Forensic Data Analytics
Key to unlocking invisible information using forensic “lookback”

Forensic Data Analytics as a topic and its adoption within the industry had long been overdue. With the advent of technology and increasing incidents of fraud, there has been a significant rise in adoption of Forensic Data Analytics. Due to this, company appointed auditors and independent directors are now seeking to implement proactive fraud-prevention solutions and are avoiding post-incident remediation processes.

Forensic Data Analytics is a science used to proactively seek opportunities to prevent and detect fraud, waste and abuse by leveraging information in corporate data assets. It enables identification of meaningful patterns and correlations in existing historic data to predict future events and assess the reasons for various fraudulent activities. Such insightful predictive information is generally “invisible,” but provides a platform on which organizations can take business decisions related to fraud, disputes and misconduct.

“The greatest value of forensic analytics is when it forces us to notice what we did not expect to see.”
Forensic Data Analytics

Big data is a reality:
The volume, variety and velocity of data coming into the organization have reached unprecedented levels. About 2.5 exabytes of data are created each day, and that number is doubling every 40 months.

Issues in managing big data:
Big data requires high performance analytics to process billions of rows of data with hundreds of millions of data combinations. The traditional data warehousing techniques may not be able to identify anomalies in the existing data set thus preventing proactive fraud management.

Recent scams in the limelight:
In the recent times, India has been hit with multi billion value scams associated with the following:
- Anti Money Laundering
- Bribery and Corruption
- Procurement fraud and collusion in bidding process
- Accounting misstatement

Adoption of forensic data analytics:
The associated risks could have been mitigated if key stakeholders would have paid attention to anomalies at an earlier stage. This could have been possible if existing data assets were analyzed from forensic perspective to avoid wrongful or criminal deception intended to result in financial or personal gain.

How does forensic data analytics help organizations?
- Proactive fraud prevention management
- Effective and focused internal controls
- Controlling the magnitude of fraud in a reactive set up
- Improving regulatory and compliance environment
Forensic Data Analytics can be used as a standalone service or in conjunction with existing practices such as investigations, audits, process review and due diligence. In the current context, data exists in structured (multiple form of databases) and unstructured forms (emails, office documents, presentations, Excel sheets, PDF files, archive files, text and image files) in organizations. Using EY’s proprietary tools, raw data can be transformed into formats that can be analyzed, and with the help of advanced analytical capabilities, anomalies can be identified that may indicate potential fraud. Some of our key offerings include, but not limited to, identification of fraud in vendor, customer and employee registration, procurement to pay, order to cash, sales and distribution, travel and entertainment, payroll disbursement.

Our forensic analytic models are developed to identify variances in data sets, which may impact an organization’s profit and loss statement. This model also touches on various aspects, from simple narration captured in a transaction to complicated sentiments and tone analysis. It also includes data within applications and data recorded on social and professional networks for further analysis. This analysis helps a company to move beyond identification of low value pilferage to implementing controls on existing and potential weak areas. Any dataset in historic, near real time and real time form can be assimilated through big data solutions to help a company improve its bottom line by checking fraudulent activities.

Figure 1: Structured output from unstructured data
Our key differentiator in forensic data analytics

At EY, data analytic techniques applied to internal or external fraud follows a four pillar approach — WHO-WHAT-WHEN-WHY. This approach looks at any situation from all possible angles and highlights key issues. This does not only help in managing risks, but also in identification of potential growth areas.

Increasing concerns about fraud and vulnerability can be alleviated by a range of forensic techniques, some of which are presented below.

Link Analysis

Link Analysis is a data-analysis technique used to evaluate relationships (connections) between nodes, including organizations, people and transactions. Key applications of this technique include analysis of EPBX data, mobile bills and user logical access records that help a company map its user footprint.

