

Domestic and International Utilization of Social Media Monitoring for Trustworthiness and Reliability Assessments: Regulations, Practices, and Targets



John E. Landers, PhD
Jessica A. Baweja, PhD
Justin R. Kinney, PhD
Marc R. Fialkoff, JD, PhD

July 2023



DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via OSTI.GOV.

Website www.osti.gov

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Website <http://classic.ntis.gov/>

Reports are available to US Department of Energy (DOE) employees, DOE contractors, Energy Technology Data Exchange representatives, and International Nuclear Information System representatives from the following source:

Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Website <https://www.osti.gov/>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Nuclear Nonproliferation Division

**DOMESTIC AND INTERNATIONAL UTILIZATION OF SOCIAL MEDIA
MONITORING FOR TRUSTWORTHINESS AND RELIABILITY ASSESSMENTS:
REGULATIONS, PRACTICES, AND TARGETS**

John E. Landers, PhD*
Jessica A. Baweja, PhD†
Justin R. Kinney, PhD*
Marc R. Fialkoff, JD, PhD*

* Oak Ridge National Laboratory

† Pacific Northwest National Laboratory

July 2023

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

EXECUTIVE SUMMARY	iv
1. INTRODUCTION	1
2. SOCIAL MEDIA: BACKGROUND.....	2
3. TRUSTWORTHINESS AND RELIABILITY: DISINHIBITION EFFECT	3
3.1 BEHAVIORS OF CONCERN.....	4
3.1.1 Personal Conduct, Alcohol Consumption, Drug Involvement, and Criminal Conduct.....	4
3.1.2 Psychological Concerns	4
3.1.3 Violations of Professional Standards and Organizational Fidelity	5
3.1.4 Other Personal Conduct	6
3.1.5 Trustworthiness and Reliability: Violence and Social Media.....	6
4. CURRENT SOCIAL MEDIA MONITORING PRACTICES	7
4.1 INDUSTRY OR SECTOR.....	8
4.1.1 Private Industry and Corporations	8
4.1.2 Government.....	9
4.1.3 Capabilities	10
4.1.4 Content.....	12
5. LEGAL AND REGULATORY OVERLAYS FOR SOCIAL MEDIA MONITORING DURING TRUSTWORTHINESS DETERMINATIONS	13
5.1 CONSTITUTIONAL CONSIDERATIONS: DUE PROCESS.....	14
5.2 STATUTORY FRAMEWORK FOR OBTAINING A SECURITY CLEARANCE IN THE UNITED STATES	14
5.3 INCLUSION OF SOCIAL MEDIA AND CYBERVETTING IN THE SECURITY CLEARANCE PROCESS	15
5.4 PUBLIC VS. PRIVATE SECTOR USE OF SOCIAL MEDIA FOR HIRING and TRUSTWORTHINESS DETERMINATION	16
6. ETHICAL AND PRACTICAL CONSIDERATIONS IN SOCIAL MEDIA MONITORING.....	17
6.1 CONCERNING CONTENT.....	17
6.2 POTENTIAL FOR BIAS.....	17
6.3 RIGHTS TO PRIVACY	17
6.4 SELECTION FOR SOCIAL MEDIA MONITORING	18
7. SOCIAL MEDIA VETTING FOR NATIONAL SECURITY POSITIONS: BEST PRACTICE RECOMMENDATIONS.....	19
8. CONCLUSION.....	21
9. REFERENCES	22

EXECUTIVE SUMMARY

This paper explores the potential benefits, challenges, and best practices of implementing social media vetting for individuals in national security positions. As social media usage continues to increase, it may offer valuable insights into an individual's trustworthiness, reliability, and suitability for positions involving sensitive information and critical infrastructure. However, several ethical, legal, and practical considerations must be addressed to ensure social media monitoring is integrated into personnel vetting processes responsibly and effectively.

Key recommendations include the following:

- Develop clear policies and procedures that comply with federal, state, and local laws and international standards. Regularly review and update vetting procedures to reflect best practices and ensure compliance with legal and ethical standards.
- Prioritize transparency, fairness, and respect for individual privacy in screening processes. Limit data collection to publicly available information, inform affected individuals about the scope and purpose of the monitoring program, and establish guidelines to protect privacy rights.
- Address the risk of intentional and unintentional bias in manual and AI-based vetting approaches. Train HR personnel and hiring managers on anti-discrimination laws, privacy concerns, and techniques to minimize unconscious biases. Continuously improve AI algorithms to reduce bias and ensure alignment with organizational values and legal requirements.
- Apply a graded approach when choosing individuals for social media monitoring by balancing the resources applied against the potential risk. Use various risk determinations, such as position risk or person-based risk, or a random-selection approach to identify monitoring targets. Make informed decisions about the number of individuals to be monitored and the frequency of monitoring, considering the workload and resources available.
- Adopt best practices for social media vetting, which include policy and procedure development, training and education, third-party involvement and verification, documentation and transparency, and relevance and fairness in decision-making. These practices ensure that the vetting process is compliant with legal requirements, promotes transparency, minimizes privacy violations, maintains fairness and objectivity, and verifies information from reliable sources.

When implemented thoughtfully and ethically, social media vetting can be a valuable tool in assessing the trustworthiness and reliability of individuals in national security positions. By adopting a mindful approach and following best practices, organizations can mitigate potential risks and enhance the effectiveness of their personnel vetting processes, thus contributing to a stronger security posture and safer environments.

1. INTRODUCTION

Trustworthiness and reliability are desirable personal attributes in any setting in which security and safety are a priority. Because of the high risk associated with mishandling and misuse of nuclear and radiological material, personnel employed in positions of trust with access to, knowledge of, and authority over such materials must be trustworthy and reliable. The International Atomic Energy Agency (IAEA) and its Member States have established best-practice guidelines to ensure the safety and security of nuclear material (IAEA 2011). The guidelines specific to trustworthiness and reliability involve vetting personnel before they are granted special access, authority, and knowledge. Although the specific means and methods for vetting vary internationally, any effective approach to ensuring individual trustworthiness and reliability must consider previous behaviors. Therefore, methods for gathering information regarding individual behaviors are of interest during the vetting process.

The formal vetting process in the nuclear industry began in the United States during the Manhattan Project (Piccolo 2022). Potential employees were subject to FBI background checks to ensure they had no known history of criminal behavior. Additionally, indicators of divided loyalties were considered, such as unauthorized associations with Axis sympathizers. The FBI considered information from multiple sources beyond public records, including interviews with previous employers, neighbors, and family members. Once employees were cleared and granted access, their behavior was continually scrutinized. For example, correspondence entering or leaving the work site was screened. After World War II, the US government established the Atomic Energy Commission (AEC), the US Department of Energy's (DOE's) predecessor, and vetting has continued in the nuclear industry since; however, the process has evolved. International safety and security standards were adopted in the nuclear industry in the 1950s, after the creation of the IAEA as an international body to regulate the use of nuclear materials and technology having been proliferated through the Atoms for Peace program. Since this time, the IAEA has recommended and its Member States individually have developed vetting programs for personnel working in the nuclear industry.

As the vetting process evolves, so do opportunities for observing behavior in and out of the workplace to verify worker trustworthiness and reliability. For example, just as correspondence was monitored during the Manhattan Project, work-related email is now subject to monitoring for employees in particular positions of trust, and since the advent of social media, consideration has continued regarding how this new media could and should be used to assess potential or current employees' trustworthiness and reliability. A recent incident involved a student at the University of Utah who posted a threat on Yik Yak, stating, "If we don't win today, I'm detonating the nuclear reactor on campus" (Tumin 2022). Yik Yak is a location-based social media app that allows people to create and view discussions within a 5-mile radius, often anonymously, popular among college and university students. Fortunately, someone who read the post reported it to local law enforcement, and a swift investigation led to the arrest of the student and elimination of the threat. In a recent study published by the FBI on violent extremism, 92% of individuals motivated by radicalized ideology to act violently publicly posted information related to their intent before acting on it (Richards et al. 2019).

Additionally, recent polling revealed that 70% of employers assess social media profiles of job applicants as a prehire screening strategy, and a majority reported having used applicants' social media activity as a basis for removing them from consideration. The poll also revealed that many employers assess current employee social media activity and have taken disciplinary action based on the content (Hayes 2018). Regarding the modern vetting process for high-security positions in the United States, the Office of the Director of National Intelligence (ODNI) and the Office of Personnel Management (OPM) have provided direction for the use of social media in background investigations in the Security Executive Agent Directive (SEAD) 5 policy statement (ODNI 2016).

Because of the increasing interest in and practice of social media monitoring, the National Nuclear Security Administration's Office of International Nuclear Security has commissioned a study of best-practice considerations for using social media to determine levels of trustworthiness and reliability for certain personnel. This paper provides a comprehensive overview of regulations and practices regarding the use of social media as a source for vetting as well as recommendations based on the findings of this project.

2. SOCIAL MEDIA: BACKGROUND

This paper defines *social media* as “forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos)” (Merriam-Webster n.d.). The origins of social media can be traced to the formative years of the internet when researchers initially developed online platforms for collaboration. However, based on the current definition, the first social media platform was launched in 1997 by SixDegrees. This platform allowed members to create personal profiles, prioritize connections to specific individuals, and communicate through a messaging system. Thereafter, some social media sites (e.g., MySpace, Facebook) were created that essentially duplicated and competed with earlier platforms, whereas others were designed for niche audiences and specific purposes (e.g., LinkedIn, Pinterest, YouTube). Social media usage has grown exponentially over the past two decades, and nearly five billion users, or approximately 75% of the world's adolescents and adults, have accounts on at least one platform (Datareportal 2022).

Several types of social media platforms exist: social networking, media sharing, discussion forums, bookmarking and content curation, consumer reviews, blogging and publishing, social shopping, and interest-based networks. Each type of social media is used differently depending on the user and their purpose. For example, a business may have accounts on Facebook and Instagram because they recognize that the user demographics differ between the platforms (Jackson 2021), and being engaged on both platforms allows for broader and targeted marketing approaches. Alternatively, an individual user may post only personal information on Facebook and professional or career-related information on LinkedIn. Platforms such as Facebook, LinkedIn, Twitter, Instagram, Snapchat, YouTube, Reddit, Pinterest, Yelp, Tumblr, Tripadvisor, Goodreads, and Etsy are familiar and generally considered safe because they are moderated. However, other social media channels are designed to implement little, if any, verification of user identity and minimal moderation. These platforms, such as Yik Yak, 4chan, Formspring, and Ask, have a potential prosocial purpose, such as providing a safe space for discussion of difficult topics, but have also become controversial because of antisocial and illicit content being shared.

