INFOLOB

SECURITY CASE STUDY

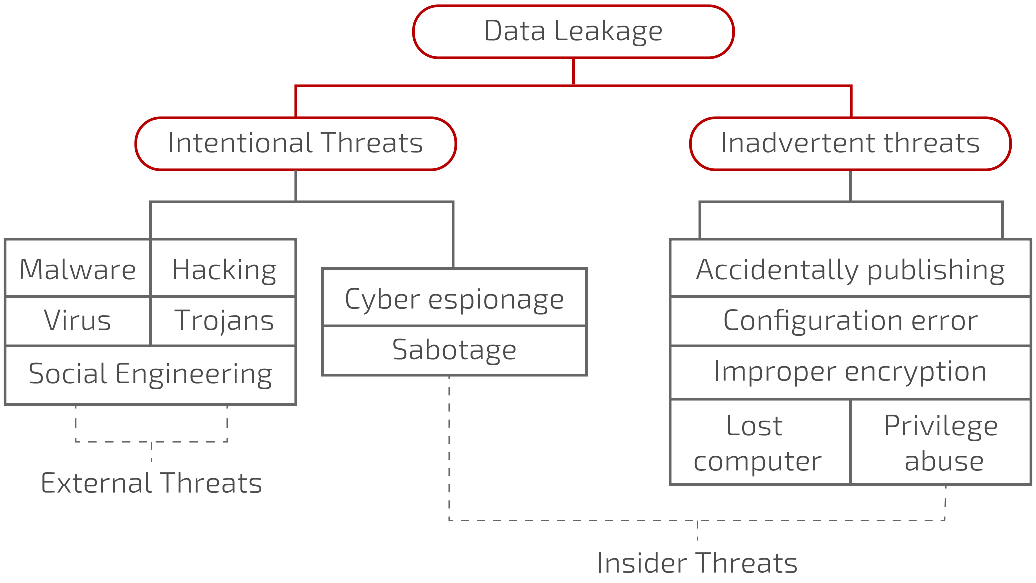
SECURITY CASE STUDY

Abstract:

This case study details Infolob's security program at one of the largest social media companies in the world. Its intent is to provide a blueprint for protecting and securing sensitive enterprise application data. The methods articulated in this study do not represent the totality of the scope involved in the project, but instead provide key considerations to take into account when undertaking a similar exercise.

Challenge:

With the proliferation of data breaches and sophisticated attacks on the rise, it is in the company's best interest to ensure internal data sets are protected from the perspective of both internal and external threats. To achieve maximum protection, the entire technology stack of the enterprise application must adhere to strict security standards and design/architecture. In addition, data governance and security analytics were key areas of concern.



Methodology:

Infolob's approach to improving the organization's data security posture starts with the following pillars:

Discover—identify what personal data is being managed and where it resides.
Manage—govern how personal data is used and accessed.
Protect—establish security controls to prevent, detect, and respond to vulnerabilities and data breaches.
Report—keep required documentation, manage data requests, and provide breach notifications.

This paper will highlight key findings in each of the pillars that helped establish security baselines going forward.

Discover:

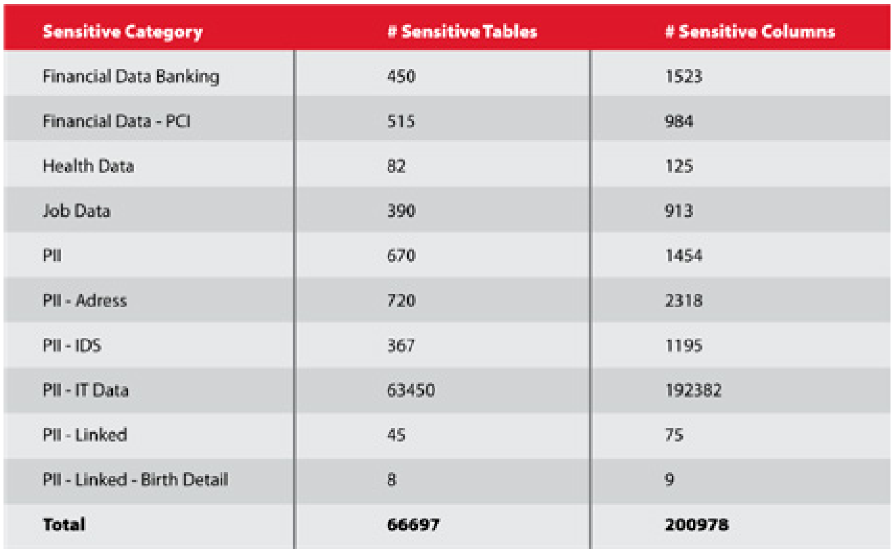
Data security begins with data discovery. In the complex world of free-flowing data, sensitive structured and un-structured data can migrate and settle in a variety of data repositories:

- Heterogeneous RDBMS systems on-prem and in the Cloud
- Big data platforms like Hadoop or DataWarehouse
- Cloud-based file sharing like Dropbox and Google Docs
- File-based, such as spreadsheets, flat-files, emails, PDFs

Before you can protect the data from being compromised, exfiltrated, or destroyed, organizations must first know what data is being stored, where it is being stored, and the context of the data.

For this exercise, Infolob identified PII data across a sensitive RDBMS system. The method for identifying sensitive data was performed in two passes: Identification of sensitive columns through names and comments, and Infolob's own deep data-scan utility which crawls through all table/row information. To identify sensitive data, we applied 70+ regular expressions that address data related to PII, PCI, HIPAA and GDPR.

In the end, we helped identify over 66,000 tables that contained sensitive data in the following categories:



Sensitive Category	# Sensitive Tables	# Sensitive Columns
Financial Data Banking	450	1523
Financial Data - PCI	515	984
Health Data	82	125
Job Data	390	913
PII	670	1454
PII - Adress	720	2318
PII - IDS	367	1195
PII - IT Data	63450	192382
PII - Linked	45	75
PII - Linked - Birth Detail	8	9
Total	**66697**	**200978**

With data discovery information in hand, the next step is to implement a data classification process to tag data according to its type, sensitivity/confidentiality, and cost/value to the organization if altered, stolen, or destroyed.

Manage:

Enterprise data is moving from one location to another at lightning speed and is being stored in countless devices and cloud storage applications. Employees, partners, and customers are accessing this data from any place and at any time, so identifying, locating, and classifying that data in order to protect it is the primary priority of data discovery security solutions.

For the tables/columns identified in the data discovery phase, the key questions that come into play are:

Who is accessing the data and by what means?

To address this question, a thorough examination was performed to identify who has access to the data from an application and database perspective. This included a functional audit from an application standpoint, as well as examination of database roles and user privileges that had access to select, insert, update, and delete the tables in question.

Where does the data reside?

It is important to take inventory of where the master data resides and where it is distributed. For tagging sensitive data, an RDBMS labeling function was adopted to appropriately identify sensitive data going forward. The label mechanism was ideal in automating various protection controls and data-masking functions.

What are the inbound/outbound data flows?

An understanding of how the data was being accessed/updated from an application and database level was needed to ensure each access point was protected as much as possible. In addition, identifying data movement was imperative for a complete data mapping. To make things more complicated, the data destination should adhere to the same type of controls and protection as the source system. In this particular case, data movement was tracked to standby database system used for disaster recovery, DataWarehouse system as part of ETL process, and an analytics system used for business intelligence. Keeping tabs on data flows is an iterative process and very difficult to automate.

Protect:

To protect the data, along with the infrastructure that makes up the enterprise application in question, Infolob developed a comprehensive security program that identified over 130 key security findings in the areas of application, database, network, OS, and operations. We will delve into the most impactful findings and remediation techniques that we used to secure the data.

Application:

The ERP system in question exposes JSP code through the application, whether you use the feature or not. This translates to roughly 16,000 exposed JSP files that can be prone to web app vulnerabilities such as XSS, SQL injection, etc. Infolob recommended implementing a feature to whitelist ONLY the JSP's that are in use. The result is dramatically reducing the attack surface of the application from 16,000 JSP's to 67 JSP's. This represented a 99.58% reduction of the application code base.

Database:

As part of the data discovery exercise, the first phase was to identify the sensitive data in question. The next step was to protect and audit the data. To protect the data, we must first encrypt the data in-transit and at-rest. For in-transit data, a security feature was implemented to ensure data transmission through client, application, and database was encrypted using SSL certificates. This ensured all traffic that passed through the application/database was secure and encrypted. For the data at-rest, it was imperative that a threat actor could not simply scan the database datafile for sensitive data. To address this concern, the RDBMS data encryption mechanism was implemented either at a tablespace or column level. Here is an example of the encryption mechanism in action:

```
1) Unencrypted test when inserting following record to table: 1234123412341234
...
Dump of memory from 0x00007F45FD9E1000 to 0x00007F45FD9E3000
7F45FD9E1000 0000A206 0300000E 000BBC44 06010000  [........D.......]
7F45FD9E1010 0000AF21 00000001 0000583E 000BBC21  [!.......>X..!...]
7F45FD9E1020 00008000 00320002 03000008 00080009  [......2.........]
7F45FD9E1030 000001B2 01401A0B 003D0042 00002001  [......@.B.=.. ..]
7F45FD9E1040 000BBC44 00000000 00000000 00000000  [D...............]
7F45FD9E1050 00000000 00000000 00000000 00000000  [................]
7F45FD9E1060 00000000 00010100 0014FFFF 1F701F84  [.............p.]
7F45FD9E1070 00001F70 1F840001 00000000 00000000  [p...............]
7F45FD9E1080 00000000 00000000 00000000 00000000  [................]
      Repeat 501 times
7F45FD9E2FE0 00000000 00000000 1001012C 34333231  [........,...1234] <<<<<< Data in Clear Text
7F45FD9E2FF0 34333231 34333231 34333231 BC440601  [123412341234..D.] <<<<<< Data in Clear Text
...

2) Enabled Transparent Data Encryption (TDE) and inserted same record to table: 1234123412341234

...
Dump of memory from 0x00007F45FD9E1000 to 0x00007F45FD9E3000
7F45FD9E1000 0000A206 0380000E 000BBC44 06010000  [........D.......]
7F45FD9E1010 00005349 00000001 00005840 000BBC37  [IS......@X..7...]
7F45FD9E1020 00008000 00320002 03800008 00080009  [......2.........]
7F45FD9E1030 000001B2 01401A0B 003F0042 00002001  [......@.B.?.. ..]
7F45FD9E1040 000BBC44 00000000 00000000 00000000  [D...............]
7F45FD9E1050 00000000 00000000 00000000 00000000  [................]
7F45FD9E1060 00000000 00010100 0014FFFF 1F3C1F50  [............P.<.]
7F45FD9E1070 00001F3C 1F500001 00000000 00000000  [<.....P.........]
7F45FD9E1080 00000000 00000000 00000000 00000000  [................]
      Repeat 498 times
7F45FD9E2FB0 00000000 4401012C FB5062D6 F34BBD00  [....,..D.bP...K.]
7F45FD9E2FC0 C27C6B67 E62CF6B2 15DF3836 04E7A5F9  [gk|...,.68......]
7F45FD9E2FD0 6C78EC9B 09B65B63 B03D5029 ACFDD9E7  [..xlc[..)P=.....]
7F45FD9E2FE0 E7AECB4A C4D20006 89654D12 CDC2E982  [J........Me.....]
7F45FD9E2FF0 BDA67ADC C5526EB9 D6E8BFFB BC440601  [.z...nR.......D.]
...
<<<<<< Did not find string '1234123412341234'
```

In addition to encryption, we must also audit and log all operations performed against the sensitive data. To achieve this, the Fine Grained Auditing (FGA) feature was implemented to track who accessed the table and what operation was performed. The FGA feature was customized so that the events being tracked were sent to the SYSLOG messages. Here is an example of a sample event:

```
SELECT on data:
Apr  9 10:53:56 infoebs01 Oracle Audit[11167]: LENGTH: "347" SESSIONID:[7] "3158797" ENTRYID:[1] "1" STATEMENT:[1] "1"
USERID:[6] "SYSTEM" USERHOST:[9] "host1" TERMINAL:[5] "pts/1" ACTION:[3] "100" RETURNCODE:[1] "0"
COMMENT$TEXT:[101] "Authenticated by: DATABASE; Client address:
(ADDRESS=(PROTOCOL=tcp)(HOST=192.x.x.x)(PORT=28791))" OS$USERID:[7] "applmgr" DBID:[9] "487768547"
PRIV$USED:[1] "5"
Apr  9 10:54:29 localhost.localdomain  Oracle Audit[11167] Object Direct Access:
SESSIONID=3158797;OBJ=USER3.ACNT_CONTRACT;SESSION_USER=SYSTEM;SQL=select * from USER3.ACNT_CONTRACT

UPDATE on data:
Apr  9 10:56:33 localhost.localdomain  Oracle Audit[11389] Object Direct Access:
SESSIONID=3158799;OBJ=USER3.ACNT_CONTRACT;SESSION_USER=USER3;SQL=UPDATE ACNT_CONTRACT SET AMOUNT
= AMOUNT*2, OPERATOR = CASE WHEN SUBSTR(SYSTIMESTAMP, 23, 1) > '5' THEN 'Peterson' ELSE 'Smith' END,
LAST_CHANGE = SYSTIMESTAMP WHERE CONTRACT_ID = :B1
```

The audit data proves to be valuable when identifying a data breach or suspicious activities.

