Developing an Ontology for Individual and Organizational Sociotechnical Indicators of Insider Threat Risk

Frank L. Greitzer[1], Muhammad Imran[2], Justin Purl[3], Elise T. Axelrad[4], Yung Mei Leong[5], D.E. (Sunny) Becker[3], Kathryn B. Laskey[2], and Paul J. Sticha[3]

[1]PsyberAnalytix, Richland WA, USA
[2]George Mason University, Fairfax, VA, USA
[3]Human Resources Research Organization, Alexandria, VA, USA
[4]Innovative Decisions, Inc., Vienna, VA, USA
[5]Independent Consultant, Hyattsville, MD, USA

Frank@PsyberAnalytix.com, mimran4@gmu.edu, JPurl@humrro.org, eaxelrad@innovativedecisions.com, y.leong03@gmail.com, sbecker@humrro.org, klaskey@gmu.edu, psticha@humrro.org

***Abstract*—Human behavioral factors are fundamental to understanding, detecting and mitigating insider threats, but to date insufficiently represented in a formal ontology. We report on the design and development of an ontology that emphasizes individual and organizational sociotechnical factors, and incorporates technical indicators from previous work. We compare our ontology with previous research and describe use cases to demonstrate how the ontology may be applied. Our work advances current efforts toward development of a comprehensive knowledge base to support advanced reasoning for insider threat mitigation.**

Keywords— insider threat; sociotechnical indicators ontology; domain knowledge representation; SME knowledge modeling; human behavioral modeling; domain knowledge sharing

I. INTRODUCTION

Government and corporate organizations alike recognize the serious threat posed by insiders who seek to destroy, steal or leak confidential information, or act in ways that expose the organization to outside attacks. A widely accepted definition of the insider threat is "a current or former employee, contractor, or other business partner who has or had authorized access to an organization's network, system, or data and who intentionally (or unintentionally) exceeds or misuses that access to negatively affect the confidentiality, integrity, or availability of the organization's information or information systems" [1]. More generally, the insider threat may be defined in terms of internal risks to physical and human assets as well as organizational information. In light of recent government initiatives, Executive Order 13587 [2], and the National Insider Threat Policy that specifies minimum standards for establishing an insider threat program, there is increasing acknowledgment of the need to develop formal frameworks to represent and analyze vast amounts of data that may be collected by insider threat monitoring and mitigation systems. There is a notable lack of standards within the insider threat domain to assist in developing, describing, testing, and sharing techniques and methods for detecting and preventing insider threats [3]. The present research is directed toward a systematic and comprehensive representation of concepts in the insider threat domain that will support reasoning and threat assessment models.

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.

II. BACKGROUND

Research on insider threat has sought to develop models and tools to identify individuals who pose increased insider threat risk. Most mitigation approaches focus more narrowly on (a) detecting unauthorized user activity and anomalous activity that may be malicious; and (b) preventing data exfiltration. Typical approaches attempt to prevent unauthorized access through the use of firewalls, passwords, and encryption. That is, they are primarily based on the tools and technology used to thwart external attacks. Unfortunately, these security measures will not prevent authorized access by an insider.

Because a key element of insider threat is a "trusted" perpetrator with authorized access to organizational assets, monitoring and analysis approaches should not only address suspicious host/network activities (identifying so-called technical indicators) but also seek to identify broader aspects of human behavior, motivation, and intent that may characterize malicious insider threats. Thus, as noted in [4], approaches should seek to identify attack-related behaviors that include deliberate markers, preparatory behaviors, correlated usage patterns, and even verbal behavior and personality traits, all of which can be pieced together to detect potential insider threats. While a number of researchers [5-9] recommend including behavioral indicators that may be accessible to organizations prior to an attack, tools and methods that incorporate formal representations of these human behavioral factors are rare (exceptions are models described in [10-12]). The research and operational security communities require a comprehensive knowledge base of technical and behavioral indicators to stimulate the development of more effective insider threat mitigation systems. Existing ontologies include a knowledge

base for technical indicators of insider threat [3][13] and a human factors oriented ontology for cybersecurity risk [14]; our work extends [13] and complements [14] by further specifying individual human and organizational sociotechnical factors.

III. OBJECTIVES

The objective of this research is to develop a formal representation of our current understanding of factors underlying insider threats, particularly relating to individual behavioral and psychological indicators and constructs reflecting organizational factors. The work to date complements and extends extant insider threat ontology frameworks. First, it adds substantial detail (depth) to existing insider threat ontology frameworks that focus on cyber/technical constructs. Second, it defines formal ontological representations of individual and organizational sociotechnical constructs, which are insufficiently represented in current ontological frameworks. The use of a formal, standardized language (ontology) for expressing knowledge about the insider threat domain facilitates information sharing across the insider threat research community and supports model development. A longer term goal is to inform the development of ontology-based reasoning systems and models to support insider threat detection and mitigation. Adopting and using more comprehensive, formal ontological representations will also facilitate the systematic construction of scenarios that may be used in exercising and validating insider threat detection models.

IV. APPROACH

Our approach consisted of (a) developing a hierarchical taxonomy for insider threat risk that can be applied generally to all types of organizations; and (b) migrating the taxonomy into a formal ontology for insider threat risk. Care was taken to compare our representation with existing frameworks (particularly the ontology developed by Carnegie Mellon University's Computer Emergency Response Team CERT [13]) to maximize consistency and interoperability among formulations across the research community. Our approach to ontology development seeks to extend the ontological framework by incorporating probabilistic methods to express and reason with uncertainty, i.e., this work will inform the development of a probabilistic ontology to support reasoning about insider threat risk.

A. Taxonomy Development

A well-defined taxonomy provides an initial hierarchy of domain concepts as a starting point for our insider threat ontology. The taxonomy is based on a systematic review, analysis and synthesis of existing research, case studies and guidelines that have been produced by the insider threat research community. Continually being expanded at the leaf nodes, the current taxonomy is 6-7 levels deep. There are 262 unique factors (leaf nodes) defined across the entire taxonomy: a total of 223 constructs defined for the individual factors and 39 for the organizational factors. Our class structure overall contains more than 350 constructs.

