New England Intergovernmental Audit Forum

Trends in data analytics – fraud detection

November 3, 2016

Building a better working world

Discussion topics

- Introduction 5 min
- EY Global Forensic Data Analytics Survey 2016 10 min
- Data analytics methodology 30 min
  - Analytics design
  - Data collection
  - Data analytics
  - Explore and risk rank
  - Findings and observations
- Insider threat analytics 20 min
- Questions and answers 10 min

Data and analytics core businesses will be disrupted!

Information is growing exponentially	
1.7 MB	New info. created every minute for every human being
80%	Data is unstructured; invisible to traditional computers
50 billion	"Things" will be connected to the Internet by 2020

Computing power is increasing, while cost is decreasing	
18 months	For computing power (analytics) to double
14 months	For cost per gigabyte of storage to go down by half

The nature of work is changing…	
30%	Of consulting will be replaced by cognitive/AI technology
47%	Of today's jobs could be automated in the next two decades, 94% probability for accounting and Tax
50%	Of professional services will be procured as "managed services"

As a result of the Open Data Initiative, government agency data is becoming more accessible.

Human beings tend to overestimate in the near term, but underestimate in the longer term.

1943	"I think there is a world market for maybe five computers."	**Thomas Watson** *Chairman of IBM*
1977	"There is no reason for any individual to have a computer in their home."	**Ken Olsen** *Co-founder and former President of Digital Equipment*
1998	"Touchscreen e-readers will never catch on."	**Bill Gates** *Co-founder of Microsoft*
1998	"The growth of the Internet will slow drastically… most people have nothing to say to each other! By 2005 or so, it will become clear that the Internet's impact on the economy has been no greater than the fax machine's".	**Paul Krugman** *Economist*

EY Global Forensic Data Analytics Survey 2016

Current and emerging risks driving demand

- The fastest-growing threat in the fraud and investigative risk universe is from cyber breaches and insider threats.
- FDA demand is also being driven by increasing government and public scrutiny of fraud risk.

"At the SEC, we have made great strides in leveraging data and technology to detect and pursue misconduct. In the enforcement arena, the Commission is using data analytics to help identify wrongdoers and conduct streamlined investigations to optimize our resources."

– Chair Mary Jo White, opening remarks at the twenty-first Annual International Institute for Securities Enforcement and Market Oversight, November 2, 2015

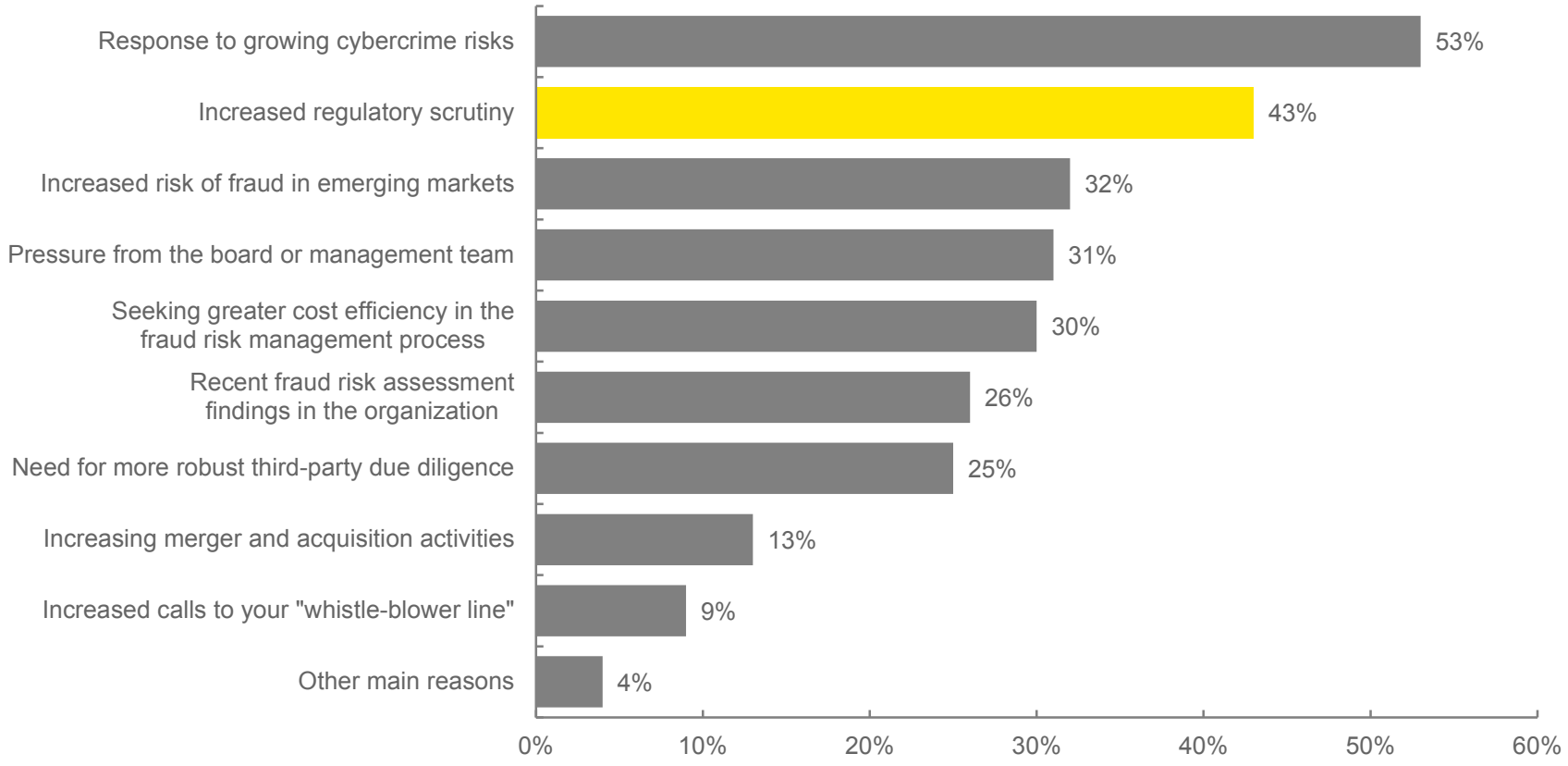
- C-suite respondents have shown a stronger tendency toward FDA adoption than other executives.
  - While 69% overall agree that, "we need to do more to improve our current anti-fraud procedures, including the use of FDA tools," this number jumps to 74% for the C-suite cohort.

74%

of C-suite respondents

"We need to do more to improve our current anti-fraud procedures, including the use of FDA tools."

FDA demand is also driven by increasing government and public scrutiny of fraud risk



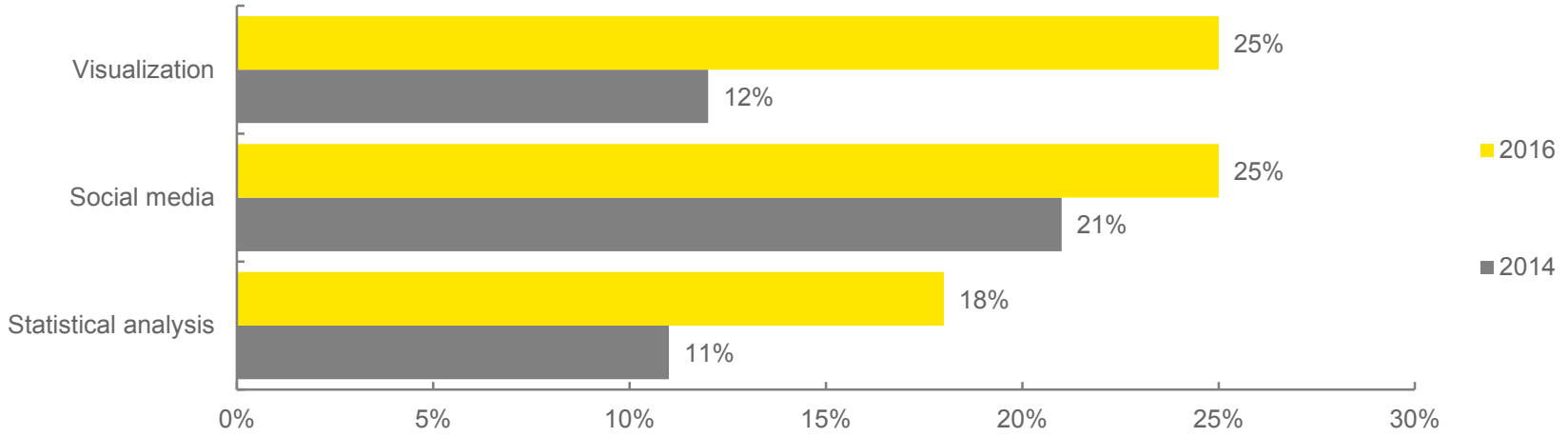
Q. What are the main reasons that you are planning to increase your investment in FDA capabilities?

