Draft - Front to Back Logical Framework or Model for Insider Threat Detection and Prevention

A preventative control framework should aim to mitigate the majority of the insider risk, a control example would be the inability to send attachments outside of the organisation without having to provide a confirmation. Files over a certain size could also be blocked, awaiting approval.

Theoretical Framework

- Risk Factors which could lead to an insider threat incident
- Questions which Literature Review will answer :

1) Which Frameworks exist?
2) Is there one dominant framework?
3) What are the common factors between frameworks?
4) Which gaps if any exist?

- Analysis of insider threat cases, which will help to determine the required data inputs
- The framework is important because it will determine the data which is required
- Certain psychological traits increase the risk of insider threat, such as low self esteem, aggressiveness

An individuals role – what are the risks associated with each role and grade, and are there mitigating controls in place? Are the controls up to date? Controls will need to be mapped to a specific insider threat. There will have to be a central controls repository, which should be updated automatically when a control runs

Data Inputs

The data inputs should map to the theoretical framework

Voice Data
- Voice Meta Data (Call start time etc..)
- Voice Recordings

E-Communications Data
- Email Meta Data
- Email Content
- Different Messaging systems meta data
- Different Messaging systems content

Social Media Data (different vendor feeds)
- As an example facebook, linked in

Human Resources Data
- Employee Relations Data
  a) Case incidents
  b) Case Outcomes
  c) Start date of case, end date of case

Performance Review Data
Bonus data – was the recipient happy, not happy
Employment Status – New joiner, leaver
Department, team, role
Skillset
School attended, university

Printer data
Alerts generated if printing patterns change for an individual

Training Data
- Training overdue

Compliance Data

Workstation sign-on and sign-off data to detect changes in working patterns

Transcription and Indexing

Natural Language Processing & indexing

This would involve transcribing voice recordings, including mobile. There has to be a measure of accuracy, associated with each transcription, for example if there is background noise this may reduce the accuracy.

The NLP engine also needs to be able to support a large number of languages and dialects.

E-Communications Processing and Indexing

There will be embedded unstructured data such as images which will need to be categorized

This would include data from messaging platforms and social media data

A large number of dialects and languages will need to be supported

Analytics Engine

Voice Recognition Search

Being able to search for a voice pattern using a sample voice recording

Sentiment Analysis

Anger can be an indication that something is wrong, this can be detected by the tone and speed of speech

Profanity can be another indictor or warning

Keywords which can be highlighted such as ‘unhappy’, ‘hate’, ‘angry’, ‘bully’

Relationship Analysis

Voice meta-data, e-mail meta-data, employment history, universities attended, schools attended, social media data can be used to capture the relationships between individuals

Behaviour Risk Analysis

Risk Engine

The risk engine would link together the data which has been processed from different sources, and graphically display the inter relationships between the data and also provide a critical, high, medium and low risk rating

The risk engine would need data which relates to existing controls by role and department

The risk engine would need data to show that the controls are up to date

The risk engine would need data for each potential insider threat and a list of the controls which mitigate that threat and also a rating of impact, such as high, medium, low if the threat occurred

The risk engine needs details of the controls which are in place by role, and how these map to a specific insider threat

What are the categories of insider threat?

There will have to be a central repository of detective controls, with users having the ability to refine and create additional controls.

Risks or Challenges

Building on Existing knowledge

- Without completing a literature review it is difficult to assess whether the proposed project will extend knowledge in this area

Ethics

- This approach involves surveillance of individuals and mapping their personal relationships, so there will be issues in terms of ethics and privacy
- Controls need to be in place to ensure that employees with access to the system, only perform relevant searches for approved reasons
- Do organisations want to uncover too many potential insider threat risks? How will this be viewed by the regulator?

Technical and Operational Challenges

- Volume of data will be significant, need to leverage big data tools and techniques
- The voice transcription will require 'tuning', and tuning will involve manual transcription of voice recordings
- Multi-language support needs to exist, which can also support different languages in one conversation, email or chat
- There needs to be an agreed calculation for voice search accuracy, which should include the number of false positives generated
- There is the challenge of representing the complexity of the data and it's inter-relationships on a UI
- Unstructured data which will need to be processed and categorized includes voice recordings, images
- It would be a significant task to create a repository of controls, which mitigate the risk of insider threats and these controls need to be kept up to date
- The controls would need to be mapped to different types of insider threat
- Performance issues
- Cost of storage
- Resources required to analyse insider threat alerts
- High number of false positives
- Is a central asset repository required, and how will this be kept up to date?

UX Design

- The UX design needs to simplify complex relationships between data and enable a 'drilldown' function, as well as workflow

How Can I Reduce the Scope?

1) Focus on just engineering

- Engineering is unique in that engineering controls will be different to controls in other divisions
- Engineering also has the skills and possibly access to cause significant issues
- Some insider threats are unique to engineering, such as copying sensitive proprietary code outside the organisation
- Engineering insider threats can be accidental or malicious, as an example a developer who releases code and floods the market with invalid options trades
- Within engineering the majority of threats should be able to be mitigated by controls, such as a change control process for migrating code into production
- Voice recording data will not be available for engineering

2) Accidental insider threat versus Malicious insider threat

3) Focus on a particular Industry or Business, such as investment banking. The industry should be high risk from a insider threat perspective in terms of impact. For investment banking there is a regulatory requirement to persist all email and voice data in WORM storage

4) Controls can be preventative or detective (after the fact)

Project Ideas

A control framework for preventing and detecting insider threats (both accidental and malicious) for an investment bank for individuals employed as technologists

- The technical elements of the project would be a logical database design and also a prototype of an insider threat UI, as well as a proposed high level architecture
- Advantages – I understand technology controls
- Challenges – how to keep controls up to date (certification, code reviews, 'heart beats' for monitoring controls etc..)
- Steps in the analysis :
  - Literature review
  - Define different types of insider threat for technology
  - Look at real examples and case studies
  - Define the controls to mitigate the threats, this would also involve a literature review
  - Apply real world examples to the proposed framework
  - Define the logical data model and test this by applying different real life scenarios
  - Develop a prototype for the insider threat UI
  - Define the different roles within technology which the threats and controls would be mapped to
  - How to calculate the probability of an attack, in terms of High, Medium and Low risk
- I have experience of voice surveillance and so shall I just focus on this area instead? But what are the gaps in terms of research in this area?

Preventative Control Unique Id	Preventative Control Description	How is the preventative control being monitored to ensure it is up to date?	Is a detective insider threat control required or has the preventative control eliminated the risk?	Detective Control
HCMData200	The Human Resources data lake does not enable any access to production data unless it is via a code release which has followed the change control process. All production data changes are applied by a separate change control team.	An alert is generated if there is any unauthorised access to production, and user names cannot be generic. If in exceptional circumstances a change does need to be applied by the development team, it has to be at senior engineer level and require managing director approval.	Yes detective control is required.	An alert is generated if there is any unauthorised access to production, and user names cannot be generic.