The accessibility of content on the internet varies depending on where it is located. The levels of accessibility have become known as open web, deep web, and dark web. Open web content can be accessed by anyone with a standard browser and clearly meets the SEAD 5 criteria for publicly available social media information. The deep web can only be accessed with specialized permissions or knowledge because the information is not searchable through regular search engines. The dark web comprises anonymously posted information that can only be found using specialized tools. Thus, deep- and dark-web information, though potentially relevant to the question of reliability and trustworthiness, is outside of the SEAD 5 permissions for use in personnel vetting. Social media information and applications exist at all structural levels on the internet.

In terms of accessibility, only sites that are associated with users' offline identity can be used for vetting. Thus, an individual subject to vetting would circumvent scrutiny by using an anonymized channel such as 4chan. Individuals may also engage in questionable behavior on personal devices (e.g., mobile phones) through a social media platform (e.g., SnapChat) that does not allow for public monitoring. However, public-facing social media users do not necessarily restrict their undesirable behavior to private social media

forums, as was demonstrated in a recent study (Sikkens et al. 2016) in which Facebook, a public platform, was used to recruit participants who had endorsed radical ideology in their posts. According to the Pew Research Center, 41% of Americans report experiencing online harassment, which is defined as offensive name-calling, purposeful embarrassment, stalking, physical threats, sustained harassment, and sexual harassment (Vogels 2021). This survey also revealed that 75% of the online harassment occurred on public-facing social media. Thus, review of public social media behavior could reveal inappropriate behavior from a percentage of those who participate in such online behavior.

The primary reported purpose for social media use is to keep in touch with friends and family; other top reasons include filling spare time, reading news stories, and looking for specific content (Datareportal 2022). Social media has certainly simplified connecting with new friends and networking with communities that share mutual interests. It also provides a platform for any individual to share their thoughts and market ideas and talents, thus leveling the playing field for the average citizen in comparison to those with more resources, including corporations. Business owners can now use social media to economize marketing, target specific audiences, and optimize messaging. However, as with any new technology, social media can be divorced from its original positive purposes and used to further a negative agenda. For example, anonymity combined with global reach enables like-minded individuals with antisocial and extremist ideas who were previously separated by physical distance to connect online. Recruitment of individuals to extremist causes is also facilitated through social media (Sikkens et al. 2016). With no corrective feedback, social media can become a self-reinforcing echo chamber that enables further radicalization and provides motivation for malicious action (Hassan et al. 2018). For this reason, social media monitoring is being considered as a means of vetting potential candidates for positions of trust.

3. TRUSTWORTHINESS AND RELIABILITY: DISINHIBITION EFFECT

Online interactions and activities differ from those that occur in person because being online lowers people's natural inhibitions, emboldening them to say and do things they otherwise would not. Although this online disinhibition effect can be used positively—for example, to create a safe environment wherein people who are shy or lack a protected in-person outlet can interact (Lapidot-Leffler and Barak 2015)—it more frequently makes people feel more comfortable violating social standards by engaging in cyberbullying and making racist, sexist, or otherwise bigoted or rude comments. Posters such as the previously mentioned University of Utah student may also make inappropriate or dangerous jokes and feel immune to consequences.

Many reasons for the online disinhibition effect have been proposed. Noting the abundance of anonymous and fake accounts, some researchers point to the relative anonymity of internet users as a potential driving factor in disinhibition (Lapidot-Leffler and Barak 2012). Adopting a new or unknown persona can make someone feel safe online from any serious repercussions for their actions. Additionally, because their tone of voice and body language are not observable, anonymous or fake users can easily misrepresent themselves and others; therefore, many view online space as a game wherein real-world rules do not apply (Suler 2004). Other researchers argue that asynchronous communication (i.e., communication that does not necessarily happen live) contributes to disinhibition. Because it can take time for messages to be received and read, the consequences of a comment can seem distanced from the comment itself, making it easier for a person not to consider how what they say will affect others. Empathy deficits and other individual personality and cultural factors have also been noted in literature as contributors to online disinhibition (Terry and Cain 2016; Antoniadou, Kokkinos, and Markos 2016).

Another proposed reason for this effect is that social media and online interaction distort several normal assumptions about communication. The internet minimizes the space between ages, levels of authority, and notoriety; people can interact with one another more as equals than they would in real life (Suler

2004). Therefore, inhibitions related to those social distances are minimized, leading to more aberrant behavior. Over time, humanity has shifted to communication that involves less face-to-face interaction—from letters to telephones to online communication. All these changes have altered the kinds of information about each other that people process; by shifting the types of interactions in which people are engaging, social perception and cognitive processing of those interactions are also transforming (Sparrow and Chatman 2013).

Regardless of the reason for this disinhibition effect, social media changes the conversational tone, encouraging people to express more extreme emotions. Anger intensifies, people say increasingly provocative things which generate interaction, and over time through that feedback loop, people learn to express more rage (Brady et al. 2001). And cruelty and anger spread faster than positivity and joy on social media (Fan, Xu, and Zhao 2016), creating more disinhibition and malicious communication. This growing online outrage can have powerful societal consequences, including sparking action for social change and producing political polarization (Spring, Cameron, and Cikara 2018). Therefore, online behavior involves different incentives, motivations, and outcomes. It produces a level of manufactured—and potentially false—privacy by removing certain normative barriers and may reveal concerning details about a person's trustworthiness and reliability that warrant further investigation.

3.1 BEHAVIORS OF CONCERN

Many types of social media behavior could be cause for concern and affect reliability and trustworthiness. In 32 CFR Part 147 (2012), “Adjudicative Guidelines for Determining Eligibility for Access to Classified Information,” 13 areas of consideration are listed, including foreign influence, personal conduct, alcohol and drug involvement, and misuse of information technology systems, some of which are addressed below.

3.1.1 Personal Conduct, Alcohol Consumption, Drug Involvement, and Criminal Conduct

Many organizations and companies treat employees' off-the-clock behavior as distinct from their performance during working hours. However, social media complicates such distinctions. Conduct on social media can create problems for companies and reveal behavior that disqualifies employees or applicants from employment. Simple examples include indications of substance abuse or criminal behavior. When evidence of alcohol or substance abuse or active criminal activity—especially internet-based criminal behavior such as hacking—appears online, it can create concerns for employers, particularly in career fields in which addictions, arrests, or blackmail material could be exploited by malicious actors. And literature suggests that once a crime has been committed, it becomes much easier to commit another, meaning past behavior could indicate future problems that may be more directly damaging to the employer or national security (Welsh et al. 2015). Social media monitoring can also reveal weak character, poor decision-making, vulnerabilities that could be exploited, or a pattern of behavior that might be inappropriately repeated on-site.

3.1.2 Psychological Concerns

Certain psychological concerns can manifest online through social media behavior. As stated in 32 CFR Part 147 (2012), certain conditions or disorders “can cause a significant deficit in an individual's psychological, social, and occupation functioning...[and] may indicate a defect in judgment, reliability, or stability.” Social media behavior can reveal current patterns of aggressive, emotionally unstable behavior or suggest that a person is not following medical advice for treatment, which could raise security concerns. When a psychological concern starts affecting off-the-clock, online behavior, the possibility that the concern will affect on-the-clock reliability increases. Such a concern could lead to poor decision-making, unpredictable behavior, or inadvertent mistakes.

Radical ideology is also a significant psychological concern. Social media has become a platform for recruitment and radicalization through deliberate efforts by malicious groups like terrorist organizations and inadvertently by creating a positive feedback loop that influences vulnerable individuals susceptible to radicalization (Ledwich and Zaitsev 2019; Schmitt et al. 2018). Social media enables connection among people who have propensities for radical thought or already possess radical beliefs they may never act on without reinforcement. Additionally, numerous social media platforms employ automated algorithms that create recommendation bubbles. These bubbles ensure that individuals primarily encounter viewpoints aligned with their preferences based on their interactions. While users intentionally shape their experience by choosing to follow certain pages, these algorithms also operate in the background, curating content that feels familiar to the user.

The relationship between social media and radicalization remains a subject of ongoing debate. Whereas some argue that it can contribute to online radicalization, others believe that its impact is exaggerated (Ledwich and Zaitsev 2019). However, the potential for radicalization varies across different platforms. Attempts have been made to exploit patterns in extremist messaging through use of pattern recognition in communications and then use these algorithms for counter-messaging and preventing radicalization (Schmitt et al. 2018). Yet, extremists have simply adapted their use of terminology and created new accounts, making it challenging to utilize data analytic approaches to detect and then counteract these extremists (Fernandez and Alani 2021). As extremist beliefs emerge and gain traction, they can pose challenges for employers, particularly in sensitive career fields. Importantly, holding a viewpoint outside of mainstream societal beliefs does not automatically indicate a threat, and an extreme, unusual, or unpopular belief is not necessarily harmful. Many unpopular beliefs never manifest in damaging consequences. However, extremism can divide individuals' loyalties, causing a split between their prosocial obligations (e.g., job, and family) and their divergent belief system.

3.1.3 Violations of Professional Standards and Organizational Fidelity

Signs of inappropriate professional conduct or foreign influence can also emerge through social media posts and messaging that may warrant monitoring. Lacking self-control regarding disclosure of secure information related to security clearance holders or professional boundaries, discussing sensitive job details in public (even if the details are not classified), or having many foreign contacts all could indicate vulnerabilities.

Inadequate personal security controls surrounding professional or work-related topics can create leverage that adversaries could use against the individual and the organization. Disclosing, even by accident, classified or official use only information can indicate potentially disqualifying problems or at the very least suggest carelessness for work-related topics. Such disclosure could be as simple as a person bragging about their level of access or power to win a social media argument, or it could be as serious as posting a photo or video taken in the workplace that reveals sensitive information or equipment. Monitoring conduct of this nature may be particularly important in career fields that involve clearances or handling delicate information, but the importance of adequate personal security controls is broadly applicable. From restaurant staff who joke about spitting in customers' food to a nuclear engineer accidentally revealing a piece of secret technology in the background of a workplace selfie, a casual disregard and disrespect for professional decorum can be problematic and quickly lead to job disqualification or revocation of clearances.

Excessive foreign influence through social media has the potential to affect an individual's professional trustworthiness significantly. Although having friends or family who are citizens of foreign nations is neither illegal nor inherently damaging, and maintaining relationships with foreign nationals is widely accepted in today's interconnected world, it is advisable to exercise caution and monitor certain types of close and continuous contact over social media. Unintentional as it may be, interacting online with

foreign nationals can create divided loyalty. Therefore, security concerns may be raised if an individual begins to exhibit a preference for a foreign country over the United States through their online activities, such as by expressing a willingness to bear arms for a foreign country, demonstrating financial or business interests abroad, or discussing attempts to influence foreign governments politically (e.g., by voting).

Moreover, close contacts and interactions with foreign nationals on social media increases the likelihood of inadvertent disclosure of sensitive information. If an individual follows numerous foreign accounts or actively engages with them on job-related topics, the potential for information leakage increases. In the United States, background investigations for security clearances encompass inquiries into close or continuing relationships with any foreign nationals. Although such relationships are not inherently disqualifying, they are examined to ensure they do not pose any potential harm to organizational or national security interests.

