APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

FY2021: Countering Insider Threat – Moving to Insider Risk
Final Reports

September 30, 2021

The Applied Research Lab for Intelligence and Security (ARLIS) is The University Affiliated Research Center (UARC) for Sociotechnical Systems and Advantage in the Human Domain at the University of Maryland

ARLIS
7005 52nd Avenue
College Park, Maryland

ACKNOWLEDGEMENTS

These reports were prepared for [Office of the Undersecretary of Defense for Intelligence and Security (OUSD(I&S)), United States Department of Defense (DoD)] under the following agreement:

HQ003420F0655, *University of Maryland,* "Insider Threat and Personnel Vetting."

DISCLAIMERS

Any views, opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of an official United States government position, policy, or decision. Additionally, neither the United States government nor any of its employees make any warranty, expressed or implied, nor assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, product, or process included in this publication.

Certain commercial entities, equipment, or materials may be identified in this document to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the Applied Research Laboratory for Intelligence and Security (ARLIS), the University of Maryland, or the United States government, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

ABOUT ARLIS

Applied Research Laboratory for Intelligence and Security (ARLIS) is a UARC based at the University of Maryland College Park and established in 2018 under the auspices of the OUSD(I&S). ARLIS is intended as a long-term strategic asset for research and development in artificial intelligence, information engineering, acquisition security, and social systems. One of only 14 designated United States Department of Defense (DoD) UARCs in the nation, ARLIS conducts both classified and unclassified research spanning from basic to applied system development and works to serve the U.S. Government as an independent and objective trusted agent.

Technical Points of Contact:

PI: Adam Russell, D.Phil.
Chief Scientist, ARLIS
301.226.8834; arussell@arlis.umd.edu

Co-PI: Kelly Jones, Ph.D.
Assistant Research Scientist, ARLIS
301.226.8850; kjones@arlis.umd.edu

Administrative Points of Contact:

Ms. Monique Anderson
Contract Officer, Office of Research Administration
Assistant Director, ARLIS
301.405.6272; manders1@umd.edu

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

FY2021: Countering Insider Threat – Moving to Insider Risk Executive Summaries of Program and All Tasks

September 30, 2021

PI: Adam Russell[1]* & Co-PI: Kelly M. Jones[1]*
[1]Applied Research Laboratory for Intelligence and security
*corresponding author, russell@umd.edu; *corresponding author, kjones@arlis.umd.edu

Table of Contents

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

FY2021: Countering Insider Threat – Moving to Insider Risk Program Executive Summary and Overview

September 30, 2021

Adam Russell[1]* & Kelly M. Jones[1]*
[1]Applied Research Laboratory for Intelligence and security
*corresponding author, russell@umd.edu; *corresponding author, kjones@arlis.umd.edu

EXECUTIVE SUMMARY

In September of 2020, the Office of the Undersecretary of Defense for Intelligence and Security (OUSD(I&S)) issued a Task Order to UMD's Applied Research Laboratory for Intelligence and Security (ARLIS) to undertake different tasks aimed at making advances in the USG's capability to address what is commonly known as Insider Threat[1]. This report details those specific tasks – their goals, approaches, outcomes, and conclusions – and provides an introductory overview of the larger Task Order, which outlines the concept of moving from a paradigm of detecting and deterring Insider Threat in complicated systems to one of "Modeling and Mitigating Insider Risk" (MInR) in complex, sociotechnical systems[2]. What might appear to be a trivial semantic difference is unpacked to reveal that there are substantive implications that follow from thinking in terms of threat vs. risk for this mission area and USG's capabilities. The introduction also provides the larger context for the Task Order's individual tasks as part of ARLIS' mission -- to help USG achieve and maintain an intelligence and security edge – in its role as the UARC for Sociotechnical Systems and the Human Domain.

Contents

[1] "The threat that an insider will use her/his authorized access, wittingly or unwittingly, to do harm to the security of United States. This threat can include damage to the United States through espionage, terrorism, unauthorized disclosure of national security information, or through the loss or degradation of departmental resources or capabilities." *NIST SP 800-53 Rev. 4 under Insider Threat Presidential Memorandum, National Insider Threat Policy and Minimum Standards for Executive Branch Insider Threat Programs*

[2] We use the term "sociotechnical" to highlight the dynamic and often non-linear behaviors that emerge from the interactions of connected hardware, software (including increasingly capable algorithms as "agents"), humans, and communities across coupled systems and networks. The implication of the term is that any effort to understand (much less optimize) such systems by only focusing on the socio- or the technical- is likely to lead to (undesirable, unintended, and counterproductive) surprises.

Modeling and Mitigating Insider Risk

Given the range and number of research tasks that were conducted under this Task Order (TO), this Introduction uses the first four questions of the Heilmeier Catechism as an organizing structure to help capture the overall goal, approach, and desired outcomes for the Task Order as a whole.

What is the Task Order trying to do?

In 2019, ARLIS was tasked by OUSD(I&S) – its new DOD sponsor -- to develop a research road map for how ARLIS, in its unique role as the UARC for Sociotechnical Systems and the Human Domain, could best leverage its applied research capabilities to positively impact the mission area generally known as "Insider Threat"[3]. In spite of the wide range of organizations and personnel already engaged in Insider Threat work, ARLIS was identified as having a somewhat unique – yet highly relevant – set of core competencies to draw from, including the social and behavioral sciences and human language and culture, but also AI/ML, autonomy, augmentation, and advanced computing. This mix of core competencies positioned ARLIS to have an equally unique perspective on the problem of Insider Threat, given ARLIS' interdisciplinary bench combining qualitative and quantitative researchers, and ability to reach out beyond itself to engage and convene other research personnel and organizations from across academia, industry, and government, as appropriate.

The resulting road map was delivered based on ARLIS' review of existing empirical literature, key policies, and strategic plans from USG stakeholders, as well as input from subject matter experts, including our core sponsor at OUSD(I&S). In the road map, ARLIS identified five areas of research that both mapped to ARLIS's core competencies and addressed research needs of the community, including focusing on Insider TRUST (Trustworthy, Resilient, and Useful Systems and Teams), and integrating and conducting both social behavioral science research and technical and cyber solutions – that is, to truly embrace and combine the "socio" and the "technical" dimensions of sociotechnical systems. This road map was designed to be a living document, meant to evolve to meet changes in the risk landscape and sponsor requirements[4].

The 2019-2020 road map then informed the 2020-2021 ARLIS research program on Countering Insider Threat: Moving to Insider Risk, in which ARLIS drew from the road map and incorporated additional sponsor needs to tackle a diverse set of programmatic goals to help OUSD(I&S) to continue its progress in addressing the hard problem of "Insider Threat", via a coordinated

[3] "The threat that an insider will use her/his authorized access, wittingly or unwittingly, to do harm to the security of United States. This threat can include damage to the United States through espionage, terrorism, unauthorized disclosure of national security information, or through the loss or degradation of departmental resources or capabilities." *NIST SP 800-53 Rev. 4 under Insider Threat Presidential Memorandum, National Insider Threat Policy and Minimum Standards for Executive Branch Insider Threat Programs*

[4] Jones, K.M., Maines, M., Gordon, R., Morrison, M., & Bunting, M. (2020, May). *FY2020: Strategic road map to leverage research to counter insider threat*. The University of Maryland, Applied Research Laboratory for Intelligence and Security (ARLIS).

program of research and development tasks that collectively seek to complement current Insider Threat efforts – focused largely on individuals -- with an "Insider Risk" paradigm that emphasizes systems thinking. Given ARLIS's unique perspective and core competencies, this Insider Risk paradigm is meant to highlight how individuals, organizations, networks, and technologies comprise increasingly complex "sociotechnical systems"[5]. Such complexity requires more risk-based approaches, thinking, and research to address the emergent nature of "failure modes" of our human-centric but technologically-mediated systems – failure modes where Insider Threats are important but not unique examples. That is, in risk-based analyses, trying to detect and deter "insider threats" makes most sense within a larger context of a sociotechnical system's vulnerabilities and the consequences of its various failure modes. Hence, approaching this challenge as one primarily of "finding and fixing" insider threats may lend itself to research and technical efforts premised on assumptions that do not reflect – and so may not wholly address – the sociotechnical complexity of today's or future insider threat challenges, nor the need for accordingly sociotechnical solutions.