Base: Respondents who plan to increase investment in FDA (405)

Multiple answers allowed, may exceed 100%

Growing sophistication in technology and the use of data

- The use of visualization tools has doubled since our 2014 survey. There has also been increasing use of social and web monitoring tools and statistical analysis and data mining packages.
- 75% of respondents routinely analyze a wide range of structured and unstructured data.



Q. Which FDA tools do you utilize in managing fraud risk?

Base: 2016 all respondents (665); 2014 all respondents (466)

Multiple answers allowed, may exceed 100%.

Key challenges
Building teams with the right skills

Successful deployment requires three distinct skillsets:

- **Technical skills** – to understand the organization's systems and advise on acquiring additional technology
- **Domain knowledge** – familiarity with the relevant risk areas in the business and the ability to interpret analytics results in the context of the organization

Don't forget the human element of your FDA program

"In 39% of leading analytics organizations – versus 12% of the rest – analytics skills are recognized, effective, efficient, monitored and clearly used to support decisions. More than one-third of the top 10% also have well-defined competencies for each role and level, along with robust training programs that address potential skills shortages."

EY and Forbes Insights, Data and Analytics Impact Index: don't forget the human element of analytics, 2015.

- **Data analytics (e.g., data science) expertise** – mathematical, computer science and business intelligence techniques, such as pattern recognition, statistical analysis, query design and data visualization

According to Gartner, the need for data scientists is growing at about three times that for statisticians and business intelligence analysts, and there is an anticipated talent shortage of 100,000 or more analytics personnel through 2020.

Many have reported positive results or recoveries from the FDA tools

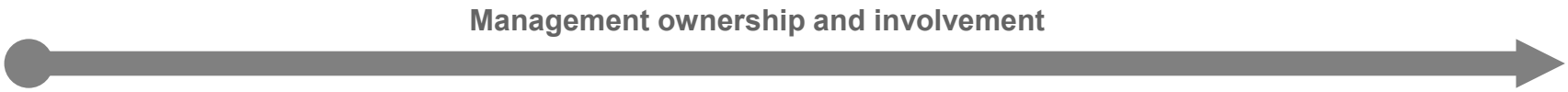
56%

agree

“We currently get positive results or recoveries from the FDA tools that we use.”

Data analytics – methodology

Analytics implementation methodology

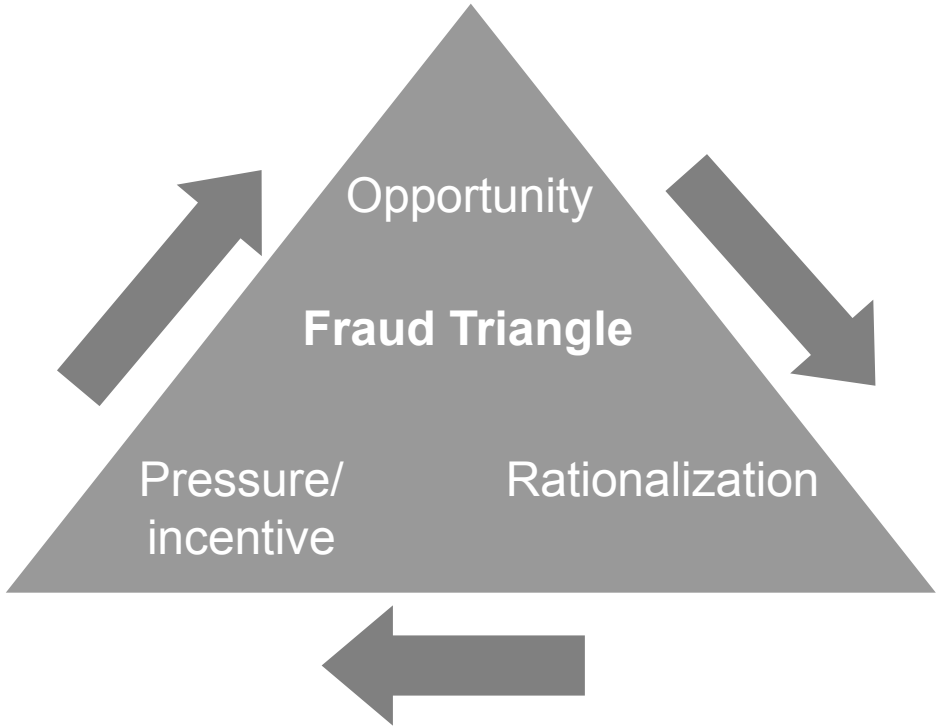


	1 Analytics design	2 Data collection	3 Data analytics	4 Explore and risk rank	5 Findings and observations
Activities	► Identify risks based on industry and client knowledge ► Understand data systems and business processes ► Develop workplan to define tests ► Refine engagement timeline	► Work with client and IT to identify, export and validate data ► Assess data integrity and completeness ► Perform data reasonableness checks ► Load, normalize, cleanse and transform data, as needed	► Execute previously identified analytics plan ► Modify analytics as appropriate based on data received, client feedback and preliminary results	► Visualize analytical results and develop dashboards ► Develop scoring model to risk rank transactions	► Identify trends and risk areas ► Quantify potential risks ► Conduct a briefing with key stakeholders ► Determine remediation actions and track dispositions
Output	► Analytics work plan	► Data requests ► Control totals and inventory	► Preliminary results for QC	► Visualized dashboards ► Risk model	► Findings and observations ► Action plan

(1) Analytics design

Why do people commit fraud?

- Many studies suggest that employees who commit fraud do so because there is opportunity, pressure (or incentive) and rationalization – The “Fraud Triangle”.



The triangle is based on Donald Cressey's hypothesis about behaviors that lead to fraud. See http://www.acfe.com/fraud-triangle.aspx.

