POL00042081

02

Driving business benefits through the consolidation of data review

Post Office Fraud Solution

NetReveal

By David Hutcheson and Jade Ferrari

18 May 2012

NRRA1207.01D001

29 pages including cover

Detica

BAE SYSTEMS

POL00042081

Version history

Version	Date	Author	Action
2.1	16 April 2012	JLF	For Post Office Review
2.2	14 May	JLF	PO adjustments included

Copyright statement

POL00042081

Executive summary

1.1 Overview

Post Office efforts in the area of fraud detection, prevention and reporting are currently limited by the lack of consolidated, data driven intelligence across the disparate IT systems. Detica was engaged to conduct a review of the data environment, understand current fraud detection and analysis processes and produce a high level plan to improve the efficiency and effectiveness of these activities.

This report provides the Post Office an overview of the data environment and a high level roadmap offering options for improving fraud detection and prevention, in different timeframes and budgets.

1.2 Key findings

Post Office operations are hindered by;

- Gaps in understanding of the Credence database - use and content;
- Day to day operations are reliant on bespoke databases;
- Unclear data ownership across the business;
- Fraud detection based on summary, rather than transactional level data;
- Employee data is sub-postmaster reliant and difficult to access; and
- A query based model for data access, so only known problems can be verified.

1.3 Recommendations

- Consolidate data access: to reduce inefficiencies and duplication of data.
- Enhance available data: to enhance profiling ability.
- Automate fraud detection: To reduce manual interrogation and improve returns.

POL00042081

List of contents

POL00042081

2 Introduction

2.1 Problem overview

Through operational experience and previous external reviews the Post Office was aware that there was evidence of fraud hidden in their data. The Security team was attempting to exploit this opportunity using existing tools. However to maximise savings they recognised the need for an enhanced fraud identification capability.

2.2 Project scope

The project aim was to provide options for how Post Office data could be better utilised for fraud detection. To achieve this we examined some of the fraud MOs that are currently affecting the Post Office and collated a holistic view of the existing data. Our principal focus was on internal fraud, as the majority of Post Office offerings are third party products and services, therefore the Post Office is not usually fiscally liable. We have, however, also considered how best to exploit this data in the future to enable the Post Office to develop a comprehensive anti-fraud solution.

Using this information we have examined the benefits that bringing in third party data, analytics software and a consolidated data view could provide to improving fraud analysis. We have also examined what benefits a consolidated data view could offer to the wider Post Office.

There are strategic workstreams within the Post Office which consider data consolidation, data access, IT architecture and the future of the general data environment. The focus of this report is the IT systems which could be used to improve fraud analysis, rather than the business processes that underpin their use. Internally the Post Office is addressing many of the issues which will be highlighted around users of the data. While they would be key to successful integration and new systems; these are dependencies for the process rather than part of the roadmap itself.

2.3 Document structure

The main body of the document sets out the approach, key findings, a system and fraud overview and recommendations, structured as follows;

- Executive Summary: Overview of the study and content
- Introduction: The context and details of the document
- Approach: The methodology used
- Key Findings: Key observations identified from the evidence collected
- System Overview: Key systems
- Fraud MOs: Collated view of fraud perpetrated
- Recommendations: Areas of potential improvement
- Solutions: Solution options
- Our recommendation: Our recommended roadmap forwards

POL00042081

3 Approach

3.1 Methodology

Detica conducted interviews with data owners and users across the Post Office over a six week period. The interviews involved a semi-structured format and focused on;

- Understanding the data environment;
- How data is extracted and analysed;
- Limitations on operational procedure which come from the data environment;
- Collating fraud MOs; and
- The impact of fraud on the business.

3.2 Identification of interviewees

Interviewees were nominated at the beginning of the project by the Security team. Those interviewed then provided more names of those who could be of use, or had specialist knowledge of a particular area of interest. Further names were obtained from Post Office literature, e.g. the risk register. Interviews (multiple in some cases) were conducted with the following individuals.

- Pete Newsom, Fujitsu
- Dawn Brooks, Product and Branch Accounting (P&BA)
- Sally Smith, Security
- John Scott, Security
- Joanne Hancock, Security
- Kim Abbotts, Security
- Andy Hayward, Security
- Matt Warren, IT & Change
- Ashley Hall , IT & Change
- Ian Trundell, IT Architecture
- Nina Trueman, Sales Planning and Analysis
- Chris Howard, Remuneration Development/Agents Pay Development
- Andy Terrett, Remuneration Advisor
- Paul Meadows, Head of Risk and Compliance
- Kevin Lenihan, Senior Information Services
- Kjetil Fuglestad, Network Design Manager
- Sean Farrow, Supply Chain
- Andrew Stevens, Operations Manager, National Stock Centre
- Graham Tiley, Retail Inventory
- Chris Furmanski, IT Architecture
- Paul Lebeter, P&BA
- Dave King, Information Security
- Lester Chine, Security
- Cathy MacDonald, P&BA
- Jason Collins, Security

POL00042081

- Phil Jeary, Post Office Model Office
- Dave Burford, Business Support Manager
- Katherine Mearman, Senior Brand & Format Manager
- Nick Fox, Market Planning
- Mark Smythe, Kings Security

3.3 Supporting literature

Our findings are also based upon documentation provided by the Post Office, including the annual security report, risk register and reports from IntelliQ and Deloitte on projected fraud losses.

POL00042081

4 Key findings

The following section summarises what we consider to be the main obstacles hindering fraud detection and prevention within the Post Office.

Gaps in understanding of the Credence database - use and content

The Credence database sits at the heart of fraud analysis in the Post Office. It holds a transaction record for everything sold in the Post Office from the till, online sales and other Post Office sales machines e.g. Post and Go. Transactional level data is available to the Post Office for 90 days and a summarised view is available for 2 years. The Credence database was designed to support the non-compliance and fraud detection work of the Product and Branch Accounting (P&BA) team. Should a branch appear to have unusual activity Credence will be used to investigate transactions. In addition, it is used throughout the business for other activities such as branch profiling, network transformation and any analysis that requires transaction level information.