This TO therefore included a collection of tasks that – taken as a whole – help move us from a focus on detecting and deterring Insider Threat to one of "Modeling and Mitigating Insider Risk" (MInR). As we hope a reader will appreciate, while each task addressed its own important questions ("Can we mitigate risk by understanding trust in organizations?" "Does online behavior reflect offline selves?" "Do certain risk-assessment technologies actually perform as advertised?" "Can we use crowds to forecast emergent insider risks?" Etc.), viewed collectively, the TO makes important contributions in moving towards the kind of R&D that will be necessary to adopt risk-based, vs. mainly threat-based, thinking in this area. In particular, this TO incorporated and explicitly embraced three key elements that will be required for tackling Insider Threat complementing it with an ability to model and mitigate Insider Risk. These were:

1. **Thinking differently.** Challenging important assumptions about how to think about Insider Threat was baked into the TO from the beginning, particularly in terms of the (to date relatively limited) roles of the social and behavioral sciences (SBS) in helping to inform how we think about "the" human in this area – noting of course that there is no such thing as "the" human, as SBS have long documented the extremely diverse, dynamic, and highly context-dependent variety of different kinds of humans and human behaviors.[6] For example, rather than just assuming that a focus on threats informed largely by case studies

[5] For more on sociotechnical systems, see for example Baxter and Sommerville's 2011 article "Socio-technical systems: From design methods to systems engineering", in which they identify key criteria that define those kinds of systems. Tellingly, a key feature of sociotechnical systems is that focusing on optimizing the technical element, OR the social element, at the expense of the other, "is likely to lead to degraded system performance utility." That is, a focus on tools OR people in a sociotechnical system will lead to unanticipated and potentially unwanted consequences. doi:10.1016/j.intcom.2010.07.003,

[6] In terms of this variability, ARLIS – as the UARC focused on sociotechnical systems for the intelligence and security communities – exists in part specifically to tackle Human Domain challenges of this kind, where the definition of the Human Domain is where the organization with the best understanding of, and capabilities to leverage, variability and diversity of human behaviors, abilities, strengths, and limitations stands to gain competitive advantage over those organizations that lack similar understanding and capabilities.

of "what went wrong" in organizations is the best approach to Insider Threat, might we stand to gain from also trying to understand "what goes right" in high performing organizations where things like building and maintaining trust among an organization and its personnel are also prioritized?

2. **Operationalizing Test and Evaluation (T&E).** Given the wide range of different technologies that make claims to helping to address some dimension of Insider Threat, these tasks emphasized putting T&E of those technologies into an Insider Risk context. That is, by "operationalizing" T&E of a technology in terms not just of a single accuracy score or a simplistic criterion of "good/bad" results, but instead operationalizing it in terms of its potential performance within the context of the mission area, this TO has helped to highlight the need to revisit existing assumptions of what constitutes substantive evidence for a technology's readiness to be adopted and deployed in pursuit of Modeling and Mitigating Insider Risk.
3. **Building the Capability Bench.** It has been said that today's problems are unlikely to be solved using the same thinking that, in part, helped to create them – or at least, that failed to prevent them from emerging. The same can be said of the tools and skillsets that we have traditionally relied upon. This is not to say that such thinking, tools, and skillsets are wrong or ineffective; only that they may be necessary but insufficient. This task has begun the process of exploring the right way to use education to help seed systems and risk-based thinking into the Insider Threat workforce, in order to help further enable USG to develop sociotechnical solutions that may be sufficient to solve sociotechnical problems.

How is it done today, and what are the limitations?

It is true that the Insider Threat paradigm has many intuitively appealing features: it highlights the potential for insiders to be threats in the first place, which can often be underappreciated; it helps to establish addressing insider threats as an important organizational mission, which can often be under-resourced; it helps to advocate for responsibilities, research, and resources for insider threat programs, which can often be lost in various bureaucratic shuffles, and; in its most mature phase, it invites thinking about how to ***counter*** Insider Threat, which should help organizations begin to think "proactively."

And yet, some of this paradigm's intuitive appeal can also be its limitation, as discussed below.

1. Insider Threat, as a term, sets up an (often unnecessarily) adversarial dynamic among an organization and its personnel, suggesting that our colleagues, and ourselves, are viewed by the organization as perpetually imminent threats, and must be treated as such by seeing people as the source of the problem.
    - This is in contrast to Insider Risk, in which an organization openly acknowledges that **everyone** in the organization is "at risk" – of making mistakes or making bad decisions, of being compromised in some way, of being influenced in the wrong way – and indeed that the organization is itself at risk, hence the need to include people as part of the solution. As citizens, we are quite used to being "at risk" of various dangers and, but that is different from messaging that we ourselves may be the dangers in the first place.
2. A focus on "Threat" necessarily invites defensive thinking, which limits the ability to think through whether, and if so to what extent, efforts to neutralize potential threats might also

hamper organizational performance. If threats must necessarily be "neutralized," the potential for negatively impacting organizational performance by doing so will often be overlooked or disregarded.
    - This is in contrast to the Insider Risk paradigm, which acknowledges that risks are always present, and that various risks therefore must be prioritized and mitigated based on an assessment of a sociotechnical system's failure modes. This will include not just different threats, but also the kinds and degrees of vulnerabilities that the system and its people may have to that threat, and the consequences of the risks should the events occur as well as the consequences to the organization of interventions designed to reduce those risks.
3. Insider Threat almost invariably leads to a narrowing of an organization's attention to the individual as the source and sink of threat. While various Insider Threat models (like the critical pathway) pay some lips service to the role of context in speeding up or slowing down people along the path, and while it is clear that individual characteristics clearly do play important roles in leading someone to be a threat, nonetheless those are still powerfully influenced by larger sociotechnical factors and contexts. Insider Threat however tends not to incorporate those larger factors and contexts, which can reduce the ability of an organization to learn the right lessons from historical failure modes.
    - Insider Risk, as a complement to Insider Threat, seeks to incorporate individuals as well as the systems in which those individuals exist and behave, and which shape that behavior. Accordingly, Insider Risk as a paradigm promotes proactive thinking in terms of modeling various potential failure modes, and assigning various potential probabilities to those, based on a wide range of variables, which would include individual behaviors and characteristics of course, but also include other "socio" and "technical" factors of the larger system (leadership, IT systems, morale, HR support, targeting, etc.) that will shape the threat, vulnerability, and consequences of various Insider outcomes and events.

What's new about this approach?

As discussed elsewhere in this report, there are some who believe that adding Insider Risk as a complement to Insider Threat is merely a semantic gesture – perhaps an effort to "soften" a possible bad reputation that Insider Threat programs might invite from employees who, as mentioned above, could interpret this as implying that they may be threats themselves, or who (understandably) might have some hesitation in reporting concerning behavior in colleagues they care about, and alongside whom they may have served years, to an "Insider Threat" program. However, this report asserts otherwise: that there is in fact power in words in driving organizational ***attention*** in terms of what an organization focuses on, what problems it sees and attends to, and consequently how the organization understands its challenges, how it chooses to address those problems, and how it seeks to research and resource certain solutions.

Accordingly, the addition of "Modeling and Mitigating" the complexity of today's sociotechnical systems and their emergent failure modes – from individuals to wholesale organizations -- have implications for the value of conventional "threat-based" approaches. Sociotechnical complexity makes detecting true threats that much more difficult, since sociotechnical behaviors are rarely simply explained by equally simple intentions. This can give rise to false alarms and false negatives

alike that can rapidly outpace, and swamp tools and methods designed to pick up on any particular signals (or "red flags"), especially since those signals themselves may change over time and with new interactions among humans, our networks, our technologies, and our missions.

Likewise, prediction in sociotechnical systems must adopt radically new approaches when compared to predicting behaviors in "merely" complicated systems. Predictions in sociotechnical systems can rarely be reasonably expected to be "point predictions" of individual events ("this person will be a threat"), since emergence means the system, the technologies, and the people will themselves change and shape each other. Hence long-term predictions based on life histories or static traits about who will prove trustworthy over months, and years, of dynamic systems evolution and human-technology interactions are likely to prove impossible and unproductive, and in the absence of strong quantitative measures of Return on Investment like those that are prioritized in more risk-based industries, could end up costing far more than the likely very limited results will probably justify.

Finally, getting ahead of failure modes in sociotechnical systems also requires a more expansive approach than trying to simply deter certain behaviors. Deterrence may have been possible in "merely" complicated systems, but in sociotechnical systems and risk-based paradigms, potential failure modes – including insider threat behaviors -- must be mitigated. That is, the system should be designed to be resilient in light of any potential failure mode, rather than premised on the idea that all given failure modes can be deterred. Instead, when interventions designed to deter behaviors in simple or complicated systems are then introduced into sociotechnical systems, they can have profound and unanticipated consequences, often negatively impacting the performance of an organization and leading to the gaming of a system or process which, in no small irony, can give rise to further emergent failure modes.

None of this is to suggest that failure modes like various Insider Threats in sociotechnical systems cannot be understood, forecasted, or mitigated, although it is to say that they can't be done perfectly and that focusing on finding individual bad apples may not be the best way forward. Importantly, other industries also concerned with failure modes related to Insider Threat have acknowledged the challenge of complex sociotechnical systems, and many have adapted to the challenge by adopting a "risk"-based paradigm, looking to first try and model the risks of various failure modes in order to better apply proportionate research, operational, and security resources, and foregoing point predictions for probabilistic forecasting. It was this kind of risk-based paradigm that motivated much of this Task Order's approach in incorporating a wide range of applied research tasks, relevant to Insider Risk as well as Insider Threat, in order to better position our national security organizations to protect themselves in increasingly complex environments while enhancing – or at least not negatively impacting -- their performance.

