AEREN LPO
...providing more with much less

CASE STUDY

RAPID PII & PHI IDENTIFICATION IN A HEALTHCARE DATA BREACH

Incident Summary

Our client was a Health institution. We analyzed the health institutions' data breach incident; our aim was to focus on how we can secure the personal data of patients under HIPPA with PII and PHI security too?

Methodological Considerations

To comply with data security norms, we do secure data with eDiscovery tools controlled by AI.

- ✓ Establish and implement a written data breach response policy.
- ✓ Review the information system, data and identify where PII and other sensitive information reside.
- ✓ Keep a close eye for any possible loss or leakage of sensitive information including PII,PHI.

Objectives

- Identify documents containing **Personally Identifiable Information (PII), Protected Health Information (PHI),** and **patient records.**
- Classify documents based on **risk level and exposure.**

Ensure **compliance with GDPR, HIPAA, and state data breach laws.**

- Notify impacted individuals within legal timeframes.

Challenges

- Large, unstructured datasets in multiple formats (emails, PDFs, Word, Excel).
- Legal and regulatory urgency (30-day breach notification deadline).
- Limited internal review resources for a document set of this size.

AI-Based Review Approach

1 Predictive Modeling and TAR Deployment

Trained AI models using sample documents to detect

- PII/PHI presence, Patient financial records
- **Deployed TAR 2.0 with Continuous Active Learning (CAL)** to dynamically prioritize documents likely containing sensitive information.

2 Reviewer-AI Hybrid Workflow

- Our reviewers validated AI outputs, focusing first on high-risk predictions.
- AI adjusted to reviewer feedback in real time, improving accuracy with each iteration.

3 QC and Notification Support

- QA team audited flagged documents using prediction scores and reviewer codes.
- Worked with legal counsel to compile notification lists for affected patients and staff.

Results

- Identified 100,000+ sensitive documents within 12 days.
- Reduced manual review burden by 70%
- Achieved 92% precision and 88%, recall across key data categories.
- Our client issued timely breach notifications, avoiding fines and reputational damage