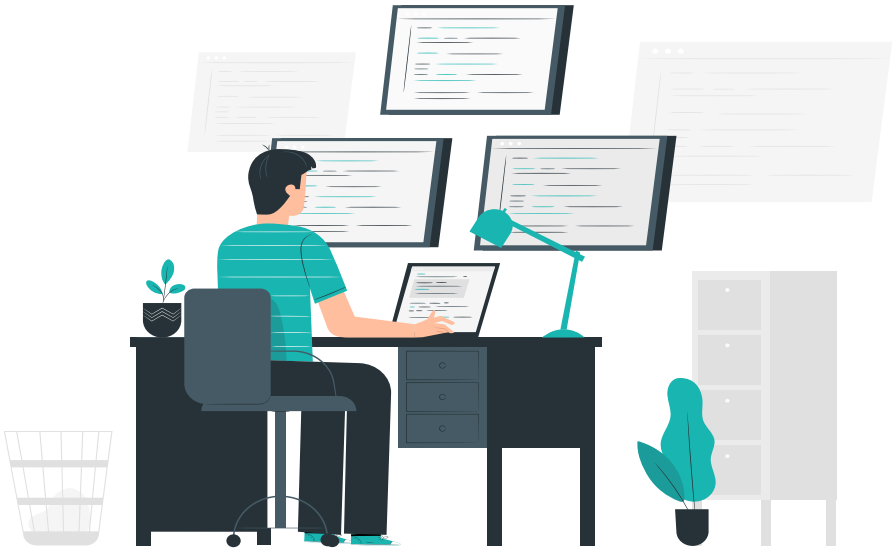

Applied Research project

Merrimack College

Data Governance, Laws and Ethics

Amol Gote

9th May 2020

ABC is a consumer personal loan financing company specializing in loans for dentistry treatments. It connects dentist networks, patients, and banks. The Customer applies for a personal loan through a mobile application, selects one of the loan options approved by credit decisioning engine, verifies his identity, and the loan is approved and funded to the customer. During the loan life cycle of the application process and servicing (loan repayment), ABC is collecting and generating a vast amount of data.

Data catalog and data capture

ABC's data catalog consists of five categories, three categories dealing with consumer data are customers PII Information, credit history, and identity verification documents. There are other two categories financial transactions for loan repayments and internal application logs for APIs and mobile application. ABC has classification defined but in a very rudimentary way by classifying data in these three buckets Confidential, Private, and Public with no detail impact analysis. Three categories related to customer's data are classified as Confidential and remaining two as Private.

Desc of Data / Data Asset	Data Category	Classification (define classification)	Impact on Loss or Leakage of data (low, medium, high)	Source	Collected By (system or group)	Collection Method (service, export)	Type of Format (JSON, CSV, etc)	Used By (group, application)	Purpose of Collection	Transfer to (if shared)	Security Control During Data Transfer	Data Repository Format (where is data stored)	Storage or Data Retention Site (physical location, domicile)	Disclosed To (which roles will have access)	Retention Policy
Customer Information (Name, SSN, Date Of Birth, Moble, Annual Income)	Customer Personal PII Information	Confidential	High	Mobile Application	Mobile Application	API	JSON	API Team, Data Analytics Team	Provide a Personal loan		HTTPS/SSL Key Field like SSN Encrypted in Database.	Database - MySql	Cloud - AWS - RDS (Releation Database Service)	Database Administrator, Chief Security Officer, Data Analyst.	6 years, then archived.
Credit History	Credit Bureau Report Data	Confidential	High	Credit Bureau - TransUnion, Experian	Credit Decisioning System	API	XML	Credit Team	Credit Decision - To Provide loan to customer	Loan Buying Bank	HTTPS/SSL Encrypted At REST	AWS S3 Bucket, Database - MySql	Cloud - AWS - S3 Cloud - AWS - RDS	Database Administrator, Credit Analyst, Data Analyst, Chief Credit Officer. Credit Team.	6 years, then archived.
Customer Identity Verification	Customer Identity Biometric Information Includes Images of Driver License / Passport	Confidential	High	Mobile Application	Mobile Application	API	JSON, Images (JPEG/PDG/PDF)	API Team, Data Analytics Team	Verify customer Identiy during loan application		HTTPS/SSL Encrypted At REST	AWS S3 Bucket, Database - MySql	Cloud - AWS - S3 Cloud - AWS - RDS	Database Administrator, Chief Security Officer, Fraud Analyst, Compliance Office, Customer Support	Varies by state laws
Monthly Payments Made By Customers	Financial Transactions	Private	Medium	Loan Servicing Vendor (Third Party)	Cron Job	ETL	CSV	Customer, Loan Servicing Team, Customer Support, Data Analytics Team	Loan Repayment details	Loan Underwriting Bank	Encrypted as REST	AWS S3 Bucket, Database - MySql	Cloud - AWS - S3 Cloud - AWS - RDS	Database Administrator, Customer Support, Head Of Transaction Servicing	6 years, then archived.
API Logs, Mobile App Logs	Logs	Private	Low	API, Mobile Application	API, Mobile Application	LogStash - Service Moves logs from Application server to Central Repository, Neurilic (Third Party) - Mobile Application logs	Log File	SRE Team	Application / API Troubleshooting		Internal system to system transfer	Elastic Search	EBS - Elastic Block Storage	SRE Team, Cloud Engineer	60 days and then archived to AWS Glacier for 6 years.