Fraud risks examples
Digital, social, news and media

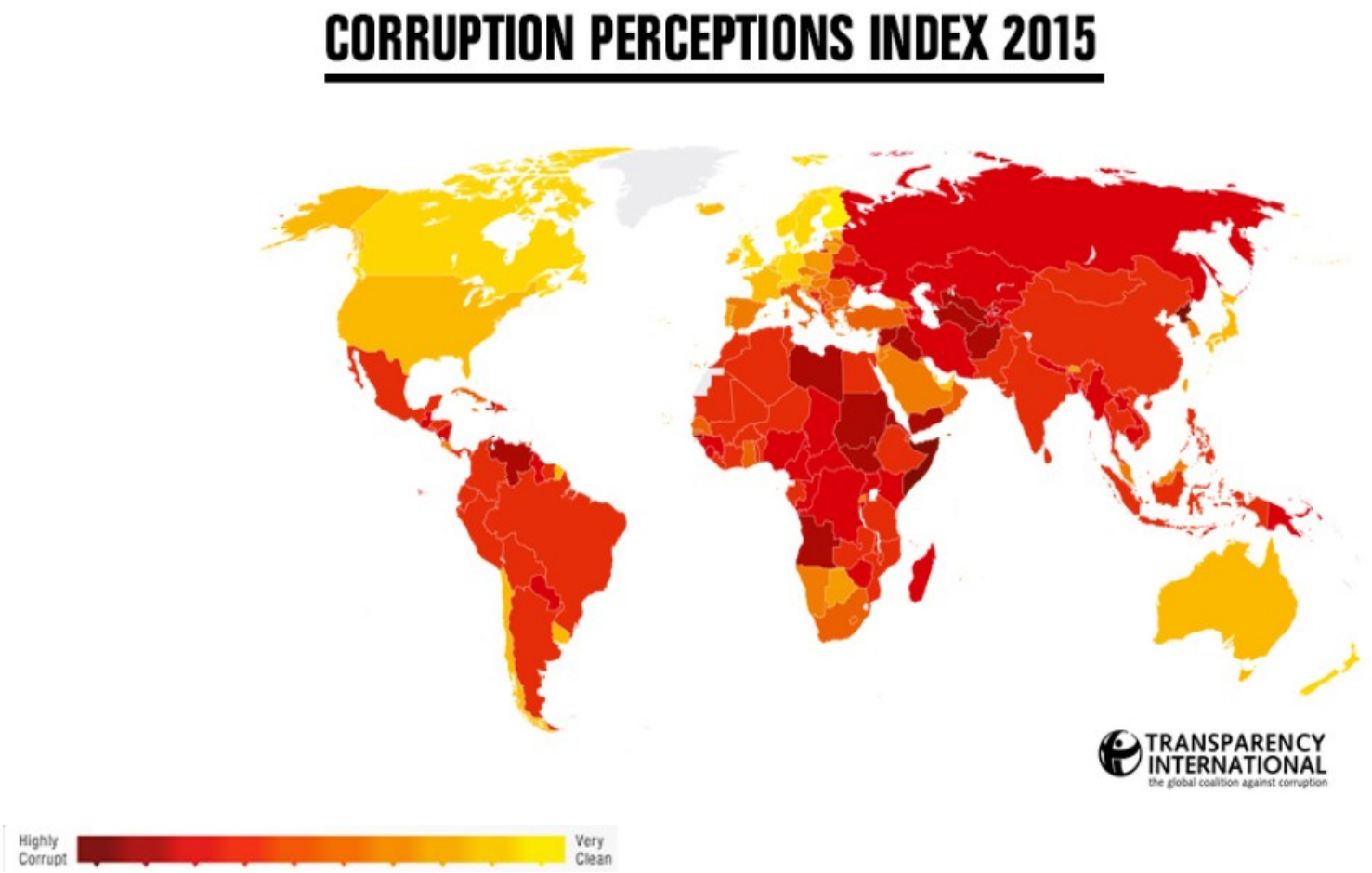
1 Analytics design

Risk factors:
- Growth focused on emerging markets
- Increasing complexity in the route to market
- Increased entrepreneurship or decentralization

Fraud risk examples		
Procurement related frauds	Bribery, inappropriate conduct with government officials and other parties	Potential corruption using third-party service providers
Overstatement of expenses/ duplicate claims	Unaccounted sales returns	Channel stuffing
Embezzlement of funds by employees	Advertising and promotion expenses	Pilferage/theft of material diversion of products
Collusion with vendors or customers	Inflated sales revenue fictitious customer/ sales orders	Conflicts of interest

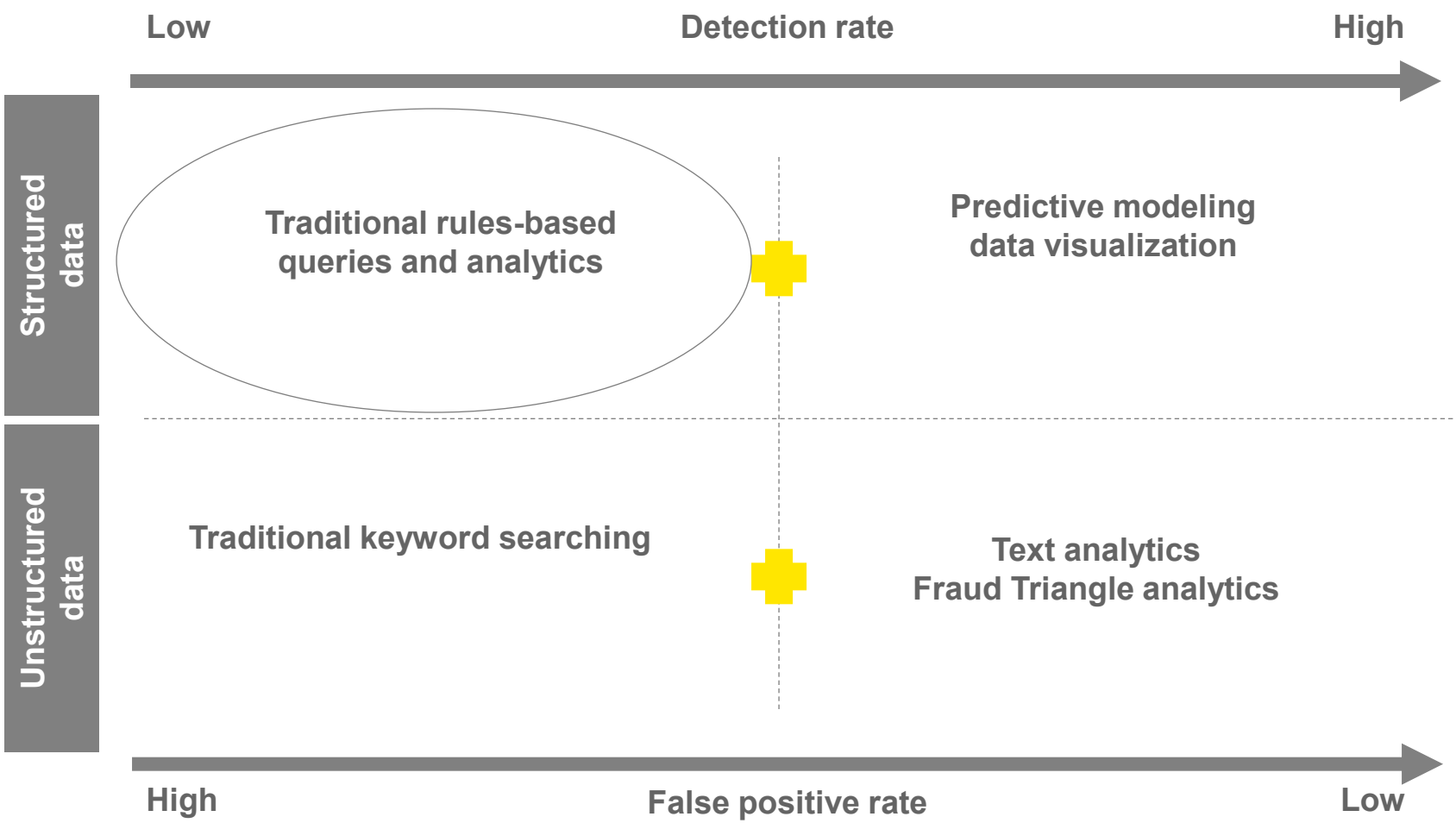
Transparency International 2015
Corruption perceptions index

1 Analytics design



Fraud detection analytics
Maturity model

1 Analytics design



EY/ACFE library of "keywords"

(More than 3,000 terms in a dozen languages so far…)