Across the 380 system users there are varying levels of understanding as to what data sits in Credence. Some users view the system as 'clunky' or believe that sophisticated queries are not possible. They informed us that queries can specify transaction data to be extracted by product type or by time period, but not both. Users are not supported by an up-to-date data dictionary (one was originally maintained by Logica, but has not been updated). Indeed whilst many users view Credence as an MI (Management Information) system, Fujitsu, who host the system, informed us this was not its original purpose.

This has led to the majority of data interrogation, research and analysis activity from Credence being conducted outside of the system itself - through the use of bespoke databases running across exports of data.

Day to day operations are reliant on bespoke databases

The limited understanding of how Credence interrogation can operate and no centralised data repository means that the majority of teams do most day to day operations on bespoke Access and Excel databases. Microsoft database products were not designed to hold multi-million volume records, nor conduct the complexity of queries that Post Office users are currently utilising them for. The Network Transformation team in particular referenced using 12 bespoke databases on a daily basis, one of which uses approximately 30 billion transaction level records as required, all of which are stored in excel format on DVDs.

These databases provide limited functionality, despite many parts of the business developing very sophisticated macros. They are not future proofed, many only being able to be operated by current employees. In addition these databases are prone to failure, as they are operating beyond capacity. Due to multiple copies of the same data no single version of the truth exists within the Post Office.

Unclear data ownership across the business

The reliance on bespoke database has led to a fractured data usage picture. Users in different parts of the business are unlikely to know who is using the data and for what purpose.

This can lead to bespoke database failures when the central or core databases are subject to system change. Examples of this include:

POL00042081

- The Security team's efforts to identify anomalous levels of spoilt postage by deriving mean branch figures from Credence data. When the data was changed, this view was no longer possible through the Security team's methods.
- The Network Transformation team's branch profiling spreadsheet began to reject batch data when change notification process failed to identify all the users of Credence when new data streams were added to it.

The prevalence of bespoke databases held within specific teams also means that analysis which could be valuable in other areas of the business is not being exposed to end users across the organisation.

Fraud detection based on summary, rather than transactional level data

Partial understanding of the transaction level data available and limited functionality of investigative tools has meant that fraud detection at a high level has focused on summary data. However this does not offer the granular view required for effective fraud detection. The Product and Branch Accounting team look for fraud by monitoring branches for unusual activity. This is completed using a data feed from POLSAP, which is summary level data offering an amalgamated branch view.

Gaps in understanding the scale, nature and impact of fraud

Although the Post Office has an understanding of how and where fraud has impacted the business, primarily through the investigation of internal frauds, there is a recognised gap in relation to the actual scale and impact (financial or otherwise) of fraud across the enterprise. Similarly, there is awareness within the Security Team and beyond of particular fraud MOs that have been used in the past, which inform current monitoring and risk mitigations efforts; however many of these are identified reactively and there is a gap in relation to the strategic analysis of fraud across the business.

Employee data is sub-postmaster reliant and difficult to access

Each individual working in the Post Office branch network has a system login ID. This is recorded in any transactions they complete and should therefore act as an identifier for fraudulent activity. However the Post Office has over 11,000 branches, more than 95% of which are run by sub-postmasters.

Sub-postmasters are checked against the electoral register, CCJs and bankruptcy when they apply to the Post Office – however their employees are not subject to any scrutiny. Sub-postmasters are required to return a list of their employees annually, and update the Post Office should this change, but this system relies on the discretion of the sub-postmaster. In addition logons are known to be used interchangeably by employees, both deliberately and accidentally. Identifying an individual who completed a transaction, even with their unique identifying code, can therefore be difficult.

A query based model for data access so only known problems can be verified

The main data source for fraud investigations is Credence, which has a three month window during which the Post Office can extract data. After this point a request has to be sent to Fujitsu. Therefore analysis on the data is limited by time period and can only be reactive. Only problems which are already identified can be verified.

POL00042081

5 The data environment

The data environment in the Post Office, in common with other large government organisations, is complex. There are multiple databases which hold different views of the same information for different processes and purposes. Included in this section are those systems which we believe have potential value in a fraud solution. Further systems which were understood, but not prioritised, can be read about in Appendix 1. In mapping the system numerous views were provided, all of which offered subtly different perspectives.

Our understanding of the Post Office data environment comes from multiple interviews with the data 'owners and users'. However as data ownership in the Post Office is not always clear, it has been challenging to identify a single version of the truth.

5.1 System view

The system diagram below was one provided to Detica by the Post Office IT department. It outlines the batch system – which feeds from the Horizon transaction data, the area this report focuses on. In the second view we have added in some databases we are aware of interacting with those systems. It also indicates those systems we consider to be a priority in developing an anti-fraud solution.

