Infosys®
Navigate your next

VIEWPOINT

OVERCOMING THE DATA PRIVACY CHALLENGE

Large corporates are in a bind. Data is becoming unmanageable even as regulators enforce strict demands on what data firms have and how it is used. But those that do data privacy well will be more trustworthy, gain more customers, mitigate reputational risk and set the benchmark for new privacy laws in the future.

Infosys® | Knowledge Institute

Systems within large corporates are often a black box. Data is created, stored and consumed with little understanding of how it moves between systems and for what purpose. This data sprawl will only get worse as edge devices come online. One estimate counts these devices in the tens of billions.[1]

This seemingly indiscriminate use of data is a great cause of concern for consumers. More than 90% are concerned about online privacy, and almost half limit the amount of business they do online because they don't trust big business.[2]

The Cambridge Analytica scandal in 2016 showed how deep the problems run at some firms. Facebook data of millions of users was harvested to build psychological profiles that may have influenced some key decisions that year. Anti-competitive practices around data usage are another big theme that have data privacy ramifications. Facebook has recently been instructed to stop combining data across its suite of products, with Amazon also in the crosshairs of regulators.[3] The fear is that this concentration of power means these companies are regulating how customers access information on the Web while tracking and profiling them to predict and influence how they behave.

A new slate of regulations has arisen in response. Customers now have the right to "be forgotten" and can transfer their data to another person. Very soon, they will have the right to exercise even more control, potentially using their online personality to trade their data for cash.[4]

This comes at a time when hackers continue to get better. Advanced analytics and machine learning can expose personally identifiable information, which can be used to steal identities and worse. An algorithm was recently shown to identify 99.98% of Americans by knowing just 15 demographic attributes per person.[5]

High profile breaches at Yahoo!, Equifax and Marriot have seen an average drop in stock value of 7.5%

For firms, reputational risk is a top strategic priority. The high-profile breaches of Yahoo! (2016), Equifax (2017) and Marriott (2018) had a widespread impact on customers, with a mean of 257 million individuals directly affected by each breach. On average, these breaches cost each company $347 million in legal fees, penalties and remediation costs, with an average drop in stock value of 7.5%.[6]

Black box systems

Firms must very quickly understand what data they have and how it is used to remain compliant, competitive and trustworthy.

Today, only 15% of the executives believe that data is helping grow their business, according to a recent survey.[7] Over four-fifths of executives are unhappy with how data travels through their organization. Overcoming this hurdle is a challenge. Firms buy and sell business units, adding to the complexity of systems. Some data lakes don't have standardized metadata, with different ways of classifying information in the same system. Shadow IT has exploded with the advent of cloud computing. Many people hoard their own versions of applications and databases without central governance. Documenting this data sprawl in real time is tricky. When asked by audit offices for financial reports, firms struggle to consolidate what data they have, using ad hoc manual accounting that is often out of date.

Often, firms don't even know how compliant they are with big regulations such as the General Data Protection Regulation and the California Consumer Privacy Act, and they end up paying up to 4% of revenues when they miss the mark. In the run-up to the CCPA, just 12% of firms had reached an adequate level of compliance. [8]

> A good data privacy strategy gets inside this black box and, when done well, acts as a business driver

Most enterprises implement temporary controls to manage data, with the best ones plowing into yearlong efforts to merge disparate datasets. These companies hope that consolidating data will generate cost savings and find new insights about customers. However, privacy is often an afterthought, and the way data is being probed to get the insights also remains hazy, with firms unaware of what data is being processed, or when and how.

A good data privacy strategy gets inside this black box and, when done well, acts as a business driver. It can be used to ensure that data is processed fairly and will enable firms to quickly tell regulators, business partners and even the media how personal data is used.

Inside the black box

Data is created, stored and then consumed by end users. Data privacy looks at each stage, in turn, to uncover what experts call "data lineage." How data travels across multiple systems then becomes apparent, and emerging technologies such as automation and AI can be used to reduce customer vulnerability, thwart attackers and ensure compliance.

Data creation

The GDPR requires data to have a business purpose. Many new regulations are coming with this caveat. It also says that all entry points into a system must have a data privacy angle, including the right to nondiscrimination and the right to disclosure.

One consumer packaged goods business Infosys worked with had over 300 websites for capturing customer information. Data resided in different geographies. Each geography chose its own form to capture information, with reasons ranging from compliance and reporting purposes to upselling opportunities. This data sprawl increased the number of input points to the system, magnifying the chance that data could be stolen or misused. There was limited understanding of whether the data served a significant business purpose, increasing the chance of noncompliance.

The firm consolidated these assets and built privacy into its working model. Data without purpose was deleted or masked using AI. All forms for capturing data were standardized and a governance model was put in place around data capture. Many manual activities were automated to reduce the number of inconsistencies and reduce the amount of data sprawl.

