

SOFIT: Sociotechnical and Organizational Factors for Insider Threat

Frank L. Greitzer,
Justin Purl,
Yung Mei Leong,
D.E. (Sunny) Becker,

PsyberAnalytix
Human Resources Research Organization
Independent Consultant
Human Resources Research Organization

Presented to:

*39th IEEE Symposium on Security & Privacy, Workshop on Research
for Insider Threat (WRIT)*

San Francisco, CA, May 24, 2018

Research reported here was supported under IARPA contract 2016-16031400006. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Government.

Motivation

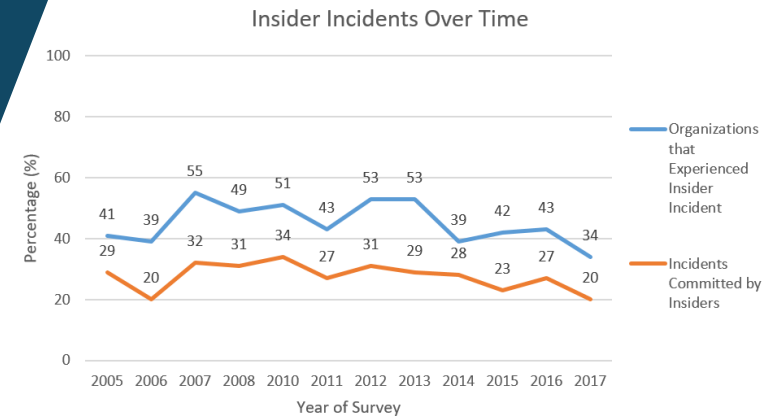
CERT/CSO Magazine annually conducts Cybercrime survey of >500 organizations that self-report on information security issues:

- 34% of reporting organizations experienced cybercrime incident
- 20% of these incidents were caused by insiders
- ~ 30% of insider attacks were more costly or damaging than outsider attacks.

ODNI and NITTF recognize need for tools to assess maturity levels of Insider Threat Program capabilities

- NITTF to conduct “independent assessments of the adequacy of agency programs to implement established policies and minimum standards.”
- All executive branch depts./agencies with national security information or classified networks are subject to NITTF independent assessments.

2017 State of Cybercrime Survey



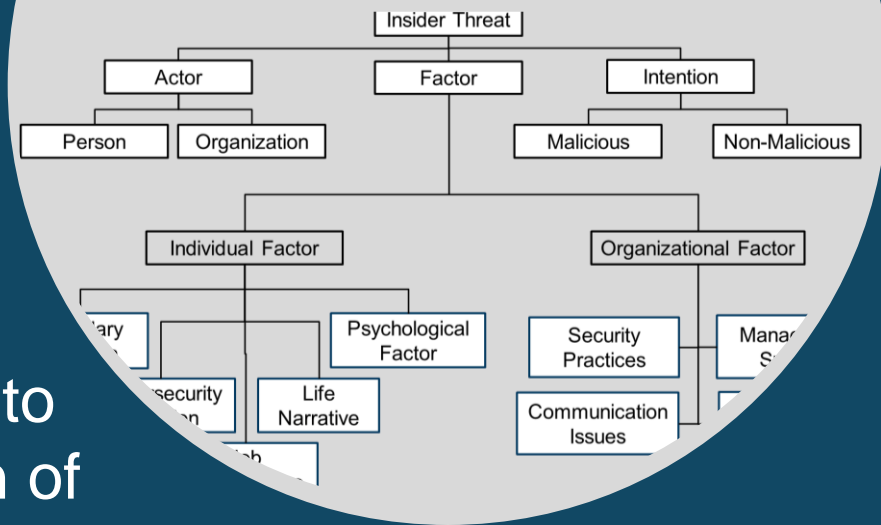
<https://insights.sei.cmu.edu/insider-threat/2018/01/2017-us-state-of-cybercrime-highlights.html>



Objectives

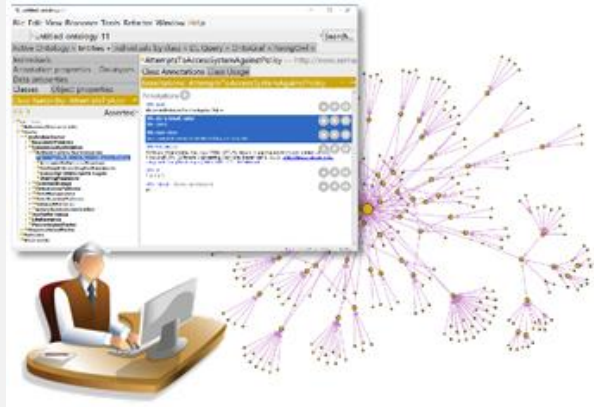
A major goal of this research is to develop a formal representation of factors underlying insider threats

- Extend current insider threat ontology frameworks by incorporating sociotechnical constructs reflecting individual/behavioral and organizational as well as cyber/technical factors
- Support modeling and reasoning approaches for insider threat assessment



Envisioned Applications

Expert Knowledge Repository for Research/Operational Communities



Aid for Evaluating Maturity Level of an Organization's Insider Threat Program

Indicator Categories	# Detected	# Indicators	% of Factors Detected	Insider Threat Monitoring Coverage				
				Boundary Viol	Job Perf	Cyber Viol	Life Narrative	Psych Factors
Boundary Violation	11	59	29%					
Job Performance	10	28	43%					
Cybersecurity Violation	36	75	48%					
Life Narrative	9	64	14%					
Psychological Factors	4	46	9%					

