software that is going to provide you with all the governance you need around data."

Preparing for Compliance With AI, Data Privacy Laws

With individuals set to gain new rights over how businesses use automated decision-making, businesses will have to ensure they're compliant with data privacy and AI regulations.

By Nathan Eddy

A growing number of data privacy laws in the United States and in the European Union (EU) mean businesses must ensure they're in compliance with regulations affecting personal data of employees and are offering clarity and consent options when it comes to the use of AI-based decision making.

Despite enforcement delays, New York's Local Law 144 will regulate the way organizations use automated employment decision tools, while in California, the Consumer Privacy Act (CCPA), recently amended by the California Privacy Rights Act (CPRA), expands data privacy law.

It will now offer protection to job applicants and current employees, as well as independent contractors and dealings between businesses.

"I strongly urge organizations to look beyond compliance," says Bart Willemsen, VP, analyst with Gartner. "There are many requirements popping up worldwide, and if you want to prevent having to ad hoc respond to all these things in detail, try to elevate your game to an ethically responsible one. Don't look at compliance. Look at risk."

He explains the CPRA explicitly includes profiling in its language, which guards against the unauthorized use of AI in employment screening tools, for example.

"Applicants must be notified of the use of technology not only during the video interview, but also in the case of intended use of AI to analyze the video interview afterwards," he says. "When you deploy or intend to deploy, always offer full transparency of both intent and technology use."

Willemsen also recommends organizations continuously monitor and manage AI risks in the development stage, training stage, and in production.

"The key items for businesses to be aware of include transparency, choice and monitoring," he says. "You can only ask an individual to make a decision after you

give clarity, transparency and the right not to be subjected to automated decision making."

Brian Platz, co-CEO and co-Founder of Fluree, says the laws underscore the need for companies to have clean and organized data that is accessible upon employee request.

"It will also be important for organizations to be aware of that data's lifetime to ensure they are providing employees with complete, comprehensive records in the event data was copied or duplicated for various purposes," he explains.

Laws Complicate Leveraging Data for AI Models

From the perspective of Muddu Sudhakar, CEO at Aisera, these laws "certainly" make it tougher to leverage valuable data for AI models.

"AI generally needs massive data sets to get effective results. Next, there is the problem that the data may have gaps," he explains. "This could lead to skewed models. There may even be potential issues with bias because the data may not be representative of the population."

He points out that another issue is that the California law has "rulemaking", which means that it is not clear what the final compliance requirements will be.

"This can add to the difficulties with building models as well as the costs," he says. "There are likely smaller organizations — who do not have strong compliance programs — that may not be aware of the new laws. There is a lack of awareness in general."

Sudhakar adds the California law applies to workers and will make privacy much more complicated for employers, raising questions as to what employee information can be deleted on request.

"However, gig companies may have the biggest challenges — especially the larger ones," he says. "They will have to manage privacy requirements across many contractors, who may not stay with the company very long."

Shira Shamban, CEO at Solvo, points out proof of compliance is not a new need.

"The interesting thing about the new regulations is that if up until now many of the frameworks we needed to comply with had to do with specific verticals, like HIPAA

for healthcare or PCI-DSS for payments, the new regulations are talking about the individual person's privacy," she says.

Like GDPR before, now other states are looking to protect their resident's data, and there isn't a single path for compliance, but what's important is to have privacy in mind.

That means security and GRC engineers should investigate existing security practices and mechanisms on the one hand, and the data their organization is storing on the other hand, and make sure they correlate.

"There are a few products out there in the market today that could help organizations to identify their private data, and from there it's the security team's job to make sure they're doing the best they can in protecting it," Shamban says.

Getting Ready for Regulatory Compliance

Even though enforcement of data privacy laws in California and New York laws have been slightly delayed, and California regulations implementing the new AI law are not yet fully baked, businesses should be employing expert consultants now to be ready when enforcement begins.

Platz notes that in the working world — and especially in an environment that is often largely remote with employees around the country and the world — these new privacy laws will affect employees beyond the states that enacted the laws if they live and work in different locations.

"With flexibility to work from virtually anywhere, this legislation will have wide reaching impact across states and sectors and will only highlight the need for employers to look closely at their path to compliance across a significant amount of data," Platz says.

Bryan Cunningham, advisory council member at Theon Technology, a provider of data security, explains California often leads the way on US privacy laws which, in turn, often are inspired by those in the European Union, and new laws and regulations around the use of artificial intelligence to process personal data are the most recent examples.

"As almost always happens, many other jurisdictions will follow suit, as New York City already has," he says. "So, businesses should be preparing to deal not just with these two new laws but, ultimately, with similar ones in most or all states and perhaps other cities."

He adds even now, businesses can take little solace in not having a California office or California resident employees, because the new law purports to protect any Californian about whom the business collects or processes data, including employees, independent contractors, and others.

"New York claims a similar reach, and also requires a not-fully-defined 'bias audit' for the use of AI in employment decision-making," Cunningham notes. "In addition, similar EU laws and regulations may well impact US-based businesses if they process data of EU citizens."

