POL00029169

Document Title: Conducting Audit Data Extractions at CSR

Document Type: Process

Abstract: This document describes the process to be followed by Post Office Internal Audit (POIA), and other groups external to Pathway as defined in Schedule A03, when requesting audit data extraction services from ICL Pathway Internal Audit. It also describes those activities carried out within ICL Pathway to handle the request, manage the data extraction and despatch the results to the original requester.

Status: APPROVED

Distribution: Jan Holmes
Martyn Bennett
Chris Paynter (POIA)
Paul Redwood (Horizon)
Library

Author: Jan Holmes

Comments to: Jan Holmes

Comments by: Soonest

POL00029169

0 Document control

0.1 Document history

Version	Date	Reason
0.1	17/02/99	Initial Draft for Comment
0.2	02/03/99	Introduction of procedure steps in process
0.3	12/05/99	Updated to remove MOT restrictions
0.4	28/05/99	Following withdrawal of DSS and incorporating comments from POIA
0.5	02/11/99	Following delivery of Cluster Determinant utility and changes to POIA personnel
1.0	04/05/00	Raised to issue status for formal sign-off

0.2 Approval authorities

Name	Position	Signature	Date
M. Bennett	Director Quality & Risk		
J. Holmes	Pathway Audit Manager		

0.3 Associated documents

	Reference	Vers	Date	Title	Source
[1]	IA/MAN/004			Horizon System Audit Manual (CSR)	PWAY
[2]	IA/REQ/002			Audit Data Retrieval Requirements	PWAY
[3]	IA/SPE/008			Audit Data Catalogue	PWAY
[4]				R-Query User Guide	PWAY

POL00029169

0.4 Abbreviations

Acronym	Meaning
AWO	Audit Workstation Operator
BA	Benefit Agency
CSR	Core System Release [Was NR2, renamed following DSS withdrawal]
CSR+	Core System Release + [Was NR2+, renamed following DSS withdrawal]
EJN	Extraction Job Number
PA	Pathway Auditor
POCL	Post Office Counters
PWAY	ICL Pathway
RFI	Request for Information
TMS	Transaction Management System

POL00029169

0.6 Table of content

POL00029169

1 Introduction

The Horizon system generates significant amounts of data that is of interest to Internal Audit and other groups. The Horizon System Audit Manual (CSR) [1], and the supporting <Product> Audit Trail Specifications provide further information on the structure, form and content of this data, referred to in this document as 'audit data'.

Subject to certain constraints the audit data must be made available to POIA or other authorised groups within timescales established in the Audit Data Retrieval Requirements [2].

This document establishes the process for requesting audit data extractions and subsequent activities undertaken to locate, retrieve, extract & filter and prepare for despatch on behalf of authorised requesters.

2 Scope

This process is for use from the start of CSR Live Trial until further notice.

Should future releases of Horizon bring about changes to the way that data is extracted this process will be updated to reflect those changes.

This process applies to ALL audit data extraction requests from outside ICL Pathway. Requests for audit data extraction from within ICL Pathway will also be subject to this process although use of the Request For Information form is optional.

The process is presented in its entirety although it is currently constrained those elements that will not be available until CSR+. *These are identified by italics..*

3 Terminology

Within this process certain terms are used which have specific meaning within the Horizon Audit Solution. They are :

Gatherer : The module responsible for collecting the audit files from the hosts, agents, correspondence servers and interface mechanisms. This module is also responsible for the application of the audit file naming policy.

Sealer : The module responsible for calculating the checksum seal of each audit data file before it is written to DLT by the **Hoarder**. This value is recalculated by the **Retriever** and compared to the original value when first sealed. Used to ensure data integrity during storage on DLT.

Hoarder : The module responsible for writing audit data files onto DLT at pre-defined intervals.

POL00029169

Retriever : The module responsible for retrieving audit data from the buffer file where it is placed by Legato when requested by the Audit Workstation.

Extractor : **Retriever** brings back complete files or groups of files from the DLTs. Further work may be required to filter out unwanted information, especially true of the TMS files, using a number of tools available on the Audit Workstation.

Legato : Legato Networker is the storage management application selected by Pathway to store and manage audit data onto DLTs.

A more complete explanation of these modules can be found in [2].