At the highest level we distinguish individual human factors from organizational factors. Individual human factors reflect behaviors, attitudes, personal issues, sociocultural or ideological factors, and various biographical factors that may indicate increased risk. The individual level also differentiates psychological traits from dynamic states, consistent with findings that these two constructs are reliably distinct despite their admitted overlap (e.g., [15-16]) and with the diverse body of psychological research that hinges on (e.g., [17-19]) or capitalizes on (e.g., [20-21]) that distinction. This detailed branch of the taxonomy reflects a substantial body of work by a diverse set of researchers and practitioners focusing on psychosocial factors underlying insider threats (e.g., [5], [7-9], [22-33]). The constructs that comprise this branch are listed in Table I, which shows the main factors (or classes) in column 1 and sub-classes (in *italics*) in column 2. Column 2 also includes illustrative descriptions or instances that reflect lower-level constructs (not exhaustive). In column 1 we also indicate a count of the total number of constructs defined at the leaf node level for each class, to provide a sense of the extensiveness of the taxonomy.

TABLE I. CONSTRUCTS COMPRISING INDIVIDUAL HUMAN FACTORS

Class[(a)]	Sub-Class and Instances
Concerning Behaviors *(140)*	*Boundary Violation* -- Concerning work habits, attendance issues, blurred personal/professional boundaries, threatening/intimidating behaviors, boundary probing, social engineering, minor policy violations, travel policy violations, unauthorized travel, unauthorized foreign travel, change in pattern of foreign travel, security violations
	Job Performance – Cyberloafing, negative evaluation
	Technical/Cyber Violation – Concerns about: authentication/ authorization, data access patterns, network patterns, data transfer patterns, command usage, data deletion/modification, suspicious communications
Life Narrative *(34)*	*Criminal Record* – Court records
	Financial Concerns – Lifestyle incongruities (unexplained affluence, etc.), risky financial profile (bankruptcy, large expenses-to-income ratio, bounced/bad checks, credit problems)
	Personal History – Demographics, employment, education background, major life events, health status, marital history, U.S. Immigration/citizenship status
Ideology *(9)*	*Disloyalty* – Behaviors or expressions of disloyalty to the organization or to the U.S. government *[2, 6]*
	Radical Beliefs – Radical political beliefs, radical religious
	Unusual Contact with Foreign Entity –Unreported contact with foreign nationals
Dynamic State *(14)*	*Affect* – Excessive anger/hostility, disengagement, mood swings
	Attitude – Lack of motivation, overly competitive, expresses feelings of disgruntlement with job, overly critical, resentful, defensive
Static Trait *(25)*	*Personality Dimensions* – Neuroticism, disagreeableness, low conscientiousness, excitement seeking, honesty-humility on six-factor personality scale
	Other Personality Traits – Characteristics associated with maliciousness or vulnerability to exploitation (Machiavellianism, narcissism, psychopathy, sadism, authoritarianism, social dominance orientation)
	Temperament – Various temperament issues that may be observed/reported by coworkers – Big ego, callousness, lack of empathy, lack of remorse, manipulativeness, rebelliousness, poor time management, preoccupation with power/grandiosity

[(a)] In parentheses is the total # of sub-classes or instances populated to date within the class

Organizational factors focus on organizational and management practices, policies, and work setting characteristics that influence worker satisfaction, attitudes, safety, or protection/vulnerabilities of assets. These factors have received much attention by organizations that publish best practices—indicating situations or conditions that contribute to an increased likelihood of insider threats within an organization. Although they may play a role in triggering malicious or unintentional insider threats, these factors have not generally been identified in insider threat ontologies to date. This branch of our taxonomy was constructed by consulting the broad and diverse literature on industrial/organizational psychology and human error research, including [34-36] and relevant discussion of these factors in the context of workplace violence and insider threat (e.g., [37-38]). Table II lists classes and sub-classes defined to date for organizational factors.

TABLE II. CONSTRUCTS COMPRISING ORGANIZATIONAL FACTORS

Class[a]	**Sub-Classes**
Security Practices *(14)*	Communication/training
	Policy clarity
	Hiring
	Monitoring
	Organizational justice
	Implementation of Security Controls
Communication Issues *(2)*	Inadequate procedures/directions
	Poor communications
Work Setting (Management Systems) *(6)*	Distractions
	Insufficient resources
	Poor management systems
	Job instability
	Lack of career advancement
	Poor physical work conditions
	Organizational changes
Work Planning and Control *(13)*	Job pressure/job stress
	Time factors/unrealistic time constraints
	Task difficulty
	Change in routine
	Heavy or prolonged workload
	Insufficient workload
	Conflict of work roles
	Work role ambiguity
	Lack of autonomy
	Lack of decision-making power
	Irregular timing of work shifts
	Extended working hours
	Lack of breaks
Mitigating Factors *(4)*	Flexible work schedule
	Employee Assistance Plan
	Effective staff training and awareness
	Reporting mechanism

[a] In parentheses is the total # sub-classes or instances populated to date within the class

B. Ontology Development Approach

To date, insider threat ontology development has focused primarily on technical factors (e.g., [13]). In contrast, our approach is grounded in an extended problem space that includes methods, motivation, psychology, and circumstances of human behavior. As noted by previous authors (e.g., [13]), behavioral aspects of insider threat can be an extraordinarily complex domain to model. There are many overlapping concepts (e.g., state and trait anger), many providing little meaning in isolation (e.g., surfing the web vs. surfing the web instead of working). Our task has been to contextualize behaviors with related concepts (e.g., underlying motivations and personality traits) that allow the cataloging of information pertaining to both the insider threat incident and the insider. Through this catalogue of information, researchers and organizations can index cases and gain further insight into common attack vectors driven by human behavior. Our ontology extends previous work [3][13][14] in two ways: (a) adding more detail to the technical indicator branch of the ontology and (b) adding material focusing on individual behavioral and organizational factors.

Our approach is to migrate our taxonomy into a formal ontology expressed in the popular OWL-DL ontology language. OWL-DL balances expressiveness (ability to represent many kinds of domain entities and relationships), computational properties (conclusions are guaranteed to be computable in finite time), and functionality for drawing inferences from asserted facts. Enumeration of (potentially hundreds of) *Competency Questions* (*CQs*) for our ontology serves as a requirements specification as well as a means of testing the ontology implementation. An example of a simple *CQ* is "What are the components of class *Attitude*?" A more complex *CQ* is "What factors are associated with the observables *attendance problems*, *unauthorized personal use of work computer*, and *hostile*? The *CQs* may be evaluated using SPARQL queries. Our OWL-DL implementation will enable automated inferences about class relationships. For example, from the assertion that an individual belongs to class *Aggressive* and class *Manipulative*, the reasoning engine can infer that the individual fulfills the membership conditions of class *Threat*.

