Mitigation Framework

Step 1. Threat Map. Identify potential threat methods for analysis.

	Subject Type		
Threat Type	Individual	Group Identity	Organization
Direct	Bullying; coordinated targeting; hateful, inflammatory, or embarrassing comments; threats of violence; upsetting content; gendered threats; sustained harassment; mob harassment; sexual harassment; stalking; doxxing; SWATing; and account takeovers/lockouts.	Tactics leveraging social cleavages (for example hate speech or dog whistles) such as race, ethnicity, socioeconomic status or class, gender, sexual orientation, religion, regional or national origin, citizenship status, occupation, employment status, age / generation, education, or political affiliation.	Coordinated targeting to organizational accounts; Denial of service or access to an organization's content;
Indirect	Spreading of false or misleading information about an individual; defamatory information; disclosure of non-consensual intimate images; impersonation; hateful, inflammatory, or embarrassing comments.	Spreading of false or misleading information about a social group; hate speech directed towards a social group; divisive speech that may be either opposed or supportive of various social groups.	Mass internet shutdowns, establishing seemingly allied organizations to share disingenuous content; establishing opposition organizations to spread opposing viewpoints; imitation of the organization's online presence(eg, typosquatting).
Ingestion	Persuasion of the individual to believe or biased towards inaccurate information.	Persuasion of groups to believe inaccurate information about other groups, sowing division or apathy or bolstering alliances.	Persuasion of the organization to use inaccurate information in decision making.

Step 2. Harm Map. Connect scenarios to potential harms for the organization or its individuals or groups of individuals.

Individual Harms	
Harms to Self Determination	**Definition**
Loss of autonomy	Loss of autonomy includes needless changes in behavior, including self-imposed restrictions on freedom of expression or assembly.
Loss of liberty	Improper exposure to arrest or detainment. Even in democratic societies, false or negative information can lead to increased scrutiny, arrest or, abuse of governmental power.
Power imbalance	Information, or threat of disclosure, can create an inappropriate power imbalance or takes unfair advantage of a power imbalance between acquirer and the individual.
Physical harm	Actual physical harm to a person, including the potential to cause death.
Psychological harm	Information can cause psychological distress to the target such as increased anxiety, fear, and depression, possibly triggering reactions to previous trauma. This distress can also contribute to physical self-harm.
Reputational Harms	
Loss of trust	The breach of implicit or explicit expectations about the character and behavior between individuals or organizations. Loss of trust can leave entities reluctant to engage in further cooperation.
Stigmatization	Information can create a stigma that can cause embarrassment, emotional distress or discrimination.
Economic Harms	
Financial losses	Harms due to a result of loss of employment, business relationships, increased government scrutiny, and imprisonment.
Group Harms	

Step 3. Threat Scenarios. Develop practical description of the threat and challenge assumptions.

	Probing Questions
Adversary	• What is the identity of the adversary responsible for the harmful information? • What are the goals (if any) of an adversary sharing the harmful information? • What resources might an adversary have at their disposal?
Content	• Does the content contain personal information? • Does the content threaten or create fear for one's safety? • What elements of "truth" are contained in the message?
Context	• How is the harmful information delivered? • When and how often are interactions taking place? • How might the harmful information affect current events or campaigns?
Audience	• Who is the intended recipient of the information? • How could various stakeholders of the organization perceive the harmful information? What social norms might be violated? • How might the audience react to the harmful information? • How might law enforcement or government regulators react to the harmful information, if known?
Legitimacy	• What might give this threat legitimacy with an influential audience? • Why might the threat's message or methods be perceived as normatively acceptable? • How might those information sources already deemed legitimate by certain audiences spread or give additional credibility to the threat? • Who in power may spread or give credibility to the threat?
Impersonation	• How might an adversary take over or share information from an account belonging to the target? • How might an adversary convince an audience that their information is being shared with the target's approval?

Step 4: Mitigation Map. Select suitable controls to mitigate potential harms.

