ROZETTA TECHNOLOGY

HOW TO BUILD AN INTEGRITY MONITORING PLATFORM TO MINIMISE FRAUD, ERROR AND WASTE WITHIN A COMPLEX MARKET?

UNLOCK THE FUTURE

RoZetta Technology is a specialist data science technology company using advanced analytics and cloud-based data management to provide a foundation to optimise value creation for its clients.

Anomaly detection is just one of the methodologies the RoZetta data science team employs to fulfil the value creation goals.

Prepared by:
Phil Anderson, Head of Product Management & Strategy
Scott Matthews, Chief Data Scientist
Dr Inigo Jauregi Unanue, Data Scientist

TABLE OF CONTENTS

1. INTRODUCTION

Some industries rely on third parties to deliver complex services to end-users, including health services, education, wealth management, capital markets, taxation services and some franchise operations.

This environment is susceptible to error, waste, abuse, and fraud as the service deliverer, or intermediary, behave in a way inconsistent with the organisation's intention and expectations.

Anomaly detection, also referred to as outlier or novelty detection, identifies events or observations that deviate significantly from the expected result, or a majority of the data, or do not conform to a well-defined notion of normal behaviour. These are often related to behavioural events of error, waste, abuse, or fraud.

This article focuses on anomaly detection, an important tool in monitoring human and systems interactions that are too complex for a human alone to comprehend.

In a services business, anomaly detection identifies behaviours or events that deviate significantly from expected outcomes in systems or processes. These anomalies should result in an investigation to determine the cause.

In an enterprise with a very high transactional environment, complex delivery channels, and an evolving compliance regime, anomaly detection discovers inefficiencies and abnormal behaviours that result in unnecessary resource consumption and expenditure.

2. DEFINITIONS

The following defines the conditions an anomaly detection system seeks to identify and measure. The system will also surface the components of each instance identified to assist with root cause analysis.

- **Error:** An innocent act caused by lack of training, carelessness or failure to comply with standard practice.
- **Waste:** Involves the end-users, the customer, not receiving reasonable value for money in connection with any service delivered. Waste includes incurring unnecessary costs resulting from inefficient or ineffective practices.
- **Abuse:** Any practice exaggerating the time and resources used to provide end-users services. It includes practices that result in unnecessary costs, like payment for unnecessary services and exaggerating the time taken to provide a service or services provided at a time or place other than expected. Abuse also relates to intentional practices that fail to meet the organisation's recognised standards for service delivery.
- **Fraud:** Any deliberate deception or misrepresentation made by a person, knowing that the deception could result in payment of unauthorised amounts to themselves or some other person. It includes any act that constitutes fraud under applicable federal or state law.

3. REQUIREMENTS

When establishing the scope, it is critical to identify the data sources and systems needed to manage the complexities of real-life relationships and behaviours.

Make it easy for users and stakeholders to share insights, and make provisions to find out why. User groups and stakeholders won't stop at one metric; they will want complete profiles of both the anomaly and the entities involved. The system must be accessible at scale, as many users will need to access it simultaneously.

Work on minimising false positives from the start. It will minimise the motivation to switch off the anomaly detection messaging because of false alerts. It is also crucial to ensure that any alerting avoids overwhelming users with a large volume of alerts leading to alert fatigue and missed critical issues. This approach will also minimise wasted effort, but most importantly, it will build the credibility of the system and the analytical outcomes.

Cluster/identify anomaly types and find new anomalies as data evolves and behaviour modifies. Some participants may seek alternate methods to disguise waste and abuse. Developing the systems to identify 'like' behaviours that result in the same outcome is important for identifying anomalies.

After finding an anomaly, those delegated to act on the findings will quickly want to know why. Consider how the system will provide guidance and access to pursue root cause analysis.

4. IMPLEMENTATION OVERVIEW

There are four stages in developing an anomaly detection system (see Figure 4.1):

