Data Security v/s Data Usage

Merrimack College

Data Governance, Laws and Ethics

Amol Gote

April 5th 2020

"The world's most valuable resource is no longer oil, but data." (The Economist 2017). Growth in the amount of data that is being collected has grown exponentially. In 2020, each person on earth will generate an average of about 1.7 MB of data per second and there will be around 40 trillion gigabytes of data (40 zettabytes). Enterprises across the world, big or small collect as much information as possible from their customers, so that they can define their advertising, marketing, product, customer support and other strategies based on data insights. This data makes these enterprises powerful, but with power there comes the responsibility of securing this vast amount of data. It has become a big challenge across all these corporate organizations on how to leverage the maximum potential of collected data through various analyses, but in doing so guarantee that the data is secure. People may argue that maintaining harmony between these two sets of activities is difficult, but in my opinion, there is a way to maintain the balance between these two processes. A lot of the data-driven companies like Amazon, Apple and others have benefited by maintaining the balance.

Typically in an organization, there is a chief data officer (CDO) or Chief information officer (CIO) who is responsible for maintaining the overall organization's data safe and secure. There are other executives from different business units within an organization who need to access the data for their data scientists to analyze the data associated with their business or application. The data insights could provide the executives inputs for deriving certain business decisions, it could help them analyze any particular issue and take corrective action to resolve the issue. It helps them in identifying patterns that can help them in organizing the inventory or adjust production quantities. It could help in deciding sales targets for a group or an individual. There are tons of benefits that can be derived from the existing data, a particular application might be collecting, so it is of prime importance that the CIO provides appropriate data access to the respective business units.

CIO's organization in process of providing secure data access must avoid any loopholes which will expose the customer confidential data, make sure proper access controls are in place, identify safeguards to protect data, maintain data quality and accuracy. To perform all these activities a data governance framework can be beneficial. *"There are four ways in which data governance and enterprise data management boost cybersecurity those are Identify data at risk, Identify and classify sensitive data, Identify sensitive data users for ensuring consistent data access processes and Ensure safer access to sensitive data." (ISACA Journal May 2017).* To secure data they first need to identify which one to secure, this is where data classification can be helpful. "*Data classification is the act of placing data into categories that will dictate the level of internal controls to protect that data against theft, compromise, and inappropriate use.*" *(www.educause.edu)*. To perform data classification, organizations can adopt the FIPS-199(Federal Information Processing Standard) framework, which is adopted by the US federal government and developed by NIST (National Institute of Standards and Technology). This framework can be easily understood, adopted and implemented. It is based on two components, the first one is security objectives (Confidentiality, Integrity and Availability) and the second one is potential impacts (Low, Moderate and High). To maintain appropriate data access RBAC (role-based access control) can be implemented. Apart from this, the CIO office also needs to warrant that they abide by all the privacy laws depending upon the industry they are aligned to, these include FTC Act (Federal Trade Commission Act), HIPPA (Health Insurance Portability and Accountability Act), Gramm–Leach–Bliley Act (Financial Services Modernization Act). Recent laws include the GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act). In a nutshell, organizations have to perform a lot of activities to provide safer data access. To prove that the perfect balancing act can lead to greater success of an organization lets illustrate some industry specific examples.

Let’s consider first the data breach example, First American Financial, one of the largest title insurers in the US, reportedly exposed more than 885 million sensitive documents online. Exposed sensitive information included bank/credit union account numbers, mortgage records, tax documents, wire transfer receipts, Social Security numbers and photos of drivers’ licenses. The breach occurred because of an incorrect web application design and without any security authentication guards in place. “*Essentially, a link to a webpage with sensitive information is created and intended to only be seen by a specific party, but there is no method to verify the identity of who is viewing the link. As a result, anyone who discovers a link to one document can view it—and can discover any of the other documents hosted on the site by simply modifying the link by incrementing the number.” (Forbes May 2019).* The company took immediate action to address the situation and shut down external access to the application. The major issue with this breach was unauthorized access and insufficient process validation which happens when a web application does not enforce appropriate business rules. In this case, a security team needs to ensure a flow control, which guarantees that each step is performed in a specific order by end user. Appropriate data governance framework especially ***Pillar 3 – Privacy, Compliance and Security*** would have been of significance over here, to avoid the breach. It talks about protecting sensitive data through access management and enforces regulatory and compliance requirements which could have helped in discovering the security lapse.

Another example that needs to be cited is Amazon, Amazon has over 750,000 employees, the largest online retailer in the world and is considered one of the big four technology companies along with Apple, Microsoft and Google. There are no notable data breaches that have happened with Amazon, barring the news around 2018 which has not been substantiated. Amazon provides recommendations in a personalized manner to its customers through various mechanisms like frequently bought together, recommendations for you based on your shopping history, recently viewed items, best-selling, etc. To come up with so many recommendations there is a lot of number crunching, data analysis of existing customer data, purchase history, products liked, viewed items and various other parameters. There would be multiple teams that build these features for their e-commerce platform, these teams need access to that data for analysis which amazon been able to provide with sound data governance practices. The same set of consumer data is required across various other business units at Amazon such as shipping, billing, and customer service. Considering the scale at which Amazon operates and that it has maintained an excellent compliance record with no data breaches is commendable. This is an excellent example of how overall profitability can be increased by using customer data and also keeping it secure. Amazon has been able to proliferate its data governance practice to other business units like its cloud business, where its customers can leverage the same data governance framework.