In a recent incident in a manufacturing company, its phone records were analyzed across different zones to determine the nexus between its employees and selected vendors on procurement and disposal of scrap. Using Link Analysis, we were able to establish “hidden” relationships and information leakage from suspected employees to identified vendors for possible “kickbacks.”

The key to identify fraud lies in the ability to comprehend what lies beneath.
Social Network Analysis

Social Network Analysis views relationships in terms of network theory, which consists of nodes and ties. Nodes represent individual “actors” within the network and ties represent relationships between individuals, e.g., friendships, kinship, organizational position, etc. Social Network Analysis, along with Link Analysis, helps to identify related parties, conflict of interest, bid rigging, among other fraud.

In a large consumer products company, the India lead had appointed his relatives as distributors, and through known vendors, managed distribution of products in key states. Social Network Analysis, followed with a background check, helped to reveal the nexus. This led to a full-blown investigation and the company now undergoes vendor due diligence before it carries out any business.

Concept Clustering

Concept Clustering involves grouping similar entities or behavior into tight semantic clusters for the purpose of identifying anomalies or red flag. It is used actively, along with an electronic data review. In this example, Concept Clustering was executed on more than a million documents to identify all the information with terms such as “gifts,” “incentive” and “facilitation.” We were able to bring these down to a sizable volume with the required criteria that was analyzed in a time-bound manner. Concept Clustering can be effectively used on structured and unstructured data.

Sentiment Analysis

Known as behavioral analysis, this refers to the application of text analytics to identify and extract subjective information including the attitudes of writers, their affective state and the intended emotional quotient. It determines whether expressed opinions in a document are positive, negative or neutral. The “fraud triangle” can be applied to categorize events into rationalization, opportunity and pressure to identify sentiments. Organizations use this data to conduct behavioral training, stem attrition, and identify disgruntled employees and potential fraud conversation.
Data Visualization — identifying the “hidden” from “not so apparent”

Data Visualization techniques have proved to be effective, since humans can better absorb large pieces of information in a visual format than that displayed in numbers or text. When the result of a fraud identification query is combined with Data Visualization, e.g., an account payable or journal entry data, a significant amount of useful and previously invisible information can be reviewed at one go.

Tag Cloud

One of the most widely used visual techniques is a Tag Cloud. This is a good example of expressing complex data that can be understood intuitively. A Tag Cloud is the visual representation of communication relating to transactional data entries. It is represented by a combination of words in varied fonts, sizes or colors. This format is useful for quickly determining the important terms to identify key fraud issues.

Interactive CXO dashboards

A useful feature of analytics is that an entire data set can be converted to a meaningful dashboard for a CXO analysis.

Such dashboards help in understanding databases and spreadsheets of any size with their easy drag and drop interface. They not only display information visually in seconds, but also create interactive maps with the click of a mouse. They can effectively analyze time series from years to months to the actual time in a day. Their most helpful feature is their capability to combine different databases to a single view and publish interactive dashboards on the Web.

Here, we have sliced the entire expense dump of an organization from four key lenses including WHERE (geography), WHAT (type of expense), HOW (expense description) and WHO (the employee who incurred the expense). Having multi-dimensional data on a common platform helps a company perform an insightful analysis to determine the tests that need to be performed on expense data.
About EY
EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization and may refer to one or more of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit ey.com.

About EY’s Fraud Investigation & Dispute Services
Dealing with complex issues of fraud, regulatory compliance and business disputes can detract from efforts to achieve your company's potential. Better management of fraud risk and compliance exposure is a critical business priority – no matter the industry sector. With our more than 2,000 fraud investigation and dispute professionals around the world, we assemble the right multidisciplinary and culturally aligned team to work with you and your legal advisors. And we work to give you the benefit of our broad sector experience, our deep subject matter knowledge and the latest insights from our work worldwide. It's how EY makes a difference.

