**A04 Notes + Case Study**

**Problems Created by AI**

**Machine Learning Biases**

AI Abuses Two Types of Data: Geographic and Demographic Data

- Demographic: Age, Sex, Race

- Geographic: Location

**Card Usage Data**

- Tracks how you spend and can lead to convenient offers or bias.

- Can be positive if more convenient deals can save you money, based on what you typically need and use.

- Can be abused for negative use, if a company sees bad spending habits of unnecessary payments they may deny requests for more credit, even if necessary.

**Lack of Transparency**

- Will not give valid or clear reasons on why certain requests were accepted or rejected.

- Can lead to confusion and more work on both sides

**Data Breaches**

- All data breaches in the last 15 years because of financial motives

- Without safeguards, if AI is given a certain prompt it can lead to a leak of data.

**Case Study: Last 10 Years of Data Breaches**

**Equifax (2017):**

- Major Attack that Affected 148 million Americans

- Compromised Names, Dates of Birth, Social Security Numbers, Driver’s License Numbers, and Credit Card Numbers.

**Cause:**

- Happened due to a Vulnerability in the Open-Source Framework “Apache Struts”, failure to update to a patch that was released for over six months.

- Poor Infrastructure allowed one access point to lead to access of multiple servers, where they were able to find usernames and passwords stored in text files.

-Able to remain undetected for months due to an unrenewed encryption certificate

**Relation to AI:**

No evidence of AI was specifically found in this case, but AI tools can easily simplify this type of attack. AI tools can log and identify released vulnerabilities and try to see if systems are not patched and also identify the text files with usernames and passwords almost instantly once they have access to unencrypted systems.

**Response:**

Equifax quickly released a new website to identify if your data was stolen, which was unsecured and a fake site with a similar domain was linked by Equifax shortly after causing more people to freak out. A settlement was made with over 400$ million dollars to affected customers, and a quick update and re-management of their systems.

**Capital One (2019):**

- Compromised Social Security Numbers, Bank Account Numbers, and Canadian Social Insurance Numbers

**Cause:**

- A Software Engineer accessed an AWS server with the company’s data and stole around one hundred million credit card applications from 2005 – 2019.

- Unsecured cloud access allowed the applications to easily be downloaded and published without any knowledge of a breach even happening.

**Relation to AI:**

No AI was used either in the attack or response.

**Response:**

Quickly worked with the FBI to capture the person who leaked the data, leading to a 190$ Million settlement.

**JPMorgan Chase Bank (2014):**

- Instead of stealing financial information, only contact information for customers was stolen.

- Compromised Names, Email Addresses, and Phone Numbers

**Cause:**

- Cyber Attackers gained access to multiple servers within JP Morgan’s systems.

- One of the servers was not set up with updated security after a recent update.

**AI Relation:**

Like the Equifax leak, unpatched vulnerabilities that attackers are aware of can be tested easily by AI tools that search for these updates and find systems on older versions.

**Response:**

Doubled it is spending on security within the next five years, and applied updates to prevent unauthorized access to servers in the future.

**First American Financial Corporation (2019)**

- Over eight hundred million Compromised Names, Email Addresses, and Phone Numbers of Customers/Employees

**Cause:**

- Due to a design error on their main website, links leading to personal and sensitive information were available to access by anybody poking around.

- Failure to implement failsafe or identification authenticators to these pages leaked the data.

**AI Relation:**

AI’s ability to quickly access and scour websites for data means simple tools could have found this information and logged it the instant it was made public.

**Response:**

After a 1$ Million-dollar settlement, the company updated their websites and redid their infrastructure to secure the leaked data.

**Experian (2020):**

- Over twenty million customers, and 700,000 Businesses were impacted by the breach.

- Compromised Work Addresses, Phone Numbers, Email Addresses, Home Addresses, and Job Information.

**Cause:**

- A fake representative for one of the company’s clients convinced an employee to give out internal data.

- The representative provided millions of user’s personal information which was then verified by Experian

**AI Relation:**

Although this case did not use AI, deepfake technology has already been used to fake voices, faces, and other information attackers can use virtually as “proof” of identity. Without proper fail safes these types of situations can become more and more common since faking your identity is easier and easier.

**Conclusion:**

From this study I learned how often these data breaches are because of employee mistakes rather than systems. Security teams that fail to update systems with patches for vulnerabilities, or employees improperly securing their login information to the point it can easily be used by malicious attackers being examples. Or when they even willingly leak information like the Experian case where an employee got tricked into revealing sensitive information. All of these breaches were caused by employee errors which could’ve been easily avoided with some proper security checks or updates, but glancing over or missing these facts led to vulnerabilities.

I also realized how proper machine learning or AI tools could have made these breaches worse or identified more leaks instantly. The ability for AI agents to browse vulnerabilities, website links, and brute force logins with available information makes these breaches more vulnerable than ever. While a human attacker might have missed specific access points, AI is able to try every available entrance with its knowledge of all vulnerabilities of a system all within a short time.

AI is also a valuable tool that these companies can use to prevent these breaches in the future. Although AI has its own issues, it can keep up to date with vulnerabilities, and available patches on all their systems rather than one person potentially missing an update or a singular access point. AI can examine everything at once and monitor suspicious activity within a company to make sure all checks are passed. Although it has its own issues where AI can be tricked, human error is just as likely to happen so it’s more of an exchange rather than being riskier.

**References**

[**10 Biggest Data Breaches in Finance | UpGuard**](https://www.upguard.com/blog/biggest-data-breaches-financial-services)

[**EPIC - Equifax Data Breach**](https://archive.epic.org/privacy/data-breach/equifax/)

[**2019 Capital One Cyber Incident | What Happened | Capital One**](https://www.capitalone.com/digital/facts2019/)

[**JPMorgan hack exposed data of 83 million, among biggest breaches in history | Reuters**](https://www.reuters.com/article/us-jpmorgan-cybersecurity-idUSKCN0HR23T20141003/)

[**What happened in the Experian data breach? | Twingate**](https://www.twingate.com/blog/tips/experian-data-breach)