Figure 1: Data Catalog

Refer the attached excel document: *Amol_Gote-Final-ProjectData-Catalog.xlsx*

Data Access

The Mobile application interacts with back end APIs over HTTPs, all the customer data is stored in MySQL database using AWS RDS service. Identity documents captured through a mobile application like scanned copies of Driver's License, Passport are stored in AWS S3 buckets and accessible only through secure APIs. Credit reports are pulled from credit bureaus and the key attributes used for credit decisions are extracted in the database and then the whole credit report is stored as an XML file in the S3 bucket. Back-end jobs and APIs generate application logs that are managed within AWS VPC (Virtual Private Cloud) and are moved internally, then indexed and stored into Elastic search. These application logs are accessible through the Kibana visualization tool to the SRE team and authorized users and roles. Mobile application logs are managed through a third-party service (New Relic) which has a web interface to view the logs. Loan servicing is done through a third-party vendor, there are 2 ways in which ABC integrates with this third-party vendor, firstly through file transfer using S3 buckets and secondly through RESTful APIs. The customer information which is been updated in the mobile application during loan servicing needs to be published to the servicing vendor, which is accomplished through APIs. ABC needs to share financial transaction data to loan underwriting bank, for that ABC leverages AWS S3 buckets to push the files periodically using a cron job. All the APIs implement OAuth token-based authentication, for B2B integration through APIs, two levels of security mechanisms are in place, first IP addresses filtering, and second token-based authentication. Production database access is very restrictive, only DBA and CSO have access to the data store.

Best Practices

ABC finances personal loans, so it must abide by the laws of FTC specifically GLBA. "*The GLBA Safeguards Rule for customers, which include: Private information must be secured against unauthorized access, Customers must be notified of private information sharing between financial institutions and third parties and can opt-out of private information sharing, User activity must be tracked, including any attempts to access protected records." (Groot , 2019).* To enable these safeguards, the organization

needs to assign dedicated CISO/CSO, who maintain data inventory and classification. All customer data must be encrypted at rest or transmitted securely to the external network. Roles based security must be enabled which guarantees access to authorized individuals. Data needs to be retained for a specific time frame as per the regulatory requirement and later needs to be deleted. Monitor and regularly test safeguards through vulnerability or penetration testing of the cloud infrastructure. Organizations in lending business have risk assessment framework, which identifies risks associated with data breaches and plans to mitigate those risks, along with it they also have an incident response plan. Lending companies use credit bureau reports for loan decisions, usage of bureau reports should be restricted within the lending organization and should not be utilized or shared for any other purpose to be FCRA compliant. ABC is fairly new in this business domain and is competing with lending company CareCredit which has been in business for 30 years, it never had any data breach, there is no publicly available information which states any wrongdoing by CareCredit as far as any privacy laws are concerned.

Points of excellence and areas of improvement

ABC is a cloud company it leverages best practices from cloud providers, AWS in this case, to protect their data. ABC has a lot of safeguards established already, it encrypts key PII (Personally Identifiable Information) database fields such as SSN, Bank account numbers, etc., uses encryption for transport (HTTPS) as well as at rest. It receives financial transactional data from loan servicing vendor and shares the same data with its partner banks, in both cases, ABC has a contractual agreement signed on the usage of the customer data. It uses role-based access through AWS IAM roles and IAM users. ABC primarily uses S3 for inbound and outbound file transfers, so in either case, S3 buckets are protected with IAM users, IAM roles, and bucket specific policies. All identity documents stored in S3 are encrypted. All the encryption keys are managed by AWS KMS (Key Management Service) which ensures the periodic key rotation. For all IAM policies, it follows the principle of granting only the permissions required to perform that task (Least Privilege). They have dedicated personnel performing the role of

