12th IIA Fraud Conference  1st Virtual

• IIA Seattle Chapter
Mr. Vona is the author of three books published by Wiley, Fraud Risk Assessment: Building a Fraud Audit Program and The Fraud Audit: Responding to the Risk of Fraud in Core Business Systems, Fraud Data Analytics Methodology: The Fraud Scenario Approach to Uncovering Fraud.

Wiley has referred to Mr. Vona as a worldwide renowned authority on internal fraud risk.

Mr. Vona’s trial experience is extensive, including appearances in federal and state courts. He is qualified as an expert witness, as a CPA and a CFE, and is cited in West Law for the successful use of circumstantial evidence.
Today’s Agenda

- Use of Fraud Data Analytics to Uncover Fraud Schemes in Core Business Systems
- Today’s Agenda
  - Strategic Plan for Robust Fraud Data Analytics Plan
  - Data Intensive Fraud Approach
  - Ten Steps: Successful Fraud Data Analytics Plan
  - Practical Illustration: Pass Through Shell Company Schemes
Could this Happen to You?

• 63 Million Theft Loss to a False Billing Scheme
• 2.3 Million Theft Spree, Purchase and Resell Scheme
• Conflict of Interest Scheme: Wife of VP of Human Resources
• Whistle Blower of Actual Fraud Committed by Senior Management: Was also accepting Kickbacks from a Vendor

• How would you explain this to your Audit Committee
Breaking The Code of Fraud
Fraud Data Analytics Plan

• The Expectation Questions:

  • Should Internal Audit find Fraud?
  • Can Internal Audit find Fraud?
  • How does Internal Audit find Fraud?
Word of Wisdom

• The world’s best audit program and the world’s best auditor cannot detect fraud unless their sample includes a fraudulent transaction.

• This is why fraud data analytics is so important to our profession.
Your Strategic Plan
Thought Provoking Words

• Five key words I want you think about!
  • Differently
  • Logic
  • Linked
  • Superior knowledge
  • Methodology
Differently

• Fraud risk identification needs to be logic driven versus professional experience driven
• The statement has five elements
  • Person committing the scheme
  • Type of entity: master file data
  • Fraud action statement: transactional data
  • Fraud impact
  • Fraud conversion
Logic

• You can compute the number of fraud risk statements in your scope with mathematical precision

• The goal of logic analysis is two fold:
  • Ensure the completeness of your analysis
  • Create time for data interpretation

• I may not know what the perpetrator is doing, but I know everything the perpetrator can do!
Linked

• The fraud risk statement is what drives the fraud data analytics plan and the fraud audit procedure
  • Person committing
  • Type of entity
  • Fraud action statement
• By understanding what you are looking for, the programming of the routine become obvious
Superior Knowledge

• Understanding data from the:
  • Data perspective
  • Industry perspective
  • Fraud perspective
There is a Methodology

1. Scope of data analysis
2. Integrating the fraud risk assessment
3. Select the data mining strategy
4. Availability, reliability, and usability
5. Understand the data
6. Develop search routines
7. Filtering techniques
8. Sample selection criteria
9. Plan to resolve false positives
10. Link the fraud audit test procedures
What is Fraud Data Analytics

• A methodology of using data mining to analyze data for the red flags that correlates to a specific fraud risk statement.
• It is about identifying transactions that has the highest probability of containing a fraudulent transaction.
Goal of Fraud Data Analytics

FRAUD RISK STATEMENT

Discreet Number

Red Flag
Red Flag
Red Flag
Circular View Of Fraud Data Analytics Plan
Three Key Considerations of the Fraud Data Analytics Plan

- **What are you looking for:**
  - What is your scope?
  - Which fraud risk statements?
- **What is your fraud data analytics strategy?**
- **How to calibrate data interrogation for the sophistication of concealment?**
The Fraud Risk Universe Establishes Scope of Project

What Fraud Risk Statements are in your audit?
The Fraud Risk Statement: The Linkage Factor

Does each fraud risk statement link to your audit plan?
Illustration of a Fraud Risk Statement

• Budget owner acting alone / cause a shell company to be set up on the vendor master file / process a contract and approves a fake invoice for goods or services not received / causing the diversion of company funds

• The statement has five elements:
  • Person committing the scheme: Budget owner
  • Type of entity: master file data: Shell company
  • Fraud action statement: transactional data: False invoice
  • Fraud Impact: Loss of company assets
  • Fraud Conversion: Budget owner receives the payment
Three Key Considerations
Fraud Data Analytics Plan

• What are you looking for?
• **Fraud data analytics strategy**
  • Data patterns?
  • Data frequencies?
• How to calibrate data interrogation for the sophistication of concealment?
Primary Data Mining Strategies