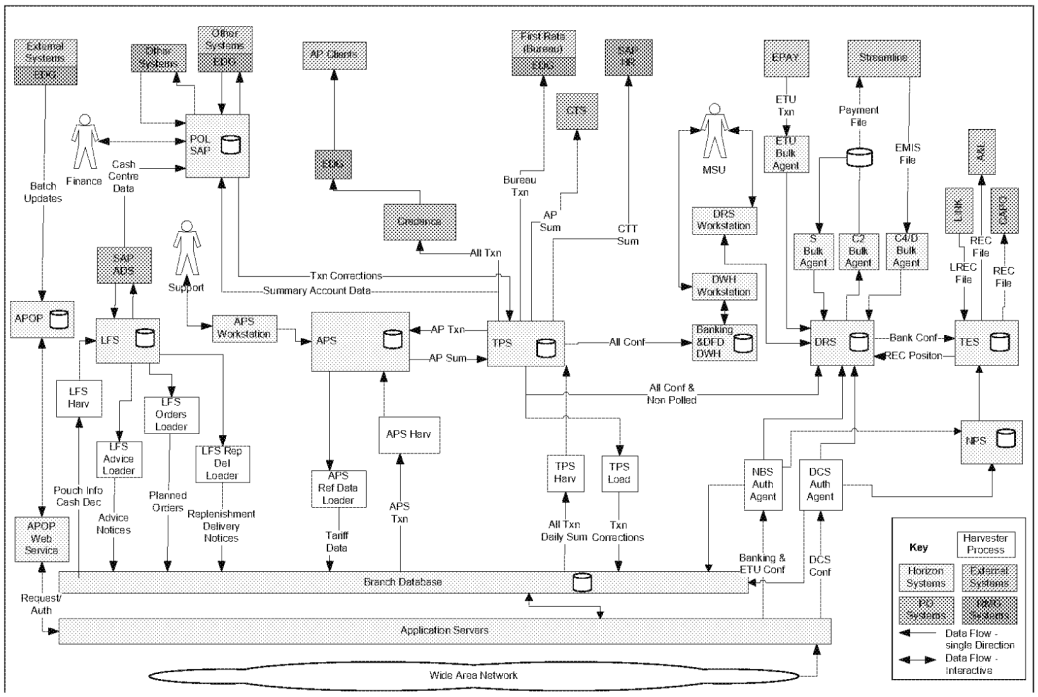

Figure 1: Post Office system

POL00042081

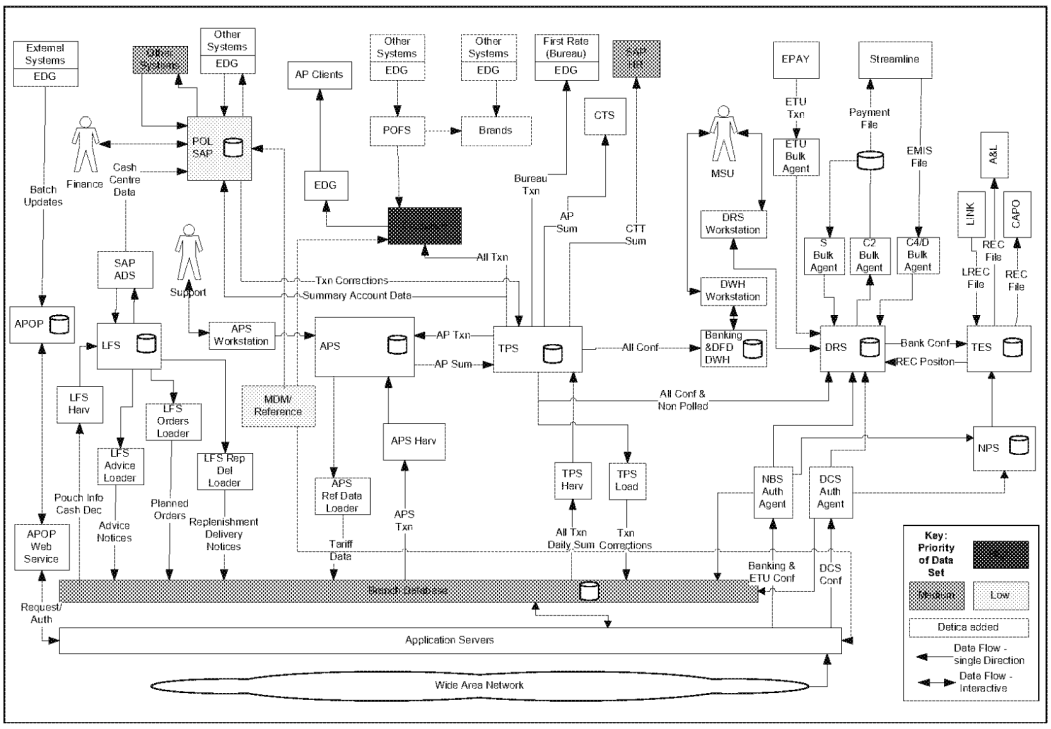

Figure 2: Post Office prioritised data sets

The Post Office is currently transitioning from the EDG data gateway to the PODG. Details can be found in Appendix 1.

5.2 Key databases

The table below outlines the current Post Office datasets which could be part of a fraud solution. These systems offer transaction information, current MI and fraud detection capability or a view of branches or employees.

Data Set	Use
Horizon	The Horizon system is the operating system on all Post Office branch terminals, with all transactions made through the till recorded there. Transactions can be sale or non-sale transactions depending on the nature of product being bought. Stock purchases are sales as the transaction is closed at that point. For some transactions where the Post Office is providing a third party good or service, then the transaction is an application. A data feed will be sent to the client (supplier of the good/service) and if this is accepted it will be separately confirmed to the Post Office as a sale. Other transaction data sources (e.g. Post Office Kiosks) do not automatically go through the Horizon system. A long-term goal within the Post Office is to rationalise the data journey so all transaction information passes into the Horizon data feed. This has begun with some transaction data sources, but does not encompass all.
Branch Database	All transactions made at Post Office counters in Horizon are stored in the Branch Database. Data is accessible for two months before it is archived. The Branch Database is hosted by Fujitsu, who also have a mirror image system that runs a

POL00042081

	few seconds behind, as a backup. The Post Office has a limited number of data queries (est. 750) annually that it can request. Beyond that point queries are £400-500 each, due to a database administrator having to extract the data over potentially multiple archives on multiple servers. This model of query based data access means that only problems already identified can be verified and on a limited time scale.
Credence	The Credence database holds transaction level data for three months and a summarised view for data after that period. It is fed from multiple sources, including a subset of Horizon data, which is chosen by the business depending on need. Other feeds come from the External Data Gateway (EDG) for third party suppliers and through Post Office Financial Services (POFS) for third party financial suppliers. At present it is not possible to marry the Horizon application transactions and the external sales transactions. The data streams that go to clients (suppliers of Post Office branded goods and services) do not necessarily include Post Office unique identifiers, therefore a transaction that is confirmed and returned as a sale cannot be easily matched to its original application. The database was originally designed to support the operations of the P&BA team for compliance and fraud detection. It has many further business functions, including being used as the main Post Office MI system. It is also used to translate data for remuneration in SAP HR, as transaction volume data is the basis for selected payments to sub-postmasters. It acts as a reconciliation point for data that comes from Horizon and third party data to ensure that expected values of transactions have been accurately reported into the Horizon system. Third party data entering Credence is usually automatically ingested via the Post Office Data Gateway, which loads third party data into various Post Office databases. However some data flows are manually entered by Post Office employees. Data from third parties largely does not require transformation, however there are instances where the data is manually transformed outside of the system before it is entered. This is usually to support the remuneration process, such is the case for over 50s life insurance. Customer data is not stored in Credence, as the system does not have sufficient security accreditation. It does however hold postcodes for those records where that information was collected. However what information can be extracted from Credence is dependant on which universe (a view of specific items of data from the business data warehouse, presented in an organised manner) the data sits in. The universes do not interact and have different user restrictions.
POLSAP	POLSAP is a summary level accounting system with two main functions - POL FS (financial services) and the stock system. It receives a summary feed from Credence, TPS (Transaction Processing Systems database) and third party financial clients such as Moneygram. POLSAP has 4 modules; • General ledger accounts (income/control/stock/liability accounts/client matching); • Vendor (accounts data for settlement with clients); • Customer (Post Offices have an account associated to a customer number, these are monitored by P&BA to recover debt); and