4 Audit Data Integrity

The integrity of audit data must be guaranteed at all times from its origination, storage and retrieval to subsequent despatch to the requester. Controls have been established to provide assurances to Post Office Internal Audit that this integrity is maintained. During audit data extractions the following controls apply :

a. Extractions can only be made through the three Audit Workstations which exist at Feltham and the 2 Data Centres. These are all subject to rigorous physical security controls appropriate to that location. Specifically, the Feltham AW – where most extractions will take place – is located in a secure room subject to proximity pass access within a secured ICL site.

b. Logical access to the AW and its functionality is controlled by dedicated Logins, password control and utilises the NT security features defined in the overall Horizon security policy.

c. All extractions are logged and supported by documented RFIs, authorised by nominated persons within POIA. This log can be scrutinised on the AW.

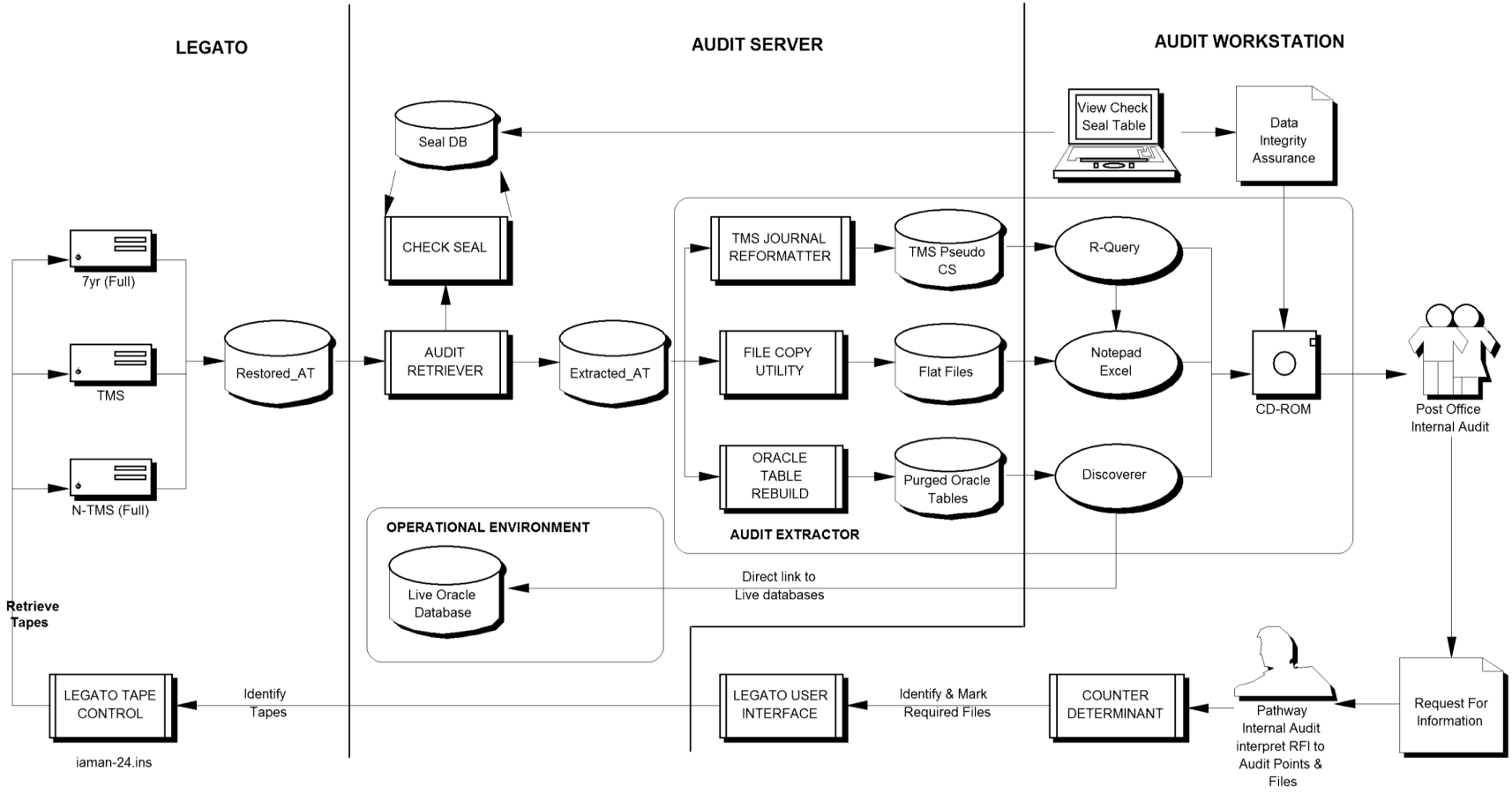
d. Extractions will only be made by individuals previously notified to POIA. Currently this is limited to the Pathway Audit Manager. Any additions will be notified to POIA .