The specific tasks therefore fell into three broad categories related to the Task Order, as described above.

Thinking Differently: These tasks brought ARLIS' core competencies and unique perspective to provide some different thinking on the Insider Threat mission area, exploring what new kinds of approaches might mean for tackling Insider Threat and, potentially, mitigating Insider Risk.

1. Task 1: Establishing Program Goals and Defining Insider TRUST
    a. Subtask 0: Insider Risk Seminar Series (IRiSS)
        i. This task organized and executed a seminar series bringing together experts and thought leaders from academia, industry, and government to discuss the pros, cons, and edge cases for operationalizing the Insider Risk (InR) paradigm.
    b. Subtask 1: Defining Insider TRUST
        i. This task scoped and defined ARLIS's Insider TRUST approach (Trustworthy, Resilient, and Useful Systems and Teams), integrating existing research across multiple disciplines to create a broad base of concepts and knowledge that are potentially applicable to countering and mitigating risks of insider events, and creating new research by drawing on ARLIS ability as a UARC to convene experts across silos of information to identify operationally focused research questions related to Insider TRUST and to the larger program of research on insider risk/threat at ARLIS.
2. Task 2: Insider Threat Seedling Program
    a. Subtask 2a: Insider Threat Hub Simulation Using the ICONS Platform for Training and Education (ICONS for MInR). This seedling explored the use of an immersive simulation platform to help engage learners in thinking about risk in various Insider Threat scenarios.
    b. Subtask 2b: Comprehensive Risk Assessment of Military and Veterans Convicted of Domestic Extremism via the PIRUS Database (CRA). This seedling collected, analyzed, and reported out on various risk factors related to active duty and veteran personnel engaging in domestic extremism.
    c. Subtask 2c: Identifying and Cataloging Datasets for Modeling and Mitigating Insider Risk (D-MInR). This seedling identified a wide range of potentially relevant datasets for modeling Insider Risk, producing a code book to assist researchers in choosing which datasets might be most useful to consider when trying to model how individual, social, and organizational factors might interact to create more or less Insider Risk.
    d. Subtask 2d: Crowdsourced Security & Intelligence Forecasting Tool for Modeling and Mitigating Insider Risk (CSIFT for MInR). This seedling explored methods for creating and decomposing organizational and decision-maker questions related to Insider Risk, which could be asked and answered using aggregative crowdsourced forecasting methods to better help DOD model, and potentially mitigate, different kinds of Insider Risk.
3. Task 3: Literature Review of the Relationship between Representations of the Self on Social Media and Representations of the Self Offline

   a. This task integrated existing evidence across research disciplines to identify whether, and -- if so -- how, online presentations of the self correlate to offline presentations of the self, which can provide more evidence of the value of using social media data to inform risk decisions in personnel vetting and security.

Operationalizing T&E: An additional area of sponsor need identified for the 2020-2021 program tapped into ARLIS's role as a trusted agent for operationalizing testing, evaluation, verification, and vetting (TEV&V) for existing and emerging technologies, tools, and methods for personnel vetting and for detecting and mitigating risks of insider threats.

4. Task 4: Technology Bakeoff – An Independent Assessment of Publicly Available Electronic Information (PAEI) Commercial Data Providers
   a. This task provided independent testing and evaluation of PAEI Scraper Companies' capabilities to collect and provide relevant data for use in personnel vetting processes, seeking to evaluate their performance using relevant operational criteria. Per the TO, we solicited companies to participate in the evaluation on a voluntary basis, and from the ten companies who volunteered, ARLIS and the sponsor jointly selected four potential performers based on preliminary information they provided based on criteria developed per the sponsor's identified operational needs.
5. Task 5: TEV&V Framework for Voice Analysis Tools that Identify Potential Risk
   a. This task developed a framework and proposed testing protocol for the independent technical evaluation of commercial products that assess risk, using a protocol that reflects operational parameters. Per the task order, this effort focused on evaluation of a particular technology from a particular company. However, the company declined to participate in the proposed experimental protocol, thus the evaluation is based on materials provided, review of the literature, and ARLIS proposed Evidence Readiness Levels framework.

Building a Capability Bench: The final area of the 2020-2021 program utilized ARLIS's role as a UARC, drawing on our university campus to develop and deploy a comprehensive educational experience on Insider Threat and Insider Risk.

6. Task 6: Summer Program for Educational Outreach
   a. This task created and executed a comprehensive overview of Insider Threat and Insider Risk as a summer course at the University of Maryland. It included the pilot run of the Task 2 Seedling – ICONS -- which developed an Insider Threat hub simulation for training and education as the capstone activity of the course, meant to highlight the complexity of today's sociotechnical systems and equip future Insider Threat personnel with some early foundational knowledge that could also inform future Insider Risk thinking.

Who cares? What difference will this make?

The Insider Threat community is enormous; the Insider Risk community, given the implications, may be even bigger. Understandably then, the wide range of tasks under this Task Order (TO) and its tasks also reflect this diversity of intelligence and security-based organizations, many of whom can be identified by asking whether they themselves must worry about sociotechnical failure modes. Accordingly, while each task has garnered attention from one or more specific organizations (identified further in this report), collectively the Task Order's products are likely to have broad relevance for national security, even as we acknowledge that these are still early days for "Modeling and Mitigating Insider Risk." That said, often research has a life of its own, and it is difficult to predict what kind of impact one, some, or any of these products may have.

However, in thinking about the impact of applied research, one should also consider not just immediate operational impact today, but also future impact on how things might be done tomorrow, to include how applied research might shape the larger Insider Threat/Insider Risk "innovation" supply chain – whether that involves research, TTPs, policies, or people. In this regard, some immediate lessons learned from this Task Order that might speak to ways to positively impact the Insider Threat/Insider Risk mission area going forward:

1. Thinking Differently:
    a) Continue to have OUSD(I&S) and ARLIS promote Risk-based thinking to complement the Threat-based thinking that remains the standard, using mechanisms like seminars, outreach, publications, etc., in particular highlighting the need to consider both "socio" and "technical" elements of risk as key to advantage
    b) Promote Risk-based thinking to inform "signal detection" approaches that go beyond conventional individual ("red flag") signals and begin to consider more sociotechnical context-based factors that might increase or decrease risks in currently unanticipated ways
    c) Use Risk-based methods to investigate and seek to quantify ROI on which signals we spend resources on detecting, and cost-benefit tradeoff from a performance perspective, in part to inform approaches, and in part to help further promote Risk-based thinking
    d) Begin a parallel research thrust on Insider TRUST to build up a further body of knowledge and capabilities related to "what goes right" in high performing organizations that create and sustain competitive advantage, and seek to promote Insider TRUST as a key factor for mitigating Insider Risk in order to enhance USG's performance
2. Operationalizing T&E:
    a) Continue to develop/promote the use of Evidence-Readiness Levels (ERLs) across the Insider Threat/Risk community, in part to encourage and help standardize the (appropriate) use of SBS theories, models, and research, which may be at various levels of maturity, as capabilities that can also speak to the "socio" aspect of sociotechnical risks (in addition to using ERLs for evaluating technologies that speak to the more "technical" aspects)

   b) Encourage and enable a "continuously updating" mindset for identifying and mitigating various Insider Risk failure modes using continual testing, red teaming, and experimentation as the norm, and not the exception, to include building up mechanisms to regularly enable the operational testing and evaluation of new capabilities through a network of ready to go, agreed upon partner-based protocols
   c) Develop modeling tools to help contextualize the impact and relevance of OT&E results under various operational conditions, moving away from a focus on a particular tool's basic "hit/miss" performance towards better understanding how that tool will work (or not) for different communities/users with different needs and parameters (that is, push tools up the ERL scale)
3. Building a Capability Bench:
   a) Invest in resources (personnel, tools) to promote formal modeling as tools to create more common knowledge bases across Insider Threat communities, incorporate new findings,
   b) Continue to explore, test, and adopt new and improved signals detection methods/tools/approaches that reflect the new challenges – but also the new opportunities – that come from complex, sociotechnical systems (crowd forecasting, organization level signals as well as individual level signals)
   c) Enable Risk-based thinking by creating and/or integrating new tools to allow I&S partners to leverage things like knowledge-graphs to begin to build enterprise level context for interpreting different signals
   d) Create and/or integrate experiential learning methods that are crucial for sensitizing Insider Threat personnel to applying risk-based approaches to low probability, high impact events in complex systems

ACKNOWLEDGEMENTS

This report was prepared for [Office of the Undersecretary of Defense for Intelligence and Security (OUSD(I&S)), United States Department of Defense (DoD)] under the following agreement:

HQ003420F0655, *University of Maryland,* "Insider Threat and Personnel Vetting."