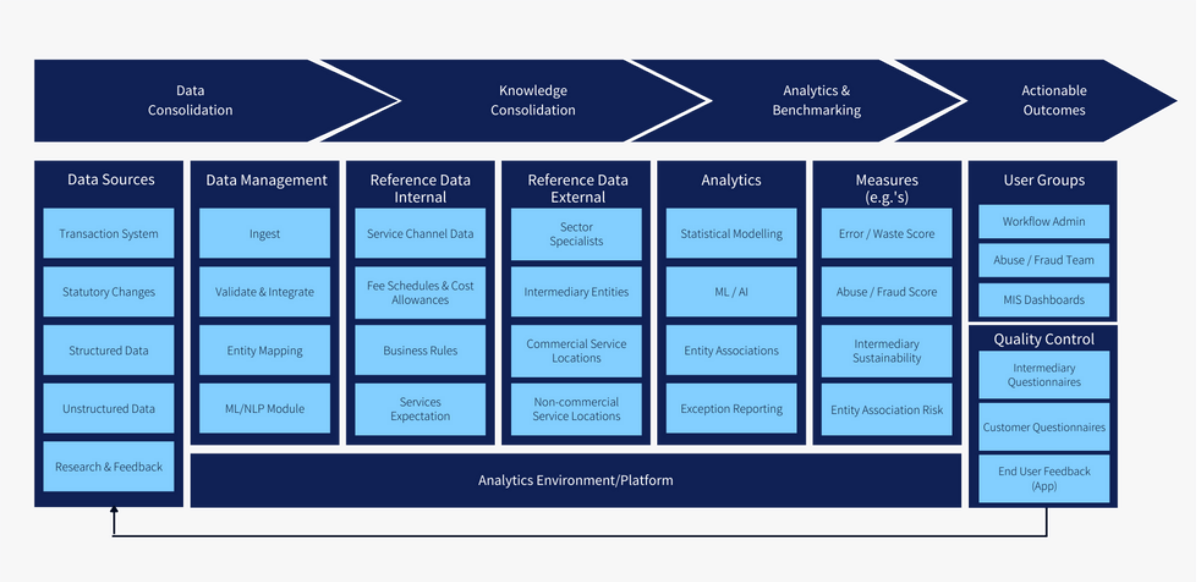
1. Data Consolidation
2. Knowledge Consolidation
3. Analysis and benchmarking
4. Actionable Outcomes

Data Consolidation and Knowledge Consolidation establish the system's foundation. Data Science plays a vital role in this phase, particularly in integrating unstructured data and developing meaningful reference data.

This is followed by Analysis & Benchmarking to establish the core capability set and baseline measures. Actionable Outcomes integrates anomaly system outputs into the organisation's MIS and workflow in the final stage.

Establish a quality control process by instituting feedback paths inherent in a closed-loop system. The feedback loop comprises customer and intermediary research. It may also include an end-user App to collect information confirming services have been provided and some commentary on the quality of the service.

Figure 4.1 Four stages of developing an anomaly detection system.



5. DATA CONSOLIDATION

In establishing an anomaly detection system, data is the foundation. Decide on the technical environment to host the solution and identify all the internal data sources used to trigger a payment to either an intermediary or an end-user. Data consolidation entails capturing every financial transaction, statutory requirements on eligibility to deliver the service, end-user entitlements, and payment limitations. Each data source needs to be updated frequently for anomaly detection to be effective. For predictive modelling, streaming transaction data would be the best option.

Capturing all the required data can be challenging if a significant volume of data is unstructured, i.e. captured from forms and reports. Introducing Data Science capability is essential to ensure the maximum utility of unstructured data sources. Integrating structured and unstructured data into the core knowledge bases is critical.

6. KNOWLEDGE CONSOLIDATION

Establishing reference data sets incorporating internal and external business rules and service expectations and integrating all these data sources moves the system to the Knowledge Management phase.

All the operational data created to manage payments is required, including detailed profiles of every intermediary. These profiles are important in the analysis to ensure intermediary comparisons are relevant. All available historical data is the basis for time-series analysis and building credibility in the modelled expectations.

This stage includes codifying business rules needed to establish benchmarks and develop ratings or rankings for each intermediary layer involved in delivering the end service to customers.

Business and compliance rules, like fee schedules, accreditation, and customer assessment guidelines, are encoded in this stage.

Business rules can be either explicit or implicit.

1. Explicit Business Rules:
An organisation sets explicit rules governing how all business areas operate. They align to compliance and operational procedures and ensure both the business and all intermediaries comply with government statutes and regulations.

2. Implicit Business Rules:
Some business rules are implicit and driven by expectations. These may include the time and place of service provision, the range and mix of services an intermediary may provide, and the likely catchment area to be covered.

In a business with a range of fees for the same service, the distribution of rates within the fee range would fit an expected pattern, i.e. a business expectation.

When administration costs make up a significant part of any entitlement, there is likely to be an implicit expectation on the proportion of total payments to cover the running costs of an intermediary. There may also be an expectation that as intermediaries become more experienced, they become more efficient.

Implicit business rules change as the business matures and experience challenges existing rules and expectations.

Developing a comprehensive knowledge base for all participants provides a means of accurately mapping out relationships between participants. It is critical to update the participants' knowledge base regularly as new participants emerge and old ones disappear, achieved by combining technology, artificial intelligence models and human supervision. A complete data set for each participant type enables the discovery of any possible collusion between participants.