• Specific identification
• Internal control avoidance
• Data interpretation
• Number anomaly
  • Patterns and frequency for:
    • Entity analysis
    • Transaction analysis
What is the Purpose
Primary Data Mining Strategies

• Methodology is designed to identify a data pattern
• Each strategy has rules
• Each strategy correlates to sophistication of concealment
• Each strategy correlates to sample selection
Creating the Score Card

- Seldom does one data element cause a sample selection
- The goal is to identify an entity or transaction that meets the criteria of the search routine: The data profile!
- Illustration of missing data test for created entity
  - How many missing data elements should cause a selection of the entity?
  - Does all missing data have the same weight?
Fraud Data Analytics Plan Usually Has Multiple Layers

Specific Identification

Internal Control Avoidance

Sample
Sample Selection is Both

Strategy

Concealment
Specific Identification
Direct Evidence Strategy

• Key words associated with strategy
  • Missing
  • Change
  • Duplicate, same file
  • Match, different file
  • Identify a specific criteria

• Sample size is based on the specific identification
Internal Control Avoidance: Inference Strategy

• Internal control avoidance
  • Circumvention of an internal control
  • Structuring multi transactions
    • One entity
    • Multiple entities
  • Transactions below a control threshold
Internal Control Avoidance
Inference Strategy

• Use of aged documents providing open authority
• Speed of transaction or illogical order of a transaction
• Off period analysis
  • Create, change, delete or void
  • Person performing task
  • Date and time
• Override transaction, by code or person
• Manual transaction
• Sample size is based on the internal control avoidance
Data Interpretation
Professional Experience Strategy

- Relies on professional experience
- Relies on filtering techniques
- Relies on visual selection
- Sample size is based on the judgment of the auditor
Number Anomaly Strategy

• Benford's Law
• Even number or a recurring number
• Contra entry

• Requires to go from anomaly to Fraud Risk Statement!
Patterns Analysis

• What is a pattern analysis in fraud data analytics?
  • Process of searching for pattern recognition
  • Pattern correlates to the fraud risk statement
    • I.e. Two vendors with different names have a duplicate address
    • I.e. Sequential pattern of vendor invoice numbers
Pattern Questions that Need to be Understood

• What are the implications of:
  • Exact match
  • Close match
  • Related match

• What are the implications of:
  • Data input error
  • Intentionally disguised
Do You Understand False Positives and False Negatives

• **False positive:** Transaction tested meets the criteria of the test, but the transaction is not fraudulent

• Improperly designed test

• Lacks a sufficient number of testing criteria

• Occurs due to data integrity

• Data field contains multiple sub set

  • i.e. Address field

  • i.e. Physical address or bank lock box
Do You Understand False Positives and False Negatives

• **False negative:** Transaction tested does not fit the criteria of the test, but the transaction is fraudulent

• Improperly designed test

• Fraud risk statement does not lend itself to data mining

• Concealment is higher than the test
Frequency Considerations

• Perpetrators seldom steal once, of course, unless the amount is huge!
• Perpetrators all have a personal risk tolerance “will I get caught?”
• Perpetrators find a comfort in a particular technique
• Perpetrators become more erratic for frauds over a long duration
The Key Considerations of the Fraud Data Analytics Plan

• What are you looking for:
• What is your fraud data analytics strategy
• How to calibrate data interrogation for the sophistication of concealment
  • How it impacts master file data
  • How it impacts transactional data
  • How it impacts your search routines
  • How it impacts your fraud testing procedure
Fraud Detection Dilemma

Fraud Risk Statement

See Something

Fraud Detection Bar

Understand Fraud Risk

Audit Program
## Fraud Concealment Linkage Master File Data

<table>
<thead>
<tr>
<th>Concealment Level</th>
<th>Details</th>
</tr>
</thead>
</table>
| High              | • Must use transaction file  
                  | • No Match |
| Medium            | • Limited Linkage between vendor and perpetrator  
                  | • Close Match |
| Low               | • Linkage between vendor and perpetrator  
                  | • Exact Match |
Illustration Using a Bank Account Number

<table>
<thead>
<tr>
<th>Concealment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>• Different bank</td>
</tr>
<tr>
<td></td>
<td>• No Match</td>
</tr>
<tr>
<td>Medium</td>
<td>• Same bank different account number</td>
</tr>
<tr>
<td></td>
<td>• Close Match</td>
</tr>
<tr>
<td>Low</td>
<td>• Same bank and same account number</td>
</tr>
<tr>
<td></td>
<td>• Exact Match</td>
</tr>
</tbody>
</table>
Fraud Concealment Linkage
Transaction Data

- **High Concealment**
  - Judgment based on visual examination
  - Related Match

- **Medium Concealment**
  - Reduces population
  - Close Match

- **Low Concealment**
  - Data pattern visible to naked eye
  - Exact Match
Low Sophistication Guidelines