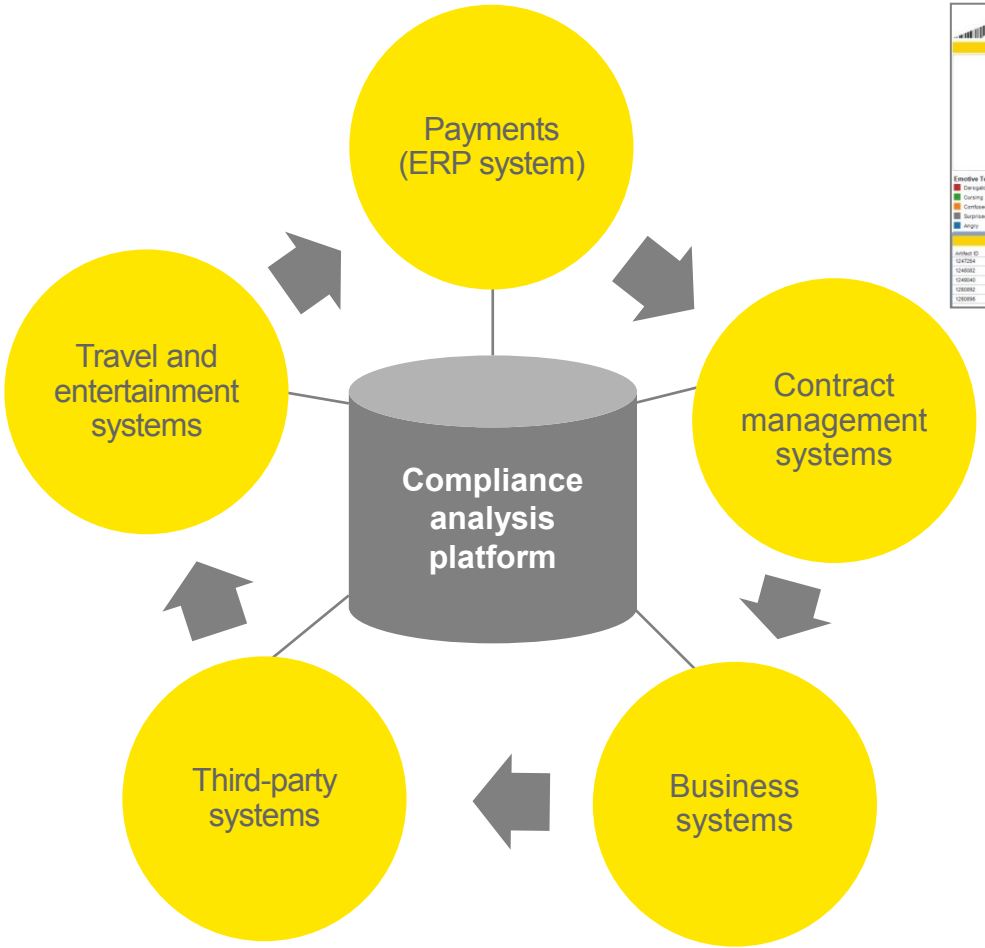
1 Analytics design

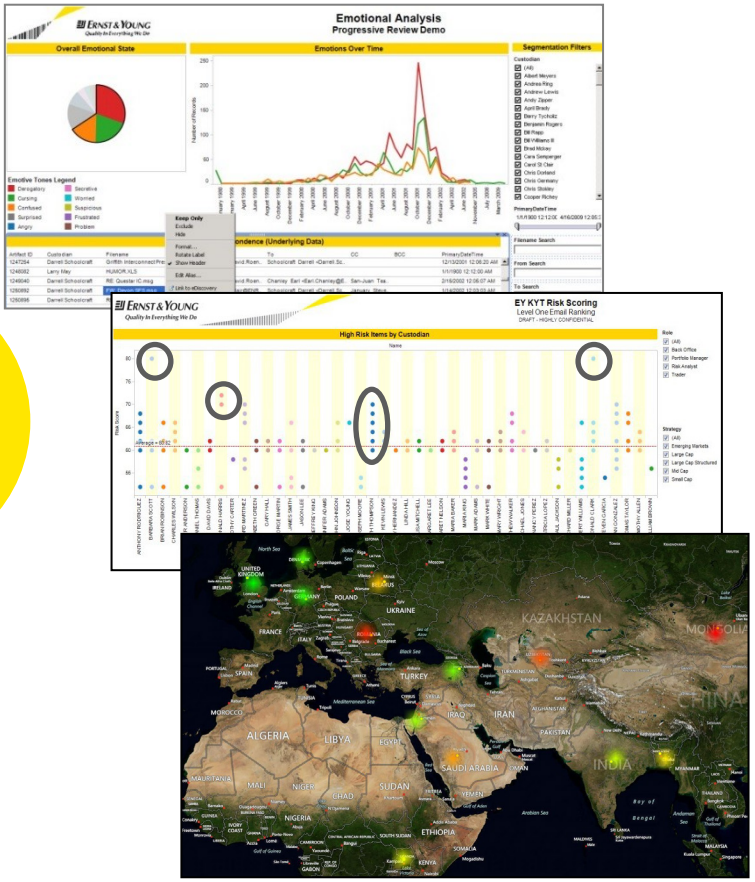
Rationalization	Incentive/pressure	Opportunity
…I deserve it	…make the number	…special fees
…nobody will find out	…don't leave a trail	…client-side storage
…gray area	…not comfortable	…off the books
…they owe it to me	…why are we doing this?	…cash advance
…everybody does it	…pull out all the stops	…side commission
…fix it later	…want no part of this	…backdate
…not hurting anyone	…only a timing difference	…no inspection
…won't miss it	…not ethical	…no receipt
…don't get paid enough		…smooth earnings
		…pull earnings forward

(2) Data collection

Addressing risks by combining multiple data sources

2 Data collection





Other external data sources

- Global watch lists, state-owned enterprises and adverse media
  - More than 1.2 million named individuals and entities
  - More than 160 sanction lists
  - Enhanced country list addressing Presidential Executive Orders (e.g., state-owned entities, subsidiaries and vessels)
  - More than 250 official lists (e.g., World Bank and Asian Development Bank)
  - People who have been formally accused, arrested, or convicted in white-collar and terrorism crimes
  - Relatives and known associates
  - Entities in the news for adverse media
- Social data

(3) Data analysis

Applied data analytics
Purpose – why?

- Assess and quantify potential areas of risk
- Analyze large volumes of data which would be time-consuming for manual review
- Identify exceptions and anomalies
- Find patterns and trends through periodic testing

Applied data analytics
Approaches – what?

Rules-based analytics	**Testing data against known behaviors/patterns**	► Suited to well-defined processes ► Effective for anomaly testing ► Used for identifying data not meeting expected behavior
Risk scoring	**Using scoring models to identify and determine areas of higher risk**	► Using rules-based tests with weighted scores ► Combining weighted scores from different datasets to determine high risk areas
Text analytics	**Analyzing natural language text:** ► Keyword searching ► Concept analysis ► Sentiment analysis	► Search against keywords, e.g., fraud, bribery and corruption, black lists ► Analyze text to determine concepts discussed ► Analyze emails for emotional tone insight
Visual analytics	**Using interactive dashboards for reporting and trend analysis**	► Visualizing facilitates understanding ► Makes comparative analysis easier ► Aids in spotting unusual trends

Applied data analytics
Applications – where?

Travel and expense	**Identify high risk employees through rules and text based analytics. Analyze findings in a dynamic risk scoring dashboard.**	► Expense reports ► Meal attendee information ► Employee master
Conflict of interest		► Employees appearing in vendor master
Procurement	**Identify high risk vendors through rules and text based analytics. Analyze findings in a dynamic risk scoring dashboard.**	► Accounts payable ledger ► Vendor master
Labor and payroll	**Identify anomalies and exceptions in labor charging and payroll.**	► Timesheets ► Contract information ► Payroll
General ledger/ accounting	**Identify anomalous trends and patterns across time, geographies, and business units.**	► General ledger ► Chart of accounts

(4) Explore and risk rank

Transactional risk scoring

4 Explore and risk rank

Per analytic:

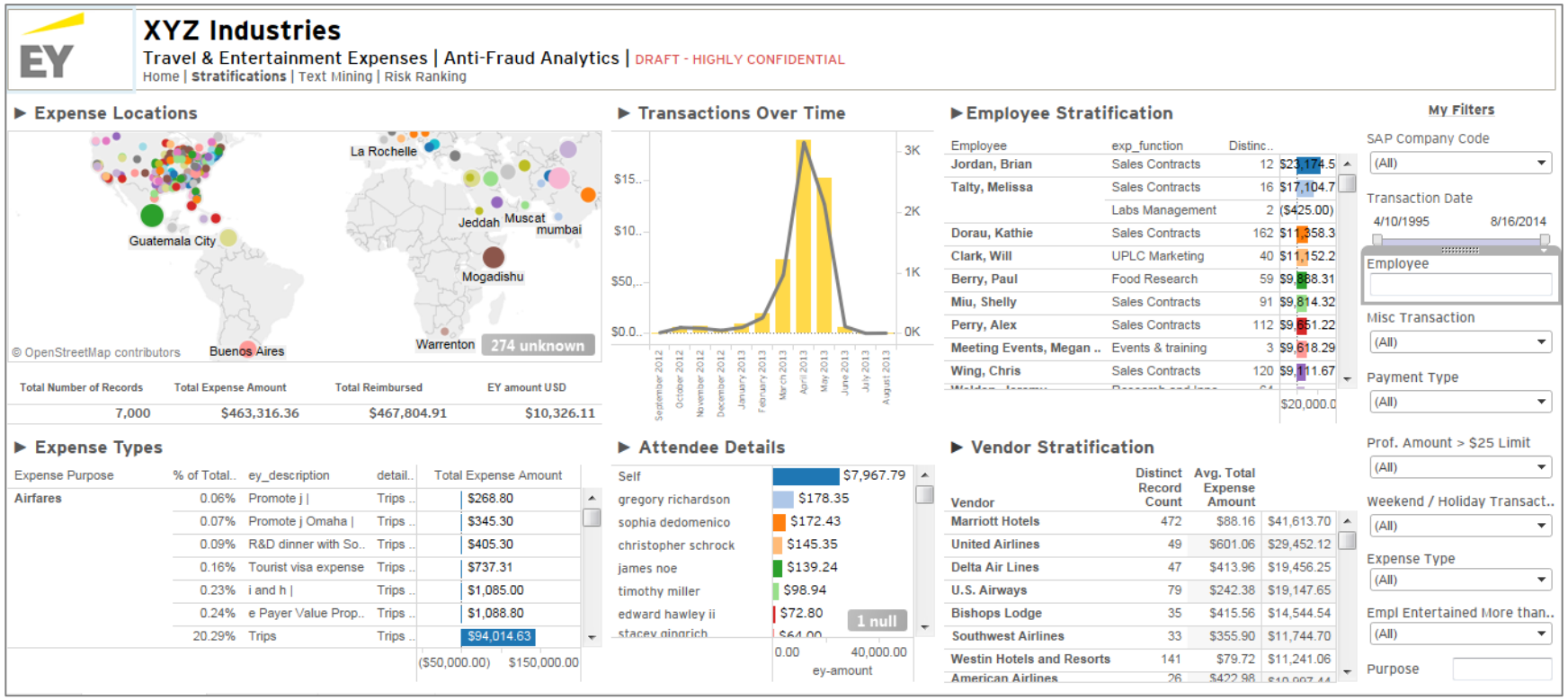
1	(yes/no flag)
x 5	(1–5 metric weight multiplier)

5	(analytic score)

A fully loaded transaction score takes into account the sum of analytic scores that all hit on that transaction, e.g., 1-duplicate, 2-wkend, 3-round payment 5+5+5=15.

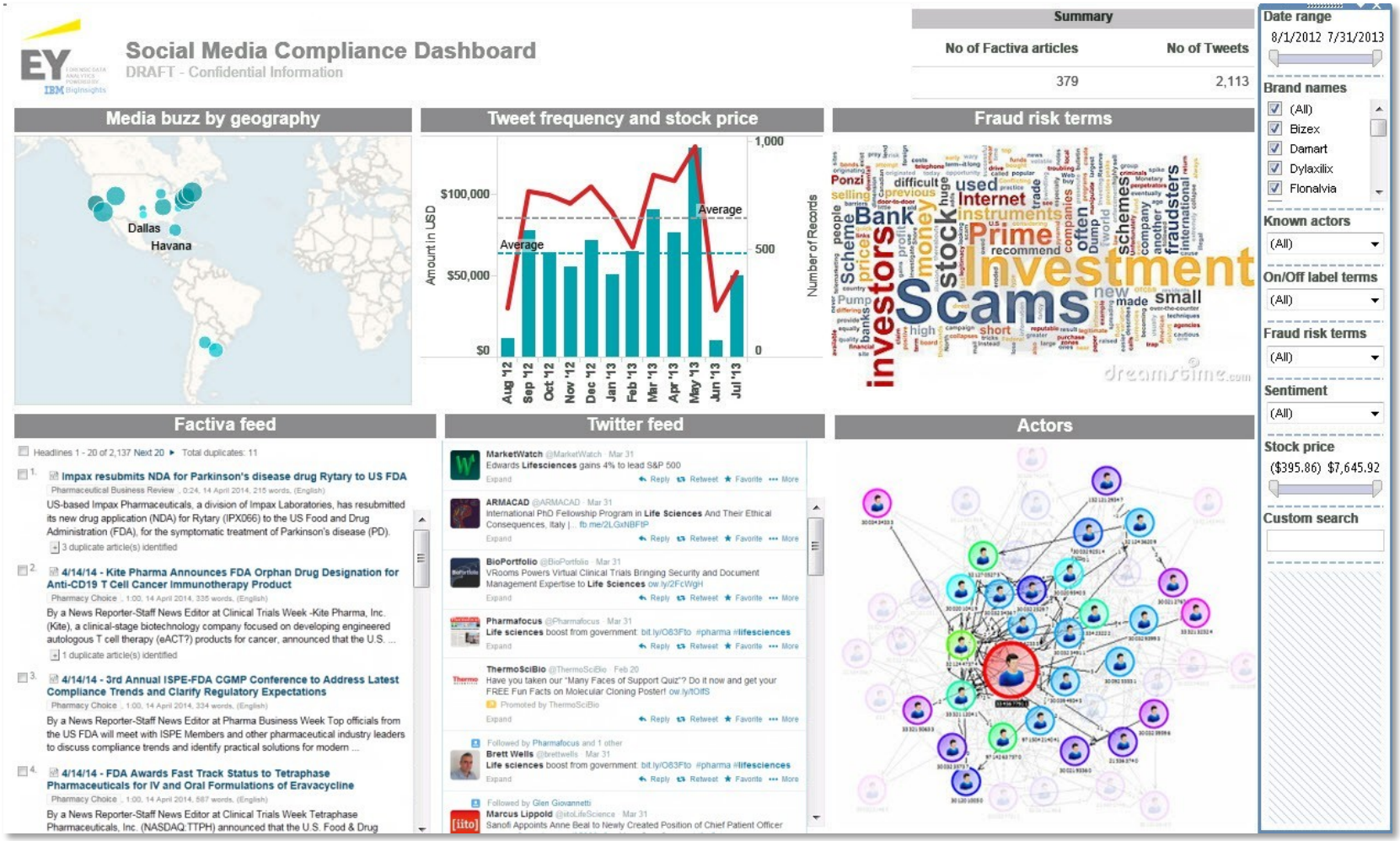
Visualization example
Travel and entertainment

4 Explore and risk rank



Visualization example
Social media

4 Explore and risk rank

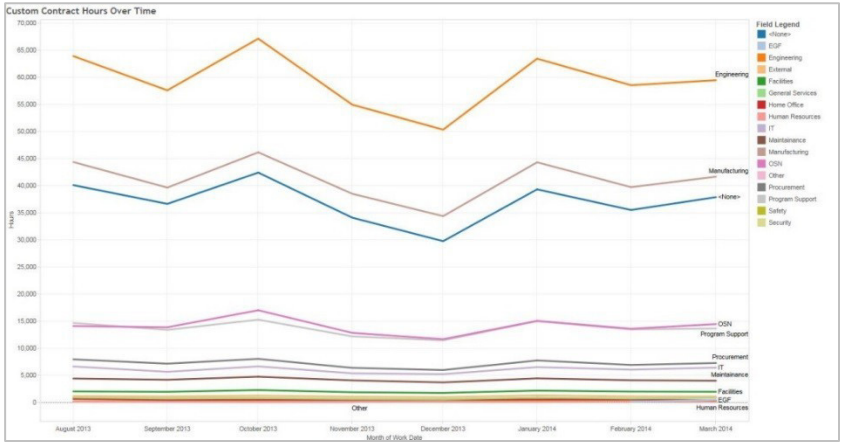


Visualization example
Contracts and time charging

4 Explore and risk rank

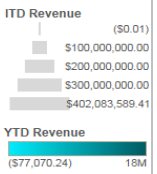
Contract analytics

- Identification of ***high risk*** contracts based on attribute-based risk ranking methodology
- Interactive tool allows for analysis and filtering by customer attributes, including:
  - Contract
  - Contract type
  - Period of performance and percent complete
  - Customer
  - Financial performance (e.g., profit, billed accounts receivable, billings in excess of revenue, revenue in excess of funding)