Our offices

<table>
<thead>
<tr>
<th>City</th>
<th>Address</th>
<th>Phone Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmedabad</td>
<td>2nd floor, Shivalik Ishaan Near, C.N Vidhyalaya Ambawadi</td>
<td>+91 80 4027 5000, +91 80 6727 5000</td>
</tr>
<tr>
<td></td>
<td>Tel: +91 80 4027 5000, +91 80 6727 5000</td>
<td>Fax: +91 80 2210 6000 (12th floor)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fax: +91 80 2224 0695 (13th floor)</td>
</tr>
<tr>
<td>Bengaluru</td>
<td>12th &amp; 13th floor “U B City” Canberra Block No.24, Vittal Mallya Road</td>
<td>+91 80 6727 5000, +91 80 2222 4112</td>
</tr>
<tr>
<td></td>
<td>Tel: +91 80 6727 5000, +91 80 2222 4112</td>
<td>Fax: +91 80 6727 7880, +91 80 172 671 7800</td>
</tr>
<tr>
<td>Chandigarh</td>
<td>1st Floor, Prestige Emerald No.4, Madras Bank Road Lavelle Road Junction</td>
<td>+91 172 671 7800, +91 172 671 7888</td>
</tr>
<tr>
<td></td>
<td>Tel: +91 172 671 7800, +91 172 671 7888</td>
<td>Fax: +91 144 6654 8100, +91 144 2254 0120</td>
</tr>
<tr>
<td>Chennai</td>
<td>Tidel Park 6th &amp; 7th Floor A Block (Module 601,701-702) No.4, Rajiv Gandhi Salai Taramani Chennai-600 113</td>
<td>+91 44 6654 8100, +91 44 2254 0120</td>
</tr>
<tr>
<td></td>
<td>Tel: +91 44 6654 8100, +91 44 2254 0120</td>
<td>Fax: +91 40 6736 2000, +91 40 6736 2200</td>
</tr>
<tr>
<td>Hyderabad</td>
<td>Oval Office 18, iLabs Centre Hitech City, Madhapur Hyderabad - 500 081</td>
<td>+91 40 6736 2000, +91 40 6736 2200</td>
</tr>
<tr>
<td>Kochi</td>
<td>9th Floor “ABAD Nucleus” Nr+49, Maradu PO Kochi - 682 304</td>
<td>Tel: +91 484 304 4000, +91 484 270 5393</td>
</tr>
</tbody>
</table>

EY contacts

<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arpinder Singh</td>
<td>Partner and National Leader</td>
<td>Direct: +91 22 6192 0160 Email: <a href="mailto:arpinder.singh@in.ey.com">arpinder.singh@in.ey.com</a></td>
</tr>
<tr>
<td>Mukul Shrivastava</td>
<td>Partner</td>
<td>Direct: +91 22 61922777 Email: <a href="mailto:mukul.shrivastava@in.ey.com">mukul.shrivastava@in.ey.com</a></td>
</tr>
<tr>
<td>Sudesh Shetty</td>
<td>Director</td>
<td>Direct: +91 22 61921957 Email: <a href="mailto:sudesh.shetty@in.ey.com">sudesh.shetty@in.ey.com</a></td>
</tr>
</tbody>
</table>

Ernst & Young LLP is one of the Indian client serving member firms of EYGM Limited. For more information about our organization, please visit www.ey.com/in.

Ernst & Young LLP is a Limited Liability Partnership, registered under the Limited Liability Partnership Act, 2008 in India, having its registered office at 22 Camac Street, 3rd Floor, Block C, Kolkata - 700016

© 2013 Ernst & Young LLP. Published in India. All Rights Reserved.

EYIN1212-108 ED None

This publication contains information in summary form and is therefore intended for general guidance only. It is not intended to be a substitute for detailed research or the exercise of professional judgment. Neither Ernst & Young LLP nor any other member of the global Ernst & Young organization can accept any responsibility for loss occasioned to any person acting or refraining from action as a result of any material in this publication. On any specific matter, reference should be made to the appropriate advisor.