• Specific identification strategies are used for both entity and transactional data.
• Entity identifying information links to the perpetrators known identifying information, for example, a specific street address.
• The false entity structure will match to another entity either in the same database or a different database.
• False entity will also reveal missing identifying information in order to reduce someone else’s ability to contact the false entity.
• The patterns associated with the transaction data will typically be overtly obvious to the naked eye.
• The pattern recognition for the transaction data allows for specific identification.
• Sample size is determined by the number of transactions that match the data profile. The sample size can be either zero because no transactions link to the data profile or a very large sample because the match criteria are not sufficiently defined.
Medium Sophistication Guidelines

• Internal control avoidance strategies tend to be more effective.
• Specific identification routines are less effective because there is no direct match to the entities data.
• Specific identification will allow for a match on some aspect of the entity information.
• Specific identification is more effective when there is an allegation that focuses on a person or department.
• Entity identifying information relates to some aspect of the perpetrators known identifying information, for example, a postal code versus a physical street address.
Medium Sophistication Guidelines

• Internal control avoidance strategies should be used for transactional data.
• Outlier patterns tend to be effective for transactional history analysis.
• Creating smaller homogeneous data groups, referred to as cluster patterns, will facilitate the auditors ability to spot an anomaly.
• Filtering techniques based on dollar magnitude are effective in reducing the number of transactions fitting the data profile.
• Sample selection is based on the entities or transactions that avoid the internal control after all relevant filtering.
High Sophistication Guidelines

• Data analytics at this level is like code breaking. There is no finite criterion that serves as identification criteria. The process tends to be judgmental selection versus a criteria selection. The key is to understand how the fraud scenario occurs in your business systems.

• The specific concealment strategies used by the perpetrator tend to be more deliberate and planned.

• Direct matches seldom occur.

• Entity identifying information has no relationship with the perpetrators known identifying information.

• Transactional data are more effective at identifying fraud scenarios versus entity data.
High Sophistication Guidelines

• The process of creating smaller homogenous data files based on geographic, transaction types, transaction codes, cost centers facilitates the data interpretation.

• Filtering techniques like drill-down analysis are effective in reducing the number of transactions fitting the data profile, thus, allowing data interpretation to be more effective.

• Sample selection relies on data interpretation skills.

• Selection process is based on understanding how the scenario operate, money trail, fraud theory, concealment theory, and professional experience of the auditor.

• **Sample size tends to be judgmentally determined based on the data interpretation.**
Sample Selection Thought Process

STRATEGY + CONCEALMENT → SAMPLE
Do you: Understand the Data?

• Data is data. Right?
• The real question is at what level do you understand the data?
• How does the data link to your fraud data analytics plan?
• Get over it. Data is not perfect.
Illustration of Fraud Risk Statement

• A fraud risk statement for a pass-through scheme (of which there are at least 15 permutations) looks as follows:

• *Budget owner, acting alone or in collusion with a direct report, causes a shell company to be set up on the master file and places orders for goods or services through the shell company. The shell company then places an order with a real supplier, and the real supplier ships directly to the budget owner’s company. The real supplier invoices the shell company, and the shell company invoices the budget owner company at an inflated price, diverting company funds.*
Pass Thru Entity Fraud Risk Statement
Version 1 -15

Your Company
- Employee is perpetrator
- Place order with shell company
- Receives goods from real company

Shell Company
- Employee owns shell company
- Place order with real company
- Invoices your company at a mark up price

Real Supplier
- Sales person is not complicit
- Ships to your company
- Invoices shell company at regular pricing
Pass Thru Entity Fraud Risk Statement
Version 1-15

Real company
- Invoice shell company for $1,000
- Not complicit in scheme

Shell Company
- Invoice real company for $1,200
- Complicit in scheme

Your Company
- Employee causes your company to be over billed by $200
- Employee perpetrator of scheme
How the Fraud Risk Statement Drives The Thought Process

• Cause a **shell company** to be set up on the vendor master file
• Cause an **invoice** to be processed

• Two aspects
  • Sample Selection
  • Audit Test
There Are Three Categories of Vendors

- Conflict of Interest
- False or Shell
- Real
Shell Company
Understanding the Data

• What is the primary type of shell company?
  • Created shell company
  • Assumed identity shell company
  • Hidden shell company
  • Limited use shell company
• Each primary category has several secondary categories
Searching for the
Created Shell Company

• The created shell company:
• Starts with the **specific identification strategy**
  • Missing information
  • Anomaly in existing information
  • Matching to other data bases
• Internal control avoidance strategy
  • Off period creation
Superior Knowledge
Categories Address Field