DISCLAIMERS

Any views, opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of an official United States government position, policy, or decision. Additionally, neither the United States government nor any of its employees make any warranty, expressed or implied, nor assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, product, or process included in this publication.

Certain commercial entities, equipment, or materials may be identified in this document to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the Applied Research Laboratory for Intelligence and Security (ARLIS), the University of Maryland, or the United States government, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

ABOUT ARLIS

Applied Research Laboratory for Intelligence and Security (ARLIS) is a UARC based at the University of Maryland College Park and established in 2018 under the auspices of the OUSD(I&S). ARLIS is intended as a long-term strategic asset for research and development in artificial intelligence, information engineering, acquisition security, and social systems. One of only 14 designated United States Department of Defense (DoD) UARCs in the nation, ARLIS conducts both classified and unclassified research spanning from basic to applied system development and works to serve the U.S. Government as an independent and objective trusted agent.

Technical Points of Contact:

PI: Adam Russell, D.Phil.
Chief Scientist, ARLIS
301.226.8834; arussell@arlis.umd.edu

Co-PI: Kelly Jones, Ph.D.
Assistant Research Scientist, ARLIS
301.226.8850; kjones@arlis.umd.edu

Administrative Points of Contact:

Ms. Monique Anderson
Contract Officer, Office of Research Administration
Assistant Director, ARLIS
301.405.6272; manders1@umd.edu

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Countering Insider Threat (CInT)
(Moving to Insider Risk)
Task 1: Insider Risk Speaker Series (IRiSS)
Final Report

September 30, 2021

Kelly Jones[1*], Shawn Janzen[1*], Joseph Kelly[1*], Bill Stephens[1], & Adam Russell[1]
[1]Applied Research Laboratory for Intelligence and Security
[*]corresponding authors email addresses: kjones@arlis.umd.edu ; sjanzen@umd.edu ; jkelly@arlis.umd.edu

"I am glad to see the term migrate from Insider Threat to Insider Risk. There are a lot of good reasons for that change, but mostly it avoids unnecessarily alienating the workforce."
– Charlie Phalen, *Principal, CS Phalen & Associates LLC*
Previous Positions
Acting Director, Defense Counterintelligence and Security Agency
Director, National Background Investigations Bureau, Office of the Program Manager
Senior Vice President for Corporate Security, Northrup Grumman
Director of Security, Central Intelligence Agency

EXECUTIVE SUMMARY

Insider Threat is a well-understood concept across the Department of Defense (DoD) and intelligence community (IC) enterprises, even if what constitutes a threat often reflects specific organizational nuances and missions. Likewise, efforts to counter Insider Threat – through deterrence, detection, and mitigation – are also increasingly common topics across DOD/IC enterprises, although reflecting again certain idiosyncrasies of various organizations. Despite these differences, what Insider Threat (InT) and Counter Insider Threat (CInT) approaches generally have in common is the tendency to focus on any given individual as the primary target for efforts to deter, detect, and/or mitigate InT. This paradigm has strong intuitive appeal – of course we want to find the "bad apples" – but has certain limitations when viewed from the perspective of modern environments of growing complexity, where traditional threat-based approaches to finding and neutralizing "bad apples" may be increasingly less effective and increasingly more reactive. This is because our organizations are "sociotechnical systems,"[7] where the complex interactions among different kinds of humans[8], different kinds of technology, and different kinds of dynamic environments mean that the threats, vulnerabilities, and consequences we seek to avoid are increasingly emergent: that is, they emerge from this complex interaction, not just from a single person. Hence, focusing on individuals as "threats" within a complex system will tend to lead us to fixate on characteristics of a given person, and consequently ignore or miss the most important

[7] Sociotechnical systems are characterized by having multiple independent parts, which adapt and pursue different goals in external environments, but which have an internal environment comprising separate but interdependent technical and social subsystems, where goals can be achieved by more than one means and thus require some kind of organizing processes to decide how to achieve goals, and where the performance of the system depends on a "joint optimization" of the technical AND the social subsystems. In sociotechnical systems, focusing on one over the other is likely to lead to degraded performance and unanticipated – and often unwanted – outcomes. By these measures, most DOD and IC organizations ought to be considered sociotechnical systems.

[8] The role of human individual differences and variability means that the challenge of Insider Threat is very much a challenge of the "Human Domain," where advantage goes to the organization that has the best understanding of, and ability to incorporate, human variability, strengths, diversity, limitations, and vulnerabilities into their systems and designs. Achieving this kind of advantage is in part why ARLIS exists to bring its core competencies to Human Domain challenges like Insider Threat.

characteristics of the larger sociotechnical system that can give rise – often unexpectedly – to unwanted behaviors and failure modes.

Other industries also concerned with insider threat have acknowledged this challenge of sociotechnical systems, and many have adapted to the challenge by adopting a "risk"-based paradigm. It is possible that DOD and the IC should also adopt an Insider Risk (InR) paradigm as a necessary supplement to InT, to better position our national security organizations to protect themselves in increasingly complex environments while enhancing their performance.

Briefly, the difference between threat and risk paradigms are summarized in the table below.

Table 1: Comparing the Insider Threat and Insider Risk Paradigms

Insider Threat		**Insider Risk**
Categorical thinking (threat or not a threat)	↔	Nuanced thinking (degrees of risk)
Static (threats do or do not exist)	↔	Dynamic (risk is always changing based on past & present factors)
Threats must be "neutralized" to be addressed	↔	Risk must be managed since risk can never go to zero.
Focus on the individual as the source of threat, minimal focus on context	↔	Risk comes with interaction of individual, contextual, organizational, systemic, and enterprise variables
People are viewed as the problem	↔	People are part of the solution
Interventions begin largely after concerning behaviors occur	↔	Interventions address identified risks and future vulnerabilities before they can be exploited
	↔	

The primary takeaway from this comparison is one of attention: InT attends to the individual. Seeking to classify individuals as a threat (or not), and consequently efforts to deter, detect, and mitigate InT will necessarily focus on individuals, but often at the expense of considering other key factors. InR, instead, attends to the characteristics of the sociotechnical systems, in which individuals operate, and of which individuals are a key part, but only a part. Concentrating on risk, vs threat, necessarily requires thinking more broadly in terms of failure modes, which is harder but, in the end, potentially much more effective as it incorporates humans and our variability in ways that also acknowledge the importance of sociotechnical contexts in shaping our behaviors.

As part of the effort to explore the value of adopting a risk-based paradigm to complement current InT efforts, ARLIS was tasked to conduct the Insider Risk Speaker Series (IRiSS) under its Countering Insider Threat (Moving to Insider Risk) contract as part of the ARLIS InR mission area. The IRiSS task was to organize and execute a seminar series bringing together experts and thought leaders to develop an approach to modeling and mitigating insider risk (MInR). IRiSS is part of ARLIS' efforts to build a capability bench to help the US government (USG) deal with emergent sociotechnical challenges and opportunities in complex systems, and draws on its capabilities as a UARC to convene groups from across government and non-government partners, experts, and thought leaders in academia, industry, and non-profits to discuss emerging issues and encourage exchange of new approaches to persistent challenges. Our goals for IRiSS were therefore two-fold:

1. Elicit and integrate diverse perspectives on InT and InR, fostering an environment that leads to better modeling (characterizing, quantifying, predicting) emergent InRs.

2. Elicit and integrate diverse perspectives on InT and InR, fostering an environment that leads to better mitigating (shaping, exploiting, preventing) emergent InRs.

This final report represents the culmination of the IRiSS program which conducted a six-month seminar series that ran from March 2022 through August 2021, wrapping in time to transition cleanly into National Insider Threat Awareness Month (NITAM). The six IRiSS events each carried a different theme and featured different guest speakers—19 in total.

Key session topics:

1. March 2021: State of Insider Threat and Insider Risk paradigms
2. April 2021: From threat to risk: Gain & loss, response, and management around insiders within academic environments
3. May 2021: Industry views – Where are we now
4. June 2021: Tools, methods, and technology -- State of the art in modeling
5. July 2021: Insider risk, human resources, and the human capital supply chain challenge
6. August 2021: Actualizing the Insider Risk Paradigm

Events were well-received by the attendees with strong, positive feedback. Attendance averaged around 200 people per event and almost double that for registration. After six months of engaging InR dialogue with experts and interested individuals, the early reaction to the two goals above posited that we now have good footing to discuss these topics. Through IRiSS and the other InR tasks, we are moving forward with efforts to improve emergent InR models and mitigation efforts. However, such change will take time. In the interim, IRiSS could continue with the popular series, return in a different format, both, explore other routes, or go dormant until needed again.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Defining Insider TRUST (Trustworthy, Reliable, and Useful Systems and Teams)
Final Report

September 30, 2021

Kelly M. Jones*[1], Bernadette Jerome[1], Jordan C. Roberts[1], Emily Pitek[1], Graduate Student 1[1+], & Adam Russell[1]

[1]Applied Research Laboratory for Intelligence and security

[+]name withheld at author's request

*Corresponding author, kjones@arlis.umd.edu

EXECUTIVE SUMMARY

Task 1 under the ARLIS FY20 Countering Insider Threat program of research focused on scoping and defining ARLIS's Insider TRUST (Trustworthy, Resilient, and Useful Systems and Teams) approach. Insider TRUST research is designed to complement the traditional approach of addressing insider threats (the detect, deter, and mitigate approach) by providing a framework of concepts and research that could be practically applied to cultivate healthy work environments that reduce the risk of insider events and/or empower individual, organizational, and mission success.