Custom Contracts Summary

Selected Field	Number of C..	Current Contract ..	Funded Value	YTD Revenue	ITD Revenue
DFJ0000118709	1	$7,313,267.00	$7,313,267.00	$40,704.65	$7,240,602.25
DFJ0000120428	1	$7,152,282.00	$7,152,282.00	$0.00	$7,152,282.00
DFJ0000129016	1	$9,000,000.00	$2,053,411.00	$0.00	$2,026,696.60
DFJ0000129946	1	$42,837,413.00	$34,180,420.00	$3,290,060.82	$24,544,238.28
DFJ0000154999	1	$6,312,584.72	$2,151,942.00	$0.00	$2,151,942.00
DFJ0000159931	1	$28,788,881.00	$28,788,881.00	$0.00	$28,788,881.00
DFJ0000163670	1	$20,597,402.00	$5,952,480.52	$0.00	$5,948,793.46
DFJ0001119323	1	$26,535,657.00	$2,164,555.00	$0.00	$2,164,555.00
DFJ0001165399	1	$7,100,000.00	$7,100,000.00	$3,426.24	$7,042,430.48
DFJ0001165511	1	$20,174,000.00	$8,906,783.34	($11,227.39)	$8,551,459.69
DFJ0001201166	1	$14,465,089.41	$14,465,089.41	$0.00	$14,447,413.38
DFJ0001215376	1	$8,247,770.00	$8,247,770.00	$0.00	$8,242,316.22
DFJ0001243632	1	$7,265,433.00	$306,106.00	$0.00	$303,604.54
DFJ0001309412	1	$180,632,584.88	$56,386,899.64	$10.22	$56,206,216.86
DFJ0001333217	1	$11,521,360.53	$4,211,808.00	$0.00	$4,046,518.27
DFJ0001350232	1	$6,986,244.00	$6,986,244.00	($0.01)	$6,943,476.82
DFJ0001403419	1	$8,621,257.32	$718,940.69	$0.00	$119,647.15
DFJ0001406782	1	$21,280,262.00	$21,280,262.00	$0.00	$21,280,262.00
DFJ0001444981	1	$6,851,628.00	$6,756,628.00	$0.00	$6,542,337.52
DFJ0001514302	1	$14,047,060.00	$14,047,060.00	$0.00	$13,890,456.00
DFJ0001516097	1	$70,179,871.69	$3,761,711.69	$0.00	$3,119,293.00
DFJ0001602020	1	$170,064,114.00	$35,811,874.00	$6,128,935.10	$11,249,227.79
DFJ0001627931	1	$8,600,000.00	$8,600,000.00	$10,957.12	$8,599,668.35
DFJ0001692378	1	$35,137,703.00	$7,068,700.00	$0.25	$7,061,524.52
DFJ0001708775	1	$12,765,189.75	$11,747,327.13	$0.00	$11,181,350.42

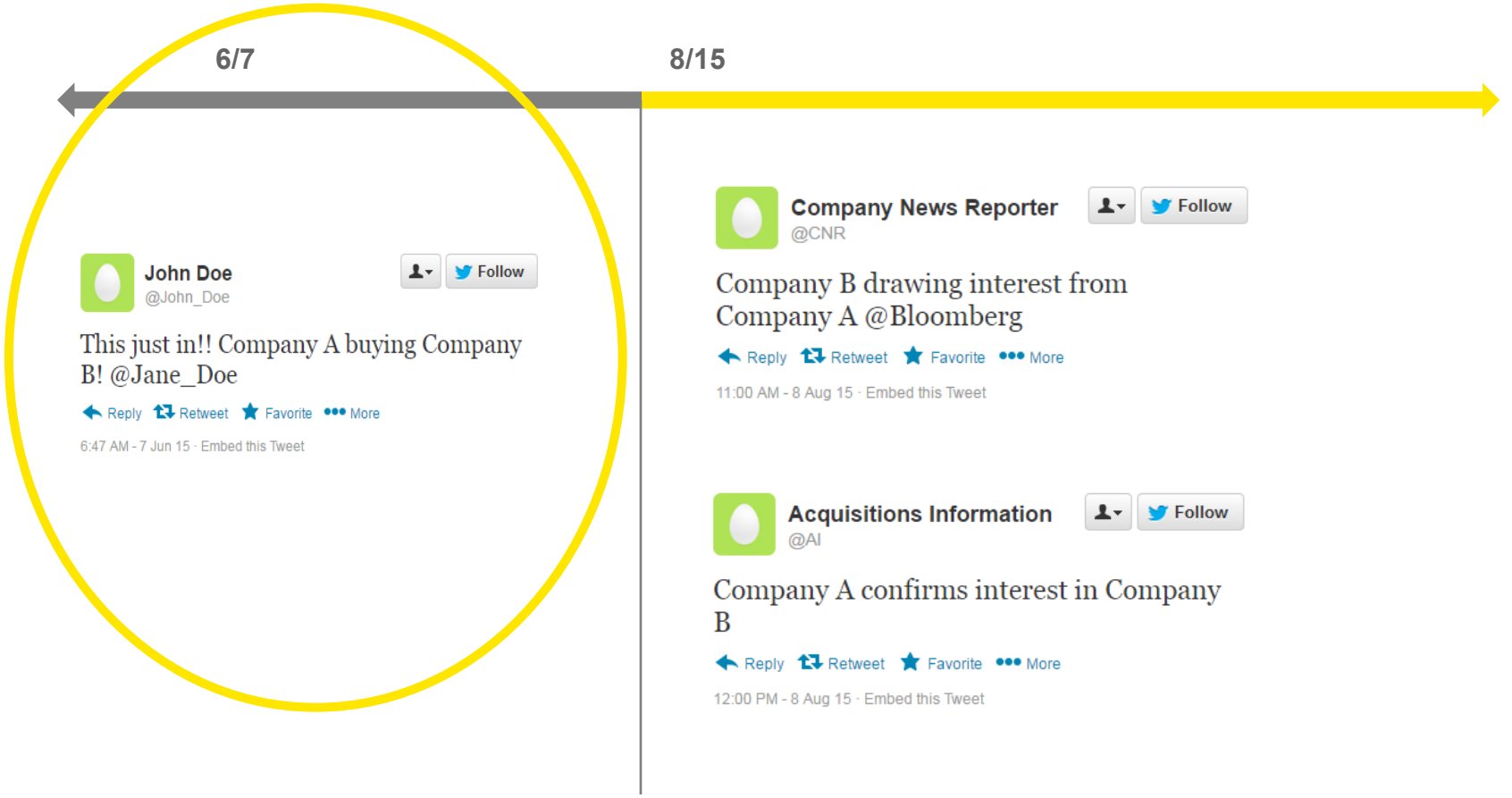


Time charging analytics

- Comparison of labor data to contract qualification requirements
- Identification of unusual trends in straight time and overtime labor charging
- Detection of charges made to multiple categories by the same employee
- Comparison of billed rates to contracted rates
- Trending of time charging practices by employee, contract, cost center and labor category