Identify		
Identify Harmful Information Risks		
Identify Harmful Information Risks	Identify Potential Threats	• Consider threats to individuals, groups, or the organization • Consider direct targeting, indirect attacks, ingestion, and generation
	Connect Threats to Potential Harms	• Identify the impact of potential threats to individuals, groups, and the organization • Consider physical, reputational, financial harms
	Create and Prioritize Threat Scenarios	• Describe threat scenarios in detail • Evaluate and prioritize scenarios based on likelihood and impact
Identify informal practices or formal policies		
Identify informal practices or formal policies	Security (Physical or Digital) or Incident Response	Identify and evaluate the following: • Evaluate security risk management abilities and training. • Consider how psychosocial risks are addressed in the risk assessment / management program. • Improve account security of organizational and personal social media accounts. • Decrease the online availability of personal information about staff members. • Other:
	Social Media Use	Identify and evaluate the following: • Acceptable social media use for

Protect		
Improve Organization-wide Digital Security		
Protect the confidentiality, integrity, and availability of the organization's and individuals' information systems	Maintaining confidentiality	• Secure accounts (personal & organizational) • Secure devices • Implement network monitoring • Other:
	Maintaining availability of information	• Implement DoS Protection • Enable Censorship Circumvention • Other:
	Maintain integrity of information	• Enable domain spoofing protection. eg DMARC • Enable DNS Hijacking protection (DNSSEC) • Register similar URLs • Other:
Minimize the Availability of Potentially Harmful Information.		
Reducing or obfuscating available open source information on organization or members.	Organizational Data Management	• Implement data minimization strategy • Conduct open source audit • Other:
	Personal Data Management	• Review Old Social Media Posts • Review Social Media Privacy Settings

Detect		
Implement Individual Detection		
Develop individual skills to identify known strategies for creating harmful information	Identify and learn how to react when in potentially compromising situations	• Verify the identity of new contacts, online and offline • Familiarize with counterintelligence tradecraft • Avoid discussing politically or culturally sensitive topics with strangers • Other:
	Improve media literacy to reduce an organization's susceptibility to its own digestion and spread of misinformation.	• Teach source checking • Implement content verification procedures • Other:
Implement Organizational Detection		
Develop organizational policies and practice for detecting harmful content	Implement manual content monitoring	• Implement and train staff on reporting harmful (or suspected) online information, including seemingly innocuous behavior • Create a plan to relieve subjects of abuse from self-monitoring • Create an emergency plan for manual monitoring of abuse campaigns by staff. • Other:
	Implement automatic content monitoring	• Set free keyword notification tools such as Google Alerts • Preset filtered feeds in tools such as TweetDeck

Respond		
Immediate Response -		
“Top 3 Things”, planned in advance.	Physical Safety and Wellbeing	• Train staff for initial shock: “breathe and connect with support, don’t handle this alone” • Plan to move to safety if credible threats • “Better to be safe than sorry” policies • Other:
	Digital Security	• Conduct Incident Response procedures • Other:
	Gather Evidence and Stay Aware of Threats	• Monitor and Archive (Tweetdeck, Dox Yourself, Hunch.ly, Archive.org, Google Alerts) • Manage manual monitoring of abuse campaigns by co-workers accounting for burn-out. • Other:
Next Stage Response		
Prevent Escalation of Harms	Respond to content on Platforms	• Engage with platforms or intermediaries for removal of harmful content or automated accounts • Use tools to identify, ignore, and/or block bots/trolls • Other:
	Execute Crisis Communication Plan	• Engage with supporters and funders to keep them informed

Recover		
Improving Safety		
Holistic Recovery	Rebuild Psychological Resilience	• Offer multiple avenues for coping • Provide counseling services for employees • Other:
	Improve Physical Protections	• Reassess physical vulnerabilities at work locations and increase protections as appropriate • Revisit personal security plans for employees • Other:
	Recover Digital Safety	• Reassess digital vulnerabilities and increase protections as appropriate • Other:
Repair Information Harms		
	Refine Communications Plan	• Adjust messaging based on counternarratives and situation • Engage with supporters and funders to keep them informed. • Inform public via media or other outlets • Other:
	Continue to use Platform-Specific Methods	• Search engine optimization • Search result downranking • Content removal processes such as Right to be Forgotten / DMCA. • Other:
	Seek Legal Remedies	• Contact legal counsel for jurisdiction-

Last update: April 1, 2020