3.1.4 Other Personal Conduct

Many of the previously mentioned factors would be viewed as problematic by many organizations and worthy of potential disqualification from a security clearance or hiring. Although each instance would be assessed on a case-by-case basis, eligibility for clearance or hiring would be considered carefully for concerns of trustworthiness and reliability when any of the previously mentioned factors are present. However, other behaviors that take place on social media platforms could raise red flags but not warrant the same level of concern as the previously mentioned factors. Put simply, being a jerk does not help someone get hired or get a security clearance but also is not necessarily criminal, violent, or otherwise disqualifying.

However, many businesses and organizations implement codes of conduct in the workplace that employees must agree to after they are hired. If violations of this code are observed on social media, then disciplinary action may be taken. Unfavorable online behavior could generate disgruntlement among coworkers, particularly if management is seen as not addressing these interpersonal concerns. Because disgruntled workers are at risk for insider behavior, addressing offensive conduct would likely be worthwhile for any organization (Bell, Rogers, and Pearce 2019). When evaluating trustworthiness and reliability, hateful or offensive speech on social media is assessed in accordance with the laws of the relevant country. In the United States, except for rare cases such as direct incitements to violence, hateful or offensive speech is generally not illegal and therefore does not warrant criminal punishment. However, this does not mean it should not raise a red flag. Whether this type of speech is predictive of the speaker's future behavior is unknown. However, evidence is emerging that hateful speech can incite violent behavior among those who agree with the sentiment of the speech (Müller and Schwarz 2021). Additionally, such online behavior can certainly damage the reputation of an employer. Thus, hateful or offensive behavior is worthy of attention because, at minimum, it signals poor judgement.

3.1.5 Trustworthiness and Reliability: Violence and Social Media

Many studies have investigated the links between the internet and acts of violence, particularly among adolescents (King, Walpole, and Lamon 2007; Tokunaga 2010). In addition to teenage bullying and interpersonal violence, the media form is a particularly effective platform for radicalization. Extremists have increasingly used various media platforms such as Twitter and 4chan in their strategy for recruiting others to their cause virtually (Jensen 2016). Extremists often create modest communities that share nominally innocent content and ideas but introduce extremist ideas over time (Torok 2010). In this way, extremists create a sense of personal community among physically distant people, gradually nudging people toward commitments to malicious acts and violence (McCormick 2003). And when a person is targeted who has experienced a personal setback like a promotion denial, tragedy, or financial instability

or who is seeking social recognition or meaning in their life, extreme behavior committed on behalf of an ideological cause becomes more likely (Kruglanski et al. 2014). This is called the *quest for significance theory*.

Fathali Moghaddam (2005) further described a staircase process of radicalization involving disinhibition of aggression and violence, which social media can introduce. Others have suggested similar models; Randy Borum (2003) discussed the dehumanization necessary for beliefs to evolve into hate and radicalization, a process facilitated by the anonymity and increased interpersonal distance inherent in social media.

If a malicious external actor successfully engages with and recruits an employee who possesses sensitive knowledge, access, or authority, that actor can exploit the employee's vulnerabilities and endanger an organization, fellow employees, or broader security interests. As radical ideas take root and evolve, the individual's divided loyalty between their newly embraced radical beliefs and their professional responsibilities may render them susceptible to external influences, potentially turning them into an insider or even a lone actor. Some researchers argue that increased social media monitoring would not be an effective vetting methodology because users can conceal meaning and context and because of the overwhelming volume of posts to monitor; however, evidence suggests that warning signs do emerge on social media and are detectable (Peterson et al., 2023).

It may seem counterproductive for people to post their radical thoughts or plans for violence online, but roughly 92% of individuals motivated by radicalized ideology to commit violent actions communicated their grievances and radical ideologies to at least one other person prior to acting, and almost half posted extremist videos or writings online (Richards et al. 2019). For example, in 2018 Nicholas Cruz posted on YouTube that he wanted to become a "professional school shooter" (McLaughlin and Park 2018). Shortly thereafter, he shot and killed 17 people at Marjory Stoneman Douglas High School. Salvador Ramos threatened to conduct school shootings on the social media app Yubo (Medina et al. 2022); in Texas in 2022, he killed over 20 people in an elementary school. Moreover, people frequently discuss their violent thoughts and issue threats online but do not act on them, perhaps because the disinhibition effect causes people to say things that they would never act upon in real life. However, it occurs often enough to suggest that individuals gain something from posting violent thoughts and threats. Sometimes individuals post threatening things on social media jokingly, like the aforementioned student who threatened to detonate the campus's nuclear reactor on Yik Yak over the outcome of a football game. Even though the chance of that happening was essentially zero, it still counted as a threat under state law.

Whether the platforms are used to cultivate radical thoughts and induce aggression, or they become outlets for potentially violent actors to express their thoughts and hint at future actions, social media and violence are intricately and complexly linked. As the total numbers of social media users and applications and platforms continue to increase, social media activity may be used as a predictor of certain harmful behaviors.

4. CURRENT SOCIAL MEDIA MONITORING PRACTICES

Industries have various approaches to social media monitoring. This section describes some of the existing practices for social media monitoring by sector, the current capabilities in social media monitoring, and the potential content of interest or concern that might be captured on social media.

4.1 INDUSTRY OR SECTOR

The type of social media monitoring conducted varies based on the industry or sector in which the screening occurs. The current practices of private industry and corporations differ from government approaches.

4.1.1 Private Industry and Corporations

Using social media for prehire screenings is becoming increasingly common in private industry (Henderson 2019; Melão and Reis 2020; SHRM 2016). A recent survey of recruiters found that roughly 53% reported reviewing social media activity as part of their hiring processes (Henderson 2019). In the United Kingdom, 65% of HR professionals surveyed reported that they Googled candidates during the hiring process (Wilson 2016). Statistics were slightly lower in a survey of HR professionals conducted in Portugal; 41% of respondents reported using social media during the employee selection process (Melão and Reis 2020).

When asked how they used social media data, many HR professionals in the United States reported that they used it to obtain information in addition to what was provided in the resume or curriculum vitae or to verify information provided by the applicant, most frequently on the professional networking website LinkedIn (SHRM 2016). Recruiters described using social media to form a better picture of the applicant (Henderson 2019). They looked for red flags such as inappropriate or offensive (e.g., racist) posts, signs of alcohol abuse, or unprofessional comments and behavior (Henderson 2019; Wilcox, Damarin, and McDonald 2022). Many hiring professionals reported that they also used social media for risk management, attempting to avoid poor hiring decisions that might lead to embarrassment or negligent hiring claims (Wilcox, Damarin, and McDonald 2022). In general, social media was viewed as a complementary source to aid in decision-making during the hiring process (Henderson 2019; Melão and Reis 2020).

Information found on social media has also been used in hiring determinations. Of the US organizations that were surveyed in 2016, 36% stated that they have disqualified applicants because of information obtained through social media, such as illegal activity (SHRM 2016). The percentage was higher in a more recent survey, in which 57% of companies reported they have found content that caused them not to hire candidates (Hayes 2018). This statistic was slightly lower in the United Kingdom, where 36% of companies surveyed reported rejecting a candidate based on social media data (Wilson 2016). A more recent survey of recruiters in the United States found that 42% have rejected a candidate because of information found on social media (Henderson 2019), most often inappropriate behavior that they felt indicated a lack of professionalism.

However, HR professionals also expressed reasons for not relying on social media, most often because of legal, ethical, and privacy concerns (SHRM 2016; Wilcox, Damarin, and McDonald 2022) or because the information found on social media is irrelevant to the employee selection process (Henderson 2019; Melão and Reis 2020). Some also reported that the time involved in assessing applicants on social media was a barrier to its use (Melão and Reis 2020; Wilcox, Damarin, and McDonald 2022). Hence, social media vetting is common but not universal, and concerns remain regarding how to implement it effectively and fairly.

Because of the challenges recruiters and HR professionals expressed in sifting through irrelevant information, the existence of many commercial companies offering social media background check services (e.g., Social Intelligence, GoodEgg) is unsurprising (Kuhn 2022). In addition to performing social media monitoring, they can also help companies comply with the potentially complex legal requirements involved in the social media monitoring process (e.g., the Fair Credit Reporting Act in the

United States) (Maurer 2018). Outsourcing data collection to a third party can also help avoid concerns regarding the potential bias introduced by using social media data because the third party does not pass along information that may be used as a basis for discrimination (e.g., race) (Maurer 2018). Although many companies offer social media checks as a service, a recent paper suggested that *cybervetting*, or social media vetting, is most often unofficial or casual (Wilcox, Damarin, and McDonald 2022) and noted that many companies lack policies regarding the collection or use of social media information during the hiring process. Cybervetting is generally defined as the use of online information, including social media, for vetting purposes and is commonly used in industry when discussing social media monitoring.

The Society for Human Resource Management (SHRM) created a how-to guide for companies interested in using social media during employee selection that outlines five steps to integrate social media into their vetting procedures (SHRM n.d.). First, they recommend developing a social media screening policy, including information on who will conduct the screening and when, for which positions, and how (i.e., on which websites) the screening will be conducted. They also suggest that the organization develop a list of screening questions prior to conducting the social media screening, generate a formal report with the results, and retain the documentation for any decisions made (SHRM n.d.). These practices can help mitigate some of the concerns regarding potential bias or inconsistency in social media vetting.

4.1.2 Government

Government entities have been somewhat slower than private industry in adopting the use of social media in employee vetting. Collecting social media data has been discussed and is being considered in some contexts. For example, in the United States, the federal government has the legal authority to examine publicly available social media data for individuals who hold specific government positions (e.g., positions with access to classified information) (SEAD 5 2016).

Although the interest and motivation exist to expand the systematic use of social media information in background investigations, a detailed approach to doing so has not yet been identified (Kyzer 2021; Losey 2021). The United States has no formal process for integrating social media information into background investigations for military personnel and security clearance holders. Similarly, using social media is not mentioned in regulations governing unescorted access to nuclear power reactor protected areas or Category I fuel cycle facility material control areas (10 CFR Part 26; 10 CFR 73.56). Likewise, the US Human Reliability Program vets personnel for certain kinds of nuclear material access (10 CFR 712.1). Of course, the use of social media by government personnel is still limited. For example, strict guidelines exist on appropriate official use of social media for public affairs purposes (DoDI 5400.17 2023). In addition, specific agencies have considered using social media as part of their vetting processes (DHS and USSS 2019). However, no processes currently exist to integrate social media into the broader background screening and investigation process for government personnel.