• More advanced analysis
• Street address
• US Mail P.O. Box
• Mail box service address
  • Public company
  • Private company
• Virtual offices
• Mail box forwarding address
• Redirect address
• Professional service provider address
Created Shell Company

• Anomaly in data elements:
  • Missing address and bank account number
  • Mail box service company in address fields
  • Mail box forwarding company in address fields
  • Email address: non business email
Correlate to Transactional Data

- Correlate master file search to dollar activity
- Dollar activity as to
  - Dollar value
  - Number of transactions
  - Average value of transaction
  - Maximum or minimum
Searching for Patterns in Transactional Data

• Identify the key tables: Purchase order, invoice and payment
• The five key fields by transaction
  • Control number
  • Control date
  • Amount
  • Description
  • General ledger account
Data Analysis
Vendor Invoice Number Test

• Pattern and frequency of invoice numbers and amounts
  • Compare beginning invoice number to ending invoice number
  • Compute invoice number range
  • Compute date range
  • Correlate to creation date
  • Search for sequential pattern or limited range pattern
  • Search for low invoice numbers
How to use the Invoice Number Report

• Pattern analysis
  • Sequential
  • Low sequential number
  • Limited range

• Frequency analysis
  • Count on number of invoices compared to the invoice number range
Date Analysis
Pattern of Date

• The primary tests for dates are:
  • Illogical order
  • Circumvent order
  • Speed of processing the transaction

• Comparison of dates between two transactions
  • Purchase order date to invoice date
  • Invoice date to payment date
Date Analysis
Pattern of Date

• Comparison of dates between two transactions to test for the **logical order**
  • Purchase order date to invoice date
  • Invoice date to payment date
• Comparison of dates to determine speed of transaction
• Invoice date to payment date
  • Payment clearing date to city / state / postal code
  • Payment clearing date to theft of stale checks
Pattern of Transaction Amount

• Below or above control threshold
• Duplicate amount, linking to control number, date or description
• Two or more in aggregate above control threshold correlate to date or control number
• Even amount
• Recurring amount
• Contra amount
Amount Analysis
One Easy Test for Amount Field

• Always, the first report I create:
• The maximum, minimum and average report
  • Vendor number and name
  • Creation date
  • Aggregate dollars
  • Frequency of records
  • Maximum, minimum and average
Pattern
Line Item Description

• Descriptions: alpha or numeric or both
  • Matching line item descriptions
  • Missing alpha
  • Missing numeric
  • String length analysis
General Ledger Account

- Links to person committing fraud risk statement:
  - Pattern by direct access
  - Pattern by budget owner
  - Pattern by senior manager
- Links to the description of the expenditure
Frequency Considerations

• Correlates to transactional analysis
• Number of transactions consistent with search routine
• Frequency of event or transaction
  • Number of instances
  • Below control range
  • Same amount
  • Same date
  • Number known to individual
Frequency: of the Event

• Correlates to person committing or entity
• Focus is on entities that have some sort of linkage
  • Name
  • Address
  • Bank account
  • Authority level
• Creator of the transaction
Remember: Circular View Of Data Profile
Remember: The Score Card

- Seldom does one data element cause a sample selection
- The goal is to identify an entity or transaction that meets the criteria of the search routine: The data profile
- Illustration of missing data test for created entity:
  - How many missing data elements should cause a selection of the entity?
  - Does all missing data have the same weight?
Remember My: Words of Wisdom

• The world’s best audit program and the world’s best auditor cannot detect fraud unless their sample includes a fraudulent transaction.

• This is why fraud data analytics is so important to our profession.
Axioms of Fraud Data Analytics

• The world’s best audit program and the world’s best auditor cannot detect fraud unless their sample includes a fraudulent transaction.

• I do not know what a perpetrator will do, but I do know everything the perpetrator can do.

• While we do not know how a perpetrator will commit a fraud or how he will conceal the fraud, we can determine the logical permutations.

• The better you can describe the fraud scenario, the more likely you will be able to find it.

• False positives will occur. You either try to resolve false positives through your fraud data analytics or through an auditor performing audit procedures.

• In fraud data analytics, fraud likelihood is based on data versus the effectiveness of internal controls.
Axioms of Fraud Data Analytics

• We search for transactions that mirror the red flag theory of the fraud scenario.
• The better we understand data, the better we can use data to search for a fraudulent transaction.
• Errors and fraud have a lot in common.
• Red flags correlate to both errors and fraud.
• Data is not perfect.
• Databases contain data errors, either caused by mistake or with intent.
• We can only search data when the data resides in our database.
• **Fraud data analytics is both a science and an art.**
Thank You

• To quote a famous person
• That’s all Folks
• Source: Porky Pig via Warner Brothers
• I am here if you have questions!
• lvona@leonardvona.com sign up for our monthly blog
• No marketing just information
• Leonardvona.com