chief security officer (CSO), who has created a data classification, which classifies data as Confidential, Private, and Public. The reason for this classification was ABC wanted to identify which database columns need to be encrypted. Database access is restrictive and MFA access enabled for root permissions. Despite all these well-established best practices, there are several areas of improvement. ABC has data classification in place, but it lacks a concrete risk framework due to which it does not have a mitigation plan against any unauthorized data access. Application logs potentially contain sensitive PII information as API team is logging the entire payload request and response, so anyone who has access to event log / Kibana console can get hold of some of the PII information. SRE team has access to production Kibana logs which is potentially a risk. Data analytics and data science team get access to a read-only snapshot of the production database, based on a certain frequency, in this snapshot key database columns like SSN, bank account numbers are encrypted, but there are still additional PII attributes like the first name, last name, address, mobile number, etc. that are accessible. ABC does not have an incident response plan in case of the data breach.

Maintaining excellence and areas of improvement.

ABC should continue with the best practices that they have incorporated in their processes and infrastructure, like encryption, secure data access through APIs, etc., but there are opportunities for improvement in certain areas. ABC needs to fix issues related to the logging framework and ensure that all the PII fields that are logged are masked especially SSN, bank account numbers. The Process of providing a read-only snapshot of the production datastore needs to change and need to incorporate data anonymization as part of the process. ABC has classified database attributes as Confidential, Private, and Public, but there is no impact analysis (High/Moderate/Low) against confidentiality, integrity, and availability. As per the NIST framework data classification needs to be defined, below are some of the classifications along with its impact.

Information Types (Domain)				Data Classification
Customer Personal PII Information	This information is collected through a mobile application and is required for loan approval.			Confidential
Financial Transactions	Monthly repayments made by a customer.			Private
Application Logs	API Logs and mobile application logs			Private
Identify Information Types	Confidentiality Impact	Integrity Impact	Availability Impact	
Customer Personal PII Information	High	High	Low	
	This could have a significant impact on customer's trust	This could impact operations and customer support	This information is only required during application processing, later most of the data is encrypted.	
Financial Transactions	High	High	Medium	
	Expose customer's financial information	This could impact the company's earnings	This impacts reporting, daily financial settlements	
Application Logs	Low	Low	Low	
	As long PII data is masked, application logs do not pose many risks	Logs have very less shelf life, also they are used by application troubleshooting	In the case of showstopper issues application logs become critical.	
Final Domain Categorization	High	High	Medium	
	Overall Impact: High			

Table 1: Data classification matrix

ABC has performed vulnerability and penetration testing only once before product launch, it needs to do it periodically and fix issues identified during those tests. ABC has not published a data retention policy document for all the data categories, for certain states biometric information needs to be deleted after a specific time, so they need to create a process that ensures timely data disposal to abide by the state laws. ABC needs to publish an incident response plan, so that in case of real incident designated employees are aware of what activities need to be performed. These actions include addressing the root cause of the data breach, notifying the stakeholders which include consumers, law enforcement, and its partners.

ABC has done a decent job in enabling good data governance practices, but there are still a lot of challenges that are at hand for the newly appointed CIO.

Bibliography

Groot, J. (2019, July 15). *What is GLBA Compliance? Understanding the Data Protection Requirements of the Gramm-Leach-Bliley Act in 2019.* Retrieved from the digital guardian: https://digitalguardian.com/blog/what-glba-compliance-understanding-data-protection-requirements-gramm-leach-bliley-act

Green, A. (2020, March 3). *We Need to Talk About Gramm-Leach-Bliley (GLB): The Safeguards Rule Will Be Changing!* Retrieved from: https://www.varonis.com/blog/we-need-to-talk-about-gramm-leach-bliley-glb-the-safeguards-rule-will-be-changing/

Experian. *Understanding the Fair Credit Reporting Act.* Retrieved from Experian: https://www.experian.com/blogs/ask-experian/credit-education/report-basics/fair-credit-reporting-act-fcra/

Taggart, C. (2019 Nov 21). *GLBA Safeguards Rule updated: FTC defines "financial institutions"*. Retrieved from plantemoran: https://www.plantemoran.com/explore-our-thinking/insight/2019/12/glba-safeguards-rule-updated-ftc-defines-financial-institutions

Skeath, C, (2016, Oct 16). *FTC Issues Guidance for Responding to Data Breaches.* Retrieved from Inside Privacy: https://www.insideprivacy.com/united-states/federal-trade-commission/ftc-issues-guidance-for-responding-to-data-breaches/