Notably, all documents that discuss social media as part of investigations in the United States specify that any data collection is limited to publicly available social media information (i.e., information not hidden behind passwords or information that does not meet SEAD 5 criteria). This suggests that, regardless of the government position the investigation is for, information collected for a background investigation will almost certainly be limited to publicly available information. Despite the public nature of the information collected, much of the hesitation regarding the use of social media in background investigations in the United States stems from concerns about protecting rights to privacy and free speech (Levinson-Waldman et al. 2022). Thus, the government may hesitate to implement social media monitoring in the interest of protecting rights and civil liberties.

Information regarding other countries' approaches to the use of social media information for trustworthiness and reliability assessments within the government is limited. One study about the use of

social networking sites in cybervetting in the Netherlands compared the approaches used by HR professionals in private companies to those used by governmental organizations (Walrave et al. 2022). The researchers found that government employees and private-sector employees reported similar uses of social media. They examined social networking sites to verify experience claimed on job applications and to investigate candidates' online images, including their social behavior (e.g., drinking alcohol, partying). The primary difference between the private sector and government related to political expression online; the participants working for government organizations used social networking information to assess candidates' discretion in discussing their political views online.

In the United Kingdom, the Baseline Personnel Security Standard is used to vet all individuals with access to government assets, with increasing levels of scrutiny for positions with higher levels of clearance (Gov.UK 2022a). Individuals whose jobs require frequent and uncontrolled access to Category I nuclear material receive additional scrutiny, which includes an internet questionnaire. The internet questionnaire asks about online behavior, including social media usage, websites frequented, and views and attitudes toward online activity (Gov.UK 2022b). However, this documentation does not make clear whether any social media information beyond what is self-reported is systematically collected. In Canada, a Site Access Security Clearance is required for unescorted access to protected areas with nuclear material (CNSC 2013). This regulatory document does not mention the use of social media data explicitly, although it also does not explicitly say social media data are not used.

In general, social media is rarely mentioned explicitly in government vetting approaches within the nuclear domain or for other sensitive positions. However, social media information could be included or discovered in the vetting processes in some circumstances, such as during reference checks or personal interviews (CNSC 2013; 10 CFR Part 712). In some countries, the use of social media is allowed in vetting for government positions (e.g., the United Kingdom and United States), but clear and consistent international practices for how that information is included do not currently exist.

International guidance is similarly vague. The IAEA's guidance on preventive and protective measures against the insider threat contained in Nuclear Security Series 8-G (2020) do not make specific recommendations regarding trustworthiness and reliability assessments. However, guidance from the IAEA for trustworthiness and reliability assessments is being drafted and may include social media information as a potential category of publicly available information that could be collected as part of the access authorization process. Overall, public documents suggest the vetting of nuclear sector employees could include publicly available social media information and do not indicate that public information discovered during vetting would be disallowed. Nevertheless, these documents contain little indication of a consistent, best-practice approach to including this information into the trustworthiness and reliability assessment process.

4.1.3 Capabilities

One of the challenges of social media monitoring is identifying relevant information. Information is plentiful, and most of it is likely irrelevant to an individual's trustworthiness and reliability (e.g., pictures of family, friends, vacations, and social activities). As already noted, surveyed HR professionals reported that reviewing social media is cumbersome and time-consuming and cited this as one of the reasons for not integrating it into their vetting processes (Henderson 2019; Melão and Reis 2020). Similarly, the SHRM how-to guide discusses the importance of developing business rules or guides to ensure relevant information is collected from social media in a consistent manner.

To address relevancy concerns, background check companies that offer social media checks (e.g., GoodEgg, Social Intelligence, CrimCheck) provide users with a list of options that screenings might include, such as drug use, obscenities, violence, or hate speech (Kuhn 2022). Their screening approaches

are usually not described in detail because they are technical and proprietary because these companies' primary product is their ability to collect, flag, and associate social media data with identified individuals. These approaches are generally described as machine-learning algorithms designed to scan the web, identify content associated with the individual under investigation, and produce a report. The goal is to automatically find the correct individual and identify content that is relevant to assessing their trustworthiness and reliability.

Notably, AI-based approaches vary in quality (Heilweil 2020; Kuhn 2022; Wilcox, Damarin, and McDonald 2022). Some fail to detect concerning content (i.e., false negatives), and some falsely flag innocuous content as concerning (i.e., false positives). In addition, the algorithms upon which they are based are not necessarily unbiased—they may be trained or developed on data that “bakes in” unintended bias (Wilcox, Damarin, and McDonald 2022). For example, to assist in hiring, Amazon created an algorithm that automatically vetted résumés and assigned them scores of 1 to 5 (Dastin 2018). Although the algorithm did not include gender, it was developed using data from hiring over the past 10 years, which inadvertently included a biased selection of males because of male dominance in the industry, particularly in software development. The algorithm was able to identify gender through language used in the résumés (e.g., women’s soccer), and because of historically biased hiring, the algorithm consistently assigned higher scores to men’s résumés than women’s. Although Amazon noted that the algorithm was never used in any selection decisions, it demonstrated the potential pitfalls of AI-based approaches to employee hiring. In another example, an algorithm was designed to identify higher-risk patients to provide them extra medical care (Obermeyer et al. 2019). Although the algorithm did not include race, it did include a variable highly correlated with race, medical cost history, which resulted in patients of color systematically receiving lower scores than white patients. Thus, white patients were differentially selected for extra medical care compared with patients of color.

Additionally, little empirical evidence exists that the information unearthed by social media vetting is predictive of the candidate’s future behavior. Whether done manually or automatically, social media vetting may not add significant value beyond traditional vetting procedures (Schroeder, Odd, and Whitaker 2020; Wilcox, Damarin, and McDonald 2022). Additionally, Americans have expressed concerns about the privacy of their information online, and many do not share information publicly, which may further reduce the value of social media information (Rivera 2022).

Overall, the collection of social media data for use in trustworthiness and reliability vetting is certainly possible, albeit challenging. When done manually, social media vetting is time-consuming and labor-intensive. It can introduce significant bias because it is often used in an informal, ad hoc way when the primary concerns being investigated are of cultural fit or professionalism. These subjective criteria may be unrelated to later job performance and could unfairly exclude qualified candidates (Wilcox, Damarin, and McDonald 2022). When done automatically, social media vetting is complex and technical has potential for error. Either approach is potentially fraught with legal and practical concerns if performed improperly. These challenges are likely part of the reasons many governments have not yet adopted social media vetting into trustworthiness and reliability assessment processes.

Social media vetting can create a paradox. When attempting to use it to assess candidate trustworthiness or reliability, organizations find themselves in a legal and ethical minefield, conducting labor-intensive and costly checks that involve sifting through large volumes of potentially irrelevant information. At the same time, there is seemingly a public imperative to conduct these checks. Organizations that do not are at risk of appearing negligent in their hiring practices if they hire problematic candidates whose past bad behavior might have been identified via a simple Google search (Kuhn 2022). Although conducting social media checks as part of trustworthiness and reliability assessments is possible, work remains to ensure they are conducted efficiently, ethically, fairly, and legally. Organizations that choose to conduct social media checks should review guidance such as that provided by SHRM (n.d.) and Wilcox, Damarin, and

McDonald (2022) to ensure consistency. They should also consult their general counsels to ensure the checks are done in compliance with any relevant laws. As outlined by Wilcox, Damarin, and McDonald (2022), any organization planning to use social media as part of their vetting processes should develop clear goals and collectively determine how they will evaluate the results.

4.1.4 Content

In addition to the technical issues associated with the collection of social media information, organizations that wish to use it as part of their trustworthiness and reliability assessments must define in detail the kind of content they wish to collect. Because the bulk of available information online is irrelevant to trustworthiness and reliability, clearly defining the kind of information to be recorded is imperative.

When creating business rules regarding what information to collect, organizations could begin by reviewing the standards by which the individual will be evaluated for the government position that they would hold. For example, in the United States, individuals who will be eligible to access classified information or hold sensitive positions (including those that involve access to nuclear material) are evaluated against the adjudicative guidelines (SEAD 4). The guidelines include 13 broad criteria that are used to establish an individual's reliability and trustworthiness. They include allegiance to the United States, foreign influence, foreign preference, sexual behavior, personal conduct, financial considerations, alcohol consumption, drug involvement, criminal conduct, handling protected information, outside activities, and use of information technology systems. Like many red flags social media background check companies screen for, these 13 criteria might encompass content relevant to an individual's reliability and trustworthiness. Thus, one approach to developing collection guidance for social media might be to use the existing guidelines against which individuals will be evaluated for their roles, such as the adjudicative guidelines or the equivalent in the country of interest.

However, a few challenges remain, even with these kinds of clear business rules for social media data collection. First, what kind of social media content indicates untrustworthiness or unreliability is very subjective. For example, in the Adjudicative Guidelines, the guideline for alcohol consumption is intended to identify individuals whose consumption of alcohol might lead to the "exercise of questionable judgment or the failure to control impulses" (SEAD 4). An individual's alcohol consumption may be evident on social media; however, gauging the seriousness of that consumption and whether it warrants concern is highly subjective. Similar issues arise when considering foreign influence or fiscal responsibility: how much evidence of foreign influence is a concern, and when does it necessitate additional investigation or disqualification for a position? Considering the difficulty and subjectivity of this and other adjudicative decisions, the shortcomings of AI-based approaches are not surprising.

Second, many behaviors could be discovered online that present ethical dilemmas for organizations to address. For instance, racism and misogyny are rampant online (Bliuc et al. 2018; Farrell et al. 2019). However, in many countries, holding potentially objectionable views on race or gender is not disqualifying for a role in the government, even for a sensitive role. Nonetheless, organizations may be tempted to disqualify an individual whose beliefs are objectionable or embarrassing to the organization. Organizations must discuss and address these kinds of situations in advance to determine what will be done when the behaviors discovered online are unpleasant but not necessarily indications that the individual is not trustworthy or reliable.

Finally, social media can be used to identify indicators of an immediate security or safety concern. For example, social media is a common platform for at-risk individuals to express thoughts of suicide (Bryan et al. 2017; Robinson et al. 2016). Risks of violence (Byars et al. 2020) or unauthorized releases of classified or sensitive information (e.g., WikiLeaks) can also be identified on social media. Organizations

collecting social media data must have protocols in place for assessing sensitive information, handling immediate threats, and deciding whether something needs to be reported. In addition, plans must be in place regarding who is responsible for making reports and to whom. In some cases, reporting the information to a local law enforcement agency may be appropriate; reporting the information to the social media platform may also be appropriate if the content violates the terms and conditions of the platform.

Unfortunately, no catchall reporting procedures exist for determining what concerning online information to report and how to report it because these processes need to be based upon local rules and laws. However, organizations that want to collect social media data should at least determine how they will address threats of violence, suicide, or unauthorized releases of sensitive information. Discussions to develop these protocols should include organization leadership, legal counsel, security personnel, and HR professionals. The individuals or organizations collecting the social media data should also be trained and empowered to report content that may not meet a reporting requirement but is still concerning. Social media is a rapidly evolving platform with the potential to reveal new and unexpected information; therefore, whatever mechanism is developed to gather social media data must be flexible to accommodate rapid changes.