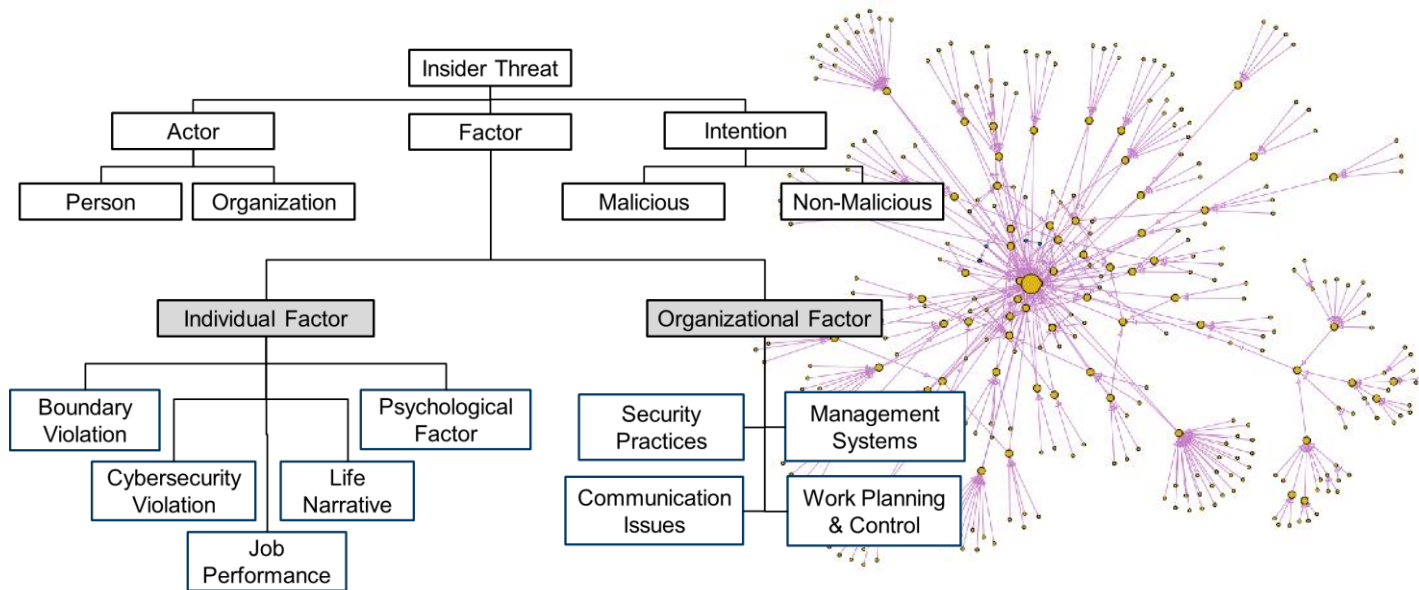
Recommendations	Definition
Emotional instability/neuroticism Low conscientiousness Unreliable Impulsivity Poor time management Disagreeableness Socially aversive Reluctant to self-discipline	<p>Manipulative</p> <p>A Machiavellian pattern of using whatever means perceived necessary in pursuit of goals, including (and especially) manipulative and/or exploitative behavior to control or play upon other persons by wily, unfair, or insidious means, to one's own advantage.</p>

Tool for Assessing Individual Insider Threat



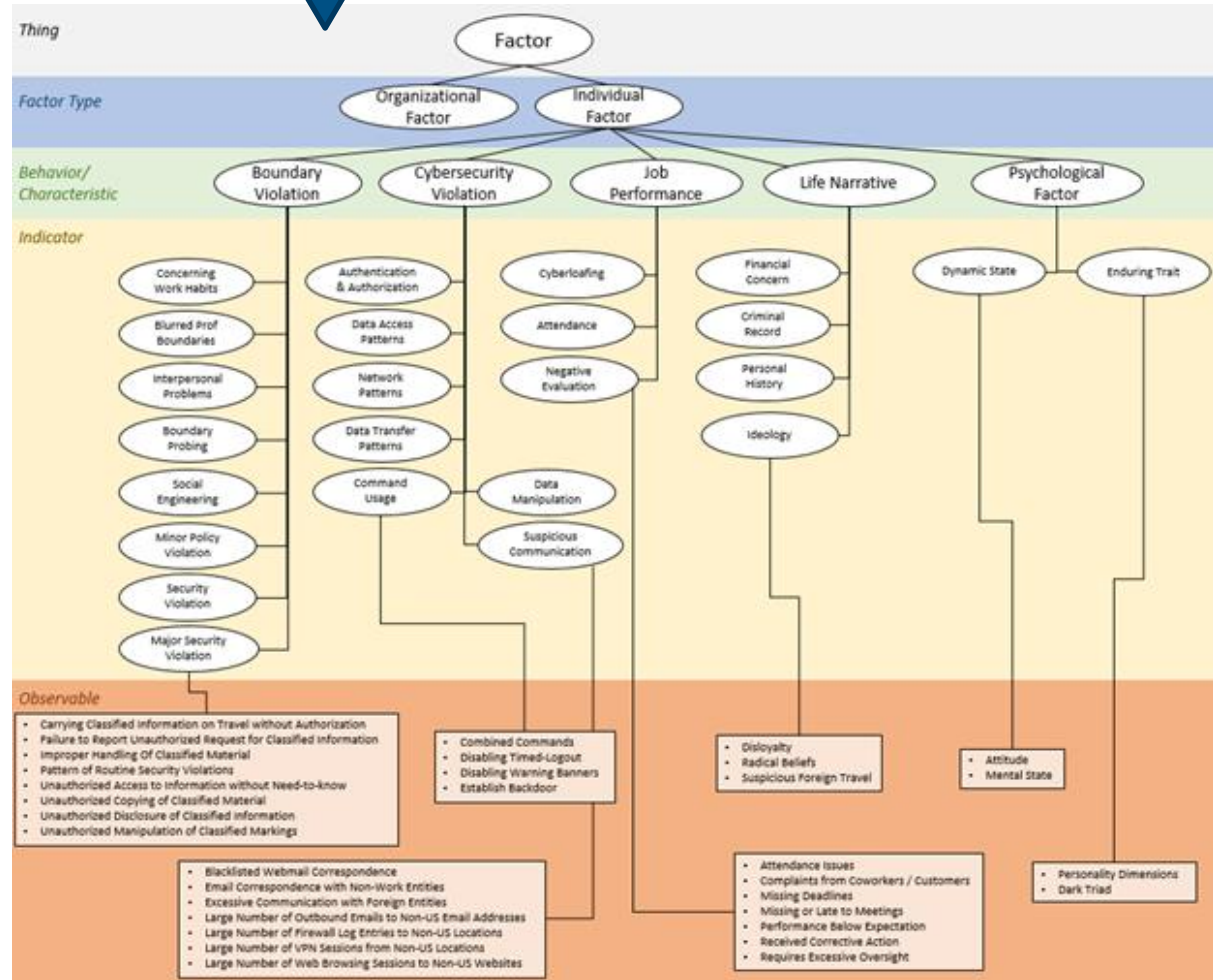
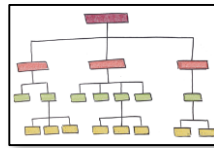
Developing SOFIT:

Sociotechnical and Organizational Factors for Insider Threat



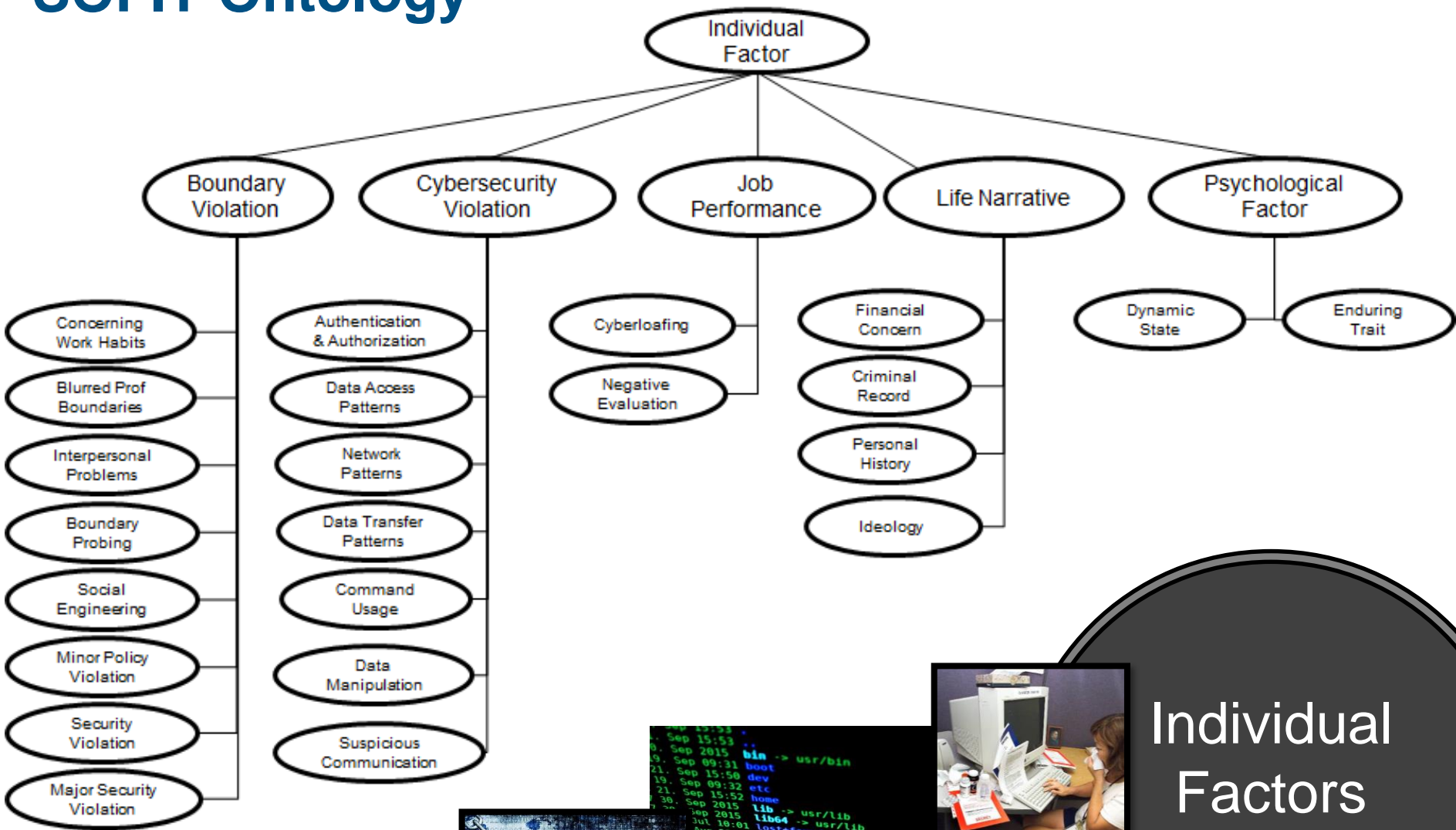
Approach

- Develop taxonomy of relevant factors based on the available knowledge contained in the research literature, case studies, and expert judgment
- Implement ontology using Protégé and the OWL-DL ontology language



SOFIT Ontology includes
 > 300 Individual and
 Organizational
 Contributing Factors

SOFIT Ontology

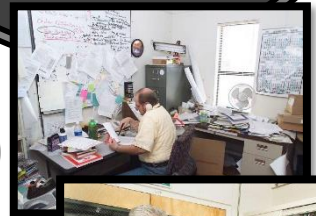
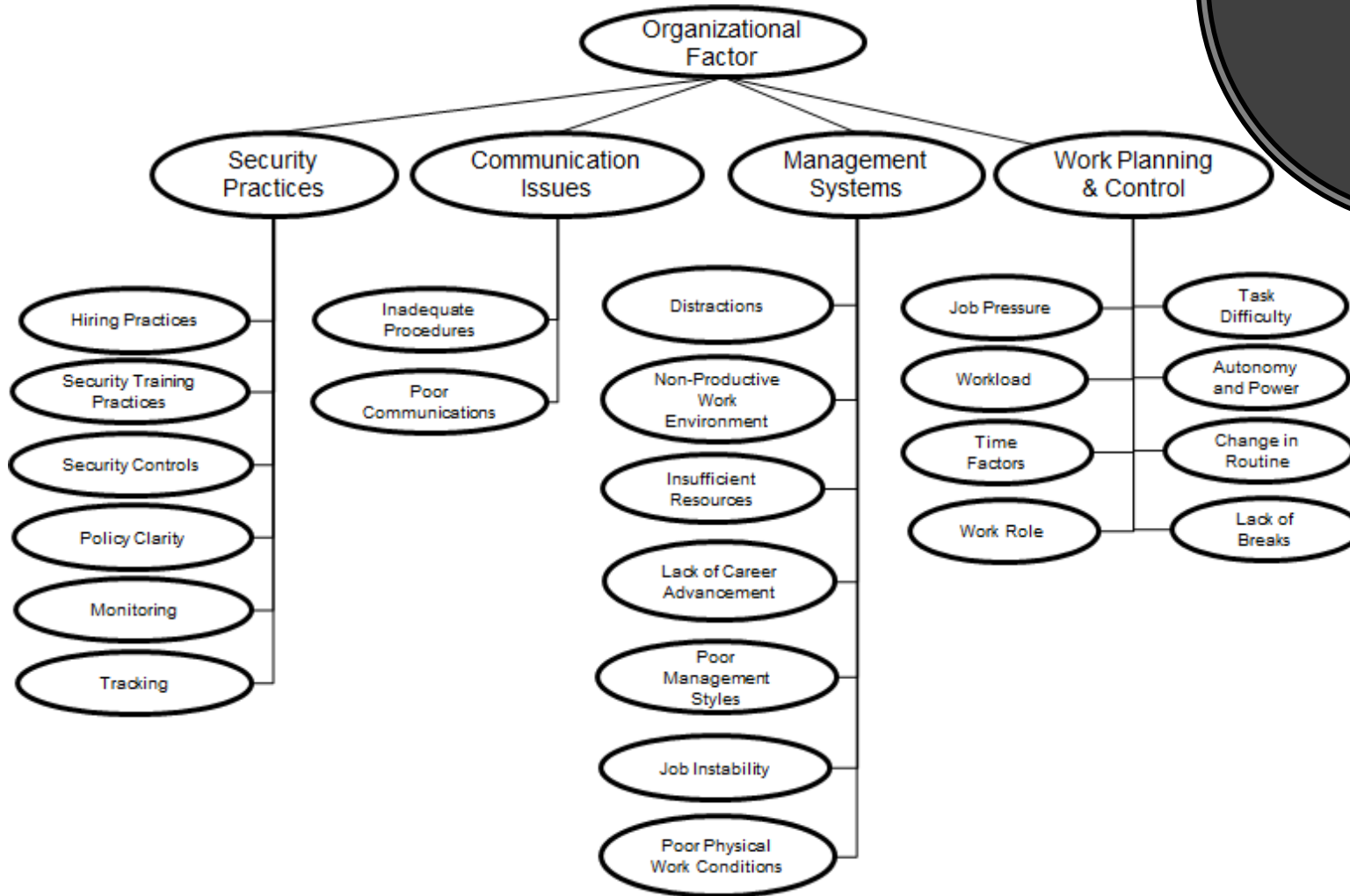


