Stage 2: Raw Data Conversion and Deduplication

Responding to a DSAR often means handling large, messy datasets. Our Raw Data Conversion & Deduplication service removes that burden by transforming raw files into a clean, consolidated, workable format.

Step 1 – Initial Dataset Assessment
You provide DPAS with:
- The dataset size (MB or GB)
- A description of file structure and types (e.g. PDFs, mailboxes, Word docs)
- From this, we provide a fixed cost for conversion and deduplication.

Step 2 – Secure Data Transfer
Once costs are agreed, you upload:
- The raw data via your secure file sharing platform
- A completed DSAR Processing Requirements spreadsheet (including eDiscovery criteria and any context)

Step 3 – Data Conversion & Deduplication
On receipt, DPAS processes the data using specialist software, performing the following tasks:
- Dataset Deduplication – exact duplicates are automatically removed
- Email Thread Deduplication – conversations grouped for easier review; unique variations retained
- File Conversion – documents converted into redactable PDFs for usability

Step 4 – Data Review (Human Input)
At an additional cost, DPAS will conduct a manual review of the deduplicated data to ensure only data relevant to the request is preserved for redaction

Step 5 – Output & Redaction Quote
You receive a single, consolidated PDF.
DPAS then provides a fixed redaction quote based on page count, or you can manage redaction internally.