5. LEGAL AND REGULATORY OVERLAYS FOR SOCIAL MEDIA MONITORING DURING TRUSTWORTHINESS DETERMINATIONS

The inclusion of social media monitoring into trustworthiness determinations for individuals seeking security clearances or access to sensitive information areas raises novel questions involving privacy and constitutional protections, including questions about whether such privacy exists within the “penumbras” of the Constitution as stated in *Griswold v. Connecticut* (1965). However, privacy and free speech concerns when submitting for a clearance raise administrated law questions more so than constitutional ones.

The right of privacy and its codification differ from other constitutional rights and are treated differently within different legal philosophies. For example, whereas the freedom of speech is clearly delineated in the First Amendment, the text of the Constitution contains no explicit right to privacy. Applications of US constitutional law during the past six decades have prompted robust discussion about the extent of the right of privacy in criminal proceedings and autonomy. Although these are heady questions that require continual debate, the question is far narrower when addressing trustworthiness determinations and the use of social media.

Processes for social media monitoring are governed by administrative procedures that do not involve any criminal proceedings or denials of liberty. The operative constitutional issue is one of due process and whether such procedures provide due process to the individuals whose social media is monitored, potentially hindering them from obtaining clearances or other types of special access.

This section will briefly identify and discuss the constitutional issue of due process within the context of administrative law and the fundamental question and intersection of privacy interests when an individual submits themselves to a security clearance request or other forms of trustworthiness determination. In the United States, the Supreme Court has clearly articulated that the government has a clear and legitimate interest in individuals conducting government work or other work that is critical to the nation and that this interest overcomes possible privacy interests for individuals submitting to investigations of trustworthiness and reliability.

Additionally, this section will identify the statutory and associated regulatory framework used in the United States for security clearances. To mitigate a possible insider threat, the US government under the ODNI issued SEAD 5, *Collection, Use, and Retention of Publicly Available Social Media Information in*

Personnel Security Background Investigations and Adjudications, outlining processes and procedures to allow cybervetting of individuals who submit for security clearances. This section will describe the practical impacts of SEAD 5 and the limitations of cybervetting to ensure balancing of possible privacy interests that are unique in the social media domain.

5.1 CONSTITUTIONAL CONSIDERATIONS: DUE PROCESS

Obtaining a security clearance and undergoing a trustworthiness determination are administrative processes governed by the rules of the executive agency executing the review. Thus, individuals seeking clearances or trustworthiness determinations under the purview of an executive agency have due process rights as expressed in the Fifth and Fourteenth Amendments of the US Constitution. This view was articulated in the Supreme Court opinion *Goldberg v. Kelly* (1970), which concluded that the due process clause in these amendments applies to federal agencies when a liberty or property interest is at stake. When an executive agency engages in decision-making, some form of notice and a hearing is required based on the procedures and processes of the agency. In *Kelly*, the Court determined that a predeprivation hearing was required before an individual could lose welfare benefits. Insofar as welfare benefits were viewed as an entitlement, they constituted a form of property that could not be taken away without some hearing by the federal agency trying to remove the benefits. Thus, §706 of the Administrative Procedures Act (APA) lays out the standards of review for courts when reviewing actions of federal agencies during the determination of whether agency actions violate their internal procedures.

Another constitutional question arises regarding whether individuals retain a right to privacy when they submit to background checks or other forms of trustworthiness determination. The Supreme Court partly resolved this issue with its decision in *NASA v. Nelson* (2011). In *Nelson*, employees of the Jet Propulsion Lab (JPL) sued the government when they were required to submit to background checks as part of their work for NASA, arguing that their right to informational privacy (i.e., protection against the collection and dissemination of individuals' data) was violated. Applying the APA, the Court determined that no informational privacy right was violated, and that the government had a legitimate reason for the workers to submit to background checks because their work was critical to support NASA and the government. Although the right to privacy was not specifically addressed in the opinion, Justice Scalia in a concurring opinion argued that a background check did not violate any constitutional right and went as far as to say that the Court should have decided that no privacy right, in the instance at bar and more broadly, existed within the Constitution.

When an individual submits for a security clearance or form of trustworthiness determination, they answer questions on a form. But what rules govern information obtained from sources other than the form such as social media? Is that information governed by the overall administrative procedure for vetting and how are due process concerns addressed? The next section will identify the statutory and regulatory framework in the United States associated with obtaining a security clearance and the introduction of SEAD 5 as part of the vetting process to include social media and the associated safeguards.

5.2 STATUTORY FRAMEWORK FOR OBTAINING A SECURITY CLEARANCE IN THE UNITED STATES

According to DOE O 472.2A (2022), obtaining a clearance in the United States is “an administrative determination an individual is eligible for access to classified information and/or access to particular types or categories of classified information...” Depending on the organization, the access control may differ, with slight differentiation in the processes. Most individuals seeking a security clearance use Standard Form No. 86 (SF-86), *Questionnaire for National Security Positions*, but individual federal agencies may have additional requirements. For example, 10 CFR Part 95 and subsequent sections contains the Nuclear

Regulatory Commission's (NRC's) requirements for individuals seeking facility security clearance and access to safeguarded information regulated by the NRC (10 CFR 95.1).

In most cases, an individual submits to a security clearance investigation upon completion of SF-86. According to the OPM, the purpose of SF-86 is to support the US government "in conducting background investigations, reinvestigations, and continuous evaluations of persons under consideration for, or retention of, national security positions" (OPM 2016). The SF-86 explains that providing the information requested is voluntary and delineates the subsequent processes after the individual submits the completed document (OPM 2016). As part of the investigative process, investigators may interview current employers and examine different answers provided in the SF-86 (OPM 2016). The investigator may choose to conduct a personal interview with the individual during which the investigator may ask the applicant to clarify or explain answers provided on the SF-86 or ask follow-up questions (OPM 2016). During the review phase, the adjudicator applies a *whole-person concept*, meaning that the whole record of an individual is examined, and no single issue is automatically disqualifying for trying to obtain a security clearance (Kyzer 2019). Only through mitigation steps taken by the individual and a review of the totality of the individual's background and circumstances does an adjudicator decide on eligibility. Once an eligibility determination has been made, the applicant is notified of the completion of the process and that they are subject to continuous evaluations which may include periodic reinvestigation.

One section of note on the SF-86 form is the disclosure of information. The section explains that the information provided is specific to determine whether the applicant should be given a national security position. The information provided is protected against unauthorized disclosure. The form explains that the collection, maintenance, and disclosure of any information gathered from the background investigation is governed by the Privacy Act of 1947 (5 USC 552a 2010). Under the Privacy Act as amended, the agency requesting information pertaining to the applicant's eligibility for a national security position developed systems for maintaining and protecting such information.

During the security clearance process, the investigator may identify a security concern. The investigator passes the security concern onto the agency adjudicator, who evaluates the concern and decides to either grant the security clearance or issue a letter of intent denying the request. If a clearance is denied, a statement of reason is provided, which is the focus of any administrative appeal. In the United States, the appeals process for government employees is slightly different than for government subcontractors. Nevertheless, the applicant is provided the letter of intent and statement of reason and is allowed to respond in writing to the concerns raised. The response is then reviewed, and a determination is made whether a security clearance can be granted or if a subsequent hearing is required under the review of an administrative law judge, who rules that the appeal has merit or that the denial is appropriate (Berry & Berry 2013).

5.3 INCLUSION OF SOCIAL MEDIA AND CYBERVETTING IN THE SECURITY CLEARANCE PROCESS

In 2016, ODNI issued SEAD 5, *Collection, Use, and Retention of Publicly Available Social Media Information in Personnel Security Background Investigations and Adjudications*. Notably, the directive came three years after "the year of the insider" (i.e., 2013), a phrase that refers to events such as Aaron Alexis killing 12 people in the Washington Navy Yard and the data breach by Edward Snowden (Kyzer 2022).

Under the directive, "[a]gencies may choose to collect publicly available social media information in the personnel security background investigation process, which pertains to the covered individuals, associations, behavior and conduct, as long as the information pertains to the adjudicative guidelines for making determinations..." (ODNI 2016). The directive is clear that "[o]nly publicly available social

media information pertaining to the covered individual under investigation shall intentionally be collected” (ODNI 2016). These individuals whose public social media information can be collected are not requested or required to provide passwords, log into private accounts, or take any action that would disclose nonpublicly available social media information (ODNI 2016). Only public information can be collected. The agency cannot ask the individual to supply private information protected by a password or that would otherwise be unavailable to the public. The directive states that the agency investigator cannot create accounts or use existing social media accounts to bypass privacy controls (e.g., by becoming the individual’s “friend” or “following” their account).

Under SEAD 5, *publicly available social media information* is defined as “[a]ny electronic social media information that has been published or broadcast for public consumption, is available on request to the public, is accessible online to the public, is available to the public by subscription or purchase, or is otherwise lawfully accessible to the public” (ODNI 2016). Publicly available social media information includes information from the following sources (Kyzer 2022):

- Social networks (Facebook, LinkedIn)
- Microblogging (Twitter)
- Blogging and forums (WordPress, Tumblr)
- Picture and video sharing (Flickr and YouTube)
- Music sharing (Spotify)
- Online commerce (eBay)
- Online dating (Match.com)
- Geosocial networks (TripAdvisor)
- News and media websites where people can comment

Applying the reasoning given by the JPL employees in *NASA v. Nelson*, one could argue that such information is protected under informational privacy and that such searches of their data and information violate their due process rights. However, the majority in *Nelson* and both concurring opinions by Justices Scalia and Thomas highlight that such informational privacy rights are not protected and perhaps do not exist as a constitutional right. Importantly, SEAD 5 does require personnel involved in the investigation or adjudicatory processes to be trained in the collection, use, and retention of related records and subsequent adjudicatory procedures (ODNI 2016).

5.4 PUBLIC VS. PRIVATE SECTOR USE OF SOCIAL MEDIA FOR HIRING AND TRUSTWORTHINESS DETERMINATION

A 2020 Harris Poll indicated that 70% of private-sector employers who responded to the survey believed that companies screen job candidates using social media profiles during the hiring process. The three main platforms identified and screened by employees are LinkedIn, Facebook, and Twitter (Cotriss 2023). Unlike in the governmental investigation process, social media profiles do not come up in traditional employment background checks, which focus on employment history, credit information, and legal problems (Cotriss 2023). However, like the government, employers are strongly discouraged from bypassing privacy controls; bypass procedures could be analogized to an employer asking inappropriate questions during interviews (Cotriss 2023). Although not illegal, bypassing privacy controls does raise questions as to the extent of someone’s privacy rights at the job requesting information.

