Image Redaction

Redaction of Historical Personal Information from Commercial Business Filings
ISSUE:

Legacy constituent data (often unrequested) potentially had latent PII and PIFI...

that could now be accessed publicly.
SOLUTION:

- GOOGLE removed all references from active search results immediately, stopped crawls.
- NASS coordinated a meeting with senior executives from Google and Smartronics.
- Smartronics developed a custom client that interacted with the Idaho SOS SQL servers and to search for PII and PIFI information. (leverages the GCP DLP solution)
  - The solution was deployed for use by Idaho in less than a week
  - Idaho has been able to process more than 2 TB’s of data in less than two months using an array of search types allowing them to check accuracy and validity of redactions
  - The application was developed with significant flexibility from a data ingestion standpoint. Practically, the Idaho SQL Admin was able to run any stored procedures necessary to collect any and all data identified for potential redaction.
STATUS:

- 10 Million images potentially impacted
- Currently running Smartronics script on SSN's Only
- Additional runs planned for Date of Birth and Bank Account Numbers
- Code can be made available to other state offices as required
Thanks To our Partners:

Google  Smartronix  NASS

For Additional Information
Chad Houck  
Chief Deputy Secretary of State  
Office: 208-332-2862  
Cell: 208-585-7370  
Email: chad.houck@sos.idaho.gov
ARCHITECTURE:

- Minimal cloud architecture configuration and administration required – no customer GCP expertise needed
- Google Cloud Services required are limited to Google DLP and Cloud Firewall
- Idaho data is encrypted and ephemeral – effectively remaining within the confines of the customer datacenter
- Solution supports multiple clients and search sets simultaneously driving flexibility and efficacy.
- The solution can be deployed to redact upon ingestion