Adding spatial data to the knowledge base will allow for analysis of the logistics involved in service delivery. Spatial data is fundamental when travel allowances are part of the cost of service delivery.

7. ANALYSIS AND BENCHMARKING

Analytics is the critical component of any anomaly detection system. The initial analysis determines the benchmarks and tolerance levels for activity and interaction.

There is no one statistical approach to anomaly detection. It is an array of capabilities, each playing a role in identifying different aspects in discovering error, waste, abuse or fraud

Analytical Approach	Anomaly Examples
Rule-Based Exception Detection	• Behaviour inconsistent with documented business rules and compliance standards. • Behaviour is inconsistent with comparable cohorts.
Anomaly Algorithms	• Outliers: Anomalous patterns appear non-systematic in the collected data. • Change in Events: Systematic or sudden changes from the previous normal behaviour. • Drifts: Slow long-term changes in intermediary or participant behaviour.
Predictive Analytics	• Discover changes in behaviour that indicate an increase in the probability of error, waste, abuse or fraud. • Discover any patterns that show a progression from error to waste and waste to abuse.
Network & Entity Relationships	• Explore and measure all participants' interactions. • Discover any probable instances of cartel behaviour between intermediaries.

A critical part of ensuring the best outcomes of analysis and benchmarking is that ongoing data collection is supported and maintained.

8. GUIDING PRINCIPLES

When designing an anomaly detection system, some guiding principles, or strategic drivers, of the design and implementation are:

Build a risk profile of each layer in the service delivery network, both internal and external.	A sustainable service delivery ecosystem needs to understand participants' financial, operational, and regulatory risks. Data collection should be broad enough to address these issues.
Eliminate alert floods and alarm fatigue. Minimise false positive alerts from the start.	Where appropriate, consolidate all related performance issues into a single actionable notification—less noise, more problem-solving. Determine how to aggregate alerts and target them to roles that have a delegation to resolve them.
Focus on the problems that matter	Not every threshold violation is a problem, and not all problems are created equal. Use the analytics to determine the level of impact an anomaly has and the actual or potential impact on the organisation and end-users or customers.
Rank and prioritise alerts that require action.	When scoring and ranking anomalies, the system should consider the seriousness of the anomaly detected. These include the potential impact on end-users, the cost of taking action/chasing the alert and if it is a fraud alert, the set of benchmarks that determine who investigates the anomaly. An appropriate communication (reminder) if the cause is likely to be error or waste.
Automate dynamic problem detection.	Integrate multidimensional baselining and predictive analytics to automatically detect when things aren't behaving as expected. Only communicate issues to specific participants where they have delegation and capability to take action.
Set benchmarks for triggering surveillance.	Create a learning system that evolves in response to ever-changing behaviours and threats. This Includes creating a feedback loop from end-user experience and activity for continuous monitoring of the provider delivery performance.
Emphasise prevention to get the best return on effort.	Research on payments in health identified 25% of unnecessary payments were attributed to error and waste. Identifying weaknesses in compliance with business rules and procedures will have the most significant impact and cause the least disruption to the business in reducing unnecessary payments.
Use choice architecture to encourage compliance.	Choice architecture is the design of different ways in which choices can be presented to users and decision-makers and measuring the impact of that presentation on decision-making.
Build a system where the inner components or logic are available for inspection.	Create a white-box learning system (as opposed to a black-box system) that allows for the interpretation of what triggered/caused the system's predictions. It helps decision-makers to take action confidently.

9. BUILD PRIORITIES INTO THE SYSTEM

Start with waste, abuse, and error. Studies of benefits systems in the US reveal that fraud makes up less than 10% of overpayments and unnecessary payments. Error and waste cause the most frequent, and costly, excessive payments, and good processes dealing with these behaviours will help protect the network from instances of fraud.

Use data to inform the choice of interventions. Each instance of suspected error, waste, abuse, and fraud should be scored on three scales:

1. Frequency of occurrence.
2. The total value of the discrepancy.
3. The volume of transactions involved.

These scores would drive the management of each anomaly detected.

Make data security, cybersecurity and identity management central to program integrity. The personal data captured in any payment system is privileged and must be protected. Employing encryption keys helps mask any sensitive information. Systems operators and administrators cannot identify individuals mentioned in any transaction, but analysts know it is the same entity across many transactions and behaviours. It may also encourage the adoption of cloud-based systems as the identity of individuals cannot be determined without the encryption key.

10. SUPPORT A COMMUNICATIONS AND FEEDBACK PROGRAM

Develop a closed-loop information system. A prevention-focused strategy can be doubly lucrative. Prevention saves the cost of overpayments and the cost of the chase.