(5) Findings and observations

Social media analytics
Sample findings – Twitter leak before acquisition

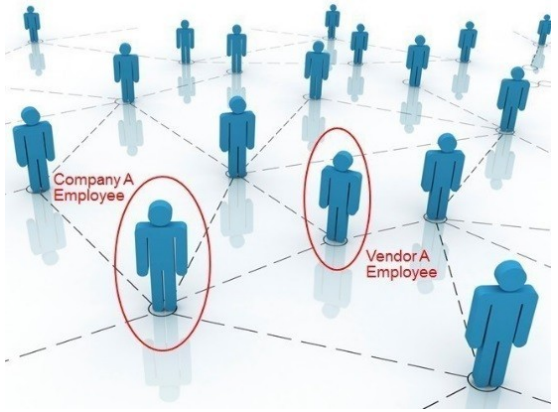


Potential improper payment
Detection of behavior and fraud risk areas

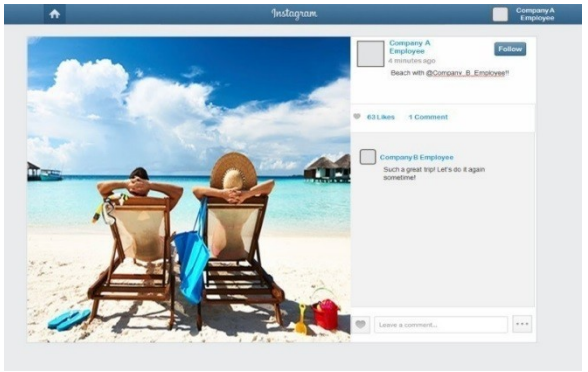
- Hypothesis: Company employees or Vendor employees should not do business and have close relationships.
- Looked through accounts payable – found large payment to Vendor A
- Found employee lists for both companies
- Searched Facebook/Instagram for employee profiles
- Observation: An employee from Company A recently went on an extravagant vacation with an Employee from Vendor A.
- Challenge: Company A employee may have secured an account with Vendor A by agreeing to take Vendor A employee on vacation.

Business impact

- Total paid in the transaction: $125,000



Company A employee and Vendor A employee are connected on social media.



Vendor A employee tagged in photo from extravagant vacation with Company A employee.

Potential improper payment
Text mining for corrupt intent

5 Findings and observations

- Found unusual spending activity in accounts payable
- Scanned Facebook posts from all of Company A's employees
- More than 350,000 distinct noun or noun phrases (concepts)
- Potentially fraudulent high risk terms derived

- Notable observations include:
  - Mention of intent to purchase luxury items
  - Mention of previously purchased luxury items
  - Employee salary does not suggest an ability to purchase such items.

Business impact

- Total paid in the transactions: $78,000

Insider threat

Insider threat
Industry perspective – where focused today

Insider threat is a major risk.

Insider fraud is the primary threat:

75% 75% of insider attacks are fraud
University of Leicester study 2013

Insider attacks take longer to resolve than any other type of cybercrime:

54.4 Average number of days to resolve a malicious insider attack
Ponemon Institute Cost of Cybercrime survey 2015

Companies remain vulnerable to the insider threat.

88% Of respondents say that they feel at least somewhat vulnerable to insider attacks. 34% felt extremely vulnerable.
Vormetric Insider Threat Report 2015

Key issues and priorities

Risk

- Identifying and profiling risks within the organization
- Developing more appropriate risk management to minimize the likelihood of a risk event resulting from authorized personnel

Stakeholder support and engagement

- Emphasis on holistic, collaborative approaches with legal, regulatory affairs, group IT security, HR and business units
- Targeted communications strategies and material for high risk user groups

Identification and inventorying of assets

- Identification and mapping of critical information assets across the enterprise to enable prioritization of control enhancements

Analytics

- The use of data analytics to better leverage existing logging and monitoring capabilities to identify malicious insider behavior
- Utilization of emerging technologies to build employee profiles and locate anomalous data use

Insider threat
Definition

- A current or former employee, contractor, or business partner who:
  - Has authorized access to an organization's network, system, or data, and intentionally exceeded or intentionally used access in a manner that negatively affected the confidentiality, integrity, or availability of the organization's information or information systems
- The insider threat can include the following:
  - Fraud, theft of intellectual property (e.g., trade secrets, strategic plans, and other confidential information), unauthorized trading, IT sabotage or espionage
  - An insider can deny, degrade, disrupt, destroy, deceive, or corrupt information or information systems.

Insider threat
Activities

- The Centre for the Protection of Critical National Infrastructure, the UK Government security authority for infrastructure protection, has identified the following five main types of insider activity:
  - Unauthorized disclosure of sensitive information
  - Process corruption
  - Facilitation of third-party access to an organization's assets
  - Physical sabotage
  - Electronic or IT sabotage
- The most frequent types of insider activity identified were unauthorized disclosure of sensitive information (47%) and process corruption (42%).

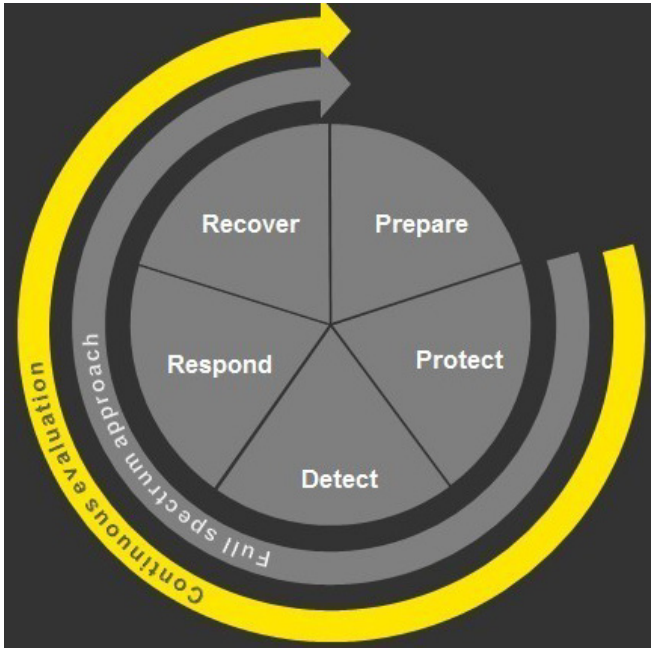
Insider threat
Insider demographics

- Available demographic information findings include:
  - Significantly more males engage in insider activity (82%) than females (18%).
  - 49% of insider cases occur within the 31–45 years' age category.
- The majority of insider acts are carried out by permanent staff (88%).
- The duration of the insider activity ranges from less than six months (41%) to more than five years (11%).
- The majority of insider cases are self-initiated (76%).
- It is common for insiders to have more than one motivation. Financial gain is the most common primary motivation (47%), followed by ideology (20%) and a desire for recognition (14%).

Insider threat
Exploiting weaknesses

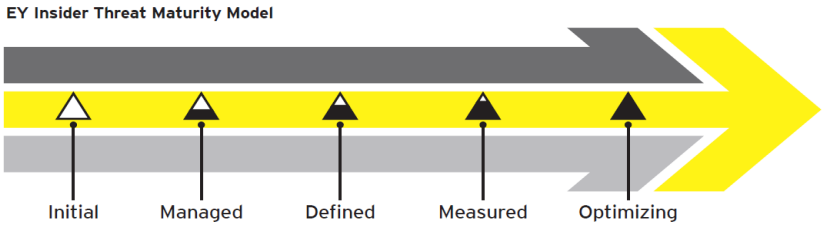
- There is a clear link between an insider act taking place and exploitable weaknesses in an employer's protective security and management processes:
  - Inadequate management practices and inadequate corporate governance
  - Insufficient use of auditing functions
  - Lack of protective security controls and a poor security culture
  - Poor pre-employment screening
  - Poor communication between business areas
  - Lack of awareness of people risk at a senior level

Insider threat
Framework and maturity model



Designed using counter-intelligence experiences and guidelines from the Intelligence and National Security Alliance, the Centre for the Protection of Critical National Infrastructure and the US/UK CERT organizations

EY has the only comprehensive insider threat framework and maturity model in the marketplace.