e. Agreement has been reached with POIA regarding their rights to witness extractions without warning or to request repeat extractions that they can witness.

f. Checksum seals are calculated for audit data files when they are written to DLT and re-calculated when the files are retrieved.

POL00029169

5 Retrieval Schematic



POL00029169

6 Overview

The process assumes that audit data has been Gathered, Sealed and Hoarded onto DLTs by the Audit Archive Server. Files will be one of three types :

a. Flattened and compressed TMS Journals from the Correspondence Servers.

b. Flattened Oracle tables output from regular OBCS database purging cycles.

c. Transaction files to and from PO systems and their associated FTMS control files.

The process is invoked through the receipt of a Request For Information (RFI) into Pathway Internal Audit. Expressed in business terms, the RFI must be interpreted into its component Audit Points and Sub-points. This then enables specific files to be identified which, through the Legato index, targets a specific DLT. Data is retrieved by the Audit Retriever, formatted as appropriate and then further Extracted against the RFI criteria. Depending on the extraction method the data can be extracted to standard MSOffice products before being placed onto CD-W or floppy disc for despatch to the RFI originator.

The following paragraphs present an overview of each step in the extraction process and are ordered to reflect the actual processing of a Request For Information (RFI) by ICL Pathway Internal Audit.

6.1 Request For Information

All requests for audit data must be made via the Request For Information form. This will contain a description, in business terms, of the times, outlets, events, items and activities that the Auditors are interested in. This request has to be interpreted by Pathway Internal Audit and mapped onto the Audit Points and Files described later in this document.

6.2 Marking Files and Tapes

Based on this interpretation as many files of audit data that are needed to satisfy the request are ‘marked’ for retrieval. Legato is notified of these files and it in turn identifies the DLTs containing these files. Legato provides system prompts for Operators to load tapes and it copies the data into a local buffer area.

6.3 Audit Track Retriever

Polls the Legato buffer area and retrieves any data files found into temporary disk storage (Export File) on the Archive Server prior to the extraction of

POL00029169

relevant data for use by the auditors. The Retriever provides a second copy of the file which is input to the Check Seal function.

6.4 Audit Data Check Seal

To assure the integrity of the audit data while on the DLT the checksum seal for the file is re-calculated by the Audit Track Sealer and compared to the original value calculated when the file was originally written to the DLT. The result is maintained in a Check Seal Table.

6.5 Audit Trail Extractor

This is a facility that uses various tools to extract or reform the retrieved audit data in accordance with the RFI. It also places the information onto a CD-ROM, or other suitable media, for despatch to the RFI originator.

POL00029169

7 Requesting Audit Data

7.1 Receiving the RFI

All requests for audit data extractions must come to Pathway Internal Audit in the form of a Request For Information. An example of this form can be found at Annex A. The RFI may be mailed, faxed or e-mailed to Pathway.

RFIs will only be accepted from the following named individual :

Chris Paynter : PO Internal Audit : Tel : GRO

If other parts of the Post Office, or other organisations, require audit data extractions they must be channelled through POIA to Pathway Internal Audit at Feltham.

7.2 Interpreting the RFI

In the early days of CSR it will be necessary to interpret the RFI by identifying the audit points and sub points that generated the records that are required and, through the Audit Point & Data Catalogue, the files produced at those audit points and sub points.

An Enquiry Catalogue of often requested extractions will accumulate which means that this part of the process would be simply about targeting the files by the date spread.