POL00042081

	• Other transactions (clearing/materials management/payment runs/ transaction corrections/ journal postings) General Ledger accounts are used by the P&BA team to monitor branch activity. It acts as the basis for their compliance and fraud detection.
MDM/Reference data	MDM data (previously referred to as Reference data) is a master data management system, supported by Logica, utilising Kalido software. It acts as the Post Office core master data, which allows all other systems to operate. The system holds data for each Post Office branch including; • Sub-postmaster • Contract type • Opening times • Branch address • Products sold • Product names (multiple) • Details of the product (vary by type) • Prices • Barcode information • Can a sale be reversed • Is there an employee discount • Product supplier This information is provided to Fujitsu (without the HR information, which is unnecessary), verified by them, and then sent to other systems including Horizon, Credence and POLSAP. Should any of this information be set up incorrectly then these systems will reject the records which conflict, or in the case of Horizon a transaction will not be able to complete. The system also underpins the website branch finder and a similar, more detailed, internal branch viewer. The information is updated daily by HR who provide it to IT and Change, who send it to Fujitsu.
SAP HR	SAP HR is the remuneration system for sub-postmasters. The system receives a monthly data feed from Credence and uses it to calculate the remunerations to sub-postmasters. This data is made up of both Horizon and third party sources and is a summarised version of both, containing an amalgamated branch view required for remuneration calculations. The system converts the data to pay, but does not add information. The remunerations team describe SAP HR as the "end" of the data line. Depending on the contract type certain sub-postmasters might not be paid by transaction, so the data will vary by branch. All transactions can be linked back to employee IDs from Credence data; who the ID corresponds to is held in SAP HR. SAP HR is the hub of employee information within the Post Office. It should contain records of all employees, sub-postmasters and branch employees.

POL00042081

5.3 Bespoke databases

Much of the analysis of data within the Post Office is completed in Excel and Access databases. In the scope of this review it would be impossible to map them all. However some of those which lie at the heart of fraud analysis are listed here.

Bespoke database	Use
P&BA POLSAP	P&BA teams investigating trends in branch holdings take a daily feed of POLSAP data into Excel. This is now semi-automated through the use of a macro and takes approximately 20 minutes daily to complete. This tracks trends for branches for 10 weeks to detect unusual behaviour.
Conformance data	Information is kept in separate databases about branches which have been investigated or reprimanded for non-conformance or fraudulent behaviour. The multiple databases do not follow the same schema, with location and branch being recorded differently in the databases.
Branch Investigations	This is a log of every branch that has been investigated or where there was some kind of unusual activity. If a situation arises which might indicate the need for an audit, the P&BA team record this in a SharePoint system, which holds approximately one year's worth of data.
Recruitment risk register	If an individual has been barred from working with the Post Office they should be recorded in the recruitment risk register.
Credence in Network Architecture	All the transaction level data from Credence since 2004 held in multiple Excel spreadsheets and run through an Access database as required. This totals something equivalent to 30 billion records and is used for branch profiling purposes.

POL00042081

6 Fraud detection

Our principal focus within this review was on internal fraud. The vast majority of Post Office offerings are third party products, thus losses to the Post Office through fraud would largely be reputational rather than fiscal.

6.1 Teams involved

Fraud detection within the Post Office falls within the remit of the Security team - however, multiple teams are involved with analysing the data in an attempt to reveal potential fraud MOs.

6.1.1 Product and Branch Accounting

P&BA monitor Post Office data to detect behaviour which may indicate fraud or non-conformance. The main areas of investigation are;

- Client settlement; transactions/over the counter/web onward sending of funds;
- Managing discrepancies between transactions;
- Debt recovery; and
- General reconciliation.

Client settlement

Client settlement teams manage the payments that are owed to the suppliers of Post Office sold goods and services. For example Santander receives weekly updates of expected deposits and the APS (Automated Payments System) provides Santander a list of expected revenue from the week. This relies on sub-postmasters conforming to Post Office accounting procedures. Any degree of non conformance could mean a disparity between actual transactions and what the client is paid thus result in penalty charges.

Managing discrepancies

Discrepancies between expected and reported cash holdings are managed by issuing transaction corrections to branches, which creates a loss or a gain in the branch cash balance. Transaction corrections at crown offices are written off as a business loss / gain. If any of the other network branches have losses which they are unable or unwilling to pay the central Post Office, the value will be added as debt from the office. This debt is then pursued through the debt recovery process.

Managing discrepancies has seen a rapid reduction since the automation of most Post Office products. Manually keyed in products have an error rate of approximately 1 in 100, whereas the scanned goods rate is 1 in 4000. This has reduced the number of staff required to monitor error from 1000 to 80.

Debt recovery

Debt recovery is either current sub-postmaster or former sub-postmaster debt owed to the Post Office due to transaction discrepancies (transaction corrections) or cash loss.

Former sub-postmaster debt for 2011/2012 will exceed £8.7million, with new debt comprising £2.3million, £1.5million of which will be written off as a loss.

POL00042081

General reconciliation

General reconciliation of expected branch holdings and transactional data is a key part of identifying internal fraudulent activity within the Post Office. Cash holdings, cheques, rejected and spoilt postage and ATMs are all monitored for unusual trends across the network. Behaviour is analysed over a three month period by the fraud and non-conformance team within P&BA to generate an expected view of the branch activity.

This has been a successful way of picking up fraud at an earlier stage while the loss is at a lower value. P&BA do not investigate fraud, but can recommend an escalation to obtain more information, recommend training for the branch, request a case to be raised via Security or apply for an audit / intervention visit.

In the 2011/2012 financial year over 104 escalations were raised due to P&BA data reviews, estimated at approximately £1.5 million in losses to the Post Office. This represents loss due to non-compliance and fraud.