6. ETHICAL AND PRACTICAL CONSIDERATIONS IN SOCIAL MEDIA MONITORING

Many different ethical and practical considerations arise when incorporating social media into a vetting and monitoring program. These considerations include how to handle concerning content, mitigate the potential for bias, and select individuals into a program for social media monitoring.

6.1 CONCERNING CONTENT

Social media can contain information that is outside of the normal purview of vetting personnel who assess trustworthiness and reliability for their organizations. For example, people might freely express opinions or thoughts, including objectionable ones, online that would not otherwise be known within a workplace. Social media could also include information that indicates a potential risk of violence against oneself or others. Before incorporating social media into a trustworthiness and reliability program, organizations should ensure that they have procedures in place to handle information that might indicate an imminent safety or security concern. These procedures should clearly identify who will be notified if information about an imminent risk is discovered and what actions will be taken. Action plans should include sufficient protection for the safety and privacy of the individuals involved.

Additionally, organizations need to consider the type of content that will be captured, with whom it will be shared, and how it will be used. This is important because individuals express beliefs online that might lead to biases during the selection process or during employment. However, organizations' reputational concerns are also important to acknowledge. To address concerns about both bias and reputational damage sufficiently, clarity about the information that needs to be collected and the process used to handle it is important.

6.2 POTENTIAL FOR BIAS

The potential for bias that arises if social media is incorporated into the personnel vetting process is discussed previously in this paper and remains whether the data are collected manually or algorithmically. If AI-based approaches are used, then methods should consider the possibility of bias and employ approaches to mitigate them. This is a nontrivial task because biases can creep into machine learning and AI in unforeseen ways. Mehrabi et al. (2019) reviewed some ways to mitigate bias in AI and to achieve fairness that should be considered by any organizations leveraging AI capabilities for social media monitoring.

Photos and videos on social media can introduce additional candidate information such as race, religion, age, and ethnicity that might lead to bias in judgment of the candidate's trustworthiness or reliability. In the United States, for instance, data on race are deliberately not collected during the security clearance process; however, integrating social media data increases the likelihood that racial information will be present in the security clearance process (Piquado et al. 2022). This is true for other characteristics for which the potential for bias exists. Organizations should have processes in place to mitigate the impacts of these biases, including clear rules and standards for determining eligibility.

6.3 RIGHTS TO PRIVACY

Regardless of the specific local laws on privacy, the privacy of the individuals who might be affected by a social media monitoring program should be considered. Just because a particular type of data collection is authorized or legal does not necessarily mean that it is defensible or reasonable from an ethical or moral perspective. To determine the value of a social media monitoring program, organizations must weigh the loss of privacy against the potential risks of failing to collect the information and decide whether the costs outweigh the benefits or vice versa.

In addition, the extent of the information collected determines the potential impact on individual rights to privacy. In the United States, SEAD 5 limits collection of information for individuals eligible to access classified information or to hold sensitive positions is limited to publicly available information. Thus, one could argue that collecting such information is no different than collecting information from a newspaper article. However, individuals may not be aware of this data collection and may express emotions and ideas on social media that they would not want to share with their employers. To ethically collect social media data, organizations must ensure that the affected individuals are sufficiently informed of the scope and purpose of the collection. Limiting data collection to public information is likely also an ethical best practice for countries developing social media monitoring programs.

Finally, the nature of the internet raises questions about anonymity. Most internet users have attempted to be anonymous or less identifiable online; in 2013, 86% of internet users reported that they attempted to conceal their identities online in some manner (Rainie et al. 2013). This raises questions as to what level of data collection on individuals who are actively seeking privacy is ethical. Although individuals who are authorized to access nuclear material have consented to a degree of surveillance given their positions, for the collection of social media data to be ethical, those individuals must clearly understand the nature, scope, and extent of the information they are authorizing employers to access. Informed consent is a necessary component of an ethical social media monitoring program.

6.4 SELECTION FOR SOCIAL MEDIA MONITORING

The types and number of individuals who are covered by social media monitoring programs must also be considered. Which positions and individuals will be monitored under the proposed program? When will that monitoring begin and when will it end? Determining the answers to these questions and informing the covered individuals are important parts of ethically collecting social media data. The choice to monitor individuals on social media can be based on access or authority (i.e., position risk). It could also be based on individual risk; for example, individuals who display specific risk factors such as histories of concerning behavior could be selected for monitoring on social media. Regardless of how monitoring targets are identified, an ethical approach requires that the individuals be informed about the monitoring program and consent to it. This consent involves transparency regarding the scope and use of the information collected.

In addition to ethical considerations, practical considerations also exist related to choosing monitoring targets. The approach used to select targets for monitoring affects the number of individuals included and the associated workload (e.g., data collection, investigation, and adjudication). These effects should be considered before beginning a monitoring program.

Finally, organizations implementing social media monitoring programs should also consider when that monitoring will occur. Perhaps monitoring will be initiated upon application to a covered position. However, questions remain about the frequency of the monitoring program and when it will be terminated. Continual monitoring of social media data involves the collection and analysis of a large volume of data that may not be feasible depending on the approach used. Moreover, monitoring should be terminated if the individual's access is terminated (e.g., upon retirement). Overall, organizations should weigh the volume of data that is likely to be created by their programs against the potential value the monitoring can provide and make informed decisions based on the resources available to them. Weighing the volume of data against value added might include an initial investigation followed by periodic reinvestigations of social media data. Organizations may also choose to monitor social media only when indications arise that an individual might be a risk (e.g., repeated security violations at work).

As with all security practices, a graded approach should be used that balances the resources applied against the potential risk. A decision to include an individual in a social media monitoring program could be made using the following risk determinations:

- Position risk. The selection for social media monitoring is based on the potential risk associated with the individual's access, authority, or knowledge.
- Individual risk. The selection for social media monitoring is based on individual indications of risk, such as a report to a behavior observation program of risky or concerning behavior (e.g., substance use issues, criminal conduct, financial concerns).
- All applicants and employees. All applicants or employees who have or will have authorized access to nuclear material and/or eligibility to access classified information could be included in a social media monitoring program.

Organizations could choose to use a combination of position and person-based risk or to monitor all employees in certain roles. They could also consider random selection as a deterrent—that is, they could periodically select a random subset of employees for social media monitoring, similar to a random drug testing program. Monitoring targets should be picked based on organizational priorities, resources, and potential risk.

7. SOCIAL MEDIA VETTING FOR NATIONAL SECURITY POSITIONS: BEST PRACTICE RECOMMENDATIONS

Utilization of public-facing social media for employment screening is common, and its use for vetting individuals for national security positions will likely become even more common (Hayes 2018). The attractiveness of available data to increase efficacy of vetting is bolstered with initial evidence that this information can be used to accurately assess job candidates' personality traits and predict academic ability and job performance (Stoughton, Thompson, and Meade 2013; Kluemper, Rosen, and Mossholder 2012). Although social media vetting offers a wealth of information, it also presents challenges and risks. Incorporating best practices can maximize the potential of social media vetting while minimizing potential negative consequences (Wilcox, Damarin, and McDonald 2022). Thus, when implementing social media vetting, several best practices are important to consider.

Recommendations	Benefits to the organization
Policy and procedure development	
Develop a clear and consistent policy for social media screening that complies with federal, state, and local laws.	<ul style="list-style-type: none"> • Ensures compliance with legal requirements • Promotes transparency and consistency in the screening process
Limit social media screening to publicly available information.	<ul style="list-style-type: none"> • Minimizes privacy violations • Reduces the risk of accessing information that may not be relevant to the position
Ensure that the vetting procedure is reviewed regularly for efficacy and updated as needed to reflect best practices.	<ul style="list-style-type: none"> • Helps to ensure that the screening process is effective and up-to-date • Ensures that the organization is using the most current and reliable information to make hiring decisions
Combine social media vetting with more traditional methods to enhance suitability assessment rather than using traditional methods alone.	<ul style="list-style-type: none"> • Provides a more comprehensive picture of the candidate • Ensures that all relevant information is taken into account when making employment decisions

Recommendations	Benefits to the organization
Training and education	
Train HR personnel and hiring managers on anti-discrimination laws and privacy concerns related to social media vetting.	<ul style="list-style-type: none"> • Helps ensure that the screening process is fair, objective, and lawful • Promotes compliance with legal and ethical standards
Ensure that the process is consistently applied to all individuals so fairness is maintained.	<ul style="list-style-type: none"> • Helps to ensure that all applicants are treated equally and that the screening process is objective and consistent • Protects against potential legal claims of discrimination or bias
Third-party involvement and verification	
Consider using a third-party vendor to conduct social media checks and ensure that they comply with all legal requirements.	<ul style="list-style-type: none"> • Reduces potential liability and ensures the use of accurate and reliable information • Protects against potential claims of bias or discrimination
Verify all information being used to make employment decisions from reliable collateral sources.	<ul style="list-style-type: none"> • Helps to ensure that the information gathered is accurate and reliable • Protects against potential legal claims or negative perceptions from employees or the public
Documentation and transparency	
Document the reasons for any adverse employment decisions based on information found on social media to demonstrate a legitimate, nondiscriminatory rationale.	<ul style="list-style-type: none"> • Provides evidence to defend against claims of discrimination or wrongful termination • Ensures that decisions are based on job-related factors and not personal biases or preferences
Maintain full transparency of the program so all employees are aware of what is being monitored and how it is being used to make decisions.	<ul style="list-style-type: none"> • Promotes openness, trust, and fairness in the screening process • Helps to avoid negative perceptions or backlash from employees or the public
Respect privacy settings so no information that employees expect to remain private is being used for vetting.	<ul style="list-style-type: none"> • Helps to promote privacy and protect against potential legal claims or negative perceptions • Ensures that the screening process is respectful and transparent
Relevance and fairness in decision-making	
Ensure that all information gathered is relevant to trustworthiness, reliability, and suitability for employment in a position of trust.	<ul style="list-style-type: none"> • Helps to ensure that the screening process is job-related and that the information gathered is relevant and reliable • Reduces the risk of legal challenges or negative perceptions from employees or the public
Consider the fact that personal and professional behaviors can be maintained separately and that private behavior does not necessarily indicate how someone will behave in the workplace.	<ul style="list-style-type: none"> • Helps to avoid unnecessary intrusions into an individual's private life • Promotes fairness, objectivity, and respect in the screening process

8. CONCLUSION

The use of social media vetting in national security positions has become an increasingly attractive option for organizations seeking to enhance the efficacy of personnel vetting processes. Social media can offer valuable insights into an individual's trustworthiness, reliability, and suitability for positions involving sensitive information and critical infrastructure. However, this paper demonstrates that implementing social media monitoring programs requires caution and diligence to address various ethical, legal, and practical considerations.