V. ONTOLOGY IMPLEMENTATION

A. Ontology Methods

Following widely recognized guidelines for ontology development [39], we used the Methontology ontology engineering methodology [40], which enables construction at the conceptual level and allows for development, re-use, or re-engineering of existing ontologies. In the *Specification* phase we defined the purpose of the ontology, its intended uses and its end users. In the *Conceptualization* phase we structured the domain knowledge into meaningful graphical models. In the *Formalization* phase we represented our conceptual models as a formal or semi-computable model. The *Implementation* phase supports the ontology development in the Web Ontology Language (OWL). Updates and corrections take place in the *Maintenance* phase. Our development also included supporting Methontology activities of *Knowledge Acquisition*, *Evaluation* (verification and validation that the ontology represents the domain), *Integration* (reuse of other available ontologies), *Documentation*, and *Configuration management*. We also adopted IDEF5 methods in conceptualization and formalization phases to acquire knowledge and develop graphical knowledge representation models. We implemented our taxonomy using an off-the-shelf ontology development tool (Protégé).

By default, the Protégé tool does not assume that classes are mutually exclusive. This is useful when concepts are most meaningful in combination. For example, high absenteeism, a weak indicator by itself, is made stronger in association with

other concerning factors [32], but the risk is mitigated when associated with documented illness, vacation or maternity leave. As another example, relaxation of the assumption of mutual exclusivity is especially useful when considering various correlated psychological or personality characteristics such as those defined in the Five Factor Model (FFM) of personality traits [41]. There are numerous well-supported relationships between dimensions of personality and various types of counterproductive work behavior [28].

B. Description of the Ontology Classes

We began by formalizing the hierarchy of concepts provided by the taxonomy discussed in Section IV-A, and translating the hierarchy into parent-child relationships of classes in our ontology. Classes represent objects with similar structure and properties Classes are arranged hierarchically; those without further subcategories are termed leaf nodes. Individuals in the ontology represent instances of classes. Class relationships other than parent-child are derived from the research literature, available material on insider threat cases, and the experience and judgment of subject-matter experts within the development team. As reuse of previous knowledge models is a key advantage of ontologies and an encouraged practice in ontology engineering, we included as much information from previous work as possible, especially the recent ontology developed by CERT [3][13]. In particular, the *Actor*, *Asset*, *Action*, *Event*, *Temporal Thing* and *Information* class structures are adopted in total. Selected classes from the Unified Cyber Security Ontology [42] were also incorporated into our ontology. For example the idea of "Consequence" class is adopted by our ontology but renamed to *Outcome* class since this terminology is more consistent with the insider threat cases scenario template used by CERT. The concepts of *Vulnerability* (e.g., [6]) and *Catalyst/Trigger* events (e.g., [43-44]) are also formalized as classes in our ontology. To capture the temporal information involved in insider threat cases, we imported the *Temporal Interval* class from the CERT ontology.

Figs. 1-3 show the hierarchy of classes in our ontology, as implemented in the Protégé tool. The ontology is derived from the extensive taxonomy described in Section IV-A. Due to space constraints we depict only selected classes with detail restricted to the 4th level of the hierarchy. A comparison of Tables I and II with Figs. 1-3, shows how the class hierarchy in the ontology represents the organization of domain concepts in the taxonomy. Fig. 1 shows how the ontology accounts for both malicious and non-malicious (unintentional) insider threats. Importantly, we distinguish between actions performed by employees (as insiders) and actions performed by organizations (which may, for example, include poor institutional policies and/or security practices as well as inadequate or exacerbating responses to potential threats). At the same root level we also include classes such as *Industry*, *Insider Threat Risk*, *Effect*, *Location* and *Outcome* as attributes of the organization. *Industry* may account for differences in organizational rules,

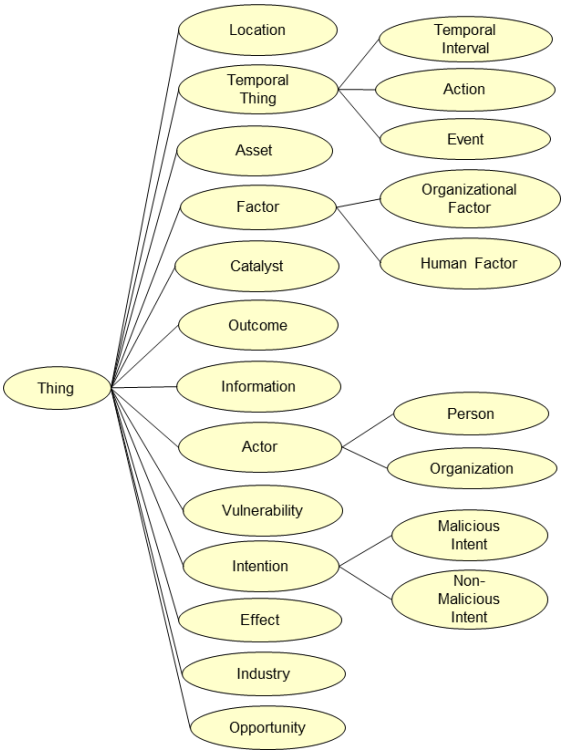

Fig. 1. Top-Level Classes

regulations and policies that differ across industry sectors. The *Effect* class captures information about the impact of the insider criminal activity on the organization(s), for example the action of injecting a virus into an enterprise network can induce a malfunction in other workstations on the network and/or a full network shutdown. The concept of the consequences of an attack is captured by the *Outcome* class, for example the shutdown of the network has an outcome of a halt of organization's operations and thousands of dollars of loss. The *Location* class encapsulates geographic information about the source of an attack. The *Insider Threat Risk* class captures the threat level that would be associated with the individuals of the *Actor* class based on the inference performed over the ontology.

Fig. 2 expands the *Human Factor* node of Fig. 1, and Fig. 3 expands the *Organizational Factor* node. Inspection of the human psychosocial factors in Fig. 2 reveal classes (and associated sub-classes) that correspond to elements of the taxonomy. Acknowledging the Capability-Motive-Opportunity (CMO) model (e.g., [4]), which postulates that the perpetrator of an attack must have the capability, motive, and opportunity to commit the attack, we include these constructs as classes in the ontology. Full implementation of CMO constructs is deferred for future efforts to define relationships among these classes.