6.1.2 Security

The Security team investigates Post Office branches which have been deemed high risk. The initial flag for investigation can come from;

- Grapevine report;
- Customer complaint;
- Internal analytics by the Security team;
- Central analytics form the P&BA team;
- Police enquiry; and/or
- Internal investigation.

Behaviours which can trigger initial interest in a branch include;

- If a Crown branch has a loss of more than £250 that they cannot account for;
- Diversion from expected patterns of sales behaviour or cash holdings;
- Losses of more than £1000 from non-Crown branches that cannot be accounted for.

Credence is used to investigate the anomalous behaviour. If the transaction period in question is outside of the 90 day Credence window then transactions will be requested from Fujitsu. These requests can be for a branch for either a one day or month long view.

The Security team owns the relationship with Fujitsu in relation to data requests. As part of their managed service they have approximately 750 free requests per annum. Should they exceed this quota then they are charged between £400-500 per data request. Any requests which come from the Police as part of a criminal investigation must be honoured, which can put a strain on this system.

In addition to the Security team's audits, random audits are conducted throughout the year. These successfully identify non-compliant and fraudulent losses; however these are largely reactive tactics. A data analytics tool could provide a more systematic profiling process, using branch views over time to highlight potential problem branches before losses occur.

POL00042081

6.2 Fraud MOs

The Post Office Security team have built up a view of the types of internal fraud that the business is exposed to. Some of these fraud types are identifiable from current data, some are not.

The value of internal fraud within the Post Office is currently an unclear figure. There is no systematic account of what has been lost through specific fraud MOs. Some fraud types will have direct fiscal impacts but also reputational impacts for the Post Office either with clients or customers.

6.2.1 False Accounting

Many of the fraud MOs could be termed as false accounting. The simplest way for a sub-postmaster to achieve this is to record less takings in Horizon than were achieved and keep the difference. This can be via the ATM, cash or cheques. This is currently believed to be the biggest direct fraud loss to the Post Office.

How to identify automatically: Anomalous/high error rate compared to other branches of similar size/profile from Branch Database or Credence

Value to business: HIGH

6.2.2 REM reversals

Sub-postmasters may use REM reversals in order to hide cash loss. When a cash remittance is entered into Horizon, the physical pouch is scanned into the system, then placed into a safe until it is collected by the central Post Office. The digital money is entered in a suspense account.

If a sub-postmaster has a loss they wish to cover up then they may scan an empty pouch, and enter a fictitious amount into Horizon to balance the amount of cash they should be physically holding. If this amount remains in suspense for more than 41 days then it is automatically re-entered into the system. The sub-postmaster may reverse the amount back into the system to avoid detection, then simply REM the same amount out at a later date.

REM reversals can be done genuinely for legitimate reasons but can also be a way to steal money and cover up the loss of the physical funds.

How to identify automatically: Large sums being 'REM'd out/in' during certain timeframes; repeated instances of this occurring.

Value to business: HIGH

6.2.3 Payment reversals

Some payments for goods are manually entered into the Horizon system. Therefore a payment of £20 could be made, and then reversed out as the sub-postmaster claims that they mistakenly entered the amount, and there should only have been £2 worth of goods sold. The sub-postmaster then keeps the extra cash.

How to identify automatically: Anomalous/high error rate compared to other branches of similar size/profile from Branch Database or Credence.

Value to business: MEDIUM

POL00042081

6.2.4 Accepting counterfeit notes

Counterfeit notes identified in Post Office branches are returned centrally to the Cash Centres, entered on to POL SAP against the branch, but with the Post Office accepting the loss. If a counterfeit note is identified in a customer's banking deposit it can be claimed from the respective customer. Therefore some branches may accept counterfeit notes from beneficial clients, as they will still be paid for the transaction and the Post Office will compensate them for the note.
How to identify automatically: N/A.

Value to business: LOW

6.2.5 Pre-paid mail acceptance

Sub-postmasters may inflate the value of the pre-paid postage they accept in order to increase the amount of remuneration they receive for the transaction. The post that departs from the branch is not monitored for accuracy.

How to identify automatically: Anomalous/high value of pre-paid postage relative to branch profile from Branch Database or Credence.

Value to business: LOW

6.2.6 Multiple transaction remuneration

If a customer asks for a transaction which is large or has multiple parts, then a sub-postmaster may do this as multiple transactions, to increase the amount of remuneration they receive, e.g. a large banking deposit may be broken down into multiple smaller deposits.

How to identify automatically: Anomalous/high rate of transaction relative to branch profile, temporal proximity of transactions from Branch Database or Credence.

Value to business: MEDIUM

6.2.7 Non-barcode driven items

Items without a barcode can be charged at an inflated price and then the sub-postmaster keeps the difference. This potentially has a negative impact to the Post Office customer reputation.

How to identify automatically: N/A.

Value to business: LOW

6.2.8 Commission theft on currency exchange

A sub-postmaster might exploit the commission on currency exchange by completing sales outside of the Horizon system.

How to identify automatically: Monitoring of frequency of exchange rate requests not leading to transaction from Branch Database.

Value to business: LOW

6.2.9 Spoilt and rejected postage

When a postage label is printed, the printer can malfunction. If the label is unusable then the transaction is rejected, and has to be processed again. The sub-postmaster can print

POL00042081

two valid labels, sell one through Horizon, and sell the second one privately. If a transaction cannot be completed after the label is printed then the label should be 'spoilt' and stored in the branch for audit purposes. A sub-postmaster can claim a transaction was spoilt, and complete the sale privately but would not have a label to support the spoilt transaction.

How to identify automatically: Anomalous/high levels relative to branch profile from Branch Database or Credence.

Value to business: MEDIUM

6.2.10 Returns from pouches

Stock which needs to be returned to the stock centre, e.g. Camelot scratch cards after a game is closed, are REM'd out of the Horizon system and sent to the stock centre. There is a slip in the pouch which the sub-postmaster completes to state what should be in the pouch; this is then checked against the stock within the pouch but it is not reconciled against the Horizon system, as the stock and Horizon system are not linked. Although all Post Office Ltd stock is checked, only a 10% random check is done on Royal Mail products. Therefore offices could keep excess goods and sell them independently. In addition sub-postmasters can claim some goods never arrived at the branch and sell them independently.

