703.520.1001 | info@sandlineglobal.com

Custom Solution for In-Place MD5 Hash Comparisons

Overview

An Employment Dispute Law firm engaged Sandline to address a critical data identification and remediation matter within a Google Workspace environment, requiring unique expertise due to Google's limitations on MD5 hash searches.

Services:

Document Review

Challenge

Sandline was approached by an Employment Dispute Law firm to provide forensic services on a fast-moving data identification and remediation matter for an end-client utilizing a Google Workspace environment. The client asked that all content in the Google Workspace have an MD5 hash obtained for the purposes of comparing to a list of known hashes of allegedly stolen documents. Google does not grant its admins the ability to search MD5 hashes, so a custom solution had to be created.

Expertise

Sandline's forensic team was able to compose a Python script to crawl through all 12+ TB of content (2 million+ files) and create an inventory of all content. This included an MD5 hash for all files present. The MD5s were then able to be compared to the known list of "bad files", and a secondary review was able to be conducted against file names and paths.

Our experts worked closely with end-client's IT team to operate within the necessary security parameters due to the sensitivity of the data. They also worked together to adjust the script as errors were encountered and customized it to ensure successful retrieval of MD5 hash values and comprehensive inventory of their environment, including custom data buckets.

Results

The client was able to certifiably say that all data was MD5 searched and identified the locations of positive hits thanks to Sandline's understanding of cloud computing systems, innovative thinking, and skilled Python developers. The client was impressed with the outcome and continues to engage Sandline for complex forensic examinations.