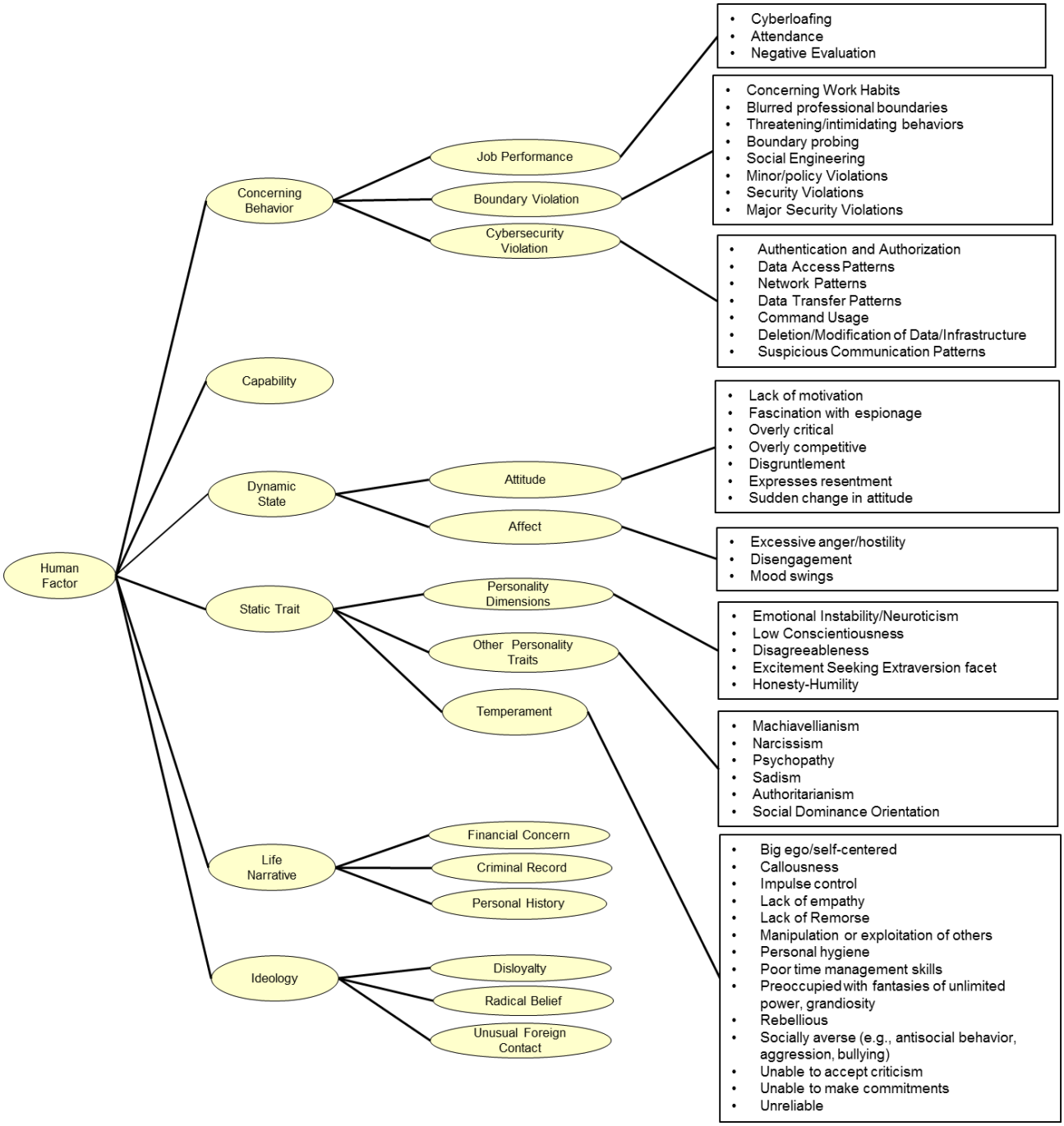

Fig. 2. *Human Factor* Classes (Lower level details for *Life Narrative* and *Ideology* classes are not shown due to space constraints)

The capability to conduct an attack is in part dependent on an individual's knowledge/skills/abilities that are represented in certain human behavioral factors (cf [14]), particularly the *Biographical Data* subclass within the *Life Narrative* factors class. Motive (or motivation) may be represented within the *Intention* class (and its *malicious* or *non-malicious* subclasses) in Fig 1; it is also related to psychological characteristics or predispositions such as *Static Traits*, *Dynamic States*, and *Life-Narrative* factors (e.g., financial or health problems that may act as stressors)—which are sub-classes of the *Human Factor* class (see Fig. 2)—as well as *Organizational Factors* (Fig. 3) that may act as stressors or triggers that can motivate an attack.

The sub-class *Concerning Behavior,* within the *Human Factor* class, contains a large set of individual actions that includes the subcategories *Job Performance, Boundary Violation*, and *Cyber Security Violation*. These in turn are broken down into more granular, lower-level constructs (shown in boxes); not shown are even lower levels of the hierarchy and individuals representing instances of the classes.

The initial structure of the ontology grew out of the detailed taxonomic structure that we developed based on subject-matter expertise and our analysis/synthesis of research literature and numerous case studies. A more robust and richer representation has been informed by exploring complex relationships among constructs (e.g., classes, sub-classes, instances) spread across multiple branches of the hierarchy. As a simple example, the ontology recognizes that different types of attack are identified from their relationships with certain aspects of the cyber/technical exploit (e.g., exfiltration requires certain actions performed on sensitive information, such as saving to external media, printing, emailing, uploading to the cloud, etc.). A more complex example may be considered in using the CMO model (mentioned above) to reason about insider risk. By incorporating knowledge of relationships among detected behaviors, individual behavioral factors, and organizational factors, the ontology allows reasoning about the risk associated

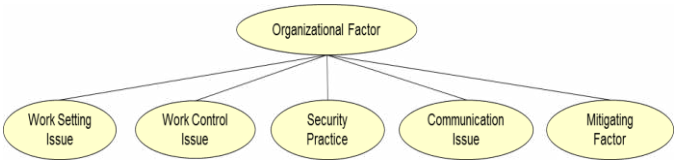

Fig. 3. *Organizational Factor* Class

with detected behaviors in the context of possible motives, capabilities, and opportunity. Relationships and gaps (missing elements in classes) were further identified by exercising the knowledge base using known or fictitious use cases.

C. Use Case and Application

Use cases help to verify the comprehensiveness of the knowledge representation and to identify missing or ill-defined classes and relationships. In this section, we demonstrate the application of the ontology to use cases that include human behavioral factors and organization factors as well as cyber/technical indicators. In the scenarios described, we use [brackets] to identify significant indicators with actions described in the scenario.

Use Case #1 (see small text box) describes a simple cyber-related insider threat incident. Use Case #2 (see large text box), which entirely subsumes the contextual and technical information regarding the insider threat incident described in the first use case, injects additional human behavioral factors.

Use Case #1

John **[PERSON: Insider X]** *is a long-time system administrator* **[LIFE NARRATIVE: PERS HISTORY] [CAPABILITY]** *with access to sensitive and classified information* **[OPPORTUNITY]** *in a company that performs government-sponsored R&D* **[ORGANIZATION: VICTIM ORGANIZATION]**.