7.3 Login Audit Workstation

Carry out following procedure to Login and obtain necessary shares
1. Login : jholm01
2. Password : *********
3. Domain : PWYDCS

The AW will present a blank desktop with a START icon in the bottom left of the screen. Using pullup <Programs> will reveal the extent of products available for any subsequent extraction work.

7.4 Preliminary Housekeeping

It is highly likely that an average RFI will need a significant number of files to satisfy it. To avoid the AW filestore becoming clogged with hundreds of files it is strongly recommended that a working directory is established on the AW to hold all files relevant to a particular RFI :

D:\audit\RFI Reference No.

POL00029169

7.5 Registering The RFI

The Pathway Auditor is responsible for processing all RFI's. These will be entered into the Audit Data Extraction Database which holds the following information :

Note : not currently available @ CSR.

AW Operator Identity	Name of Audit Workstation Operator
Access Reason	Reason for access to Audit Workstation
RFI Reference	Reference Number from RFI or other unique identifier
Catalogue Entry (Opt)	Enquiry reference from catalogue of standard enquiries.
RFI Requester Identity	Name of person requesting extraction
RFI Requester phone (Opt)	Contact number of requester
Date RFI Received	Date RFI received
Required by date	Date by which data extraction is required
Delivery Date (Opt)	Actual date extraction despatched to requester
Comments (Opt)	Anything else of interest

7.6 Counter Determinant

The architecture of the Correspondence Server has Post Office outlets spread across 4 separate clusters. These clusters are separately archived thus it is necessary to identify the particular cluster that services the outlet, or outlets, for which audit data is being extracted.

1.	Select <Counter_Determinant.CMD> from the main program menu.
2.	Type 6 character FAD code <enter>.
3.	Type <ctrl & z> <enter>
4.	Cluster identity will be displayed.
5.	Note that multiple FAD codes can be entered at the same time.

7.7 Identifying Closed Outlets

The architecture enables Outlets to operate while detached from the Correspondence for up to n days. While this is good for service continuity at the outlet it means that TMS audit data is not made available to the Gatherer - and subsequently Hoarded - on the same day as the records were created. The consequence is when a <start date> or <end date> on a search is invalidated because the outlet's audit data was not archived on the appropriate day.

CSR+ : This is not available until CSR+.

POL00029169

7.8 Targeting the Data Files

It is highly unlikely that a single file will hold the information required by the RFI. Indeed, the broader the date spread or complexity of request the greater the number of files that will have to be retrieved from DLT.

The default Legato approach, where the primary search index is the instance of a DLT hoard, does not allow for quick and easy identification of the required files. If files to be retrieved are spread across more than 1 hoarding instance then they have to be retrieved on a hoard instance basis. For example, if 3 hoarding instances happened in a day and all 3 contained files of interest to a particular RFI there would have to be 3 separate retrieval runs.

1. Select <Legato_Bootle_Client.CMD> from main program menu
2. Select <Directed Recovery> from Operations drop down menu
3. Confirm <mboarc01> as Source in dialogue. <OK>
4. Confirm <mboarc01> as Destination dialogue. <OK>

Note that Bootle is assumed as the primary retrieval location. There is no difference in the audit data held at each Data Centre. If Wigan is selected then the <Legato Wigan Client CMD> should be selected and <mwiarc01> used to confirm Source and Destination dialogues.

5. Select <Change Browse Time> from View drop down menu
6. Select appropriate date button
7. Select appropriate Hoard time (note 7:30p)
8. Locate files through Legato directory structure and naming convention [3]
9. <Mark> files using <✓> button on toolbar
10. Select <Recover Options> from Options drop down menu
11. Enter d:\Archiveserver\INTERFACES\RESTORED_AT into dialogue box
12. Select <traffic lights> button on toolbar

7.9 Targeting the DLTs

Most Retrievals will be made from the TMS18Mnth and NonTMS18Mnth tape pools. DLTs in these pools are replaced every 5 days or when full and despatched to the DataVault (DV) offsite storage facility for safekeeping. Each file that is marked has an associated Volume Name and this must be notified to OSD if the DLT has already been sent to DV. In order to achieve next day delivery OSD must be notified before 1200hrs.

Emergency recoveries can be organised but these incur an extra charge and should be avoided where possible.

Use the OCR mechanism via email: PATHWAYSM ICLCFM to invoke DLT recovery from DV.