How to identify automatically: Linking stock system and Horizon to reveal discrepancies at branch level.

Value to business: MEDIUM

6.2.11 Post Office card account

When a withdrawal is made from a Post Office card the sub-postmaster can claim the transaction did not go through and ask the individual to re-enter their pin details. They then keep the duplicate cash. This is particularly targeted at those collecting pension and benefits, who are vulnerable and less likely to notice the money has been taken. This potentially has a negative impact to the Post Office customer reputation.

How to identify automatically: Duplicate transactions and temporal proximity from Branch Database or Credence.

Value to business: MEDIUM

6.2.12 Annual leave

When a sub-postmaster takes annual or sick leave they are compensated by the Post Office to have someone cover their position. Sub-postmasters can claim to be on holiday and keep the compensation, whilst covering the position themselves.

This MO will diminish over the medium to long term as this payment is no longer paid with the new contracts signed under the Network Transformation Programme.

How to identify automatically: Horizon logon profile compared with holiday information from HR data.

Value to business: MEDIUM

POL00042081

7 Recommendations

The following recommendations are designed to enable the Post Office to achieve its core aim of improved fraud detection, through better use of its data.

7.1 Consolidation of data access and usage

Users across the organisation would benefit from enhanced ability to search, sort and analyse Post Office data.

As is discussed in the key recommendations section, a lack of understanding of Credence has meant much of the data analysis in the Post Office occurs in bespoke databases. This has created a number of problems for the organisation, particularly;

- Multiple versions of data means no single version of truth;
- The databases are operating beyond capacity and are therefore prone to failure;
- Databases are not future proofed, many are only usable by their current operators; and
- Limited sharing of data analysis across the business.

The Post Office therefore requires an effective data fusion capability which provides a single view of the branches, employees and customers. This would be realised through the creation of a new database including;

- Transactional information from Credence or Branch Database;
- Customer information from the Post Office or third party suppliers;
- Employee information from HR SAP; and
- Stock movements from the Galaxy system.

Ideally the new database would also have a powerful data analytics tool and a new, more user friendly, front end.

7.2 Enrich available data

Effective data analytics relies on good quality data. When its purpose is to underpin an automated fraud detection system it is vital that this data is rich in entity information (names, addresses, telephone numbers, bank accounts and other individual identifiers). Improved data quantity, quality and access would enhance the Post Office's ability to;

Profile branches and products

Profiling can be achieved through the unique identifiers for products and branches held in Horizon transactional data. Access to this data, coupled with an enhanced data analytics tool and new user friendly interface, would allow analysts and investigators to navigate through data easily and intuitively. This would have widespread application across the Post Office. The addition of a transactional analysis capability would allow the identification of potential indicators of non-compliance and fraud. This would be reliant upon established fraud MOs that were identifiable purely through transactional patterns.

POL00042081

Identify 'internal' fraud and single view of the employee

The initial set of data analytics would enable profiling of individual Horizon IDs. This could provide a single view of an employee's interaction with all Post Office products at transactional and summary level. Matching the Horizon IDs to individual names and personal details through the use of HR SAP would allow for improved analysis of risk, particularly in conjunction with data from third parties. Third party data sets of particular value would be;

- Credit reference, e.g. Equifax or Experian;
- Identity data, e.g. GB Group;
- Fraud data (confirmed or suspected) from the wider fraud community, e.g. National Fraud Intelligence Bureau (NFIB) and Insurance Fraud Bureau (IFB); and
- Other government agencies and industry partners.

Identify 'external' fraud and single view of the customer

The Post Office does not bear the financial loss for third party financial and insurance products. Nevertheless it is important for the organisation to understand if and where fraudsters are targeting Post Office-branded products and branches.

The third party product providers are likely to have their own anti-fraud measures in place. However fraudsters are adept at identifying and exploiting weaknesses within fraud detection capabilities. Therefore it is important for both the Post Office and the third party provider to have a view of how their Post Office-derived customers are interacting with third party providers.

As the Post Office does not generally hold personal customer data, claims data and post-account creation transactional data for these products, it is difficult to understand the scale or nature of this fraud. Identifying and matching data across multiple data sets is vital in building a full picture of an individual or group's activities and contacts. Automating the building of these 'social networks' is in turn critical to understanding fraud activity associated with an individual or account – and indeed across a network. To develop this view, data must have entity values, e.g. names, addresses, phone number, account information, email address etc.

Any comprehensive anti-fraud solution must be supported by customer data, allowing for a single view of the customer to enhance efforts at understanding and mitigating fraud.

Enhancing the relevant datasets in this way will pose some challenges, principally in capturing and storing customer data. However it is critical for the Post Office to develop a comprehensive analytical and investigative platform. With the enriched data a system can provide automated fraud detection, single view of customers, employees, branches, detailed management information and multiple analysis and research capabilities to benefit the wider user community.

7.3 Automated identification of risk

With a richer set of data a sophisticated fraud detection, analysis and investigation platform could be implemented that would achieve;

POL00042081

- Automated identification of risk and potential fraud linked to branches or employees;
- Prioritisation for further investigation through the use of sophisticated scoring models (to identify entity and network behaviours of interest);
- Proactive identification of fraud trends including, for example, specific geographical areas and products that are susceptible to fraud;
- Enterprise search capability to allow users to rapidly identify all data of interest; and
- Provision of Management Information (MI) across all relevant areas of the business.

In essence, this type of capability would provide the Security team and the wider organisation with a comprehensive analytical and investigative tool that would realise a wide range of benefits. Automated identification and prioritisation of risk removes laborious research and analysis processes that currently occur every day. In addition the ability to easily search and sort data will allow for user driven analysis and investigation across the business.

POL00042081

8 Solutions

There a number of options the Post Office could pursue to improve their current fraud detection. These offer different levels of capability, at different costs, timescales and project prerequisites. Each option can be combined in full or in part with the other options. The strengths and weaknesses of each option refers to the situation in the Post Office today, they are not innate weaknesses of the systems themselves.

