Semele Saves National Leader in Healthcare Several Million Dollars Through Audit Automation

Problem

A national leader in healthcare, serving more than 62 million people, engaged **Semele** to audit their lower environments to find instances of live production data, such as Protected Health Information (PHI) and Personally Identifiable Information (PII). The client was concerned about the presence of sensitive data in non-production environments where it is far more susceptible to compromise by internal than external forces. In a complex environment with over 4,000 tables and an excess of 100 billion records, this audit would be no small task. It was also likely that combinations of previously obfuscated data and copies of live data would be present. Tools to scan for data that resembled PHI/PII, but may not actually be, would produce many false positives and be ineffective. Because of our expertise in dealing with challenging problems like these, **Semele** was asked to take on the project.

Solution

Semele automated the process of first analyzing the records to find combinations of PHI and PII and then comparing them back to production. Positive matches identified the exact data that needed remediation to secure the environment. **Semele** produced comprehensive results reports of toxic data requiring remediation and configurations were automated for reuse in scheduled and ad-hoc intervals to alleviate risk exposure in remediated datasets.

Result

Through **Semele's** automated audit process, we reduced the time and expense of a PHI audit and remediation by over 80%, **taking several man-years off the project and saving the client several million dollars**. Exposure risks were mitigated much faster than otherwise possible and at a significant financial savings to the client. The second phase of this project will be to transform toxic data to protect the identities and personal details of their subscribers. Using **Semele's** subsetting and obfuscation engines, this capability is already in place for the client to easily remediate the problems and significantly improve sensitive data security.