Insider threat detection program
Prepare

	Plan and collect	Review and analyze	Develop strategic framework
Key activities	► Identify key stakeholders ► Current access management policies and procedures ► Examine corporate hiring and screening procedures ► Review legal and privacy considerations and requirements ► Gather sampling of historical insider threat use cases ► Inventory current user behavioral monitoring tools and techniques ► Obtain consequence protocols and procedures	► Analyze key cybersecurity roles and responsibilities ► Highlight critical business stakeholders (e.g., HR, Legal and Global Security) ► Create key use cases to assist in scoping the insider threat detection program ► Identify legal, privacy and cultural considerations for adoption of insider threat detection program ► Document current insider threat strategic considerations and scope definitions (e.g., corporate communications, awareness and training, employee screening, consequence management)	► Draft initial insider threat strategic roadmap ► Validate key program stakeholders ► Define governance structure ► Identify integration with relevant, existing programs (e.g., employee background verification, Keep It Tight security awareness, incident response) ► Document insider threat detection framework and high-level process flows ► Define insider threat detection team to include experience and skillsets ► Draft sample communications plan ► Socialize strategic roadmap with key stakeholders ► Refine and finalize insider threat detection strategic roadmap based on stakeholder feedback ► Explore behavior analytics/implementation
Reports		► **Insider threat detection strategic considerations and scope definition**	► **Insider threat detection strategic roadmap**

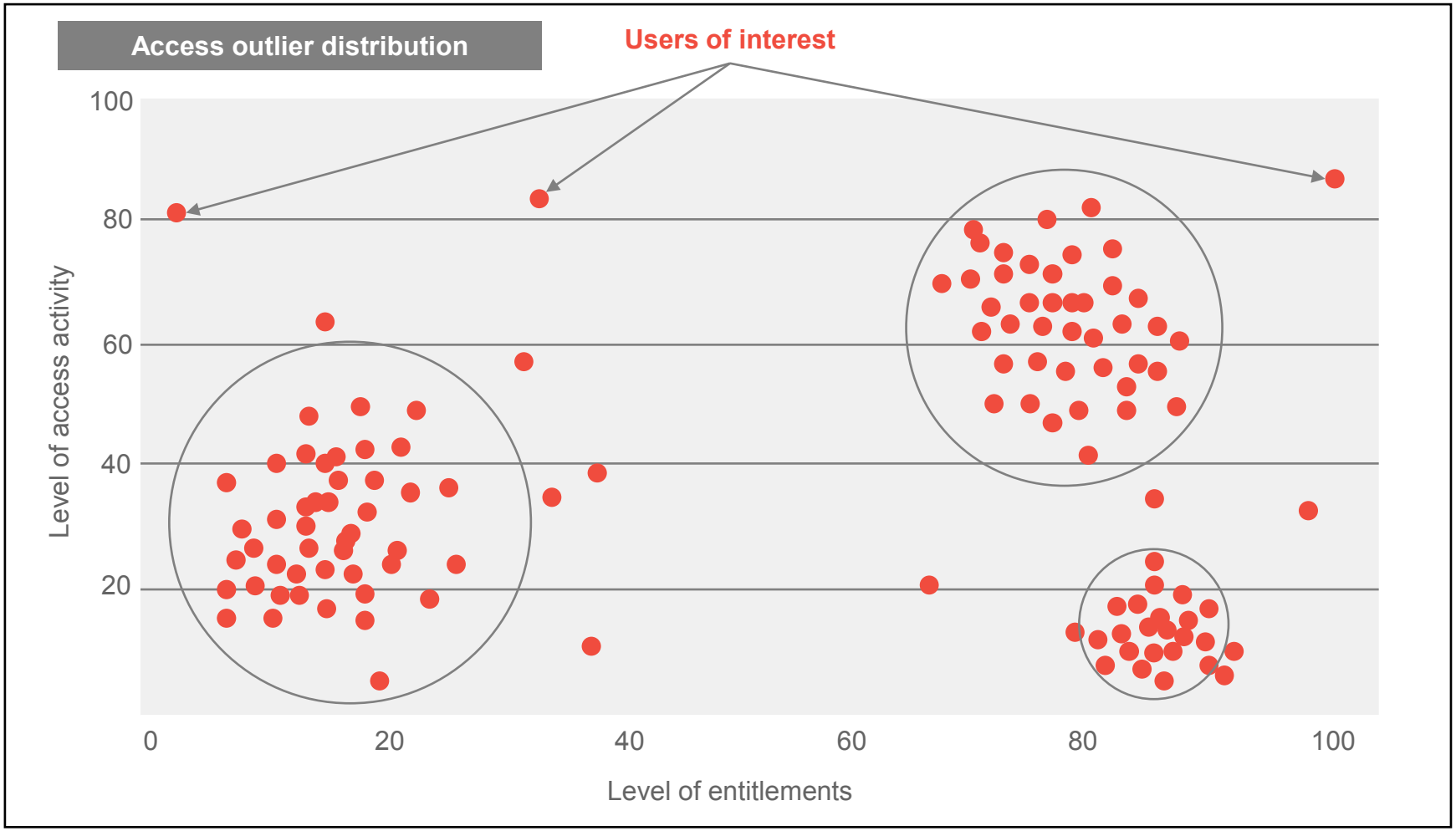
Insider threat detection program
Protect

- Establish or refine policies and procedures
- Establish or refine rules development and management
- Set in place awareness and training
- Evaluate access controls
- Employee background screening
- Formulate key performance indicators (KPIs)

Insider threat detection program
Detect

- Advanced technology – proactive monitoring and behavioral analytics
  - User behavior analytics
  - Data loss prevention
- Identify and leverage technical and non-technical data
- Review current employee reporting programs
- Develop playbook and response plans

Insider threat detection program
Behavioral access outlier analysis



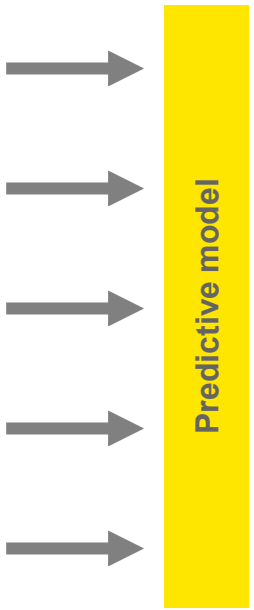
Insider threat detection program
Prioritized data sources

Data source	Indicators
HR	Terminations, layoffs, performance issues
Employee HR/ethics hotline	Complaints of hostile, abnormal or illegal behavior, living beyond means
Criminal records	Identify type of criminal record
Financial issues	Financial hardships, bankruptcy
Legal issues	Subject or party to lawsuit
Phone logs	Calls to/from countries known for IP theft or hosting competitor
Security/employee badge logs	Identify anomalies in employee work hours
Travel data	Travel to countries known for IP theft, or hosting competitors
Data loss prevention logs	Alerts for movement of data containing proprietary markings, sensitive materials, PII
Linguistic analysis of email/IM content	Identify individuals exhibiting malicious behavior
Endpoint/network data	Web crawling, data hoarding, copying from internal sites/repositories, connections to foreign countries
Open source data	Leaks via social networking, indications of employee leaving to competitor

Insider threat detection program
Machine learning models

Sample threat vectors

- Emails to competitors
- Frequencies of email
- Time stamps of email headers
- GeoIP of email headers
- Poor performance rating
- Angry emails or instant messages (IM)
- Secretive emails or IMs
- Copying from SharePoint/repository
- Copying to USB drive
- Frequency and location of large volume of file deletions
- Crawling through network
- Transfers to cloud storage
- Transfer to personal emails



Risk categorization

Insider threat detection program
Respond/recover

- Respond:
  - Execution of response procedures (Playbook)
  - Mitigate damage and contain threats
  - Notification and escalation procedures
  - Evidence collection and handling
- Recover:
  - Damage assessment/mitigation/remediation
  - After action reporting/lessons learned
  - Business continuity and resiliency
  - Review KPI effectiveness
  - Coordination with law enforcement or other external stakeholders

Questions and answers

Contact information

Ken Feinstein
Senior Manager, Fraud Investigations & Dispute Services
+1 203 674 3177
kenneth.feinstein@ey.com

EY | Assurance | Tax | Transactions | Advisory

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.

Ernst & Young LLP is a client-serving member firm of Ernst & Young Global Limited operating in the US.

1609-2049242
ED None

ey.com