EXAM PROFESSIONAL CLOUD SECURITY ENGINEER TOPIC 1 QUESTION 94 DISCUSSION

Actual exam question from Google's Professional Cloud Security Engineer
Question #: 94
Topic #: 1
[All Professional Cloud Security Engineer Questions]

You are responsible for protecting highly sensitive data in BigQuery. Your operations teams need access to this data, but given privacy regulations, you want to ensure that they cannot read the sensitive fields such as email addresses and first names. These specific sensitive fields should only be available on a need-to- know basis to the Human Resources team. What should you do?

A. Perform data masking with the Cloud Data Loss Prevention API, and store that data in BigQuery for later use.

B. Perform data redaction with the Cloud Data Loss Prevention API, and store that data in BigQuery for later use.

C. Perform data inspection with the Cloud Data Loss Prevention API, and store that data in BigQuery for later use.

D. Perform tokenization for Pseudonymization with the Cloud Data Loss Prevention API, and store that data in BigQuery for later use.

Show Suggested Answer

by Random_Mane at *Sept. 20, 2022, 10:36 a.m.*

Comments

Type your comment...

Submit

AwesomeGCP Highly Voted 1 year, 9 months ago

Selected Answer: D

D. Perform tokenization for Pseudonymization with the Cloud Data Loss Prevention API, and store that data in BigQuery for later use.

upvoted 5 times

zellck Highly Voted 1 year, 10 months ago

Selected Answer: D

D is the answer as tokenization can support re-identification for use by HR.

https://cloud.google.com/dlp/docs/pseudonymization

upvoted 5 times

Sammydp202020 Most Recent 1 year, 5 months ago

Selected Answer: D

Both A & D will do the job. But, A is preferred as the data is PII and needs to be secure.

https://cloud.google.com/dlp/docs/pseudonymization#how-tokenization-works

Why A is not a apt response:
https://cloud.google.com/bigquery/docs/column-data-masking-intro
The SHA-256 function used in data masking is type preserving, so the hash value it returns has the same data type as the column value.

SHA-256 is a deterministic hashing function; an initial value always resolves to the same hash value. However, it does not require encryption keys. This makes it possible for a malicious actor to use a brute force attack to determine the original value, by running all possible original values through the SHA-256 algorithm and seeing which one produces a hash that matches the hash returned by data masking.

upvoted 1 times

pedrojorge 1 year, 6 months ago

Selected Answer: D

D, as tokenization supports re-identification for the HR team

upvoted 2 times

therealsohail 1 year, 6 months ago

B is okay
Data redaction, as opposed to data masking or tokenization, completely removes or replaces the sensitive fields, making it so that the operations teams cannot see the sensitive information. This ensures that the sensitive data is only available to the Human Resources team on a need-to-know basis, as per the privacy regulations. The Cloud Data Loss Prevention API is able to inspect and redact data, making it a suitable choice for this task.

upvoted 2 times

AzureDP900 1 year, 8 months ago

D is correct
Pseudonymization is a de-identification technique that replaces sensitive data values with cryptographically generated tokens. Pseudonymization is widely used in industries like finance and healthcare to help reduce the risk of data in use, narrow compliance scope, and minimize the exposure of sensitive data to systems while preserving data utility and accuracy.

upvoted 4 times

Random_Mane 1 year, 10 months ago

Selected Answer: A

A https://cloud.google.com/bigquery/docs/column-data-masking-intro

upvoted 3 times

heftjustice 1 year, 6 months ago

Data masking doesn't need DLP.

upvoted 2 times