John uses his personal web-based email account from his work computer to communicate with prospective employers **[DIGITAL ACTION: EMAIL ACTION]**. *Then he uses his administrative privileges to access some sensitive intellectual property information* **[BUSINESS INFORMATION: INTELLECTUAL PROPERTY]** *that will be of interest to a competitor. John saves these files to his computer* **[COMPUTER ASSET: WORK PC]** *and copies the files to a thumb drive* **[CONCERNING BEHAVIOR: TECH/CYBER VIOLATION–DIGITAL ACTION/COPY ACTION] [PHYSICAL ASSET: USB DRIVE]**, *which he then sneaks out of the office with the intention of using the information to leverage a job offer with a competitor* **[THEFT EVENT: DATA THEFT]**. *Subsequently John resigns and accepts a job offer from a competitor.*

It is evident that Use Case #1 lacks substantial contextual information described in Use Case #2 regarding possible contributing or mitigating factors, relevant personal predispositions, or concerning behaviors that may be associated with this individual's insider threat risk. Fig. 4 is a concept map depicting Use Case #2, showing all the behavioral and technical concepts and their associated relations. The dashed

Use Case #2

John **[PERSON: Insider X]** *is a long-time system administrator* **[LIFE NARRATIVE: PERS HISTORY] [CAPABILITY]** *with access to sensitive and classified information* **[OPPORTUNITY]** *in a company that performs government-sponsored R&D* **[ORGANIZATION: VICTIM ORGANIZATION]**. *The following input was recorded in his personnel file: (1) One colleague states that John discounts the opinions of colleagues and he becomes hostile when colleagues discuss and critique his ideas* **[STATIC TRAIT: TEMPERAMENT: RESISTS CRITICISM] [DYNAMIC STATE: AFFECT—HOSTILE].** *(2) A different colleague states that John seeks to control all aspects of a project and often insists on dominating the conversation about project tasks and approach* **[STATIC TRAIT: OTHER PERSONALITY DIMENSIONS—AUTHORITARIANISM]**. *(3) His manager corroborates these inputs and adds that John tends to become argumentative and irritated, and defensively cites his superior knowledge of industry best practices when others criticize his rigid protocols* **[DYNAMIC STATE: AFFECT–HOSTILE] [STATIC TRAIT: TEMPERAMENT—BIG EGO]**. *Staff development/performance review assessment includes criticism by colleagues that portions of his protocols are idiosyncratic with weak rationale, and that his rigid protocols have impacted company projects* **[CONCERNING BEHAVIORS: JOB PERF—NEGATIVE PERF EVALUATION].**

John was passed over for a promotion to manage a new, prestigious project **[LIFE NARRATIVE: PERS HISTORY: EMPLOYMENT–PASSED OVER FOR PROMOTION].** *He files a complaint with HR claiming unfair treatment and his manager, compelled to meet with him, comes away with the impression that John still harbors resentment over not being promoted. John's most recent evaluation cited a decline in performance* **[CONCERNING BEHAVIORS: JOB PERF—NEGATIVE PERF EVALUATION]**; *since being denied the promotion his attitude has been increasingly disgruntled* **[DYNAMIC STATE: ATTITUDE—DISGRUNTLEMENT]**; *and that there were multiple complaints from coworkers about frequent tardiness* **[CONCERNING BEHAVIORS: BOUNDARY VIOLATION—ATTENDANCE]**. The attendance problem led to *a formal, written warning* **[CONCERNING BEHAVIORS: BOUNDARY VIOLATION–POLICY VIOLATION].** *After getting the warning, John talks to his manager and loses his cool—storming out of the office* **[DYNAMIC STATE: AFFECT–HOSTILE].** *A colleague hears John's outburst and tells the manager about John's recent marital separation to provide some context to Johns behavior* **[LIFE NARRATIVE: PERS HISTORY—MAJOR LIFE EVENTS/RECENT CHANGE IN MARITAL STATUS (MARITAL SEPARATION)**]. *The incident prompts the manager to contact the company Security Office. The Security Office checks the local court records to learn that three weeks ago, John was arrested for allegedly driving under the influence (his first contact with the criminal justice system)* **[LIFE NARRATIVE: CRIMINAL RECORD—DUI]**.

Faced with these job and personal stressors, John begins to seek work with a competitor. John contacts a competitor to see if they are interested in him and in proprietary information he can provide. To avoid being noticed, John carries out email dialogue with the competitor by logging into his personal Yahoo web mail account from his work computer **[CONCERNING BEHAVIORS: JOB PERFORMANCE—CYBERLOAFING]**. *Next, John carries out the insider threat attack and resigns, as described in second paragraph of Use Case #1.*

