Great Expectations:
How to Detect and Prevent Fraud using Data Analysis

Peter Millar
Director, Technology Application
ACL Services Ltd

peter_millar@acl.com | twitter.com/PBMillar
called myself Pip, and came to be called Pip.'

- Great Expectations, a novel Charles Dickens (1861)
- Major themes are of crime, and ambition
- Starts with a 6-year old boy stealing food to help an escaped convict,
  - who threatens him not to tell a soul or he will cut out his heart and liver

- Not exactly a bed-time story but one that aptly suits a discussion on fraud how to live up to the great expectations of management in detecting and preventing fraud
Using Data Analysis for Fraud Detection

1. Test against 100% of transactions
2. Automate testing to enable:
   - Continuous assessment of problem areas
   - Scheduled repetitive monitoring of other risk areas
   - Increased efficiencies in identifying indicators of fraud
3. Access and relate data from virtually any source
   - Internal or external to organization
   - Without moving sensitive data outside of the secure data center
4. Identify where automated system-based controls:
   - Are not functioning effectively
   - Do not apply to the business process (manual controls only)
### Application areas for data analytics in fraud

#### Standard Business Processes
- Procure-to-Pay
- Travel & Entertainment
- Corporate Cards
- Order-to-Cash
- Payroll
- Inventory and Materials Management
- Capital Assets

#### Vertical Business Processes
- Insurance Claims
- Healthcare
- Financial Services
- Manufacturing
- Retail
- Construction/Engineering Contracts
- Telco

#### Financial Statements and Reporting
- General Ledger
- Revenue Recognition

#### Information Systems
- Segregation of Duties
- Systems Access
- Master Data Files
- Configuration Settings
How...

...to integrate analytics into fraud detection work
It’s a Journey
Think in Terms of **People & Process**
Analytic Capability Model

- Level 1: Basic
- Level 2: Applied
- Level 3: Managed
- Level 4: Automated
- Level 5: Monitoring

Foresight: ad hoc → repetitive → continuous
Insight: basic → managed → automated
Hindsight: applied → monitored
Level 1: Basic

Characteristics
- Leverages technology
- Ad hoc query and analysis
- Used to perform analysis of large data sets

Benefits
- Rapid insight into entire data populations
- Increased ability to detect fraud, errors and inefficiencies
- Increased quality and levels of assurance
Level 1: Basic – Fraud Management Role

1. Encourage experimentation
2. Ensure analytic objectives are realistic
3. Facilitate the access to data
4. Apply technology that can support growth
5. Provide support for training, as necessary
Level 2: Applied

Characteristics
- Analytics fully integrated into fraud detection process
- Comprehensive suite of repeatable tests are designed and developed
- Applied to a range of fraud detection and control objectives

Benefits
- Gain greater insight into targeted areas
- Improved quality and reliability of tests
- Improve efficiency & effectiveness of anti-fraud activities
Level 2: Applied – Fraud Management Role

1. Clearly set your expectation for using analytics
2. Define roles for both technical and non-technical staff
   - Brainstorming
   - Analytic development
   - Testing & QA
3. Integrate analytic planning as part of the normal processes
4. Allow time for data acquisition & profiling
5. Build analytic expectations into investigation and review processes
6. Provide effective team incentives
– assessing the risk of fraud through Vendor Master Data

Scripps Health is a non-profit, community-based health care delivery network in San Diego, California, that includes:

- Four acute-care hospitals on five campuses
- More than 2,600 affiliated physicians
- Extensive outpatient care network
- Home health care
- Associated support services
Vendor Master Analysis

- **5,941**
  - vendors with no Taxpayer Identification Number (TIN)
- **15,268**
  - vendor records with no phone number.
- **11,497**
  - vendor records with no Vendor Contact.
- **23,391**
  - vendors with no email address or website entered.
- **548**
  - Active Vendors in the Vendor Master Not Existing in the Vendor Address Table
Vendor Master/EE File Match

Comparison of the Vendor Master and Employee File revealed

- **211**
  - Employee Names established as Vendors in the Vendor Master.

- **418**
  - Addresses in the Employee File that were also in the Vendor Master.

- **79**
  - Social Security Numbers that are also in the Vendor Master.
### Employee Master File