Solution	Purpose	Estimated Cost	Prerequisites	Project timescale	Strengths	Weaknesses
1) Stand-alone data analytics	Understand current trends in the transactional data for the Post Office, from Branch Database or Credence. Provide a breakdown of the data by branches, employees (Horizon IDs) and products.	75-100k	Provision of data from Post Office/third party providers; Decide on what analysis would be valuable to obtain.	1-2 months	This data analysis would present usable intelligence from complex, large-scale transactional data, from easily obtainable sources with no additional customer or employee data required. It would clarify what can or cannot be derived from the available data and provide a foundation for future implementation of a more comprehensive solution.	Provides a one time view of the data, without any ability for the Post Office to continue their interaction with it. It would be expensive to do continually and of limited value for fraud detection without investigation and further data.
2) Record-based enterprise search and analysis	Understand current trends in the transactional data for the Post Office, from Branch Database or Credence. Provide a breakdown of the data by branches, employees (Horizon IDs) and products.	75-100k	Provision of data from Post Office/third party data providers; Decide on how the search tool should categorise data and how it should be to interact with.	1-2 months	A low cost option to sort data and improve trend analysis. Enables quick and easy access for Post Office employees and could provide alerts for perceived fraudulent behaviour.	Without initial data analytics as described in option 1, it would be difficult to define risky behaviour to search for without encountering a high rate of false positives, as there would be no baseline. Enterprise searches work best on entity based data, but much of that will be missing here.

POL00042081

3) Transactional analysis	To allow user-driven interrogation of transactions to identify specific behaviour of interest (in relation to specific branches, employees, and products). Automated search across all data for specific fraud MOs visible in transactional data.	100-500k	Provision of data from Post Office/third party data providers; Identification of key users to contribute to the building, development and maintenance of the system; Understanding how fraudulent behaviour appears in transactional data.	2-3 months	Allows for user-driven interrogation of transactional data to reveal potential non-compliance or fraudulent behaviour. Automatically identifies data of interest across high volume of transactions much more rapidly and effectively.	Transaction-based, so limited in terms of an entity-based fraud analysis and investigation capability; more expensive from a product licensing and implementation perspective than options 1 and 2. It also requires more commitment from Post Office during scoping, implementation and day-to-day use and maintenance of the system.
4) Network-based automated detection of fraud and risk, including enhanced analysis and investigation capability	Provision of a comprehensive data fusion capability that will provide automated risk/fraud detection, enterprise search, management information and complementary types of analytical and investigative functionality for use across the Post Office.	1-2 million for product licence and implement ation*	Provision of all relevant data sets including customer/employee data from across the Post Office enterprise; Detailed understanding of Post Office business and system requirements from across the user community; Identification of key users to contribute to the building, development and maintenance of the system.	6-12 months	This option will provide the capability for proactive fraud identification, investigation and prevention. The breadth of automation and user driven access will allow Security and the wider business to achieve savings. This will include significant financial benefits (through direct savings and efficiencies) and through wider business benefits (increased reassurance to third party government and commercial partners, improved reputation for combating fraud/loss).	Relies on entity-rich data not currently available within the Post Office environment; longer implementation time; greater costs attached to licensing and implementation of the solution.

*Potentially significantly more if multiple products required for integration

NB: Each option description is dependent on the specific scope of the project, but should provide an indicator to support decision making in relation to next steps. Prices are not including any costs associated with extracting data from Fujitsu or other managed service providers.

POL00042081

8.1 Our recommendation

Our recommendation is that the Post Office begins its data journey by completing option 1, combined with a pilot of option 2. The Post Office is currently not sufficiently mature to take advantage of the full benefits of option 3 or 4 – although options 2,3, or 4 could all be possibilities for the future.

As the true value of fraud in the Post Office remains an unknown and there is no clear understanding of how it can be best identified in data, the correct tool to tackle fraud can not yet confidently be selected. A clear view of normal business operations is key to both writing rules and successful transaction monitoring. If a comprehensive data analysis is undertaken it will expose the picture of typical branches, employees and products. This will provide the basis for a more comprehensive on going solution. In addition it will clarify the need for further, possibly external, data.

Simultaneously an investigative tool should be piloted to allow the exploration of historic fraud cases and known MOs to show how fraud appears in data. Combined this knowledge will clarify what is the right solution for the Post Office to progress with.

To achieve desired results the Post Office should address the issues highlighted in the recommendations section of this report. To achieve options 1 and 2 the Post Office will need to;

- Obtain a data sample from its hosting providers; and
- Agreement from the wider business on the incorporation of employee information into the core dataset, including plugging any gaps in employee data (such as sub-postmaster employees not currently captured by HR).
- Establish appropriate project management and oversight functions for the life of the implementation (and beyond).

If appropriate, to move forwards into option 3 or 4 the Post Office will also have to;

- Construct a model of suspicious behaviour with data experts and those involved in fraud detection.

For option 4 there will be two further tasks;

- Identify core users for the analysis and investigation function that would be driven by this capability, potentially including the recruitment of new staff for this purpose; and
- Investigation of options for incorporating customer data from third party product providers; and

Oversight of the project is unlikely to be a significant issue for the Post Office given its experience of managing large data/IT projects. However the other tasks will undoubtedly present some challenges.

We believe the best way to approach implementation of a full solution is in a phased manner. This approach has the benefit of providing a strong foundation for the development and implementation of a comprehensive anti-fraud management tool, whilst providing benefits to the potential user community as the project progresses.

POL00042081

8.2 Solution implementation roadmap

8.2.1 Phase 1 - Initial data analytics and the provision of search and analysis capability (supplier); preparation for future implementation, including scoping and resolution of data quality issues (Post Office)

Combining solutions 1 and 2 in a "proof of concept" environment will reveal multiple key findings from Post Office data. Work in this phase will effectively profile the business at all levels and provide useful intelligence for the Security team and beyond.

In order to do this successfully any supplier is likely to require data from at least Credence, Branch Database, HR SAP, POLSAP and MDM Reference Data. As no single database can identify the journey and impact of a transaction, (from counter creation to remuneration) it is important that a slice of data from across these databases is available for analysis.

During this phase the Post Office can work to fill any recognised gaps related to data quality/availability, business processes, staffing and any other areas identified during pre-project discussions.