Individual Factors



SOFIT Ontology

Organizational Factors



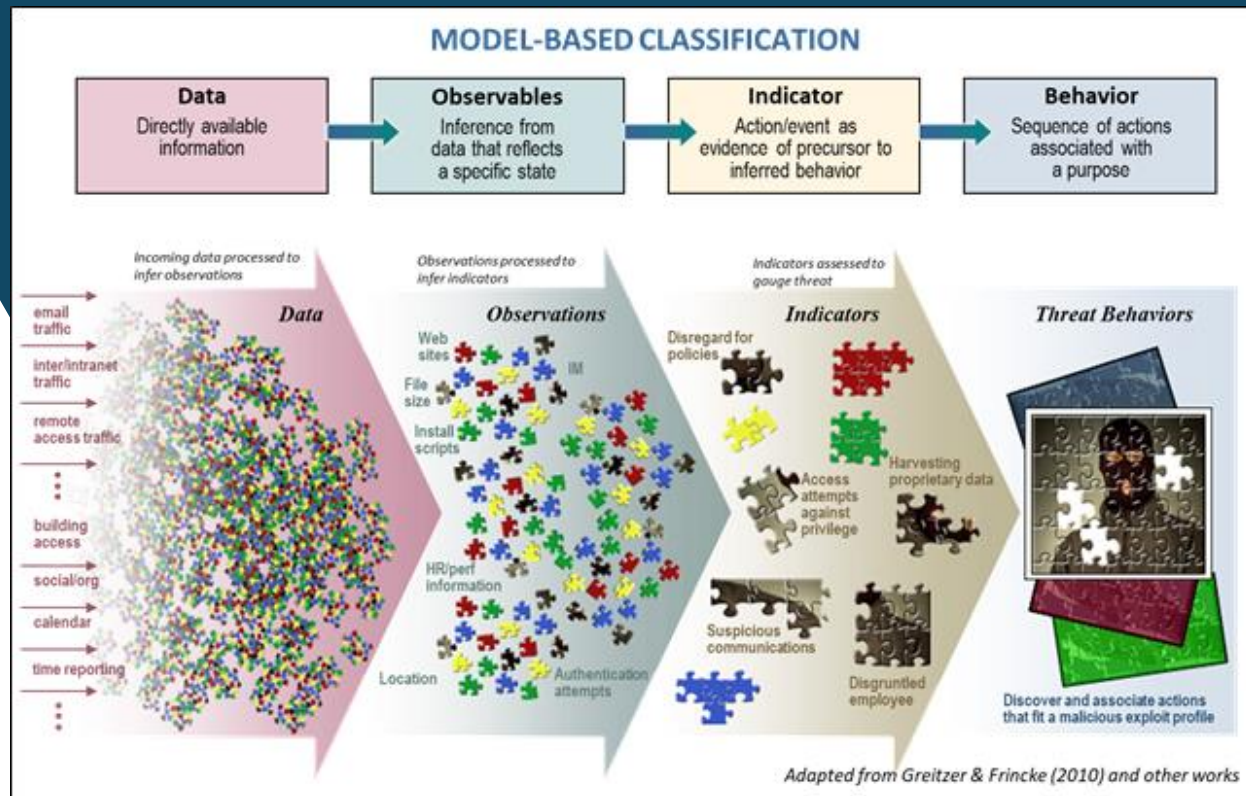
Security awareness training



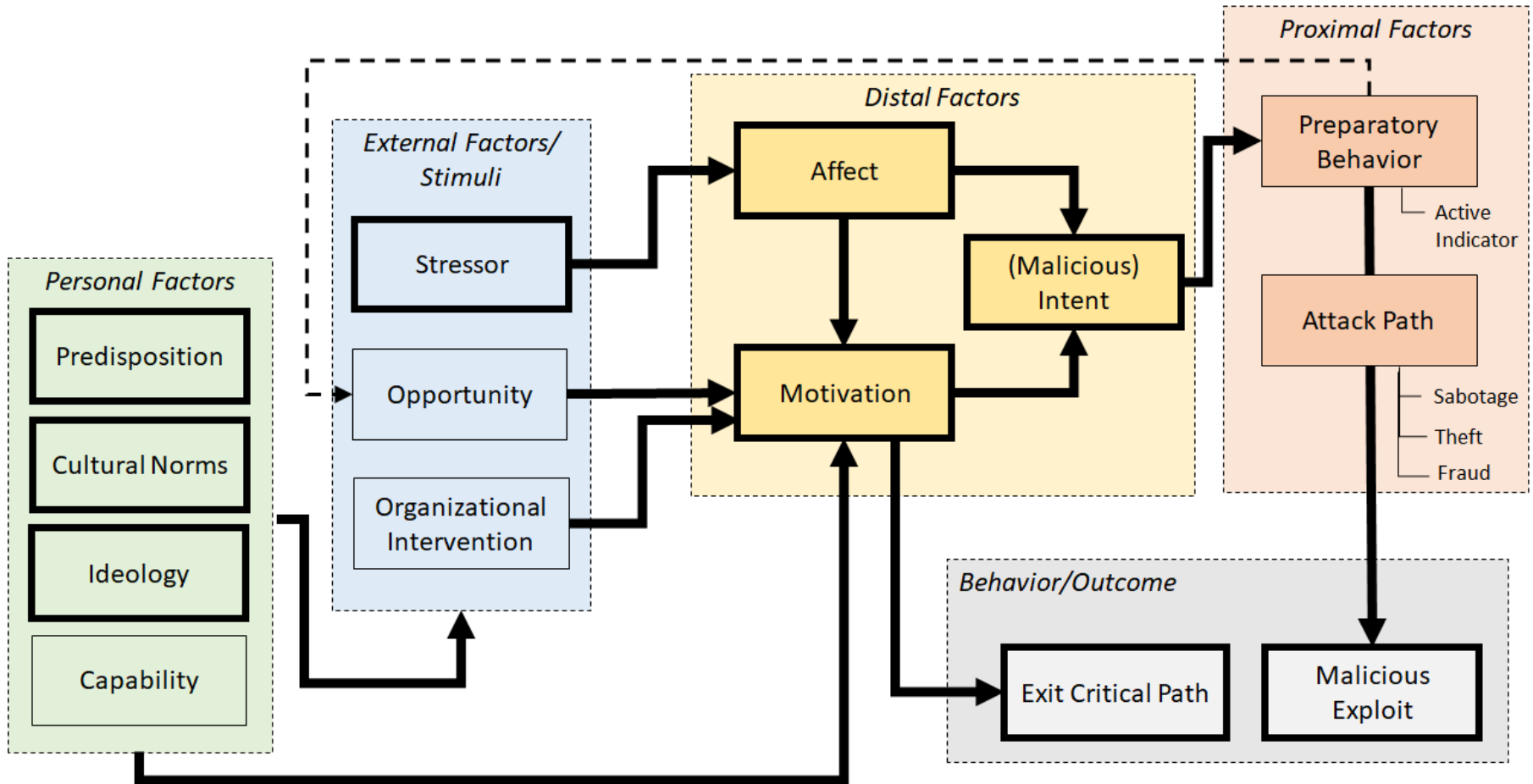
Related Ontology Frameworks

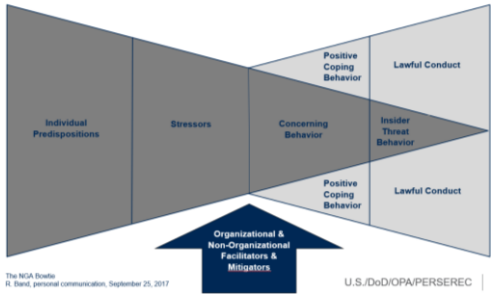
Summary of Current Ontology Representations in Cybersecurity/Insider Threat				
Ontology/Reference	Domain/Scope	Types of Constructs Represented		
		Technical/ Cyber	Human/ Behavioral	Organizational
CERT ITIO	Insider Threat	✓	-	-
MITRE (STIX)	Cyber Security	✓	-	-
MITRE (CAPEC)	Cyber Security - Attack Patterns	✓	-	-
MITRE (CWE)	Cyber Security - Weaknesses	✓	-	-
MAEC	Cyber Security - Malware	✓	-	-
CRATELO	Cyber Security	✓	-	-
HUFO	Cyber Security - Trust	✓	✓	-
SOFIT	Insider Threat	✓	✓	✓

Model Development

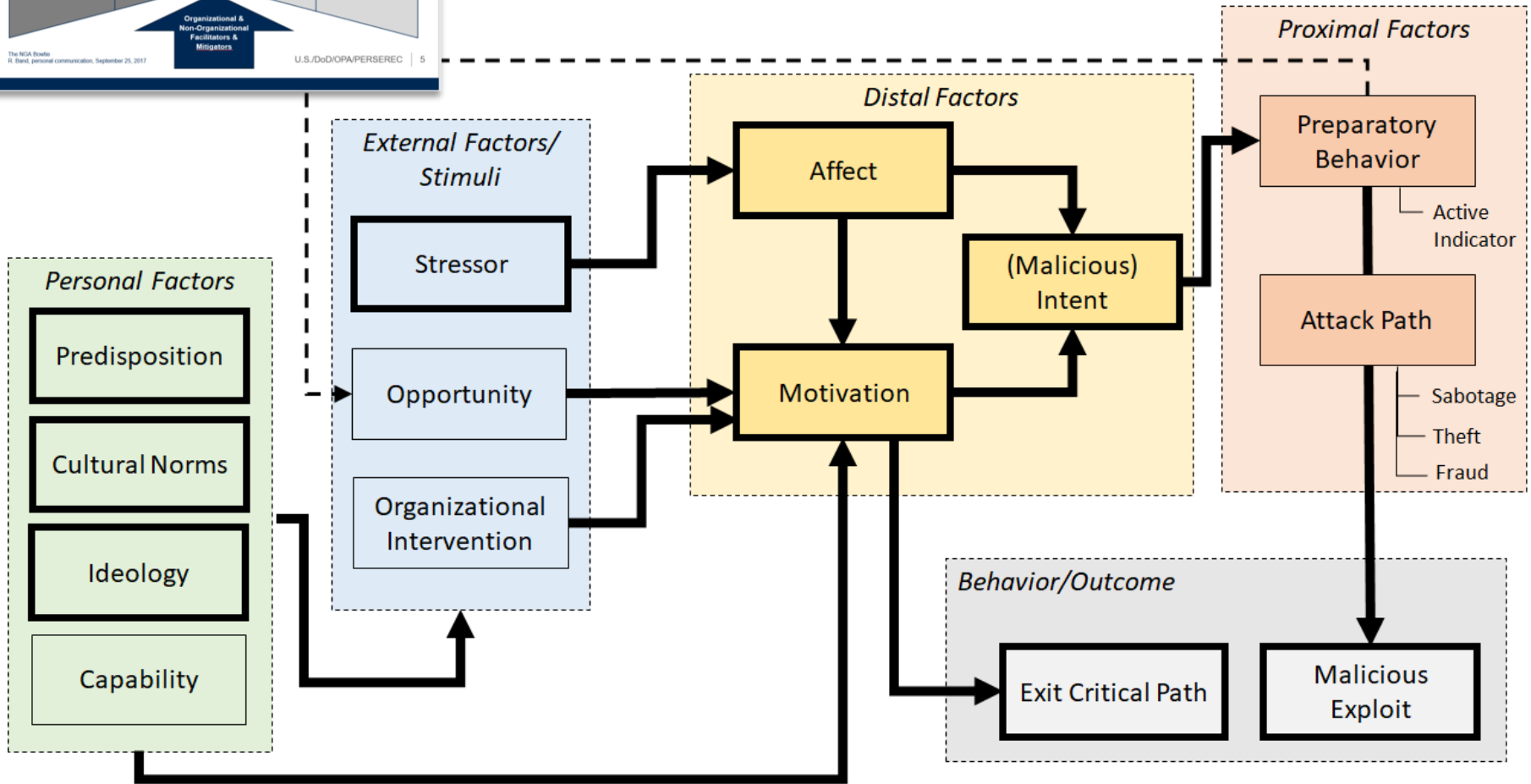


General Framework for Malicious Insider Threat Constructs





Work for Malicious Insider Threat Constructs



Complex Relationships among Constructs

We are considering possible relationships between the insider threat indicators and additional constructs: *threat types* and an indicator's *role* in the insider threat exploit:

Threat Types

- **Insider Sabotage.** An act by an insider to direct specific harm toward an organization or its assets.
- **Insider Data Theft/Exfiltration.** Theft of an organization's intellectual property by an insider.
- **Insider Fraud.** Modification, addition, deletion, or theft, of an organization's data for personal gain, leading to an identity crime (e.g., identity theft, credit card fraud).
- **Unintentional Insider Threat (UIT).** An act or failure to act by an insider, without malicious intent, that causes harm or substantially increases the probability of future harm to an organization or its assets.
- **Workplace Violence.** Any act or threat of physical violence, harassment, intimidation, or other threatening disruptive behavior that occurs at the work site.