Apple is another shining star in this space, Apple essentially has a lot more personal data, be it photos or documents on iCloud, credit card information, and all the health related data that it collects from its wearable technology devices. There is a plethora of data that Apple collects from its consumer, but so far there has never been any data leaks with an exception of celeb-gate. Consumer privacy is in Apple’s DNA and Apple pay is just a prime example. When you add a credit card to Apple Pay, that information that you enter on your device is encrypted and sent to Apple servers to be passed to the payment provider. After your card is approved, your bank creates a device-specific Device Account Number, encrypts it, and sends it along with other data (such as the key used to generate dynamic security codes that are unique to each transaction) to Apple. The Device Account Number can't be decrypted by Apple but is stored in the Secure Element—an industry-standard, certified chip designed to store your payment information safely. Apple doesn't store or have access to the original card numbers of credit, debit, or prepaid cards that you add to Apple Pay. Apple Pay stores only a portion of your actual card numbers and a portion of your Device Account Numbers, along with a card description. It doesn't store or have access to the original card numbers of credit, debit, or prepaid cards that you add to Apple Pay. Apple Pay stores only a portion of your actual card numbers and a portion of your Device Account Numbers, along with a card description. This form of data governance makes payment processes more secure because the user's financial information is not available to Apple or the merchant at any point. This way Apple mitigates the risk of storing user's sensitive financial information, which in turn would make it easy for it to distribute Apple pay data internally for all kinds of analysis including fraud analysis. Apart from this, Apple collects lots of data related to its applications with which it can discover how people are using apps and alter future designs. The same application related data concept can be applied to the devices it manufactures like iPhones, watches, MacBook's, etc. The Apple Watch is another example of how it safeguards customer data and uses it to its benefit, by providing data to researchers through a research app. It offers researchers the ability to conduct large-scale health studies in a way that hasn't been possible. The approach of sharing personal health data is in user's control, data is encrypted, is not sold and that research studies have to inform users how their data will support the research. Participants also can withdraw at any time. *"Privacy is a fundamental human right. It's also one of the core values. Your devices are important to so many parts of your life. What you share from those experiences, and who you share it with, should be up to you. We design Apple products to protect your privacy and give you control over your information. It's not always easy. But that's the kind of innovation we believe in."* *(Apple Privacy Statement).* In summary, Apple has achieved the perfect agreement between data usage and data security because of that it is one of the top 5 most valuable companies in the world.

Finally, to summarize based on the examples mentioned above, organizations can collect data, draw insights from it for its own profits and in doing so warrant that data is secure and protected. Some exceptions cause data breaches, but most often than not these organizations take corrective actions. Finding the perfect balance is difficult and challenging, but with a data governance framework in place, this could be easy to achieve with the right set of technology, tools, processes and most importantly the right personnel.

Bibliography

- Data is the new oil
    - https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

- Big data statistics
    - https://kommandotech.com/statistics/big-data-statistics/
    - https://hostingtribunal.com/blog/big-data-stats/#gref

- CIO role
    - https://www.cio.com/article/3234884/what-is-a-chief-data-officer.html
    - https://www.isaca.org/resources/isaca-journal/issues/2017/volume-3/boosting-cyber-security-with-data-governance-and-enterprise-data-management

- Data classification
    - https://www.educause.edu/focus-areas-and-initiatives/policy-and-security/cybersecurity-program/resources/information-security-guide/toolkits/electronic-records-management-toolkit

- First American Financial breach
    - https://www.forbes.com/sites/ajdellinger/2019/05/26/understanding-the-first-american-financial-data-leak-how-did-it-happen-and-what-does-it-mean/#5449a1e5567f
    - https://www.bankinfosecurity.com/report-sec-investigates-first-american-data-exposure-a-12910

- Amazon
    - www.amazon.com, https://en.wikipedia.org/wiki/Amazon_(company)

- Amazon recommendation engine
    - http://rejoiner.com/resources/amazon-recommendations-secret-selling-online/
    - https://www.martechadvisor.com/articles/customer-experience-2/recommendation-engines-how-amazon-and-netflix-are-winning-the-personalization-battle/

- AWS governance at scale
    - https://d1.awsstatic.com/whitepapers/Security/AWS_Governance_at_Scale.pdf

- Apple Pay
    - https://support.apple.com/en-us/HT203027

- Apple Watch Research App
    - https://techcrunch.com/2019/11/14/apple-research-app-arrives-on-iphone-and-apple-watch-with-three-opt-in-health-studies/