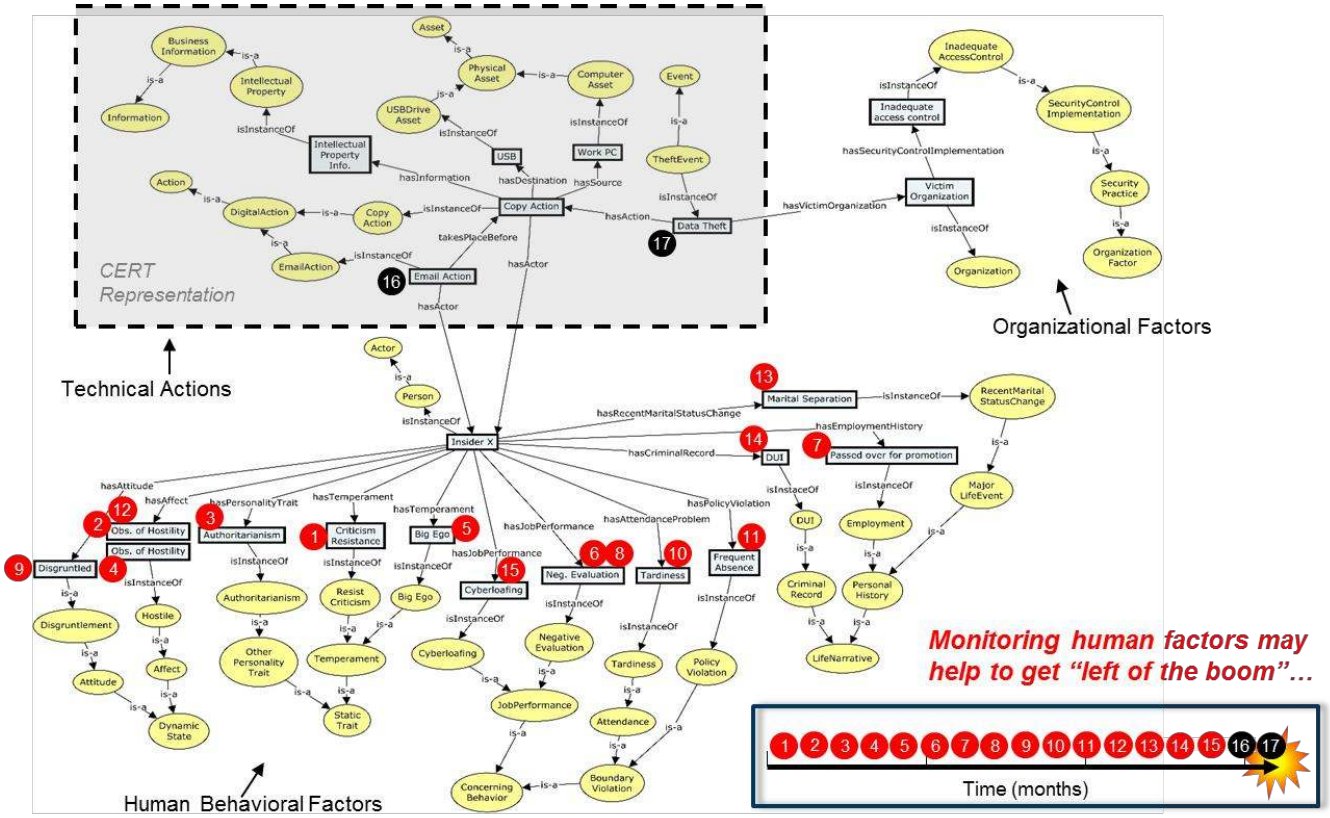

Fig. 4. Concept map representation of Use Case #2 (Case #1 is within dashed box). Not all concepts and relations are shown due to space limitations.

box in Fig. 4 represents Use Case #1 (due to space limitations, not all details are shown). In a real scenario, detecting concerning behaviors or other factors may require multiple factors to meet threshold requirements for alerts—these are not described or represented here due to space constraints. Events depicted in the use case scenarios are numbered chronologically. Shown in the lower right side of the figure is a timeline (spanning several months for illustrative purposes) suggesting that monitoring of sociotechnical factors may help achieve proactive mitigation goals (getting "left of the boom").

VI. Comparison with Related Work

The focus of our effort is to express and represent individual and organizational sociotechnical factors in an ontological characterization of insider threat risk (e.g., [13-14]). Our ontology provides a more robust, richer description of not only the nature of the attack but also possible contributing factors that more fully describe the insider threat to the organization. CERT [13] began with a database of insider threat case descriptions. The information framework underlying this database informed the vocabulary in the ontology. Namely, organizations grant access to persons that perpetrate events that harm the organization. Persons and Organizations are the actors in the CERT ontology, and their actions culminate in events (i.e., insider threat incidents). Instead of a focus on events, our ontology focuses on the insider. Our taxonomy and ontology are based on theories and models of insider threat in the literature that incorporate human behavioral as well as technical indicators of threat (e.g., [10-12]). While the current CERT ontology only describes technical/cyber events, our ontology also includes non-technical or sociotechnical constructs that reflect actions and psychosocial indicators of persons of interest. As a specific example, consider the class *Concerning Behaviors*. A concerning behavior such as "Leaving a classified security container unlocked and unattended" can be described using two concepts in the CERT ontology: an Asset (e.g., Classified file) and an Action (e.g., Unlock). However, this may not be the focal event, or a precipitating event, in a case description, and there may be other related contributing factors. For example, a previous condition (e.g., organizational reduction in force/layoffs) or individual predispositions (e.g., personality traits, personal stress) may lead to actions that reflect a lack of diligence or motivation in an actor who later commits an act of insider threat (these contributing factors are in part identified in the cybersecurity human factors ontology (HUFO) by [14]. The CERT ontology, in particular, does not connect these behavioral constructs to technical/cyber actions that comprise the actual exploit.

At a basic level, the *Factor* class, which contains much of the vocabulary in our ontology, can be placed alongside *Assets* in the CERT ontology. Both are non-temporal classes that a person can possess (i.e., *Things*). We integrated the two ontologies and eliminated duplications. All CERT ontology classes were incorporated in this way. There are, however, stark differences between the extent and scope of the CERT ontology and our ontology. The CERT ontology contains a standardized and well-defined vocabulary for describing the actions of insider threats. It contains 31 actions (e.g., Copy), along with six action modifiers (e.g., Suspicious), organized under four major

classes to describe digital, financial, and job-related insider threat behavior. These actions can be taken on 26 assets (e.g., USB drive) in three major categories (i.e., Physical, Financial, and Digital) and/or 16 types of information (e.g., Password) organized in seven major categories (i.e., National Security, Technology, Financial, Medical, Classified, Business, and Uniquely Identifiable). Eleven focal events are also captured as classes in the ontology (e.g., Theft), for a total of 125 constructs within their class structure. In contrast to the CERT ontology, our framework is broader and deeper. In addition to containing these constructs, our ontology represents a knowledge base that is six to seven layers deep, comprising a total of over 350 constructs. In sum, we have greatly expanded the CERT ontology by adding classes representing human behavioral and organizational factors of insider threat.

While not specifically addressing insider threat, the cybersecurity HUFO presented by [14], which focuses on trust, is similar to and largely compatible with our ontology; it defines roughly 48 human factors classes that address characteristics such as motivation, integrity, rationality, benevolence, personality, ideology, ethics, and risk posture, as well as knowledge, skills and abilities. In comparison, our ontology probes several levels deeper than the HUFO ontology. Further work is planned to integrate relevant features of these ontologies.

VII. Conclusions and Future Work

Our work addresses two major challenges. First, due to the large number of concepts and their complex interrelationships, the insider threat domain is cumbersome to model. Second, there is a need to establish a common terminology and shared understanding of the complex insider threat domain. We used an exhaustive approach that incorporates into our taxonomy most of the concepts we have encountered in the insider threat literature. We then developed a mapping that transforms the taxonomy into an ontology, and added relationships to the ontology to produce a formal representation of concepts and their interrelationships. By synthesizing the contributions of a diverse set of experts, we developed a knowledge representation that more fully characterizes insider threat indicators—from the perspective of human behavior as well as cyber/technical indicators—and that can be made available in a shareable knowledge base to facilitate reuse and collaboration.