Indicator Roles

- **Precipitating Event.** An event that triggers or motivates the insider to carry out an insider crime
- **Personal Predisposition.** A characteristic historically linked to a propensity to exhibit malicious insider behavior.
- **Behavioral Precursor.** An individual action, event, or condition that involves personal or interpersonal behaviors and that precedes and is associated with insider activity.
- **Technical Precursor.** An individual action, event, or condition that involves computer or electronic media and that precedes and is associated with malicious insider activity.
- **Access Path.** Sequence of one or more access points that occur within an attack or exploit—also known as "attack vector" or "kill chain."
- **Contextual Variable.** Factor that adds context; not necessarily predictive.

Qualitative Threat Assessment

Case #1:

- Depression
- Misses or late for meetings
- Recent change in marital status
- Receiving large email attachments
- Requires excessive oversight

Case #2:

- Terminated
- Extreme discontent
- Establish backdoor
- Transfer large amount of data
- Strong reaction to organizational sanctions

Characterization of Case #1	Characterization of Case #2
<p>Precipitating Event</p> <ul style="list-style-type: none"> • Recent change in marital status 	<p>Precipitating Event</p> <ul style="list-style-type: none"> • Terminated
<p>Behavioral Precursor</p> <ul style="list-style-type: none"> • Misses or late for meetings 	<p>Behavioral Precursor</p> <ul style="list-style-type: none"> • Extreme discontent • Strong reaction to organizational sanctions
<p>Contextual Variable</p> <ul style="list-style-type: none"> • Depression • Receiving large email attachments • Requires excessive oversight 	<p>Technical Precursor</p> <ul style="list-style-type: none"> • Establish backdoor • Transfer large amount of data
<p>Indication of Insider Threat: None</p> <p>While there are contextual factors of concern about this employee that may indicate a need for follow-up, there is no indication that this person represents an insider threat risk.</p>	<p>Indication of Insider Threat: Strong</p> <p>The presence of both behavioral and technical precursors, as well as a precipitating event associated with insider threat risk, yields a high level of concern that justifies further analysis by insider threat team.</p>

Quantitative Models

Counting Model

A simple approach is to count the number of indicators observed (n), irrespective of the level of concern associated with any indicator. Formally, the counting model risk score is $R = n$, where n is the number of indicators.

Sum of Risk Model

A simple elaboration of the counting model is obtained by adding the ratings to form a risk score. This model takes account of the variability revealed in the rating task in the most basic way possible. Formally, the risk score is simply the sum of the individual risk values for the reported indicators (x_i) within a given combination:

$$R = \sum_{i=1}^n x_i$$

Linear Regression Weight Model

One such method would be to have analysts make judgments about a sample of (or the full set of) indicator combinations, and regress (using linear regression) the presence of an indicator on the judgments of risk.

$$R = \sum_{i=1}^n b_i.$$

Sequential Weighted Model

An aggregated risk score is obtained by adding increments for each indicator (r_i), based on the indicator's unique risk value (i.e., the individual indicator risk judgment [x_i]), with the constraint that an upper limit is imposed on the aggregated total risk for the set of reported indicators (X). For a case with n indicators that represent k classes, the risk computation is given by:

$$R = \sum_{i=1}^n (r_i),$$

$$r_i = x_1, \text{ for } i = 1$$

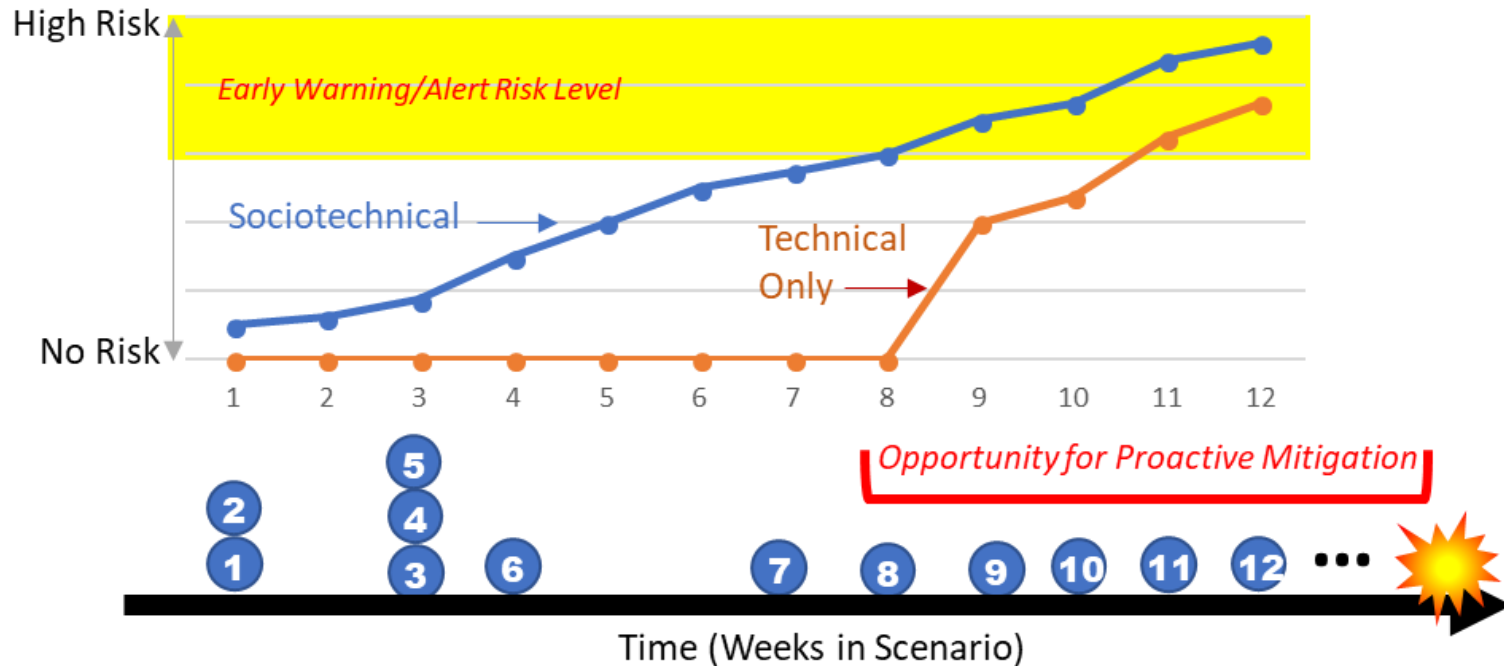
$$r_i = \left(X - \sum_{i=1}^{i-1} r_i \right) * \left(\frac{x_i}{X} \right) * w, \text{ for } i = 2, \dots, n$$