POL00029169

7.10 Reformatting Retrieved Data

Before detailed extractions can take place using R-Query, Wordpad, Discoverer or other appropriate tools it is necessary to 're-format' the retrieved data into a format suitable for access. There are three options :

a. TMS Re-formatter to rebuild a pseudo Correspondence Server.

b. Winzip for flat files that were zipped prior to Hoarding.

c. Oracle Table Re-formatter to rebuild Oracle tables.

7.10.1 Reformatting TMS Journals

Once the TMS Archive files have been deposited in EXTRACTED_AT they must be 'built' into a pseudo Correspondence Server for R-Query to access. Further filtering is available to restrict the number of Outlet records that are included in the re-build activity based on the original RFI.

This utility is invoked at the Data Centre. Before requesting prepare a text file containing the FAD codes to be included [fads.txt]. If all Outlets are to be included the text file should contain 'ALL'.
Notify DC Operator of file name and location and ask for <Restore_Audit_Data> to be run. It has a parameter that defines the date range :
Restore_Audit_Data –d 19990521:19990522 –f c:\temp\fads.txt
This module can be found in c:\Program Files\Pathway Agents on the Audit Server.
Request via the OSD OCR process @ email: PATHWAYSM ICLCFM asking for the CS re-build utility to be run. Provide details of date range required and the FAD code(s) to be included.

7.10.2 Unzipping Zipped Flat Files

It is strongly recommended that files to be unzipped are transferred from the AS to the AW in their zipped state and unzipped on the AW.

1. Select <Winzip.CMD> from main program Menu.
2. Select <Open> and identify zipped file through dialogue screen.
3. Select <Extract> and establish a new 'Unzipped' directory for unzipped datafiles.
4. Unzipped file will be placed into new Directory
5. Open unzipped files using the <Wordpad.CMD> utility from main program menu

7.10.3 Rebuilding Oracle Archive Tables

To be completed.

7.11 Checking the Seals

When Legato recovers a file from DLT a copy is made and subjected to a re-calculation of the integrity seal. This value is compared to the original value on the Seal Database and an entry made in the Check Seal table of MatchOK, MatchNOTOK or MatchFAIL. This activity carries on independently of any further extraction or filtering activity on the part of the AW Operator.

POL00029169

1. Select <Microsoft_Access.CMD> from main program menu.
2. Using **File/Open Database…** open the share'd 'Audit_Seal_DB.mdb' database that exists on the mapped drive (F:)
3. A list of 4 database 'tables' will be displayed.
4. Position the mouse cursor on the **<QUERIES>** tab and click.
5. A list of 2 database 'queries' will be displayed.
6. Double click on the <Seals Match Check – Normal> icon.
7. You will obtain an extract of the data that is in the <Check Seal Table> of the database.

(Note: only 5 of the available fields from this table, will be displayed. These are:

Request ID Audit Track Match? On At

8. From this point on, all of the 'Access' facilities to: sort, filter, export to spread sheet etc. are available.
9. Should you need to examine the records in the 'No Initial track table' i.e. the exceptions, then you will have to double click on the 'Seals Match Check – Exceptions' icon.
10. You will obtain an extract of the data that is in the 'No Initial Track Table' of the database.

(Note: as above, only 5 of the available fields from this table, will be displayed. These are:

Request ID Audit Track Match? On At

11. From this point on, all of the 'Access' facilities to: sort, filter, export to spread sheet etc. are available.

7.12 Despatch of Audit Data

Despatch of the extract data is by the most appropriate means depending on the nature and volume of the extracted data, and subject to any special requests made on the RFI.

The return part of the RFI should be completed with details of the media used for despatch as well as the date and time of despatch.

The Audit Data Extraction Database must be updated to record the date that the extraction activity was completed.

1. Select <CD_Writer_Software.CMD> from main program menu
2. Maximise dialogue box
3. Select files required in top dialogue box
4. Drag & drop to bottom dialogue box
5. When complete select RED dot <Red>
6. Save layout as RFI_id

The media despatched using Royal Mail Special Delivery. This ensures that a receipt is provided to Pathway confirming delivery.

POL00029169

8 Annex A - Request for Information (RFI)

I. To be provided based on examples used during Audit Acceptance Review Segment 3.