Under such laws, individuals gain new rights over how businesses use automated decision-making, including notification, transparency, opt-out, and correction rights.

"In conjunction with expert lawyers and consultants, businesses should first identify, catalog, and map personal data they hold and any automated or AI-based decision-making tools they use," he says. "Then they should determine which of the new and emerging laws apply to them. And they cannot begin too soon."

10 Actionable Tips for Managing/Governing Data

Here is a quick look at things you can do to help govern and manage your data in the most practical sense.

By Pam Baker

If you believe the marketing hype, you'd think data management and governance is a snap. An easy peasey, automated to the hilt, set-it and forget-it, little cleanup task on the prepping end of the serious work: data analysis.

But today it's more like mapping a mine field while trying not to step on one of the many camouflaged dangers. If you mess this part up, the aftermath will be even messier.

To put it simply: If the data is wrong or incomplete, the analysis will rank somewhere between useless and dangerous. If the data slips through the cracks unnoticed, your company could be at risk of hefty fines and penalties.

Let's skip the hype then and get down to what works best in terms of practices and processes. Here is a quick look at 10 things you can do to help you govern and manage your data in the most practical sense.

1. Check for Hidden Constraints

There is a natural tendency to consider work constraints but overlook everything else.

"We tend to focus on all of the facets of the work: data ownership, access, security, quality, and so on," says David Allen, Senior Director of Developer Relations at Neo4j, a producer of a graph database management system. "However, all those things are constrained by the company context they reside in such as data owners who are organizational actors with incentives, pressures, challenges, limitations, and so forth."

So where else should you look to find constraints on your efforts to manage and govern data?

"In short, pay some, but not too much attention to frameworks and technologies — and never lose sight of the human and organizational element. The practitioner's job is to do the best they can within a real context, and that almost always looks different than what the textbooks say," Allen adds.

2. Balance the Conflicts

Managing and governing data is rarely a straightforward, unencumbered exercise. It's usually a mesh of entanglements built of conflicts within and between demands on the business.

"Consumers are simultaneously requesting personalization and privacy, and that's why businesses are now placing much more value in their own customer data," says Keyvan Mohajer, CEO of SoundHound, an audio and speech recognition company that develops speech recognition, natural language understanding, sound recognition and search technologies. "First-party data allows brands to create great experiences, but it also puts them in control when it comes to data transparency and privacy."

Data management and governance becomes much trickier when you lose full control of the data.

"Brands looking to use voice AI are becoming increasingly aware of the risk of handing this data control over to Big Tech voice assistant providers. Having an intermediary not only obstructs a business' view of valuable user feedback, but it also prevents them from reassuring customers about what their data is used for — and allowing them to opt out," Mohajer adds.

3. Track Data Lineage

Given deep fake attacks and increasing regulatory demands, it's better to know the origin and the trail for every data set, if not every data point. Without a clear and uncorrupted data trail, you'll never know whether the data is trustworthy — and neither will auditors, cybersecurity pros, or regulators.

"Less than one third of companies are able to trace their data to the source and ensure that it's visible to only the authorized parties. At scale this requires 'guardrails,' basically reinforcement mechanisms, to combat and prevent regulatory lapses, while still enabling you to use AI to make workflows more efficient," says Seth Dobrin, IBM's Global Chief AI Officer.

"These are not insignificant challenges and solving them requires five key technological building blocks to help simplify how we integrate and improve data management and governance: AI-augmented data cataloging, automated metadata generation, automated governance, data virtualization, and reporting and auditing," he adds.

4. Consider a 'Product Management' Approach

Organizing data into safe and servable portions per domain use can be a practical way to managing it well.

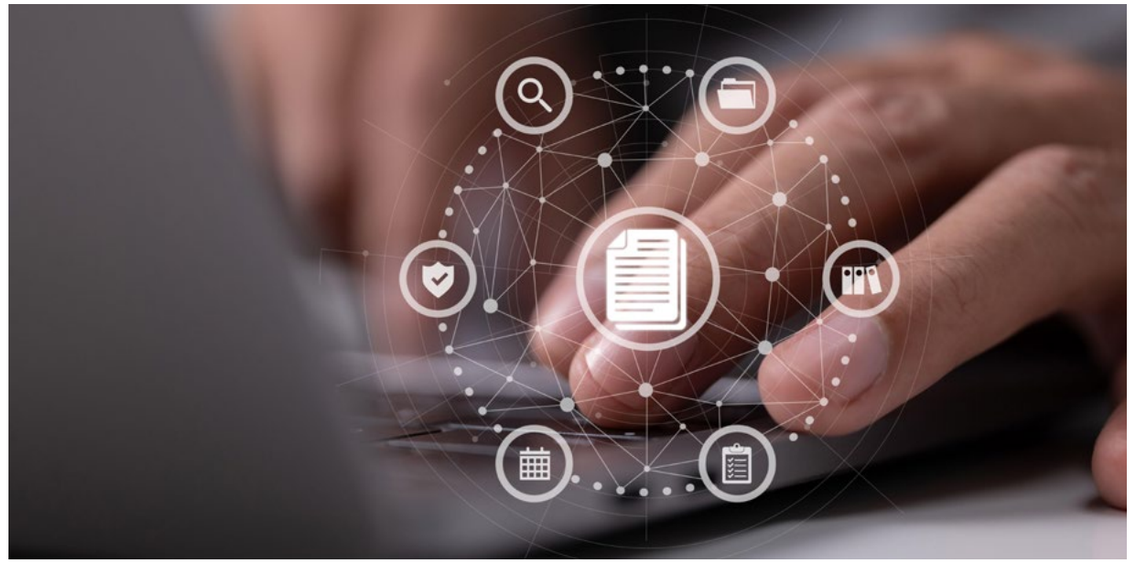
"Data management is increasingly becoming more of a 'product management' practice — curated data sets, built from a number of data sources from across application and business areas become data domains that benefit from the formal requirements gathering, roadmap planning, quality assurance, build automation, and ongoing change management associated with more traditional product development practices," says James Fairweather, Chief Innovation Officer of Pitney Bowes, a 100+ year old, global shipping and mailing company.