The goal of this task has been to provide a starting point for addressing complex sociotechnical problems of the *human domain* by *integrating* existing research across multiple disciplines (with particular emphasis on drawing from social and behavioral science disciplines not often incorporated into security research) to create a broad base of concepts and knowledge that are potentially applicable to countering and mitigating risks of insider events, and *creating* new research by drawing on ARLIS ability as a UARC to *convene* experts across silos of information to identify the best ways forward in developing future operationally focused research questions related to the larger program of research on insider risk/threat at ARLIS.

We conducted a literature review of concepts related to Insider TRUST (e.g., resilience, burnout, organizational behaviors), and built conceptual models of the relationships between constructs based on the literature. The team developed a list of questions we then posed to subject matter experts (SMEs) in a semi-structured interview, regarding the constructs identified as potentially relevant to understanding and modeling Insider TRUST, additional constructs to be included, as well as potential relationships between these constructs. The goals of the research were to 1) identify common themes and questions among insider threat operators, researchers, and policymakers, 2) develop, refine, and expand the conceptual model of relevant constructs, and 3) develop future operationally focused research questions related to the larger program of research on insider risk/threat at ARLIS.

We noted four major recurring themes across the majority of our work.

1. *The importance of trust.* The need for solid foundational research on the relevance of trust to reducing insider risk, how to best assess trust, how best to increase trust, and the importance of trust as a part of leadership and workplace relationships were all raised as key questions for future research.
2. *The importance of good leadership and good relationships with colleagues*. The importance of cultivating and maintaining strong, healthy, and supportive relationships between colleagues, between immediate supervisors and supervisees, and between organizational leadership and employees was repeatedly highlighted. However, the greatest emphasis was given to the importance of good leadership at every level throughout an entire organization. Key future applied research questions focused on effectiveness and content training to organizational leaders to build TRUST through best leadership practices (e.g., demonstrating respect and valuing employees, setting and living up to organizational standards and expectations, clear communication).

3. *De-emphasis on tracking individual characteristics, and the need for greater emphasis on organizational characteristics.* Individual factors (e.g., personality, individual motivation, burnout) which currently receive great attention, were largely de-emphasized in favor of greater focus on organizational level factors such as leadership and the perceptions of organizational policies (including insider threat programs). When individual characteristics were mentioned by SMEs, they focused more on positive characteristics that can be fostered (e.g., resilience, well-being), rather than on the innate characteristics which can sometimes be negative (e.g., personality, pessimism).
4. *Recognition of the difficulty in, but ultimate importance of, getting ahead of threats before they manifest.* The security challenges, and opportunities, of insider risk/threat and personnel vetting are problems of the human domain, requiring successfully integrating the social and behavioral sciences in concert with technical solutions to better measure and model (characterize, quantify, predict) and mitigate (shape, exploit, prevent) emergent insider risks/threats. Gaining and maintaining advantage in human domains requires proactive, rather than reactive, interventions. By recognizing and operationalizing new predictors and indicators of increased risk, the security community will be able to build a resilient and trusted workforce – one that is much less likely to produce detrimental surprise. For example, rather than just assuming that a focus on threats informed largely by case studies of "what went wrong" in organizations is the best approach to Insider Threat, might we stand to gain from also trying to understand "what goes right" in high performing organizations where things like building and maintaining trust among an organization and its personnel are also prioritized?

A sample of possible applied research questions proposed either by the team or SMEs:

1. How do we take into account trustworthiness variables in our calculations of risk? How do we build systems that consider trustworthiness (and training, resilience, and other social behavioral factors) in risk assessments?
2. How do insider threat programs, leaders, and organizations, best build and communicate shared visions of addressing insider risk and building Insider TRUST? How effective are such messages in decreasing risk? Who are the best messengers of shared visions – leaders?
3. Is an effective way to build Insider TRUST through identifying and training leaders (and future leaders) in best leadership practices (e.g., demonstrating respect and valuing employees, setting and living up to organizational standards and expectations, clear communication)?
4. What elements of best leadership practices will remain constant over time and what will need to change in order to retain our advantage in the human domain?
5. What theoretical models for decision making can be leveraged or built for insider risk and Insider TRUST? How can we effectively draw research from other fields (e.g., game theory, Industrial/Organizational psychology, management, communications, risk-based decision making), making sure that what we draw is relevant and applicable for insider risk/threat contexts (both the theoretical model and its foundational concepts)?
6. What is the effect of the information environment on insider risk/threats and Insider TRUST? Do belief polarization and disinformation contribute to increased risk of insider

events? If they do, what are the best techniques for decreasing risks of insider events that are influenced by disinformation?

7. What is the larger organizational ecosystem from which insider risks emerge? What is the best way to model that ecosystem?
8. How do we link research, and operational best practices, that addresses the immutably sociotechnical nature of the human domain challenges we face across the entire interdependent supply chain of products, services, workforces, and workplaces?

The fourth key theme our SMEs identified is that gaining and maintaining advantage in human domains requires proactive, rather than reactive, interventions. The above applied research questions represent one approach to proactive work - identifying key questions in a research portfolio on addressing insider risk/threat and creating Insider TRUST.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Radicalization in the Ranks: An Assessment of the Scope and Nature of Criminal Extremism in the United States Military

Final Report

September 30, 2021

Michael A. Jensen[1]*, Elizabeth Yates[1], Sheehan Kane[1]
[1]National Consortium for the Study of Terrorism and Responses to Terrorism (START)
*Corresponding author, majensen@umd.edu

Profiles of Individual Radicalization in the United States (PIRUS)

RaD_UMD

The Applied Research Lab for Intelligence and Security (ARLIS) is The University Affiliated Research Center (UARC) for Sociotechnical Systems and Advantage in the Human Domain at the University of Maryland

ARLIS
7005 52nd Avenue
College Park, Maryland

EXECUTIVE SUMMARY

This project finds that criminal extremism with a nexus to the United States military is a limited, but possibly growing, problem that is primarily centered in the veteran community. From 1990 through the first eight months of 2021, 411 individuals with U.S. military backgrounds committed criminal acts that were motivated by their political, economic, social, or religious goals. Subjects with U.S. military backgrounds represent a small portion (11.5%) of the broader set of extremists who have committed criminal offenses in the United States since 1990. Moreover, the majority (83.5%) of these subjects were no longer serving in the U.S. military when they committed extremist crimes. However, there has been an upward trend in recent cases of criminal extremists with military backgrounds, suggesting that extremism in the ranks may be a growing concern. For example, from 1990-2010, an average of 6.9 subjects per year with U.S. military backgrounds committed extremist crimes. Over the last decade, that number has more than tripled to 24.3 subjects per year.

In addition to these aggregate trends, this study finds that:

- Approximately 14% (89 subjects) of the individuals who have been charged for participating in the Capitol breach on January 6, 2021, have U.S. military backgrounds.
- Just over 16% (68 subjects) of the extremists with military backgrounds who committed crimes in the United States since 1990 were actively serving at the time of their offenses or arrests.
- Approximately 79% of criminal extremists with military backgrounds served in the U.S. Army or Marine Corps, including Reserve and National Guard units.
- Nearly half of criminal extremists with military backgrounds espoused anti-government views or were members of organized militias. An additional 34.3% of the subjects promoted views of white supremacy and/or xenophobia, while 11% were connected to, or inspired by, Salafi Jihadist groups, including al-Qaeda and its affiliated movements and the Islamic State of Iraq and Syria (ISIS).
- Radicalization processes among active service members are likely to involve risk factors related to military service, including combat deployments and membership in extremist cliques with fellow service members. Veterans, on the other hand, often face age-related risk factors for radicalization, such as failed relationships, unemployment, and previous encounters with the criminal justice system, as well as psychological vulnerabilities tied to their military service, including high rates of post-traumatic stress disorder.

A public health model that focuses on education, prevention, treatment, and evaluation provides the best opportunity for the long-term mitigation of the risks associated with extremism in the armed forces. A public health model should prioritize:

- Data collection and scientific discovery on the scope and nature of extremism in the ranks.
- Prevention programs that (1) inoculate incoming service members (and future veterans) against extremist recruitment; (2) disseminate tailored awareness briefs about extremist narratives and recruitment techniques; (3) devise non-punitive responses to extremism

that increase the likelihood that concerning behaviors will be reported; and (4) form partnerships with the VA and community-based veterans' organizations to counter radicalization among past service members.