where the increments are weighted by $w = (k/n^2)$

$$= \left(X - \sum_{i=1}^{i-1} r_i \right) * \left(\frac{x_i k}{X n^2} \right), \text{ for } i = 2, \dots, n$$

substituting (k/n^2) for w

Illustrative Timeline Highlighting Potential Proactive Impact of Monitoring Sociotechnical Indicators



Indicators in Scenario Use Case

- Sociotechnical
- Technical

1. [PSYCHOLOGICAL FACTOR: ENDURING TRAIT: NARCISSISM]
2. [PSYCHOLOGICAL FACTOR: ENDURING TRAIT: MANIPULATIVE]
3. [PSYCHOLOGICAL FACTOR: ENDURING TRAIT: CALLOUSNESS]
4. [BOUNDARY VIOLATIONS: INTERPERSONAL PROBLEMS: VERBAL ABUSE]
5. [BOUNDARY VIOLATIONS: INTERPERSONAL PROBLEMS: INTIMIDATING]
6. [BOUNDARY VIOLATIONS: INTERPERSONAL PROBLEMS: INTIMIDATING]
7. [JOB PERFORMANCE: NEGATIVE EVALUATION: ATTENDANCE]
8. [PSYCHOLOGICAL FACTOR: DYNAMIC STATE: ATTITUDE: OVERLY CRITICAL]
9. [CYBERSECURITY VIOLATION: ATTEMPTED ACCESS AGAINST POLICY]
10. [JOB PERFORMANCE: WORKING AT UNUSUAL HOURS]
11. [CYBERSECURITY VIOLATION: PROHIBITED FILESHARING WEB SITE]
12. [BOUNDARY VIOLATION: POLICY VIOLATION: UNREPORTED CONTACT WITH FOREIGN NATIONALS]

Expert Knowledge Elicitation Study



We conducted two distinct studies:

- An initial proof of concept study, narrowly focused on obtaining expert judgments for a small number of indicators (nine experts rated 24 indicators selected from the ontology)
- A broader study seeking expert judgments on all individual indicators (14 experts rated 203 indicators).

Method:

1. Obtain judgments of individual indicator risk to estimate risk scores x_i (level of concern, 0-100)
2. Obtain expert judgment rankings of cases comprising multiple indicators
3. Test ability of alternative models to predict expert judgments of rankings of cases, based on estimates of x_i

Materials Used for Expert Knowledge Elicitation

Study Part II: Estimating Threat/Risk Values for Factors

INSTRUCTIONS: An insider threat Indicator Alerting system monitors a large set of factors (indicators) that are potentially tied to insider threats. The system can ALERT the threat analyst when it detects indicators of sufficient concern (i.e., egregious actions, behavioral patterns, or characteristics) that warrant additional monitoring and analysis by an insider threat assessment team. In the sheet below, for each Class of indicators that you assessed in Part I, there is a list of representative members of the class. In Part I, you estimated the range of insider threat/risk concern for these classes—your range judgments for each indicator Class are displayed as a horizontal blue bar as a reminder for you. Now, for each factor in the class, please estimate a Level of Concern for the indicator's insider threat/risk by entering a value in the corresponding yellow cell or by moving the adjacent slider to an appropriate value. You may use your range estimate as a guide, but you are not required to stay within the previously specified range. The values should be interpreted using the same Indicator Alert scale (0-100) as before, where 0 = "No Concern at all" and 100 = "Gravest concern about an actual exploit or strong inclination/likelihood of committing an exploit." Descriptions/examples of the indicators are shown at right for reference.

Indicator Class/Factor	Range of Insider Threat/Risk Concern	Threat Risk Score	Description
Data Access Patterns		0 - 100	
Attempt unauthorized access to files not backed up		50	Seeking to gain undue access to files that are not backed up
Granting unauthorized access to sensitive data		50	Granting access to sensitive documents to persons) without a need to know
Attempt unauthorized access to sensitive data		50	Seeking to gain undue access to sensitive data or documents without a need to
Attempts to access new workstation		50	Login to a new workstation (physically or remotely)
Attempts to change file permissions		50	Changing or attempting to change file permissions
Document document control		50	Disabling document control safeguards
Request unauthorized access to sensitive data		50	Requesting unapproved access (e.g., access to documents for which one has n
Network Patterns		0 - 100	
Rooting from local media		50	Using local media (e.g., CD or USB-drive) to boot a separate operating system
Compromised machine		50	An individual or network of work computers infected with malicious software (bot
Duplicate log file backup		50	Duplicating log files
Extra backups		50	Making extra backups of network files
Failure to join machine to domain		50	Use or attempt a computer that is disconnected from domain
High activity on high target machine		50	Unusually high activity on a machine containing sensitive data
Printing documents of others		50	Printing documents that are owned by others
Using multiple printers simultaneously		50	Concurrent use of multiple printers
Use of unusual printer		50	Printing to locations unrelated to one's work location
Search computer libraries		50	Unusual search of computer libraries
Search own name		50	Searching logs or security data for own name

Estimating Individual Indicator Level of Concern

Ranking of cases comprising multiple indicators

Categorization

Cases to Assign

- [Unintentional Breach] [Excessive Early departures]
- [Failure to join machine to domain] [Avoids Meetings]
- [Tamper with document classification] [Large data transfers via VPN]
- [Encrypted protocols] [Mislabeling documents]
- [Disengaged Socially] [Working at unusual hours]
- [Rejects criticism] [Search own name]
- [Disable VM protection] [Obscure report of foreign contact]
- [Circumvent document control] [Unusual registry entries or configuration]
- [Violence Outside Workplace] [Sudden Negative Attitude]
- [Grant unauthorized access to classified information] [Use of unusual printer]
- [Emotional Instability/Neurosis] [Attempt unauthorized access to files not backed up]
- [Unreliable] [Attempted unauthorized access]
- [Apply improper security classification markings] [Attempts to attain facility documentation or logs]
- [Excessive curiosity] [Excitement-seeking]
- [Surreptitious rumormongering in internet activity] [Use of covert channels]

Low Concern | Low-Moderate Concern | Moderate Concern | Moderate-High Concern | Extreme Concern

Assign to Low Concern [Remove]

Instructions: To assign the listed cases to one of the five Level of Concern categories, click the appropriate tab to view the desired category "bin" and use the "Assign to" button to deposit one or more selected cases into the desired bin. Once you have assigned two or more cases to a bin, please use the arrow buttons to the right of each bin to rank cases from highest concern (top) to lowest concern (bottom). At any time, you may move a case out of a bin and deposit it in a different bin. When you are finished, please use the "Finish" button to exit the task.