<table>
<thead>
<tr>
<th>EMPNO</th>
<th>LAST</th>
<th>FIRST</th>
<th>ADDRESS</th>
<th>CITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>200170</td>
<td>ESSEX</td>
<td>CATHERINE</td>
<td>1130 WEST RIVER</td>
<td>VANCOUVER</td>
</tr>
<tr>
<td>200200</td>
<td>HARRIS</td>
<td>CHARLES</td>
<td>534 GERY STREET</td>
<td>ANDERSON</td>
</tr>
<tr>
<td>200240</td>
<td>ESTRIN</td>
<td>GERALD</td>
<td>PO BOX 2189</td>
<td>COLUMBUS</td>
</tr>
<tr>
<td>200250</td>
<td>ALLAN</td>
<td>AMY</td>
<td>58 EAST BROAD STREET</td>
<td>MUSKOGON</td>
</tr>
<tr>
<td>200300</td>
<td>ASTOFONI</td>
<td>JULIAN</td>
<td>DR. JORG SMI 197</td>
<td>CAMPANA</td>
</tr>
<tr>
<td>200330</td>
<td>SCARLATTA</td>
<td>VINCE</td>
<td>SUITE 6 - 635 NORTH MICHIGAN AVENUE</td>
<td>BOSTON</td>
</tr>
<tr>
<td>200350</td>
<td>SEIBER</td>
<td>ALEXANDRO</td>
<td>125-145</td>
<td>MUSKOGON</td>
</tr>
</tbody>
</table>

### Vendor Master File

**JOIN on Matching Addresses:**

<table>
<thead>
<tr>
<th>EMPNO</th>
<th>LAST</th>
<th>FIRST</th>
<th>ADDRESS</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>200280</td>
<td>WILLIAM</td>
<td>WILSON</td>
<td>88 EAST BROAD STREET</td>
<td>Gusswerk US, 88 East Broad Street</td>
</tr>
</tbody>
</table>

**Vendor Master (Name 1)**

<table>
<thead>
<tr>
<th>Vendor Master (Name 1)</th>
<th>Vendor Master (General Section)</th>
<th>Vendor Master (General Section)</th>
<th>Vendor Master (General Section)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gusswerk US</td>
<td>88 East Broad Street</td>
<td>Muskegon</td>
<td>MI</td>
</tr>
</tbody>
</table>
Level 3: Managed

Characteristics

- Centralized processing
- Secure environment to store data
- Analytic content shared within the team

Benefits

- Manage all analytic content in one secure place
- Do more with faster analysis
- Risk mitigation / sustainability
Level 3: Managed – Fraud Management Role

1. Champion the move to secure, collaborative, centralized technology
2. Define processes for standardizing, and sharing analytics/data
   – Consider nominating a librarian or ‘knowledge czar’
3. Facilitate partnership with IT to refresh of key data files into the central repository
4. Formalize QA & evaluation into analytic processes
5. Provide strategic program oversight, ensuring that tactical analytics meet bigger-picture fraud detection & prevention objectives
Case Study: Managed

“...helped us overcome our data access challenges with an organized roll-out. Now we know how to prioritize our time and resources.”

Gina Anonuevo
First Hawaiian Bank

Highlights:

• Secure script sharing within a centralized repository
• Permission based access to data and scripts
• Ability to test complete data populations
Level 4: Automated

Characteristics
- Recurring and scheduled analysis
- Fraud detection analytics change from cyclical to continuous
- A broader set of users applying analytics across their fraud detection processes

Benefits
- Achieve timely insight into fraud and control issues
- Improve team efficiency through automation
- Scope increased through multiple business processes
Level 4: Automated– Fraud Management Role

1. Rethink traditional analytic procedures
   - Reporting cycle & frequency
   - Issue follow-up and resolution procedures
2. Communicate goals, objectives, and timelines for processes to be moved to continuous fraud detection
3. Provide sufficient technology training for staff that will be developing fully automated testing procedures
4. Allow time to convert analytics previously designed for ad-hoc mode to automated
   - Focus on testing only new/updated transactions vs. all data
5. Ensure appropriate conversations with IT are occurring around the proper scheduling of analytics
Case Study: Automated

"In these times, it is not good enough to look back over six months at audit time and discover wrongdoing or process inefficiency that may have cost the company money."

Marius Kies,
AngloGold Ashanti

Highlights:

- Identifying deviations, weaknesses and breaches as they happen
- Moving from a reactive model, to an immediate proactive one
- Automated data collection, analysis and reporting
Level 5: Monitoring

Characteristics
- Exceptions routed to business process owners for review and remediation
- Reporting and monitoring of the results and trends identified
- Increased use of analytics across the organization

Benefits
- Develop a partnership with business process owners to provide insight into their operational performance
- Achieve timely resolution of exceptions
- Provide the organization with a clearer picture of risks in business processes
Case Study: Monitoring

...fundamentally transform how we manage spending and evaluate compliance. Reporting that was previously non-existent will now provide senior leadership with a clear picture of our control environment.

Hal Laughlin

Highlights:

- Continuous monitoring with a common set of analytics in all operations
- Consistent reporting and tracking of exceptions across the business
- Improved regulatory compliance with Sarbanes Oxley and FCPA
Questions you should ask about data analysis & fraud

- Where is my highest risk of fraud?
- What indicators – if any – would I expect to see in the data?
- What systems do I need to access to highlight suspected fraud?
- Can I get access to this data?
- What techniques (matching, grouping, filtering) should I apply?
- Can I automate these analytics to drive efficiency and immediacy of results?
Questions?

peter_millar@acl.com | twitter.com/PBMillar