To incorporate social media monitoring into a trustworthiness and reliability program, organizations need to develop robust policies and procedures ensuring compliance with federal, state, and local laws as well as international standards. Organizations should prioritize transparency, fairness, and respect for individual privacy in their screening processes by limiting data collection to publicly available information, offering clear explanations to affected individuals about the scope and purpose of their monitoring programs, and ensuring informed consent, among other actions.

Explicit and implicit bias are significant concerns when integrating social media into personnel vetting processes. Organizations must acknowledge the risk of bias in both manual and AI-based approaches and employ strategies to mitigate biases. This effort involves training HR personnel and hiring managers on anti-discrimination laws, using third-party vendors that comply with legal requirements, and reviewing and updating vetting procedures regularly to reflect best practices.

When choosing individuals for social media monitoring, organizations should apply a graded approach balancing the resources applied against the potential risk. They can use various risk determinations, such as position risk or person-based risk, or a random selection approach to identify monitoring targets. Organizational priorities, resources, and potential risks should guide the chosen approach.

Adopting best practices for social media vetting in national security positions is crucial for organizations to maximize potential benefits while minimizing negative consequences. By developing clear policies and procedures, ensuring training and education, involving third-party vendors, maintaining documentation and transparency, and ensuring relevance and fairness in decision-making, organizations can enhance the vetting process while respecting individual rights and avoiding potential pitfalls.

In conclusion, when implemented thoughtfully and ethically, social media vetting can be a valuable tool in assessing the trustworthiness and reliability of individuals in national security positions. Organizations must carefully weigh the benefits against the costs and challenges associated with social media monitoring programs. By adopting a mindful approach and following best practices, organizations can mitigate potential risks and enhance the effectiveness of their personnel vetting processes, ultimately contributing to a stronger security posture and safer environments.

9. REFERENCES

- 10 CFR Part 26. (2008). "Fitness for Duty Drug Testing Requirements." Nuclear Regulatory Commission. Available: <https://www.govinfo.gov/content/pkg/FR-2022-11-22/pdf/2022-24903.pdf>
- 10 CFR Part 712. (2004). "Human Reliability Program." US Department of Energy. Available: <https://www.govinfo.gov/content/pkg/FR-2004-01-23/pdf/04-1316.pdf>
- 10 CFR Part 73. (1973). "Physical Protection of Plants and Materials." Available: <https://www.ecfr.gov/current/title-10/chapter-I/part-73>
- 10 CFR Part 95. (2023). "Facility Security Clearance and Safeguarding of National Security Information and Restricted Data." Available: <https://www.ecfr.gov/current/title-10/chapter-I/part-95>
- 32 CFR Part 147. (2012). "Adjudicative Guidelines for Determining Eligibility for Access to Classified Information."
- 5 USC 552a. (2010). Privacy Act of 1947. Available: <https://www.govinfo.gov/content/pkg/USCODE-2018-title5/pdf/USCODE-2018-title5-partI-chap5-subchapII-sec552a.pdf>
- Antoniadou, Nafsika, Constantinos M. Kokkinos, and Angelos Markos. (2016). "Possible Common Correlates between Bullying and Cyber-Bullying among Adolescents." *Psicologia Educativa* 22 no. 1: 27–38. DOI: <https://doi.org/10.1016/j.pse.2016.01.003>
- Bell, A. J., Rogers, M. B., & Pearce, J. M. (2019). The insider threat: Behavioral indicators and factors influencing likelihood of intervention. *International Journal of Critical Infrastructure Protection*, 24, 166-176.
- Berry & Berry. (2013). "A summary of the security clearance appeals process." Resources. *Berry & Berry, PLLC*. Available: <https://www.berrylegal.com/resources/security-clearance/a-summary-of-the-security-clearance-appeals-process/>
- Bliuc, Ana-Maria, Nicholas Faulkner, Andrew Jakubowicz, and Craig McGarty. (2018). "Online networks of racial hate: A systematic review of 10 years of research on cyber-racism." *Computers in Human Behavior* 87 (October): 75–86. DOI: <https://doi.org/10.1016/j.chb.2018.05.026>
- Borum, Randy. (2003). "Understanding the Terrorist Mind-Set." *Criminal Justice Periodicals. FBI Law Enforcement Bulletin* 72 no. 7 (July): 7–10. Available: <https://www.ojp.gov/pdffiles1/nij/grants/201462.pdf>
- Brady, William J., Killian McLoughlin, Tuan N. Doan, Molly J. Crockett. (2021) "How social learning amplifies moral outrage expression in online social networks." *Science Advances* 7 no. 33. DOI: <https://doi.org/10.1126/sciadv.abe5641>
- Bryan, Craig J., Jonathan E. Butner, Sungchoon Sinclair, Anna Belle O. Bryan, Christina M. Hesse, and Andree E. Rose. (2018). "Predictors of emerging suicide death among military personnel on social media networks." *Suicide and Life-Threatening Behavior* 48 no. 4 (August): 413–430. DOI: <https://doi.org/10.1111/sltb.12370>
- Byars, Jason, Emily Graybill, Quynh Wellons, and Lonny Harper. (2020). "Monitoring social media and technology use to prevent youth suicide and school violence." *Contemporary School Psychology* 24 no. 3 (September): 318–326. DOI: <http://dx.doi.org/10.1007/s40688-020-00277-x>
- CNSC (Canadian Nuclear Safety Commission). (2013). REGDOG-2.12.2, Security: *Site Access Security Clearance*. Available: <https://nuclearsafety.gc.ca/eng/acts-and-regulations/regulatory-documents/published/html/regdoc2-12-2/index.cfm>

- Cotriss, David. (2023). “Keep it clean: social media screenings gain in popularity.” *Business News Daily*. Available: <https://www.businessnewsdaily.com/2377-social-media-hiring.html>
- Dastin, Jeffrey. (2018). “Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women.” *Reuters*. Available: <https://tinyurl.com/49f69hw2>
- Datareportal. (2022). “Global Social Media Statistics.” *Kepios*. Available: tinyurl.com/v2hp379w
- DHS (Department of Homeland Security) and USSS (United States Secret Service). (2019). Privacy Impact Assessment for the Social Media Screening. DHS/USSS/PIA-026. Available: <https://www.dhs.gov/sites/default/files/publications/privacy-pia-uss-026-socialmedia-december2019.pdf>
- DoDI (Department of Defense Instruction) 5400.17. (2023). *Official Use of Social Media for Public Affairs Purposes*. Office of the Assistant to the Secretary of Defense for Public Affairs. Available: <http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/540017p.PDF>
- DOE O 472.2A. (2022). *Personnel Security*, Washington, DC: US Department of Energy. Available: <https://www.directives.doe.gov/directives-documents/400-series/0472.2-BOrder-A>
- Fan, Rui; Ke Xu, and Jichang Zhao. (2016). “Higher contagion and weaker ties make anger spread faster than joy in social media.” arXiv: Computer Science, Social and Information Networks. Available through arXiv: https://ui.adsabs.harvard.edu/link_gateway/2016arXiv160803656F/arXiv:1608.03656
- Farrell, Tracie, Miriam Fernandez, Jakub Novotny, and Harith Alani. (2019). “Exploring misogyny across the manosphere in reddit.” *Proceedings of the Web of Science 10th ACM Conference* (June): 87–96. DOI: <http://dx.doi.org/10.1145/3292522.3326045>
- Fernandez, Miriam and Harith Alani. (2021). “Artificial intelligence and online extremism: challenges and opportunities.” In *Predictive Policing and Artificial Intelligence*, edited by John McDaniel and Ken Pease, 132–162. Abingdon: Routledge. DOI: <https://doi.org/10.4324/9780429265365-7>
- Goldberg v. Kelly, 397 U.S. 254 (1970).
- Gov.UK. (2022a). “National security vetting: clearance levels.” Guidance. *Gov.UK*. Available: <https://www.gov.uk/government/publications/united-kingdom-security-vetting-clearance-levels/national-security-vetting-clearance-levels>
- Gov.UK. (2022b). “Demystifying vetting.” Guidance. *Gov.UK*. Available: <https://www.gov.uk/government/publications/demystifying-vetting>
- Hassan, Ghayda, Sebastien Brouillette-Alarie, Seraphin Alava, Divina Frau-Meigs, Lysiane Lavoie, Arber Fetiu, Wynnnpaul Varela, Evgueni Borokhovski, Vivek Venkatesh, Cecile Rousseau, and Stijn Sieckelinck. (2018). “Exposure to extremist online content could lead to violent radicalization: A systematic review of empirical evidence.” *International Journal of Developmental Science* 12 no. 1–2: 71–88. DOI: <https://psycnet.apa.org/doi/10.3233/DEV-170233>
- Hayes, Ladan N. (2018). “More than half of employers have found content on social media that caused them not to hire a candidate, according to recent CareerBuilder survey.” *CareerBuilder*. Available: tinyurl.com/bb8ha23u
- Heilweil, Rebecca. (2020). “Beware of these Futuristic Background Checks.” *Vox*. Available: <https://www.vox.com/recode/2020/5/11/21166291/artificial-intelligence-ai-background-check-checkr-fama>
- Henderson, Kevin E. (2019). “They posted what? Recruiter use of social media for selection.” *Organizational Dynamics* 48 no. 4: 100663. DOI: <https://psycnet.apa.org/doi/10.1016/j.orgdyn.2018.05.005>