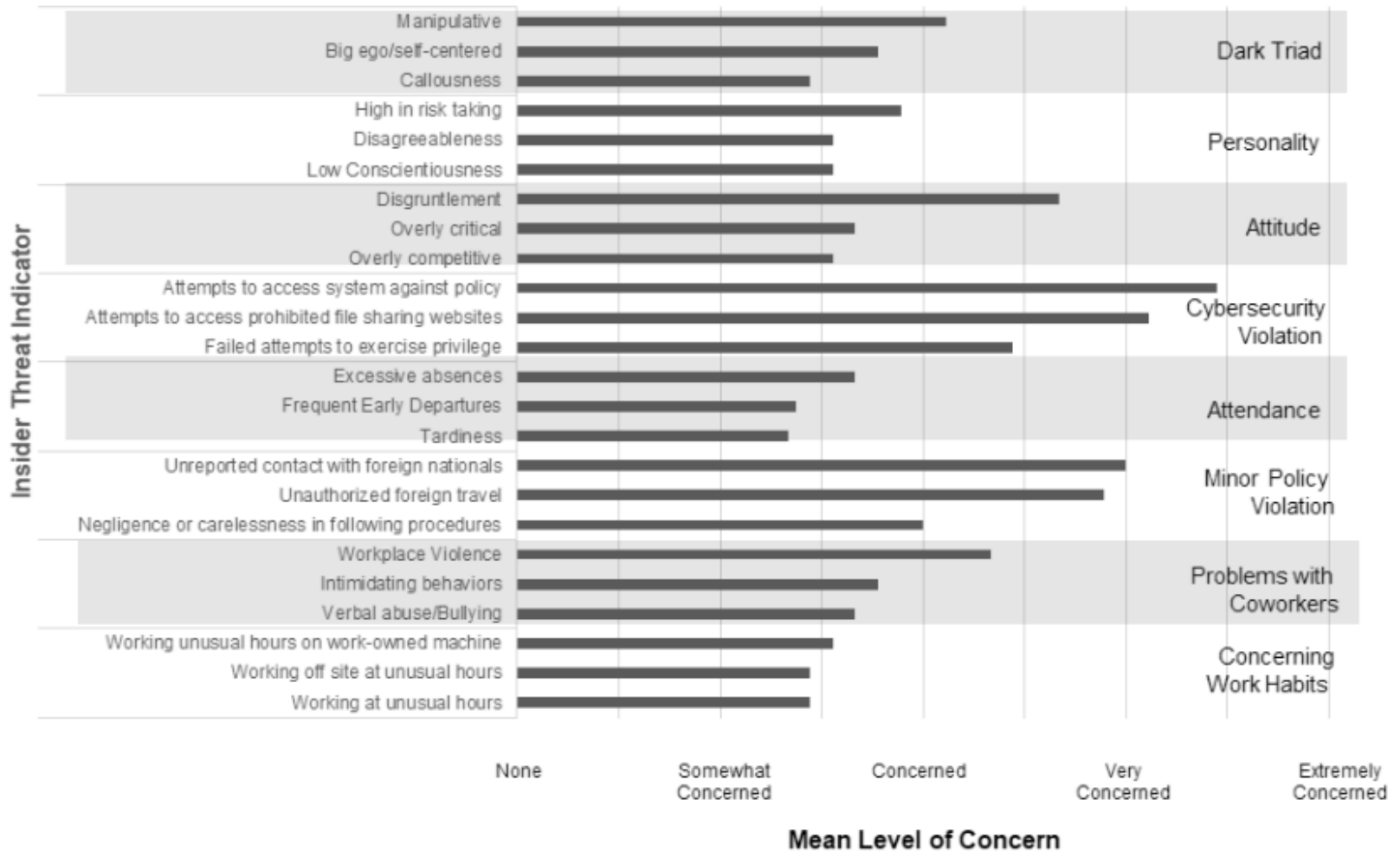
Bin Definitions

- Low Concern:** There is little or no association or potential for insider threat risk
- Low-Moderate Concern:** There is a weak association, potential or inclination/likelihood of an insider exploit
- Moderate Concern:** There is a moderate association, potential or inclination/likelihood of an insider exploit
- Moderate-High Concern:** There is a strong association, potential or inclination/likelihood of an insider exploit
- Extreme Concern:** There is extreme or grave concern about an actual exploit or strong inclination/likelihood of an insider exploit

[Save and Exit] [Finish]

Initial Results – Study 1

Individual Indicator Mean Concern/Risk Scores



Preliminary Results from Study 1

Amount of variance accounted for by alternative models in predicting ranking of cases used in Study 1

Model	R^2	Study2
(1) Counting Model	0.12	0.26
(2) Sum of Risks Model	0.55	0.48
(3) Linear Regression Weight Model	0.85	0.62
(4) Sequential Weighted Model	0.68	0.45

- The simple counting model is clearly inadequate
- The Linear regression model empirically derives indicator weights from the ranking data and therefore represents an optimal (though data-intensive) prediction of the data
- The Sequential weighted model performed reasonably well in Study 1, but not as well in Study 2. Given these results, and the comparative simplicity of the Sum of Risks model, the latest results tend to provide the greatest support for the Sum of Risks model. (We are currently exploring other variations of models).

[Study 2 data are still being analyzed]

Summary & Contributions

- The SOFIT knowledge representation substantially advances the specification of human/social/behavioral and organizational indicators of insider threat.
- The knowledge base is shareable to facilitate reuse and collaboration with the research community.
- The SOFIT ontology can serve as a foundation for assessments of an organization's insider threat mitigation program, and thus can help to inform the technology maturation assessment of existing programs and approaches, identifying gaps in coverage that would be the most productive areas for improvement.

This paper received the Best Paper Award!



**Thank You for
your Attention**

Contact:

Frank L. Greitzer, *PsyberAnalytix*
Frank@PsyberAnalytix.com

The research reported
here was supported
by IARPA contract
2016-16031400006.



The content is solely the responsibility of the author and does not necessarily represent the official views of the U.S. Government.