Beyond its immediate use in providing a common, shareable knowledge base of insider threat problem space constructs, the present research will help to advance efforts to model and mitigate insider threats. Informed by extant research on human and organizational factors associated with insider threats, the constructs and indicators represented in the present ontology can be used to develop models to assess individual risk and organizational vulnerability, as well as to inform operational risk management practices. In addition, by specifying a more comprehensive knowledge base, our ontology facilitates the generation of diverse scenarios for use in red teaming and testing of more holistic insider threat models. Finally, the knowledge base provided here may have further operational impact by informing the structure of data to be captured by enterprises for effective insider threat monitoring and analysis.

A brief discussion of some limitations of the research reported here may be useful in interpreting progress to date as well as motivating future work. First, our choice to define a taxonomy as a foundation for the ontology meant that the initial structure only specified hierarchical parent-child relationships among constructs. Other relationships were then defined as part of the process of transforming the taxonomy into an ontology. Because our primary interest (and recognized need in modeling insider threats) was to incorporate sociotechnical factors that have been suggested in research literature, there was also an inherent limitation in the ability to specify robust axioms that reflect more complex relationships among constructs. Ultimately this more complete specification will be required to support inferences about classes and individuals. There is a tradeoff between implementing the asserted classes and individuals versus the inferred constructs. While some of the classes in our ontology are defined by certain inference rules and axioms (e.g., the class *Capability* categorizes instances based on specified rules), much more work is needed to more fully specify relationships that will ultimately be required to support inferences about insider threat risks. A second limitation is that, while the current ontology has captured salient constructs in the literature, there are certainly more constructs that can and should be added to the ontology. Research should continue the process of encapsulating the entirety of constructs related to insider threat. We are continually populating the individual and organizational classes of ontology with relevant instances (informed by use cases); we plan to further develop the *Capabilities* and *Opportunities* classes and associated relationships, building upon recent related work [14]. Future research should also focus on addressing the need to represent temporal relationships among constructs.

We use the present forum and others to share these results with the research community. We also plan to extend our ontology into a probabilistic ontology by incorporating information about uncertainty in the insider threat domain. The resulting probabilistic ontology will support reasoning under uncertainty [45]. Probabilistic ontologies combine semantically rich representations that support interoperability and automated reasoning with mathematically well-founded uncertainty management. Advancing research and development of probabilistic ontologies for insider threats will facilitate modeling and tool development. Our ontology provides a rich foundation for logical and probabilistic inferences necessary for protection against insider attacks.

References

[1] D. M. Cappelli, A. P. Moore, and R. F. Trzeciak, *The CERT guide to insider threats: How to prevent, detect, and respond to information technology crimes (theft, sabotage, fraud)*. Addison-Wesley, 2012.

[2] The White House. Executive Order 13587—Structural Reforms to Improve the Security of Classified Networks and the Responsible Sharing and Safeguarding of Classified Information, October 2011. http://www.whitehouse.gov/the-press-office/2011/10/07/executive-order-structural-reforms-improve-security-classified-networks-

[3] D. L. Costa, M. Collins, J. S. Perl, J. M. Albrethsen, J.G. Silowash, and D. Spooner. (2014). An Ontology for Insider Threat Indicators. In K. B. Laskey, I. Emmons and P C.G. Costa (Eds.), *Proceedings of the Ninth Conference on Semantic Technologies for Intelligence, Defense, and Security* (STIDS 2014), 2014, 48–53.

[4] E. E. Schultz, "A framework for understanding and predicting insider attacks." *Computers & Security*, 2002, vol. 21, 526–531.

[5] E. D. Shaw, J. M. Post, and K. G. Ruby, "Inside the mind of the insider." *Security Management*, 1999, vol 43 (12), 34-42.

[6] M. R. Randazzo, M. M. Keeney, E. F. Kowalski, D. M. Cappelli, and A. P. Moore. I*nsider threat study: illicit cyber activity in the banking and financial sector*. Carnegie-Mellon University. Software Engineering Institute. CMU/SEI-2004-TR-021, 2012.

[7] S. R. Band, D. M. Cappelli, L. F. Fischer, A. P. Moore, E. D. Shaw, and R. F. Trzeciak. *Comparing insider IT sabotage and espionage: a model-based analysis*. Carnegie-Mellon University. Software Engineering Institute. CERT Coordination Center. CMU/SEI-2006-TR-026, 2006.

[8] F. L. Greitzer, L. J. Kangas, C. F. Noonan, C. R. Brown, and T. Ferryman. Psychosocial modeling of insider threat risk based on behavioral and word use analysis. *e-Service Journal*, 2013, 9(1), 106-138. http://www.jstor.org/stable/10.2979/eservicej.9.1.106

[9] M. Maasberg, J. Warren, and N. L.Beebe. The dark side of the insider: Detecting the insider threat through examination of dark triad personality traits. IEEE. *48th Hawaii International Conference on System Sciences,* 2015*,* 3518-3526. DOI 10.1109/HICSS.2015.423

[10] F. L. Greitzer and R. E. Hohimer. "Modeling Human Behavior to Anticipate Insider Attacks." *Journal of Strategic Security,* 2011*, 4(2)*:25-48. http://scholarcommons.usf.edu/jss/vol4/iss2/

[11] R. E. Hohimer, F. L. Greitzer, C. F. Noonan, and J. D. Strasburg. "CHAMPION: Intelligent Hierarchical Reasoning Agents for Enhanced Decision Support." In *Semantic Technology for Intelligence, Defense, and Security (STIDS 2011*). 2011, 36-43

[12] E. T. Axelrad, P. J. Sticha, O. Brdiczka, and J. Shen, "A Bayesian network model for predicting insider threats." IEEE SPW *Workshop on Research for Insider Threat (WRIT), San Francisco, CA,* 2013, 82-89.

[13] D. L.Costa, M. J. Albrethsen, M. L. Collins, S. J. Perl, G. J. Silowash, and D. L. Spooner. *An Insider Threat Indicator Ontology*. TECHNICAL REPORT CMU/SEI-2016-TR-007. Pittsburgh, PA: SEI, 2016.

[14] A.Oltramari, D. H. Henshel, M. Cains, and B. Hoffman. "Towards a human factors ontology for cyber security." In *Semantic Technology for Intelligence, Defense, and Security (STIDS 2015*). 2015, 26-33.

[15] W. E. Chaplin, O. P. John, and L. R. Goldberg. "Conceptions of states and traits: Dimensional attributes with ideals as prototypes." *Journal of Personality and Social Psychology*, 1988, *54(4)*, 541-557.

[16] R. Steyer, A. Mayer, C. Geiser, and D. A. Cole. "A theory of states and traits—Revised." *Annual Review of Clinical Psychology*, 2015, *11*, 71-98.