- IAEA (International Atomic Energy Agency). (2011). *Nuclear Security Recommendations on Physical Protection of Nuclear Material and Nuclear Facilities (INFCIRC/225/Revision 5)*. Recommendations. IAEA Nuclear Security Series No. 13. STI/PUB/1481. Available: https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1481_web.pdf
- IAEA (International Atomic Energy Agency). (2020). *Preventive and Protective Measures Against Insider Threats*. Implementation Guide. IAEA Nuclear Security Series No. 8-G (Rev. 1). STI/PUB/1858. Available: https://www-pub.iaea.org/MTCD/Publications/PDF/PUB1858_web.pdf
- Jackson, Dominique. (2021). “Instagram vs facebook: Which is best for your Brand's strategy?” *Sprout Social*. Available: <https://sproutsocial.com/insights/instagram-vs-facebook/>
- Jensen, M. (2016). Use of social media by US extremists - UMD. The Use of Social Media by United States Extremists. https://www.start.umd.edu/pubs/START_PIRUS_UseOfSocialMediaByUSExtremists_ResearchBrief_July2018.pdf
- King, Jonathan E., Carolyn E. Walpole, and Kristi Lamon. (2007) “Surf and Turf Wars Online—Growing Implications of Internet Gang Violence.” *Journal of Adolescent Health* 41 no. 6 (December): S66–S68. DOI: <https://doi.org/10.1016/j.jadohealth.2007.09.001>
- Kluemper, Donald Herbert, Peter A. Rosen, and Kevin W. Mossholder. (2012). “Social networking websites, personality ratings, and the organizational context: More than meets the eye?” 1. *Journal of Applied Social Psychology* 42 no. 5 (May): 1143–1172. DOI: <http://dx.doi.org/10.1111/j.1559-1816.2011.00881.x>
- Kruglanski, Arie W., Michele J. Gelfand, Jocelyn J. Belanger, Anna Sheveland, Malkanthi Hetiarachchi, and Rohan Gunaratna. (2014). “The Psychology of Radicalization and Deradicalization: How Significance Quest Impacts Violent Extremism.” *Political Psychology* 35 no. S1 (January): 69–93. DOI: <https://doi.org/10.1111/pops.12163>
- Kruglanski, Arie W., Rohan Gunaratna, Molly Ellenberg, and Anne Speckhard. (2020). “Terrorism in time of the pandemic: exploiting mayhem.” *Global Security: Health, Science and Policy* 5 no. 1 (October): 121–132. DOI: <https://doi.org/10.1080/23779497.2020.1832903>
- Kuhn, Kristine M. (2022). “The business of cybervetting.” *Industrial and Organizational Psychology* 15 no. 3 (September): 382–384.
- Kyzer, Lindy. (2019). “What is the whole person concept?” *ClearanceJobs*. Available: <https://news.clearancejobs.com/2019/09/04/what-is-the-whole-person-concept/>
- Kyzer, Lindy. (2022). “Social media monitoring in the security clearance process.” *Government Executive*. Available: <https://www.govexec.com/workforce/2022/12/social-media-monitoring-security-clearance-process/381024/>
- Lapidot-Lefler, Noam and Azy Barak. (2012). “Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition.” *Computers in Human Behavior* 28 no. 2 (March): 434–443. DOI: <https://doi.org/10.1016/j.chb.2011.10.014>
- Lapidot-Lefler, Noam and Azy Barak. (2015). “The benign online disinhibition effect: Could situational factors induce self-disclosure and prosocial behaviors?” *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 9 no. 2 (July): Article 3. DOI: <http://dx.doi.org/10.5817/CP2015-2-3>
- Ledwich, Mark and Anna Zaitsev. (2019). “Algorithmic extremism: Examining YouTube's rabbit hole of radicalization.” arXiv: Computer Science, Social and Information Networks. Available through arXiv: <https://arxiv.org/abs/1912.11211>

- Levinson-Waldman, Rachel, Harsha Panduranga, and Faiza Patel. (2022). "Social Media Surveillance by the U.S. Government." *Brennan Center for Justice*. Available: <https://www.brennancenter.org/our-work/research-reports/social-media-surveillance-us-government>
- McLaughlin, E. C. and Park, M. (2018, February 16). Social Media Paints Picture of Racist "professional school shooter." CNN. <https://www.cnn.com/2018/02/14/us/nikolas-cruz-florida-shooting-suspect/index.html>
- Medina, D. A., Chapman, I., Winter, J., & Tolan, C. (2022, May 28). Uvalde gunman threatened rapes and school shootings on social media app Yubo in weeks leading up to the massacre, users say. CNN. <https://www.cnn.com/2022/05/27/us/yubo-app-salvador-ramos-threats-invs/index.html>
- Melão, Nuno and João Reis. (2020). "Using social networks in personnel selection: a survey of human resource professionals." *Proceeding of the 2020 15th Iberian Conference on Information Systems and Technologies (CISTI)* (June): 1-6. IEEE. DOI: <https://doi.org/10.23919/CISTI49556.2020.9140854>
- Merriam-Webster. (n.d.). *Social media*. In Merriam-Webster.com dictionary. Available: <https://www.merriam-webster.com/dictionary/social%20media>
- Moghaddam, Fathali. (2005). "The Staircase to Terrorism: A Psychological Exploration." *American Psychologist* 60 no. 2: 161–169. DOI: <https://psycnet.apa.org/doi/10.1037/0003-066X.60.2.161>
- Müller, K., & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4), 2131-2167.
- NASA v. Nelson, 562 U.S. 134 (2011).
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. (2019). "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366 no. 6464 (October): 447–453. DOI: <https://doi.org/10.1126/science.aax2342>
- ODNI (Office of the Director of National Intelligence). (2016). "Collection, Use, and Retention of Publicly Available Social Media Information in Personnel Security Background Investigations and Adjudications." *Security Executive Agent Directive 5*, version 5.4. Available: <https://www.dni.gov/files/documents/Newsroom/Press%20Releases/SEAD5-12May2016.pdf>, accessed 7 August 2021
- OPM (Office of Personnel Management). (2016). *Questionnaire for National Security Positions*. Standard Form-86. Available: https://www.opm.gov/forms/pdf_fill/sf86.pdf
- Peterson, J., Densley, J., Spaulding, J., & Higgins, S. (2023). How Mass Public Shooters Use Social Media: Exploring Themes and Future Directions. *Social Media+ Society*, 9(1), 20563051231155101.
- Piccolo, Angie. (2022). "Guarding Science." *The Vault*. Published by Los Alamos National Laboratory. Available: <https://discover.lanl.gov/publications/the-vault/the-vault-2022/guarding-science/>
- Rainie, Lee, Sara Kiesler, Ruogu Kang, and Mary Madden. (2013). "Anonymity, Privacy, and Security Online." *Pew Research Center*. Available: <https://www.pewresearch.org/internet/2013/09/05/anonymity-privacy-and-security-online/>
- Richards, Lauren, Peter Molinaro, John Wyman, and Sarah Craun. (2019). *Lone Offender: A Study of Lone Offender Terrorism in the United States (1972-2015)*. National Center for the Analysis of Violent Crime, Federal Bureau of Investigation, Behavioral Analysis Unit, Behavioral Threat Assessment Center. Available: <https://www.fbi.gov/file-repository/lone-offender-terrorism-report-111319.pdf/view>
- Rivera, Alex. (2022). "Go.Verizon Social Privacy Survey Report 2022." *Verizon Resource Center*. Available: <https://go.verizon.com/resources/social-media-personal-data-privacy-survey/>

- Robinson, Jo, Georgina Cox, Eleanor Bailey, Sarah Hetrick, Maria Rodrigues, Steve Fisher, and Helen Herrman. (2016). "Social media and suicide prevention: a systematic review." *Early Intervention in Psychiatry* 10 no. 2 (April): 103–121. DOI: <https://doi.org/10.1111/eip.12229>
- Schmitt, Josephine B., Diana Rieger, Olivia Rutkowski, and Julian Ernst. (2018). "Counter-messages as Prevention or Promotion of Extremism?! The Potential Role of YouTube: Recommendation Algorithms." *Journal of Communication* 68 no. 4 (June): 780–808. DOI: <https://doi.org/10.1093/joc/jqy029>
- Schroeder, Amber, Kaleena Odd, and Julia H. Whitaker. (2020). "Agree to disagree: Examining the psychometrics of cybervetting." *Journal of Managerial Psychology* 35 no. 5 (June): 435–450. DOI: <http://dx.doi.org/10.1108/JMP-09-2018-0420>
- SHRM (Society for Human Resource Management). (2016). *SHRM Survey Findings: Using Social Media for Talent Acquisition—Recruitment and Screening*. Available: <https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/documents/shrm-social-media-recruiting-screening-2015.pdf>
- SHRM. (n.d.). "How to Use Social Media for Applicant Screening." How-To Guides. Available: <https://www.shrm.org/resourcesandtools/tools-and-samples/how-to-guides/pages/howtousesocialmediaforapplicantscreening.aspx>
- Sikkens, Elga, Marion van San, Stijn Sieckelinck, Hennie Boeije, and Micha de Winter. (2016). "Participant recruitment through social media: Lessons learned from a qualitative radicalization study using Facebook." *Field Methods* 29 no. 2 (September): 130–139. DOI: <http://dx.doi.org/10.1177/1525822X16663146>
- Sparrow, Betsy and Ljubica Chatman. (2013). "Social cognition in the Internet age: Same as it ever was?" *Psychological Inquiry* 24 no. 4: 273–292. DOI: <https://psycnet.apa.org/doi/10.1080/1047840X.2013.827079>
- Spring, Victoria L., C. Daryl Cameron, and Mina Cikara. (2018). "The upside of outrage." *Trends in Cognitive Science* 22 no. 12 (December): 1067–1069. DOI: <https://doi.org/10.1016/j.tics.2018.09.006>
- Stoughton, J. William, Lori Foster Thompson, and Adam W. Meade. (2013). "Big five personality traits reflected in job applicants' social media postings." *Cyberpsychology, Behavior, and Social Networking* 16 no. 11 (November): 800–805. DOI: <https://doi.org/10.1089/cyber.2012.0163>
- Suler, John. "The online disinhibition effect." *Cyberpsychology & behavior* 7.3 (2004): 321–326.
- Terry, Christopher and Jeff Cain. (2016). "The Emerging Issue of Digital Empathy." *American Journal of Pharmaceutical Education* 80 no. 4 (May): 58. DOI: <https://doi.org/10.5688/ajpe80458>
- Tokunaga, Robert S. (2010). "Following you home from school: A critical review and synthesis of research on cyberbullying victimization." *Computers in Human Behavior* 26 no. 3 (May): 277–287. DOI: <https://doi.org/10.1016/j.chb.2009.11.014>
- Torok, Robyn. (2010). "'Make a Bomb in Your Mum's Kitchen': Cyber Recruiting and Socialization of 'White Moors' and Home Grown Jihadists." *Proceedings of the 1st Australian Counter Terrorism Conference*, 54–61. Available: <https://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1005&context=act>
- Tumin, Remy. (2022). "University of Utah student made threat about campus nuclear reactor, officials say." *The New York Times*. Available: <https://www.nytimes.com/2022/09/23/us/utah-nuclear-threat-football-game.html>
- Vogels, Emily A. (2021). The state of online harassment. Pew Research Center. Available: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/01/PI_2021.01.13_Online-Harassment_FINAL-1.pdf

- Walrave, Michel, Joris Van Ouytsel, Kay Diederer, and Koen Ponnet. (2022). "Checked and Approved? Human Resources Managers' Uses of Social Media for Cybervetting." *Journal of Cybersecurity and Privacy* 2 no. 2: 402–417. DOI: <http://dx.doi.org/10.3390/jcp2020021>
- Welsh, David T., Lisa D. Ordóñez, Deirdre G. Snyder, and Michael S. Christian. (2015). "The slippery slope: How small ethical transgressions pave the way for larger future transgressions." *Journal of Applied Psychology* 100 no. 1 (January): 114–127. DOI: <https://doi.org/10.1037/a0036950>
- Wilcox, Annika, Amanda K. Damarin, and Steve McDonald. (2022). "Is cybervetting valuable?" *Industrial and Organizational Psychology* 15 no. 3 (September): 315–333. DOI: <https://doi.org/10.1017/iop.2022.28>