- Interventions for at-risk service members that address a variety of concerns, including mental health, substance use disorders, anti-social relationships, previous criminality, and unemployment.

Finally, this study argues that while it might be appealing to use military separations as a quick fix to the problem of extremism in the ranks, military discharges could result in transferring risk to local law enforcement agencies if they are not accompanied by the provision of rehabilitation services. Furthermore, as an all-volunteer force that depends upon willing recruits, the DoD should be aware that veterans who engage in extremist crime cause significant damage the reputation of military service and undermine U.S. national security as a result. Simply put, separations from the military neither address the underlying issues that cause individuals to radicalize, nor shield the military from blame when violence occurs in U.S. communities. Thus, when military separations are used to counter extremism in the ranks, they should be paired with referrals for support services, and potential risks to community safety should be effectively communicated to law enforcement partners.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Crowdsourced Security and Intelligence Forecasting Tool (CSIFT) Mitigating Insider Risk (MInR)
Seedling
Final Report

September 30, 2021

Monique Beaudoin, PhD[1]*, Thalia Baumgarten[2], Sean Kucer[2], Adam Siegel[2]
[1]Applied Research Laboratory for Intelligence and Security
[2]Cultivate Labs
*corresponding author, mbeaudoin@arlis.umd.edu

EXECUTIVE SUMMARY

Insider threats are a long-standing concern within the Department of Defense (DoD), where individuals can use inside knowledge of their workplace, employer, or coworkers to threaten the security of U.S. national security interests, whether intentionally or unintentionally. As prior case studies have shown, insider threat events are difficult to predict, due to the complex interaction of multiple factors, combined with limitless unknown variables in an ever-changing world context.

To assist the DoD and USG in its mission to anticipate, mitigate, and prevent such threats, ARLIS performed research and feasibility analyses to determine how to apply the Crowdsourced Security & Intelligence Forecasting Tool (CSIFT) and its quantitative methodology specifically to this critical IC/DoD challenge area. Combining literature research with inputs from government stakeholders and subject matter experts, The ARLIS and Cultivate Labs team identified three unique insider threat/risk use cases and decomposed key issues into specific forecasting questions for use in a future launch of the CSIFT tool for the IC/DoD, to anticipate and mitigate insider risk. A draft of the dashboard design was also created, and suggested forecasting populations were identified.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Datasets for Modeling and Mitigating Insider Risk Final Report and Codebook

September 30, 2021

Steve S. Sin[1]*, Kathryn A. Lindquist[1]
[1]National Consortium for the Study of Terrorism and Responses to Terrorism
*corresponding author, sinss@umd.edu

UWT
UNCONVENTIONAL **WEAPONS AND TECHNOLOGY**

EXECUTIVE SUMMARY

The breadth and scope of different variables that surround or are integral to understanding the risks that insiders may pose creates a complexity that requires additional applied and theoretical research across a wide range of topics and disciplines. Though data availability and privacy issues remain central issues in the field, data valuable for the study of insider threats, insider risk, detection, and mitigation do exist. The seedling project "Datasets for Modeling and Mitigating Insider Risk" (D-MInR) was designed specifically to identify extant datasets that may be useful for future Insider Risk efforts for the US government (USG) and its allies and partners, and to characterize those datasets in terms of their contents, location, relevance, accessibility, and other key parameters. D-MInR is the first known effort to track, catalogue, or characterize these potential resources across a wide variety of attributes and within multiple disciplines.

The result of the project was the development and completion of D-MInR Catalog Matrix and its accompanying final report and codebook. The catalog matrix lists the available data resources for insider threat and insider risk studies. The searchable and filterable catalog contains 107 total entries, with 45 in the main catalog and 62 in additional tabs to map more fully the data landscape. Information about each data source, including links and information about access as well as contents, data format, and other attributes, are coded for each resource in the catalog. The accompanying final report and codebook provides the approach and methodology used to develop the D-MInR Catalog Matrix, a quick start guide for using the catalog, and the codebook for the catalog.

One of the major findings from the catalog development was that there is a limited number of high quality, public-use datasets for the study of insider threat/ insider risk. While limited in numbers, these datasets do possess some important strengths and have contributed meaningfully to the advancement of the field; however, many are over a decade old and few include psychosocial measures alongside technical data, making them only partially suited to addressing the challenges of insider threat and insider risk circa 2021.

Another finding was that both the USG and a wide variety of other organizations have access to data on their personnel but limited capacity for turning that data into useful information about insider threat/insider risk. This is a problem shared by organizations writ large: Two in three organizations (66%) report a struggle with turning volumes of security activity and event data being collected into "intelligent, actionable insights." Many commercial products purportedly provide services to address this challenge, but the problem of integrating psychosocial information into tools that emphasize cyber security is largely unresolved.

As we anticipate that D-MInR Catalog Matrix will help researchers to improve their research designs and strategies, by leveraging these datasets, to conduct exploratory research (e.g., develop new hypotheses, discover new patterns, build new models or tools) and/or confirmatory research (e.g., test and validate existing hypotheses, tools, and models) that are crucial to enhancing the USG's understanding of and ability to mitigate insider risk, one should entertain the continuation of this seedling project so that the Catalog Matrix can be maintained and kept up to date. Additionally, to increase the accessibility of the D-MInR Catalog Matrix and its associated products to the countering insider threat enterprise, one should consider transitioning the catalog into a format and onto a platform that is conducive to wider research access.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Countering Insider Threat (CInT) Simulation

September 30, 2021

Ron Capps[1]* and Devin Ellis[1]*
[1]National Consortium for the Study of Terrorism and Responses to Terrorism
*corresponding authors, rcapps@umd.edu and ellisd@umd.edu

Created by:
The ICONS Project

EXECUTIVE SUMMARY

As part of ARLIS' work to support the Department of Defense through the Countering Insider Threat program, the University of Maryland's ICONS Project was funded to develop and implement an online training exercise. The materials presented here are the content of that exercise which can function as a stand-alone, in-person tabletop, or can be conducted over the ICONSnet online platform for distributed use.

The design of the exercise sees participants taking the roles of five key C-Suite executives (or their teams) at a fictional defense contractor, Kings Bay. As the exercise unfolds, they are presented with a series of vignettes that highlight the challenges of dealing with different types of insider threats. A defense contractor was selected as the base for the scenario because of its connectivity to the various stages of work on programs that present differing profiles for insider threats: basic research, classified R&D, production, an interface/embedding directly with the government.

Each of the vignettes that the participants receive focuses on a slightly different dilemma, ranging from potential espionage to information exposed due to human error, to the potential for workplace violence. For each, the differing C-Suite executives have their own individual goals and objectives based on their functions within the corporation - but also must have the wellbeing of the company, the employees, and the national security firmly in mind. Their objective is to propose a course of action to mitigate or resolve the issues raised in each vignette and present them to the facilitator. The learning objective is not to get a 'correct' answer to each dilemma - in some cases there might be multiple good solutions, or none. The objective, rather is to think more about how to tackle the hard problems in this space, and have an opportunity to learn through review and debriefing of the decisions they make.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Literature Review of the Relationship between Representations of the Self on Social Media and Representations of the Self Offline
Final Report

September 30, 2021

Long Doan[1]*, Paige Miller[1]
[1] Department of Sociology, University of Maryland, College Park
*corresponding author, longdoan@umd.edu

EXECUTIVE SUMMARY

Social media usage has been and continues to be on the rise (Auxier and Anderson 2021). Corresponding to this growth is increasing interest in better understanding how to use user-generated content like social media posts to infer information about the users that posted this content. Organizational scholars seek to understand how social media posts can foretell potential insider threats (Legg et al. 2013). However, little is known about the relationship between online content and offline selves. There are compelling reasons to expect that online and offline selves are different (Kozyreva et al. 2020). Online spaces are inherently different from and more constrained in many ways than offline spaces. Accordingly, unexplored discrepancy between online and offline selves may lead some factors to have inflated correlations to risk while other factors may have depressed correlations. Reviewing the work on the relationship between online and offline selves allows programs to better evaluate social media information and correct for distortions in the relationship between online selves and risk. To do this, we provide:

1. A systematic overview of the existing research across security, psychology, business, organizations, sociology and computer science disciplines,
2. Translation and synthesis of these disparate strands of research, and
3. Identification of feasible research questions that remain unanswered given this cumulative knowledge.

Three main findings emerge from our systematic review of the literature on online versus offline selves:

1. Online content can be used to determine valuable information about people's personalities,
2. Online self-presentations are generally accurate, and instances of deception are limited in scope and are often reactions to perceived barriers to social interactions rather than a desire to deceive, and
3. Online behaviors influence offline behaviors and vice versa.

