Office of the Principal Accountant General (Audit-I), West Bengal
2, Govt. Place (W), Treasury Buildings, Kolkata – 700 001.
****

EDP SECTION

OFFICE ORDER

O.O.No.EDP/80/IS Security Guidelines and Instructions/2025-26/43 Dated: 12.12.2025

Subject: Standard Operating Procedure (SOP) for Efficient Data Storage and Management – Reg.

To establish a standardized and disciplined approach to the management of stored digital files across all Office systems, ensuring periodic review, elimination of redundant or obsolete data/files, and efficient utilization of available storage space for future analytics, audits, and operational needs, a Standard Operating Procedure (SOP) has been formulated.

"Title: SOP for Data Management and Storage Optimization"

****

1. **Purpose**: The purpose of this SOP is to outline protocols and practices for efficient storage, management, and utilization of resources within this Office, with an aim to enhance productivity, reduce redundancy, and ensure optimal storage utilization.
2. **Applicability**: This SOP applies to all forms of –

DATA as defined in the **Digital Personal Data Protection (DPDP) Act**, 2023 include '*transactions, records, books, accounts, papers etc. to comprehensively address whichever form data, information and documents are maintained by the auditable entity or within the IA&AD or collected through primary survey like a beneficiary survey during the field audit, an audio/video file obtained during field verification, and whether stored in a computer system or moved through any analog, printed or digital media*'; and

INFORMATION as defined in **Regulation of Audit & Accounts, 2020** as '*any material in any form, including records, transactions, documents, memos, e-mails, opinions, advice, press releases, circulars, orders, logbooks, contracts, reports, papers, samples, models, data or other material held in any electronic form and information relating to any private body which can be accessed by a public authority under any law for the time being in force*'.

3. **Periodic Data Review, Data Compression & Data Deletion:** All personnel involved in the storage, management, retrieval, and disposal of records and information assets –

i. Should conduct periodic reviews (at least quarterly) of all data stored on servers, NAS drives, or external HDDs and IT systems under their control.
ii. Identify and delete duplicate, obsolete, redundant, or unused files to free up storage space. The user should, as far as possible, refrain from creating duplicate copies of the same files to maintain storage space efficiency.
iii. Employ data compression techniques (using compressing software like WinRAR, WinZip, 7zip etc.,) wherever feasible to reduce storage space requirements.

iv. Large-sized data dumps, database or files of unknown origin/ownership that are not actively used by the user, or whose whereabouts cannot be traced, or purpose is unknown to the user, or whose purpose is no longer valid or has been superseded may be deleted after proper documentation and obtaining permission form the Branch Officer concerned.

4. **Retention of Large Data Dumps:** While large data dumps may occasionally require retention beyond their immediate utility, prior written permission must be obtained from the concerned Group Officer/Data Protection Officer, accompanied by a detailed justification for their continued storage, under intimation to the EDP Section with details, for keeping the record updated.

5. **Responsibility of Branch Officers:**

i. Branch Officers shall prepare/update list of Data, held within its jurisdiction.

ii. Branch Officers shall be responsible for ensuring compliance of the SOP by subordinate officers/officials within their jurisdiction.

iii. Branch Officers should submit Quarterly Reports to EDP Section detailing the amount of storage space (in GB) freed up through these data management exercises in the manner mentioned in **Para 10**.

6. **Responsibility of Field Audit Parties:**

i. CAG India adopts cryptographic hashing with SHA-256 as the primary method for data authentication in audits, ensuring data integrity and non-repudiation during collection and verification. This approach generates a fixed 256-bit digest from any input data, enabling auditors to detect tampering through hash comparison. SHA-256 provides deterministic output, pre-image resistance, collision resistance, and avalanche effect for robust verification.

Hash Generation Process

- Apply SHA-256 algorithm to the exact original file or dataset using standardized tools: PowerShell (Get-FileHash -Algorithm SHA256).
- Command example: certutil -hashfile filename.xlsx SHA256 (Windows) or shasum -a 256 filename.xlsx (Linux/Mac).
- Store the 64-character hexadecimal hash alongside metadata in a tamper-evident audit worksheet, signed digitally by the collecting officer.

ii. Before commencement of data analysis, it is vital to ensure that the audited entity does not, at a later point of time, raise doubts on the integrity of the data set provided or repudiate the data set. For this purpose, it is desirable that suitable controls are adopted, such as obtaining a letter from the audited entity which specifies the data source (through reference to time stamp of generation of the data set/ hash value for the data set), along with the details of the parameters for extraction used to create the data set, i.e. queries/ scripts executed. If such a letter is not forthcoming from the audited entity, internal documentation may be generated by the field audit Offices for the above purpose, noting the date of receipt of the data set, the name of the personnel of the audited entity who handed over the data set, and the hash value for the data set and communicating/ getting confirmation on the same from the audited entity before commencing any data analysis. To the extent feasible, control totals for the data set may be reviewed and cross-verification from other sources may be carried out to derive assurance regarding the accuracy and completeness of the data set respectively.