[17] S. C. Roesch, A. A. Aldridge, S. N. Stocking, F. Villodas, Q. Leung, C. E. Bartley, and L. J. Black. "Multilevel factor analysis and structural equation modeling of daily diary coping data: Modeling trait and state variation. *Multivariate Behavioral Research,* 2010*, 45(5),* 767-789.

[18] L. Van Gelder and R. E. De Vries. "Traits and states at work: Lure, risk and personality as predictors of occupational crime." *Psychology, Crime & Law*, 2016, *22(7)*, 701-720. DOI 10.1080/1068316X.2016. 1174863

[19] D. F. Gro¨s, L. J. Simms, M. M. Antony, and R. E. McCabe. "Psychometric properties of the State–Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State–Trait Anxiety Inventory (STAI)." *Psychological Assessment*. 2007, *19(4)*, 369–381.

[20] K. S. Douglas, S. D. Hart, C. D. Webster, and H. Belfrage. *HCR-20V3: Assessing risk of violence – User guide.* 2013. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

[21] J. R. Meloy, S. G. White, and S. Hart. "Workplace assessment of targeted violence risk: The development and reliability of the WAVR-21". *Journal of Forensic Sciences*, 2013, *58(5)*, 1353-1358.

[22] E. D. Shaw and L. F. Fischer. Ten Tales of Betrayal: The Threat to Corporate Infrastructures by Information Technology Insiders. Report 1—Overview and General Observations. Technical Report 05-04, April 2005. Monterey, CA: Defense Personnel Security Research Center.

[23] M. Gelles, M. Exploring the mind of the spy. In Online *Employees' Guide to Security Responsibilities: Treason 101*. 2005. Retrieved from Texas A&M University Research Foundation website: http://www.dss.mil/search-dir/training/csg/security/Treason/Mind.htm

[24] J. L. Krofcheck and M. G. Gelles. *Behavioral Consultation in Personnel Security: Training and Reference Manual for Personnel Security Professionals.* Yarrow and Associates, 2005.

[25] D. Bulling, M. Scalora, R Borum, J Panuzio, and A Donica. *Behavioral science guidelines for assessing insider threats*. Publications of the University of Nebraska Public Policy Center. Paper 37. 2008. http://digitalcommons.unl.edu/publicpolicypublications/37

[26] D. B. Parker. *Fighting computer crime: A new framework for protecting information*. New York, NY: John Wiley & Sons, Inc., 1998.

[27] F. L. Greitzer, A. P. Moore, D. M. Cappelli, D. H. Andrews, L. A. Carroll, and T. D. Hull. Combating the insider threat. (2008). *IEEE Security & Privacy*, January/February 2008, 61-64.

[28] E. D. Shaw, L. F. Fischer, and A. E. Rose. *Insider risk evaluation and audit* (No. TR-09-02). Monterey, CA: Defense Personnel Security Research Center, 2009.

[29] B. Zadeh and F. L. Greitzer. "Motivation and Capability Modeling for Threat Anticipation." *OSD Human Social Culture Behavior (HSCB) Modeling Program Conference.* Chantilly, VA, 5-7 August 2009.

[30] F. L. Greitzer and D. A. Frincke. D.A. "Combining traditional cyber security audit data with psychosocial data: towards predictive modeling for insider threat," in *Insider Threats in Cyber Security*. vol. 49, C. W. Probst*, et al.*, Eds., Springer US, 2010, 85–114.

[31] F. L. Greitzer, L. J. Kangas, C. F. Noonan, and A. Dalton. *Identifying at-risk employees: A behavioral model for predicting potential insider threats.* PNNL-19665, Richland, WA: Pacific NW National Laboratory, 2010. http://www.pnl.gov/main/publications/external/technical_reports/PNNL-19665.pdf.

[32] F. L. Greitzer, L. J. Kangas, C. F. Noonan, A. Dalton, and R. E. Hohimer. "Identifying at-risk employees: a behavioral model for predicting potential insider threats." *Hawaii International Conference on System Sciences*. Maui, HI, Jan 4-7, 2012.

[33] Software Engineering Institute (SEI). *Analytic approaches to detect insider threats*. White Paper, SEI, December 9, 2015. http://resources.sei.cmu.edu/asset_files/WhitePaper/2015_019_001_451069.pdf

[34] S. Dekker. *The field guide to human error investigations*. Burlington, VT: Ashgate, 2002.

[35] D. J. Pond and K. R. Leifheit. "End of an error." *Security Management*, 2003, *47(5)*. 113–117.

[36] D. J. Pond and F. L. Greitzer. "Error-based accidents and security incidents in nuclear materials management." *Institute of Nuclear Materials Management 46th Annual Meeting*, Phoenix, AZ, 2005. http://www.osti.gov/scitech/biblio/966022

[37] R. Baron and J. Neuman. Workplace violence and workplace aggression: Evidence on their relative frequency and potential causes. *Aggressive Behavior*, 1996, vol. 22, no. 3, 161–173.

[38] F. L. Greitzer, J. Strozer, S. Cohen, J. Bergey, J. Cowley, A. Moore, and D. Mundie. "Unintentional insider threat: contributing factors, observables, and mitigation strategies." *47th Hawaii International Conference on Systems Sciences (HICSS-47)*, Big Island, Hawaii, 2014.

[39] N. F. Noy and D. L. McGuinness. *Ontology Development 101: A Guide to Creating Your First Ontology.* (SMI-2001-0880 (also available as KSL Technical Report KSL-01-05)) 2001

[40] M. Fernández-López, and A. Gómez-Pérez. Overview and analysis of methodologies for building ontologies. *The Knowledge Engineering Review, 2002*, *17(2)*, 129– 156.

[41] L. R. Goldberg, "The structure of phenotypic personality traits." *American Psychologist,* 1993, vol. 48, 26-34.

[42] Z. Syed., A. Padia., T. Finin, L. Mathews, and A. Joshi. *UCO: A Unified Cybersecurity Ontology* (Tech.). Baltimore, MD, 2016.

[43] Claycomb, W. R., Huth, C. L., Flynn, L., McIntire, D. M., Lewellen, T. B., & Center, C. I. T. (2012). Chronological Examination of Insider Threat Sabotage: Preliminary Observations. JoWUA, 3(4), 4-20.

[44] J. R. C. Nurse, O. Buckley, P.A. Legg, M. Goldsmith, S. Creese, G. R. T. Wright, and M. Whitty. *Understanding Insider Threat: A Framework for Characterising Attacks,* 2014.

[45] R. N. Carvalho, K. B. Laskey, and P. C. Costa. "Uncertainty modeling process for semantic technology." *PeerJ Computer Science, 2016,* 2:e77 https://doi.org/10.7717/peerj-cs.77