Articulating direct financial benefit (through loss/fraud savings) is challenging given the recognised gaps in understanding the scale, nature and impact of fraud (as mentioned previously in this report). It is, however, anticipated that there will be numerous tangible business benefits realised by the Post Office by the end of phase 1.

During this period the Post Office will benefit from the findings of the data analytics workstream and also allow users to become comfortable with a new user interface which will provide access to, and categorisation of, data in ways that have previously been difficult (if not impossible) to achieve through current databases. It also provides a project window within which the Post Office, in conjunction with data and industry partners (where applicable) and through support from their chosen supplier(s), can create the conditions required to progress to phase 2. This should facilitate an easier and more successful implementation of a full analysis and investigation capability.

8.2.2 Phase 2 - Currently unknown

Phase 1 may expose that a complete picture of the Post Office data with a comprehensive sort and search capability it is all that is required to prevent the majority of internal fraud, if coupled with proactive investigation. In this case formalising option 2 may be sufficient to tackle fraud.

Alternatively, should there be strong evidence of fraud by deviation from an accurately calculated, branch, product or individual profile; a transaction monitoring tool may be appropriate.

Finally, if an initial analysis indicates that a transaction based view would be overly prescriptive and give insufficient context to anomalous behaviour, then a full network solution utilising third party data may provide the optimum solution.

POL00042081

A Appendix 1

Further system components which were identified but are unlikely to play a part in a fraud solution are included here.

A.1 Databases

A.1.1 Salesforce

Salesforce is currently only available to Financial Services experts in branches on specific laptops. As part of the Post Office 3 year IT strategy it is intended to hold a holistic view of the customer's holdings within the Post Office offerings.

A.1.2 Post Office Financial Services (POFS)

POFS is an Oracle database that is the first point of contact for financial services provider's data to enter the system. For certain product feeds (such as BT home bill payments) the data is transformed before it enters Credence so it is usable for remuneration purposes. Largely, it just transfers data to Credence and sends customer data to the Brands database.

A.1.3 Brands

The Brands database is used by the Post Office for marketing and advertising. It is fed by third party data feeds from Post Office Financial Services. It is being developed to provide a single view of the customer for the commercial and advertising teams. It currently holds some transactional level data, but sporadically. Coherence of content will be part of its on-going development.

A.1.4 Stock system

The stock system operates the warehouse control system and manages transtrack/different delivery companies. It generates consignment item numbers, and is exclusively used by the national stock team.

Currently the stock system is not linked to the Horizon system, so the warehouse cannot track stock movements from branch to the warehouse. The business relies on sub-postmasters accurately reporting to the warehouse what they have returned, as they cannot verify that with what has been removed or 'REM'd out' from the branch system.

A.1.5 Grapevine

The Grapevine system allows Post Office staff to report crimes to Kings Security, e.g. use of false notes in branch. If they think there is a risk of duplicate crimes occurring in the local area they will notify other high street businesses in the area via text blasts. Post Office branches can choose to join this system.

Currently the Post Office and Kings are in the process of developing multiple activities that tackle physical fraud within the Post Office network.

A.1.6 Customer database

The customer database shows all of the products the Post Office hold. It is currently used by the sales team.

POL00042081

A.1.7 Sales centre

The Sales centre shows trends in sales and the market for Crown Offices. It is currently used by Brands.

A.1.8 ICOW database

The ICOW database holds signage, floor plans, design manuals and employee data for all the Crown offices. It is currently used by Brands.

A.1.9 APS (Automated Payments System)

Converts a branch database feed into expected payments which are owed to each client. This feed is also sent to TPS.

A.1.10 TPS (Transaction Processing Systems)

TPS takes the client payment feed from APS and compiles reports for the Post Office on sales and marketing. A feed of those sales figures are summarised and sent to POLSAP and a transactional view is sent to Credence.

A.1.11 DRS (Data Reconciliation Service)

DRS reconciles Branch Database transaction data and incoming banking data. If the two data sets do not match then a report is sent to P&BA who identify the cause of the error.

A.1.12 Online Mapinfo

In the near future the Network Transformation team intends to launch an intranet mapping view of Post Office data. It currently exists but is largely outdated.

The new version will be overlaid data onto Google maps with some of their most common information queries available, including appropriate security controls.

A.2 Batch feeds

These data sources provide Post Office transactions that were completed out of the till, therefore out of the Horizon system. They batch feed into the Horizon data in the system, either going into Branch Database or Credence.

A.2.1 Post and Go

Post and Go machines offer goods and services as an alternative to the counter till. This data is owned by Wincor and provided to the Post Office. The Post Office will then send the expected revenue amount back to the branch for confirmation, and the takings are then put into the till. Post and Go has experienced some problems completing transactions meaning duplicate transactions have had to be completed in Horizon. The Post and Go data feed now joins Horizon before it goes into Credence and Branch database.

A.2.2 Paystation

Paystation terminals can process bill payments when Post Office counters are shut. The system records the transaction which is owned by Ingenico. This information is then sent in a batch to the Horizon feed.

POL00042081

A.2.3 AEI

Holds the biometric data services which the Post Office provides for the DVLA and others. The system is owned by Cogent and like the Post and Go machines the transaction feed enters the Horizon feed. However the biometric information itself is securely stored separately.

A.3 Data gateways

The Post Office system has many incoming data streams, APOP and EDG act as a conformance check and to transform and load data.

A.3.1 APOP (Automated Payments Out Payments)

APOP translates incoming data to the Post Office and acts as a processing bucket. It checks data for conformance to rules and reject anything that does not conform before it reaches the Post Office Data Gateway. Any data which is rejected will be sent to the Post Office admin team to resolve the issue and have the data resubmitted.

A.3.2 EDG (External Data Gateway)– transitioning to PODG (Post Office Data Gateway)

The EDG extracts data from non-Post Office systems, transforms it into a useable format for internal databases, and loads it into the final destination. It is also able to send data to clients. Largely it is focused on transforming and loading incoming third party data for systems, particularly Credence and POLSAP. It provides the ability to send different sub-sections of the data to multiple destinations, using the same ID number. An example of this is the data captured in a Moneygram transaction. Anti-money laundering legislation means some data collected in a Moneygram transaction is not for the client; instead it is captured and provided to the government. Data can be transferred via web, FTP, connect direct and secure email (for internal data transfers only).