However, there is emerging evidence as well as well-documented pitfalls with that comes with any operationalization of these findings. Specifically, many studies in this literature rely on self-selected and convenience samples to draw inferences and either treat demographic differences as unimportant or nonexistent. However, there is evidence that online users are not a monolithic group and there are indeed important demographic differences among users. For example, the research shows that:

- Gender, race, culture, age, and cohort all create "digital divides" that shape how and how much information people post online
- There are large personality differences in the meaning and purpose of social media posts

As such, using social media information, no matter how accurate, for personnel vetting and risk assessment should be balanced with the very real and easy to imagine ethical and practical dilemmas:

- Personnel need to provide consent to their social media information being used, potentially against them. Although this consent is relatively easy to obtain from the federal workforce, consent is harder, if not impossible, to obtain from those who are connected, knowingly or unknowingly, to the focal person. This presents at least two problems:
    - Studies demonstrating the effectiveness of social media information in predicting personality are based on complete data. Censored data provided to the government due to the lack of third-party consent may be limited in predictability at best and even potentially misleading.
    - even if the data is adequately censored by vendors collecting social media information before passing them on to the government, it is difficult to monitor what these companies can and will do with the raw, uncensored data.
- There is a large potential for the unequal policing of certain subgroups compared to others. Existing data suggests that there will be strong differences in the likelihood of some groups being flagged by a risk assessment tool. Given that two people are of equal risk of being a malicious insider, the one who posts more frequently on social media will be more likely to be detected than the one who posts less frequently.
- Studies demonstrating that social media information can be used to predict personality, including "dark" personalities shown to be predictive of insider risk, are based on population aggregates. It is a stretch, then, to apply these findings to assessment and detection tools designed to predict individual-level risk.
- Machine learning and artificial intelligence, although predictive, cannot provide clear and convincing rationales for flagging someone as high risk. Black-box theorizing based on outputs leads to post-hoc explanations that are neither satisfying from an academic perspective nor concrete enough from a legal perspective. An algorithm can flag someone as being high risk, but if it cannot tell us *why* that person is high risk, it is of limited value.

Given the patterns identified in our literature review, we propose several key directions for future research:

1. Gaining a deeper understanding of subgroup differences and how various groups differentially use social media is needed to better contextualize and understand research findings.
2. The rise of social networks like Instagram and others focused on pictures and videos presents challenges that need to be addressed. Whether and how people represent themselves on these newer platforms compared to both more text-based platforms and offline interactions needs to be better understood.
3. The rise of "alternative" social media platforms that are not technically different but caters to specific subgroups as opposed to the general population, like Parlor, also complicates the relationship between online and offline selves.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Personnel Vetting Vendor Evaluation: Evaluating Publicly Available Electronic Information (PAEI) Vendors to Enhance Personnel Security Vetting
Fiscal Year 2021 Final Report

September 30, 2021

Valerie Novak[1]*, C. Anton Rytting[1]*, Judy Philipson[1]
Sara McConnell[1], Victor Frank[1], Bernadette Jerome[1], Kelly Jones[1], Adam Russell[1]

[1]Applied Research Laboratory for Intelligence and Security
*corresponding author, vnovak@arlis.umd.edu;
*corresponding author 2, crytting@arlis.umd.edu

Note: the executive summary of this report is UNCLASSIFIED, thus is included in this summary document. However, the main body of the report contains CUI//PROPIN.

EXECUTIVE SUMMARY

Hiring and sustaining an uncompromised and, ideally, uncompromisable workforce requires USG to be able to collect and analyze different types and sources of data rapidly, accurately, reliably, and fairly, which when combined, may reveal something about a person's trustworthiness and suitability for a career in national security. However, this requirement is premised on many assumptions, one of which is that the tools and methods used to collect and analyze those data are themselves accurate, reliable, and unbiased.

In May of 2016, the Security Executive Agent Direction-5 (SEAD-5) authorized the collection, use and preservation of social media information for purposes of security clearance vetting (Office of the Director of National Intelligence, 2016). Yet despite the advantages that this opportunity may present for personnel vetting and Trusted Workforce 2.0 (Center for Development of Security Excellence, 2021), identifying the right signals within the vast amounts of noise has proven to be a significant challenge. Most notably, employing this data in a real-world vetting scenario would require data orders of magnitude greater in scale than current practices.

Publicly Available Electronic Information (PAEI) vendor companies that perform web scraping services claim that their tools are accurate, scalable, unbiased, and operationally applicable. However, these assertions have yet to be tested in an impartial, objective, reproducible and scientific way. ARLIS has designed a study and performed the data collection portion of the study to evaluate three vendors and their tools for the task of personnel vetting.

In conjunction with the USG, ARLIS researchers designed a study to evaluate the ability of PAEI vendors to support and enhance personnel vetting. The six areas of evaluation include: 1) the ability of PAEI vendors to accurately identify the correct subject; 2) the ability of vendors to identify derogatory data in accordance with the Security Executive Agent Directive-4 (SEAD-4)'s Adjudicative Guidelines (Office of the Director of National Intelligence, 2017) and derived "Business Rules" about that subject; 3) the usability of the documents provided by the vendors; 4) the ability of the vendors to redact personally identifiable information (PII) of *non*-subjects; 5) the ability of vendors to access multiple and varied data sources; and 6) the potential for vendors to scale this process to accommodate large number of subjects in a relatively short period of time. Another significant criterion that could be considered in future Periods of Performance (PoP) would be to explore the correlation between the USG's Adjudicative Guidelines/Business Rules ("derogatory information") with the credibility of the individual, and therefore validate through research the effectiveness of the Adjudicative Guidelines as the basis for personnel vetting.

Per the TO, we solicited companies to participate in the evaluation on a voluntary basis, and from the ten companies who volunteered, ARLIS and the sponsor jointly selected four potential

performers based on preliminary information they provided based on criteria developed per the sponsor's identified operational needs. Due to unanticipated delays in receiving the data, ARLIS was unable to annotate the data in the current Period of Performance (PoP). Because of this limitation, we provide descriptive statistics on data returned by each of the three vendors, as comparisons between vendors cannot be made until annotation occurs. ARLIS proposes to annotate the vendor data in a subsequent POP and as a result, may gain further insights.

In the main body of this report, ARLIS presents preliminary findings and recommendations about the next phase, which may require changes to evaluation methods. For instance, the methodology for identity resolution may be strengthened to compare the subject account information (i.e., related email addresses, Facebook account, etc.) across vendors, rather than an evaluation on the level of document or social media post as originally planned. No vendor chose to redact their PDFs of non-subject PII, so the evaluation of redaction can either be discarded or altered to indicate the amount of non-subject PII found in vendors' returned PDFs.

Additionally, population sampling and, consequently, power analysis of the dataset will also need to be re-evaluated. ARLIS created a data sampling plan, in the case that vendors provided more relevant data than could be annotated within the timeframe, which found that the annotation of 130 subjects' data could reasonably demonstrate differences of performance metrics within and across vendors, *so long as each of those 130 subjects had some adjudicatively relevant document or information returned for them*. This number was based on *(1) limitations in prior research, (2) initial expectations of data received from vendors, (3) perceived background of sample subjects, and (4) ARLIS researchers' assumptions based on vendor's demonstrations*. The limited output of adjudicatively relevant information returned so far for 3,000 cleared subjects, and the uncertainty of identity resolution, makes it difficult to draw any conclusions at this point. Lastly, the evaluation of scalability will need to be reconsidered and may be out of scope for ARLIS researchers, based on preliminary inquiry of results.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

Technology Evaluation for Voice Analysis Tools That Identify Potential Risk
Final Report

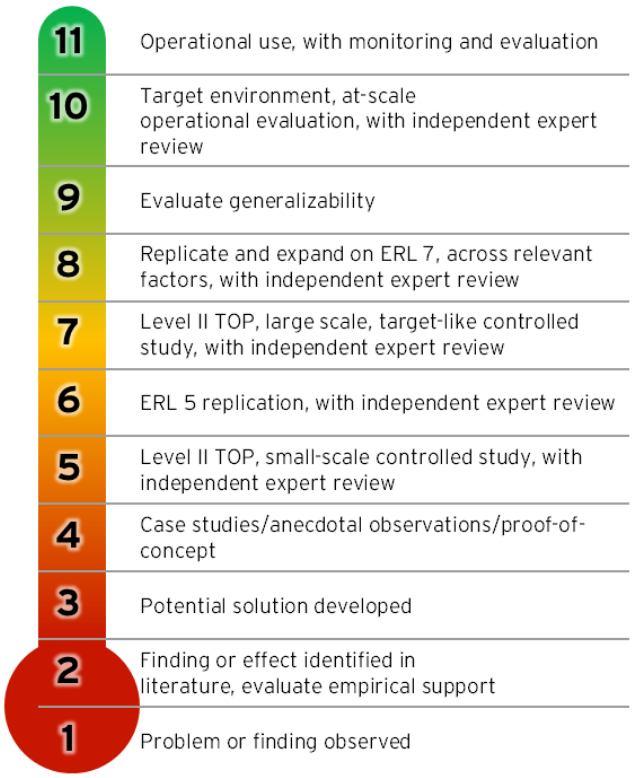
September 30, 2021

Meredith Hughes, MA[1]*, Alexandra Maddox, PhD[1], Carol Espy-Wilson[2], Polly O'Rourke, PhD[1], Breana Carter-Browne, PhD[1], Kelly Jones, PhD[1], Adam Russell, PhD[1]
[1]Applied Research Laboratory for Intelligence and Security
[2] Electrical and Computer Engineering, The Institute for Systems Research
*Meredith Hughes, mhughes@arlis.umd.edu

Note: the executive summary of this report is UNCLASSIFIED, thus is included in this summary document. The main report is also UNCLASSIFIED; however the Addendum contains UMD and Company PROPIN.