Data storage

Is data in the cloud? Which country holds backup data, and what regulation is in place in that domain? Is encryption in place when data is migrated? Firms must know how controlled and secure the mechanisms are for protecting data. Even if a cloud provider has secure storage facilities, such as Amazon Web Services, due diligence must be carried out by the data owner to ensure that corporate strategies, industry best practices and regulations are adhered to. To help in this, large corporates should understand where data is located and how sensitive that data is in different geolocations. Personal data should be identified and data sovereignty policies enforced. Client-based encryption is a good practice as it reduces the risk of data loss if a cloud service is hacked. It also cuts the risk of losing data in transit.

Data consumption

Retail companies often share customer data with market analysts, and health care organizations share patient information with medical researchers. This data must be "private by design." Code must be written that captures user consent, and technologies such as data masking should be used to hide unique attributes of users, such as Social Security information. Data can also be pseudonymized, anonymized or even deleted at this stage of the data life cycle.

Encryption, tamper-proof logs and secure authorization should also be used as a "zero trust" measure. This paradigm means that data privacy isn't a one-time activity. Every time a new product is launched or a firm onboards a new customer, compliance and regulation norms must be adhered to.

AI as a protective weapon

An insight into data sprawl and how data is created, stored and consumed opens up use cases for AI.

AI, using fuzzy logic, can be used to run through software and find PII. It can then prepare and cleanse data through entity resolution, auto cleansing, deletion or data masking. With the rise of chatbots and automated call centers, AI detects when customers are sharing PII and either masks it immediately or deletes it.

AI can also be used to surface system vulnerabilities and patch code in real time. This is very similar to how AI is used in the fraud analytics space at major banks. For instance, businesses can use network intrusion technologies to identify points of failure in business architecture and patch software before a breach happens.

AI is also used to reduce the amount of data that firms process and retain. The SkyTeam alliance, made up of major airline providers, is one such use case. Delta Air Lines Inc. and Air France-KLM use AI to comply with GDPR; if they send baggage information that contains PII, the AI tool immediately strips that data away. The tool also monitors cybersecurity threats in real time, detecting abnormal patterns in less than a second.[9]

What firms need to do now

These proactive data privacy practices also act as a catalyst toward competitive advantage. Those that do them well can reduce the chance of reputational damage. They can also increase security, data effectiveness and customer trust.

> Companies can remain compliant across jurisdictions by working together with governments

First, firms must instate a data privacy officer who is accountable when breaches occur. This person must also ensure there's consistency of data standards in policies and procedures for capturing and storing data. Second, the amount spent on new technology must be based on the volume and complexity of data. Controls must be linked to the risk profile of assets and up-to-date appraisals of security gaps. Third, DevSecOps should be used to bring security and privacy functions closer together. Privacy and security experts must make up software development teams and be embedded into the software life cycle as early as possible. This emphasis on privacy and security by design also enables firms to continuously track their attack surface, conducting simulations where hackers actively probe networks and monitor the security posture.

Once firms are on par with the best in the business, they can think about adding to regulatory policies and provisions. Big tech companies such as Microsoft and Apple have taken the lead to show support for privacy protection of customers, increasing customer trust in the process. Microsoft extended the CCPA's consumer rights to all of its U.S. customers and the GDPR's data subject rights to customers across the globe. iPhone advertisements at the 2019 World Series were not focused on processing speed or camera ability, but rather on data privacy.[10]

By working together with governments, companies can remain compliant across jurisdictions where data is stored. One federal law might work across America, which, combined with GDPR, could birth a holistic data privacy policy to ease firms into the future.

Whatever happens, firms must now keep customers front and center. By focusing on customer privacy, compliance will come as a matter of course. End users now have the ability to quickly jump ship and move their business to a new company if trust is diminished. While firms should give them the capability to ease this transition, they should give them no reason to do so.

References

1. ‘Internet of things’ connected devices to almost triple to over 38 billion units by 2020, Juniper Research
2. Consumer-data privacy and personalization at scale: How leading retailers and consumer brands can strategize for both, McKinsey
3. German Court Rules Against Facebook on Data Protection, The Wall Street Journal
4. Why It’s So Hard for Users to Control Their Data, Harvard Business Review
5. Estimating the success of re-identifications in incomplete datasets using generative models, nature communications
6. Companies’ Stock Value Dropped 7.5% after Data Breaches, Infosecurity
7. Organizations need to improve data protection and compliance protocols, Information Age
8. Study Shows Only 12% of Companies Are Ready For New CCPA Data Privacy Regulation, CPO Magazine
9. AI Helps SkyTeam Comply with EU Privacy Rules, WSJ Pro
10. 9 Data Privacy Trends to Watch in 2020, Focal Point Insights

Authors

Gaurav Bhandari

AVP and Head of Consulting, Data & Analytics – Infosys
Gaurav_Bhandari@infosys.com

Harry Keir Hughes

Senior Consultant – Infosys Knowledge Institute
Harrykeir.Hughes@infosys.com

About Infosys Knowledge Institute

The Infosys Knowledge Institute helps industry leaders develop a deeper understanding of business and technology trends through compelling thought leadership. Our researchers and subject matter experts provide a fact base that aids decision making on critical business and technology issues.
To view our research, visit Infosys Knowledge Institute at infosys.com/IKI

For more information, contact askus@infosys.com

Infosys®
Navigate your next