Emphasise prevention to get the best return on effort. Incorporate feedback loops at all levels to collect data on performance, timing and quality of service delivery.

Systematic surveillance will improve compliance, reduce error and abuse, and make possible fraud more obvious.

Part of the feedback loop is to incorporate a manual review of a sample of the anomalies to ensure the continuous training of the various models used in the analysis.

11. BEHAVIOURAL NUDGES BUILD PROGRAM INTEGRITY

("Speed camera with no camera")

Behavioural economics provides a growing source of ideas where direct feedback can “nudge” participants in any service delivery program. Participants, internal or external, can be nudged to comply more fully.

Nudges delivered through carefully worded communications can achieve impressive results at a much lower cost. In the form of personalised feedback on deviations from expected behaviour, evaluations against benchmarks, or even a simple checklist of desirable action steps, Nudges can be used effectively during the in-line processing to prevent error, waste, abuse and fraud.

12. ACTIONING THE OUTCOMES

The benefits materialise when the outputs of the anomaly detection system drive a responsive action. Three basic workflows target different players in the service delivery process, each providing measurable gains for the organisation.

The first is using errors and waste, depending on the frequency of occurrence, as an opportunity for retraining. The organisation will need to develop metrics for measuring how effective the training is in reducing identified errors and waste.

Not all anomalies are negative. A continuous lower cost of administration may impact an intermediaries sustainability. In these instances, training in better managing a business may result in a more stable business model. This is particularly important where the intermediary is part of an independent delivery channel.

The second aspect of workflow integration is using the frequency, value and volume scores to allocate the follow-up and assessment of an anomaly or set of anomalies. This approach ensures allocation to specific workgroups with the appropriate skills. These scores also allow for consistent priorities. For example, every instance of excessive payments is not automatically considered fraud and not every fraud consists of large amounts of expenditure over a short period.

Dashboards and MIS are the third workflow integration. Developing a summary dashboard for all intermediaries to access is important for the nudging process. The dashboards would show the relative position of each intermediary with intermediaries with a similar profile. These dashboards continuously remind intermediaries of the surveillance and monitoring of their delivery channel.

13. QUALITY CONTROL

Establishing consistent feedback loops is essential for assessing the effectiveness and quality of the services provided.

A mobile application will establish this feedback from end-users. Best practice, particularly where the end-user may not be familiar with mobile apps, would be to augment the mobile application with research projects and questionnaires targeting each step of the delivery value chain.

This approach is vital in determining the validity of some business rules and establishing qualitative benchmarks that may drive adjustments to existing business rules and standards.

14. ANOMALY DETECTION IN OTHER ENVIRONMENTS

RoZetta, using its DataHex platform, has delivered anomaly detection solutions in different environments. They continue to add value to energy, utilities, health sectors, and digital messaging organisations.

In a water supply network, sensors measure the pressure, directional flow and volumes of water. The first benefit was immediately detecting supply interruptions by location and severity. In the agricultural sector, sensors measured the volume and purity of a liquid product as it moved through the logistics chain. The system could identify changes in volume and detect pollutants in the transfer of liquid from one container to another. An anomaly system identified possible fraud in digital messaging, comprising SMS, email and social media platforms

ABOUT ROZETTA TECHNOLOGY

RoZetta Technology's core belief is that fusing data science, technology, and data management is the path that amplifies human experience and knowledge.

RoZetta's clients have an understanding of the challenges they face. RoZetta brings proven capability, experience and a mindset to create products and systems that resolve these challenges in a way that optimises the value created.

The mix of good data, the right technology, the right design and smart people can solve complex problems.

REFERENCES

In developing this paper, RoZetta's data science and management team have called on their experience and research. The following are some of the sources used in creating, reviewing and validating the content of this paper:

- 5 Insights on AI to Detect FWA in Healthcare from AHIP, Brighterion AI, 2021
- Using Symbolic Data Analysis to Detect Fraud, Waste, and Abuse in Healthcare Insurance Claims Data, Federick Jonathan Reynolds, Auburn University, Alabama, 2, 2020
- Shutting down fraud, waste, and abuse, Deloitte University Press, 2016
- A Better Approach to Avoiding Misconduct, Harvard Business Review, 2022
- Cracking down on government fraud with data analytics, McKinsey & Co, 2018
- OIG Oversight of the Unemployment Insurance Program, 2022
- Background information: Fraud and error in the benefit system statistics, 2021 to 2022 estimates, UK Department for Works and Pensions, 2022 MS, email and social media platforms

ROZETTA
TECHNOLOGY

Published by RoZetta Technology June 2022
www.rozettatechnology.com
enquiries@rozettatechnology.com