Network:

From a systems architecture point of view, the network topology must be designed in a way to protect key access points and limit traffic to an explicit set of hosts and ports. Infolob's contributions led to recommendations on VLAN segmentation design, use of bastion hosts and database listener configuration for SSL, and whitelisting of hosts that could connect to the database.

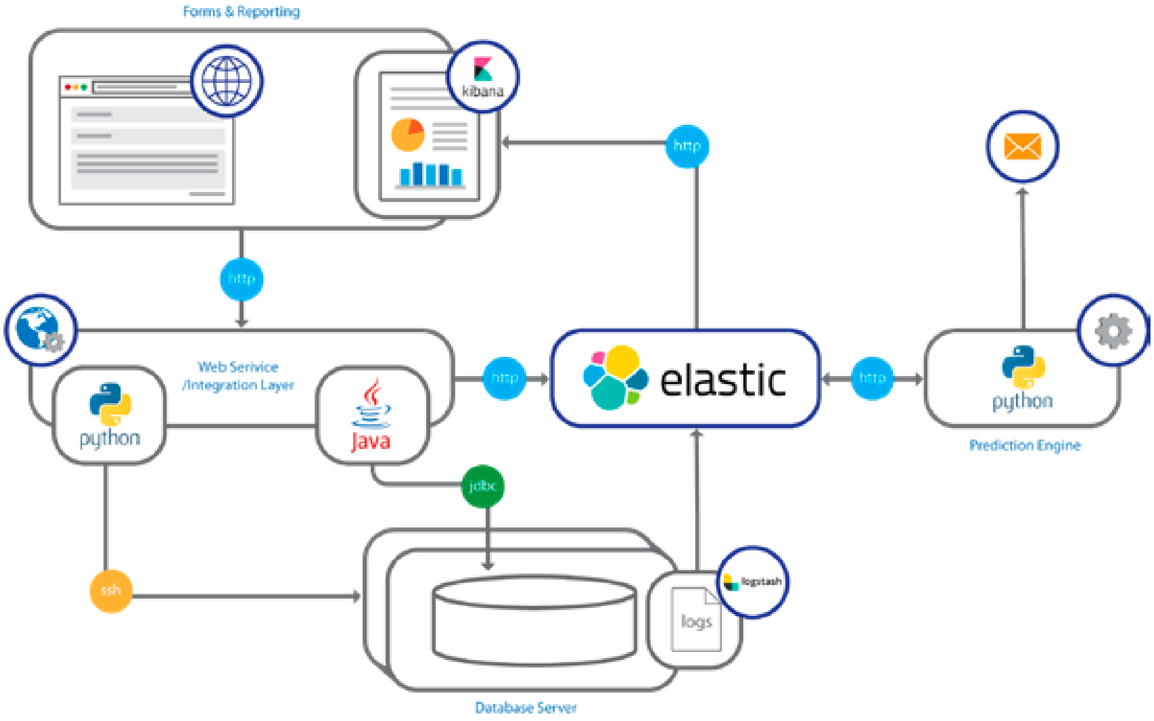
OS:

For best-practice approach in OS hardening, CIS (https://www.cisecurity.org/cis-benchmarks) and NIST (https://nvd.nist.gov/ncp/repository) benchmark guides were used to set proper baseline standards.

Report:

While big data can clearly lead to cybersecurity improvement, it is often challenging to handle. From our database auditing example earlier, how do you improve the signal to noise ratio of events? Adding machine learning into the equation might just be the answer to using big data more effectively and improving cybersecurity beyond measure. Machine learning solutions can quickly scan data to generate a picture of historical patterns of positive and negative behaviors. Businesses can use these capabilities to detect vulnerabilities, identify a breach as it's happening, and correlate information from multiple sources. By uniting these tools, organizations can successfully thwart attacks and decrease the chance of experiencing breaches.

For security analytics, Infolob recommended the following open source solution and design:



Conclusion:

Whether it takes the form of intellectual property, healthcare, financial, credit card, or customer information, data is the lifeblood of the enterprise. To protect this data, organizations must take a holistic approach in securing the infrastructure, application, and database.

Allow Infolob to help your organization build and run a successful cybersecurity program and help mitigate your risks.

For more information about Infolob's services, visit www.infolob.com.