EXECUTIVE SUMMARY

Identifying those who may pose an insider threat/risk is costly, both in terms of detection efforts and the potential costs when detection fails. Innovations in physiological measurement using new technologies (e.g., novel or more affordable measurement and analysis technology, application of artificial intelligence/machine learning analysis) have the potential to positively influence the insider treat/risk domain by providing additional metrics in a cost-effective and at-scale manner. To support the Office of the Undersecretary of Defense for Intelligence and Security, OUSD(I&S), in evaluating insider risk detection tools, ARLIS was tasked to develop and provide a framework for future independent testing and evaluation (T&E) of new and emerging technologies in insider threat detection, deterrence, and mitigation, along with a T&E protocol for such technology that would assess its strengths and limitations for use within the DoD. Per the task order, this effort focused on evaluation of a particular technology from a particular company. However, the company declined to participate in the proposed experimental protocol, thus the evaluation is based on materials provided, review of the literature, and ARLIS proposed Evidence Readiness Levels framework.



Evidence Readiness Levels

The Evidence Readiness Levels (ERL) framework (**Figure 1**) is proposed as a tool to enable and encourage intentional and thoughtful innovation—for the Insider Risk/Threat domain, but also for the social sciences in which it is more difficult to define and evaluate explicit measurement criteria (e.g., psychology, political science). While the milestones may look different for these sciences, it is important to define them so potential users can judge the maturity of tools, technologies, and theories for their purposes: they are crucial to understanding the *Human Domain*, where understanding human sociotechnical diversity, networks, systems, beliefs, strengths, limitations, and vulnerabilities confers advantage (and, likewise, not understanding confers a disadvantage).

Figure 1. Evidence Readiness Level (ERL) Scale.
***Note*. TOP = Open Science Framework (OSF) Transparency and Openness Promotion guidelines**

The proposed ERL framework draws on pertinent elements of the NASA Technology Readiness Levels[9] as well as several other scales and guidelines that have been proposed for social/behavioral sciences. The framework emphasizes (1) open science practices and (2) successful replication of results. We refer to the Open Science Framework (OSF) Transparency and Openness Promotion (TOP) guidelines,[10] such that advancing ERLs requires transparency and openness about the science performed: to truly evaluate the evidence supporting solutions, **evaluators and potential users must be able to read and understand the research** that has been performed. Replication across relevant factors stresses the importance of testing and communicating the limits/boundaries of the current science: **understanding the knowns and the (so far) unknowns** about the conditions under which the tool can perform.

Background Research

The company states that their tool indicates risk using indicators of 'acute stress' and 'cognitive effort.' In a review of literature several key acoustic parameters were found that could be strong indicators of cognitive load and emotional stress in the voice. Publicly available information was analyzed to assess the ability of promising parameters, and analyses showed that there was support for several parameters to have reliable, observable differences under varying degree of stress and/or cognitive effort. Overall, these were taken as positive indications that a voice analytic tool could potentially be used for making risk-assessments, if the assumption is that deceptive responses are correlated with higher cognitive load is shown to be justifiable. Note, that because the evaluated tool's input parameters and processing were protected from ARLIS as trade secrets, it not clear how relevant our findings are to this specific tool.

Major Recommendations

1. Adopt usage of the Evidence Readiness Levels.

We encourage USG and other agencies to adopt the use of the proposed ERLs when assessing the maturity of technological solutions in this domain. A common, objective scale may be used as a common short-hand for potential users of technology in understanding what level of evidence has been gathered so far in supporting the use of the tool. Potential users can then weigh the evidence against their needs and potential risks in deploying a less-mature tool for operational use.

2. Continue research in voice analytic risk-assessment tools

Background research in the use of voice/speech analysis for Insider Risk/Threat detection is promising: there is evidence of observable voice/speech changes associated with increased

---

9 See https://www.nasa.gov/directorates/heo/scan/engineering/technology/technology_readiness_level

10 The Level II TOP guidelines are to (1) cite data and materials appropriately; (2) post data, code, and materials to a repository or note exceptions; (3) adhere to publication design transparency standards; (4) state whether preregistration—with or without analysis plan—exists and allows verification if so; (5) encourages replication studies and results-blind review. https://www.cos.io/initiatives/top-guidelines.

cognitive effort and when under stress. While additional research is warranted, it is also necessary: the next recommended step to advance the body of evidence supporting the tool is a controlled, experimental study, following transparency and openness guidelines to support external evaluation of its evidence. A protocol for a study which would provide such evidence is proposed, including supporting research materials.

3. Expand research on innovative Insider Risk/Threat physiological tools

Finally, we recommend expanding the scope of the Insider Risk/Threat detection tool assessments. While voice/speech analytics have empirical support, there are additional types of physiological measurement that may be used separately or together to improve current procedures. Widening the scope of tool types, particularly if different tool types are compared across more similar protocols, may better inform funding and use decisions by allowing for more direct comparison of results.

APPLIED RESEARCH LAB
FOR INTELLIGENCE
AND SECURITY

A Course on Understanding Insider Threat: From Threat to Risk and Trust Final Report

September 30, 2021

Steve S. Sin[1]*, Judith A. Philipson[2], Juliet R. Aiken[3], Liberty Day[1]
[1]National Consortium for the Study of Terrorism and Responses to Terrorism
[2]Advanced Research Lab for Intelligence and Security
[3]Conducere, LLC
*corresponding author, sinss@umd.edu

UNCONVENTIONAL **WEAPONS AND TECHNOLOGY**

EXECUTIVE SUMMARY

The National Consortium for the Study of Terrorism and Responses to Terrorism (START), Applied Research Laboratory for Intelligence and Security (ARLIS), and Conducere, LLC developed and delivered a fully online course that blends synchronous and asynchronous components on insider threat that is approximately 60 contact hours – granting 6 continuing education units (CEUs). The course included live and recorded lectures; live class discussions and activities; group in-class simulations; participation in guided online discussion forums; completion of reading materials; and participation in course evaluation processes.

Insider threat has become a common lexicon in our society today – most commonly linked with information leak and active shooter incidents. While insider threat is not a new phenomenon, there is a need for a better understanding of the phenomenon and the ways to mitigate it. This course took the form of a survey course where we examined individual, organizational, and social stressors that could contribute to insider behaviors. The course then discussed past and present trends of insider threat, response and mitigation challenges, as well as policies, procedures, and practices currently implemented (within the government and industry) to respond to and mitigate it. Finally, the course explored the systems approach to manage insider threat vis-à-vis risk assessments and risk management. In doing so, the course exposed the students to the new paradigm of thinking - shifting from insider threat to insider risk and from countering insider threat to mitigating insider risk.

By the start of course, 25 individuals had registered for the course and 63 individuals had been put on the waiting list. For the 25 individuals registered for the course, 16 individuals were U.S. students while nine were international students. For the nine international students, two were identified as being affiliated with foreign governments; three were identified as being affiliated with international organizations; three were identified as being affiliated with private industry; and one was a full-time student.

The overall impression from the post-course survey was that the course conveyed new information and that the students learned the material well, and were leaving the course with new perspectives on how to think about insider threat and how to mitigate insider threat in the future. Specifically, 100-percent of the students responding to the post-course survey strongly agreed with the statement, “I would recommend this course to my peers.” Additionally, 87.5-percent stated that they strongly agreed with the statement “overall, the course met my needs and expectations.” Finally, 87.5-percent stated that they strongly agreed with the statement, “Overall, the course increased my knowledge, skills, and abilities relevant to the course topics.”

Finally, the participants showed improvements in its knowledge on insider threat and mitigation thereof. The class average for the pre-test at the beginning of the course was 9.13 out of 15 (63.42%). The class average for the post-test, on the other hand, was 13.23 out of 15 (88.21%), showing a 24.79-percent increase in the average score and a general increase in the students’

knowledge level on topics covered during the course. Of note, all participants scored higher on their post-tests compared to their pre-test scores. These results indicate the course contributed to student knowledge gain in the subject matter.