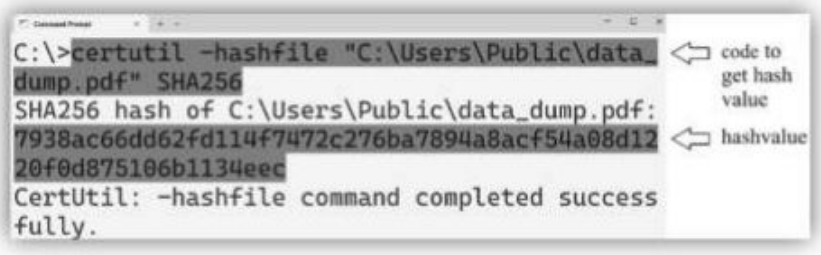

Fig. An illustration for generation of a SHA256 hash value of a (.pdf) file using command prompt.

iii. While collecting data, the authenticity, integrity, relevance, usability and security of the data sets should be ensured. For ensuring the integrity of data (i.e. – that some data is not lost), checks such as counting the total number of records or sum of numeric columns adding up to total (hash totals) may be undertaken. For ensuring that data is complete, completeness control measures should be undertaken, e.g., taxes collected by individual taxpayers should add up to the total tax collected in the Tax office. The auditor should obtain a certificate stating that the data is complete and the same as in the IT system of the audited entity at the time of receiving data. An indicative template of such certificate is provided below. It should be ensured that only authorized personnel handle data transfers from the data sources to the auditors. The access to such data should be through appropriate access controls to prevent any unauthorized access to data.

Annexure 2 (Refer para 2.13)

Indicative Template of Certificate for completeness, consistency and integrity of data

(To be collected from audited entity while receiving data)

The data dump provided to the O/o _____ (name of audit office) in respect of _____ (name of database) for the period _____ to _____ maintained by Ministry/Department/ _____ (Name of entity providing data) on an external storage device/provided online duly marked as <XXXXX> (in case of external device) and signed/authorised by <XXXXX> (name and designation of nodal officer providing the data) on <date>.

It is certified that:-

(i) Officials are authorised by the _____ (name of audited entity) for sharing this data with audit and they understand relevant provision of the Information Technology Act 2008.

(ii) The data dump is full, complete and whole of actual data.

(iii) There is no erasure, tampering or overwriting of original data.

(iv) There is no data inconsistency and there was no loss of data during data migration from one system to another or backup or due to theft/hacking etc.

(v) There is no damage of data i.e. by destruction, alteration, modification, deletion or re-arrangement of any computer resources by any means.

Summary information on key parameters – total number of transactions, date and details of first and last transactions and hash totals of various numeric data fields is also provided to ensure the completeness, consistency and integrity of data.

(Name, designation, e-mail & signatures of authorised officials)

Date:

Place:

iv. The Supervising Officer of FAP should submit a copy of the requisition of data from the Auditee and Government Departments and agencies, where the dataset involved is of high volume and electronic in nature, to the concerned Branch Officer/Branch Officer (Coordination) for review by the DPO prior to placing it with the data providing authority.

v. FAPs should transfer all data, received during the field audit, including in personal devices like mobile phones, from the portable devices to a dedicated system(s) after the end of an audit assignment.

7. **Role of EDP Section:**

i. EDP Section shall act as the central authority for managing data on the Server, NAS drives and other storage media.

ii. It shall initiate the process of purging or deletion of DATA proposed for deletion by the Branch Officers after obtaining approval of the Data Protection Officer.

iii. EDP Section is authorized to run duplicate file search on Central Storage Media (Server, NAS drives) and delete them after obtaining approval from the Data Protection Officer, unless specific request for their retention is received.

8. **Role of Data Protection Officer:**

i. Grant permission for retaining large-sized data dumps, database etc. provided a detailed justification is submitted explaining their continued relevance to audit or data analytical purposes by the concerned Group/Section.

ii. Reviews and approves deletion requests where necessary.

iii. Carry out periodic assessments of the data governance, privacy, and security aspects.

9. **Compliance & Reporting:**

i. The first review and purge exercise shall be conducted immediately upon issuance of this Circular.

ii. A formal compliance report detailing the total storage space freed (in GB) through this initial exercise shall be submitted to the EDP Section latest by 31st December 2025.

iii. Subsequent reviews shall be carried out at the end of every quarter (i.e., December, March, June, September), and shall be reported to the EDP Section within 7 working days after each review cycle for record-keeping and analysis purposes.

10. **Reporting Format:** Branch Officers, after concluding each periodic review, should forward a compliance report to EDP Section for consolidation and submission to the Data Protection Office. The Compliance Report shall include:

a. Name of Group/Section
b. Date of Review
c. Total storage space freed up (in GB)
d. Any large data dumps retained with justification and approval details

This SOP applies to all officers/officials of this Office, including LAD, and is circulated for reference, adherence, and strict compliance by all concerned.

This SOP will be reviewed annually or upon significant changes in departmental operations, data types, or government directives.

Group Officers are requested to ensure implementation of this SOP and strict compliance by all concerned.

Authority: PAG's order dated 12.12.2025 kept in e-file no. EDP/80/IS Security Guidelines and Instructions.

Digitally signed by Avnindra Kumar Rai
Date: 12-12-2025 20:03:05

Dy. Accountant General (Admn)

1. For all circulation.
2. For publication in the Official Website for wide publicity.
3. Rajbhasha Section for translation into Hindi.