"For example, Pitney Bowes has begun building data domains using concepts associated with data fabric and data virtualization to provide well curated data products for use in analytics, data science modeling, and reporting." Fairweather says his company uses "tools like SelectStar for data governance, and MonteCarlo to detect anomalies by improving data observability in our pipelines."

5. Know Thy Data Extremely Well

Yes, data is huge and getting bigger. Yes, it's pouring in from an ever-growing number of sources. Even so, you must understand it well and truly know what information your company has.

"The best thing corporations can do to manage and govern their data is to intimately know their data," says Chida Sadayappan, Cloud AI/ML Offering Leader, at Deloitte Consulting. "Understanding data creation, processing, consumption, and retention will help them find appropriate tools and processes to manage and govern their data well."



6. Don't Forget Data Coming Out the Other Side

Companies tend to think of managing data to be ingested and analyzed. But data coming out of the analysis also must be managed, governed and its lineage clearly documented. In other words, make sure you're managing ALL the data — not just some of it. Unfortunately, that can be quite the challenge.

"Make sure you are taking the time to regularly engage with and understand exactly how your users are currently accessing and utilizing your data," says Christopher Goranson, service professor at Carnegie Mellon University's Heinz College. "Understand what they do with the data once they access it — do they aggregate it further? Do they combine it with other datasets? Can they understand what the data represents, and any limitations based on your existing documentation? If

your organization provides publicly accessible datasets, how are those used? What questions are they trying to answer?"

"These can often be clues you can use to improve the value of the data you manage to your organization," Goranson explains.

7. Connect the Fragments

Complying with data privacy regulations can break chains of knowledge needed to resolve pressing issues. Consider using technologies that can protect privacy without fragmenting shared data chains needed for collective wins.

"A fundamental issue in data governance is the fragmented nature of data across multiple silos — both internally across borders and externally between firms," says Michael Hughes, Chief Business Officer at Duality Technologies, a provider of Privacy Enhancing Technologies (PETs). "This creates a challenge for enterprises that need to share and collaborate on this data to derive insights,"

"Banks, for example, rely on collaboration in the fight against fraud, cybercrimes, and money laundering because data exists across providers and jurisdictions. Healthcare research also depends on the sharing of clinical and genomic data to advance treatments. The problem is they can only share data if they can preserve privacy and confidentiality, while maintaining compliance in an increasingly complex regulatory environment and many existing approaches fall short," Hughes adds.

8. Always Name the Problem

As the adage goes, you can't manage it unless you can manage it. However, you can't measure it either, unless you can name it. In other words, to err is to be vague. To name it is to define it.

"The easy parts of the equation are financing the governance process and the creation of data management policies," says Stefan Thorpe, Chief Engineering Officer at Cherre, a real estate data and analytics platform based in New York. "The real challenge comes from enforcing the data management policies, especially when the enterprise is relatively complex in structure. Even simple tasks such as defining and monitoring key performance metrics can be complex when the processes are not well-defined."

9. Remove the Blinders, Bring on More Eyes

AI can do a lot but it can't outright replace human workers. At least not yet.

"Data governance is essential to any organization's data blueprint," says Manish Sood, Founder and CTO of Reltio, a master data management (MDM) platform. "One of the ways to ensure better governance is by finding ways to put data into the hands of more users but doing so with processes that scale with an organization and create that alignment across teams. It's simple: the more eyeballs on the data, the better the quality and the more thorough the governance. Or put in even simpler terms, you don't fix what you can't see."

10. Send More Data to the Morgue

Ok, not to the morgue exactly, but certainly to cheaper cold storage. In other words, data is hot until it's not and there's no reason to keep it in a warmer when its fine chilled.

"Be aggressive in culling data that you don't need. Also, minimize the amount of data that is stored in expensive 'hot' or 'warm' storage. Kick things that you need to keep to cheap 'cold' storage as soon as you can," says Matt Shea, Head of Federal at MixMode AI, an AI powered cybersecurity platform.