DIGITAL DEFENSE ALLIANCE | AHEF DEFENSE LIST

Deepfake or Synthetic Media Abuse

KB-AK019

1. Reduce discoverable personal context before an adversary can weaponize it. Audit public profiles, family posts, data-broker listings, school/activity rosters, old breach exposure, and household routines; remove or limit unnecessary details and teach family members to treat unexpected personalized contact as a verification trigger.
2. Avoid revealing where you bank, invest, work, receive benefits, store crypto, or manage important accounts. Keep financial and account details out of public posts and casual chats, and use account alerts so unusual access or payment requests are caught quickly. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
3. Separate claims from evidence. Verify identities, emergencies, brands, documents, and stories through independent sources before acting, and preserve screenshots or message headers when something seems staged or inconsistent.
4. End the call and restart through a trusted number from a bill, card, portal, school directory, or official website. Do not provide codes, payment, remote access, or personal details to anyone who reached you first. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
5. Treat pressure as evidence that verification is needed. Refuse secrecy, deadlines, threats, and fear-based demands until the claim is checked through official channels and a trusted supporter has reviewed it.
6. Create a pause-and-verify rule for Fear, Shame, and Reputation Leverage. Before acting, confirm the claim through an independent source, involve a trusted reviewer, preserve evidence, and refuse secrecy, payment, credentials, data, or device access until the request is verified. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
7. Keep the relationship on-platform until trust is independently established, refuse financial requests and intimate-image pressure, and ask a trusted person to review inconsistencies, secrecy, or rapid emotional escalation.
8. Break isolation early. Keep sensitive decisions inside a circle of trusted people, use family code words for emergencies, and make it normal to ask for help before sending money, data, photos, or access.
9. Use a pre-agreed family verification phrase, call the person directly, and contact local emergency or institutional channels yourself. Do not send money, codes, travel details, or documents while panic is driving the decision. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
10. Break the secrecy instruction immediately by telling one trusted person and moving the decision into a shared review process. Legitimate helpers will not require isolation from family, caregivers, schools, banks, or advocates.

11. Protect accounts, devices, and finances with layered safeguards. Use password managers, passkeys or MFA, device updates, account alerts, transaction limits, and a rule that no one receives codes, remote access, or payment instructions through an unsolicited conversation.
12. Prevent the adversary from shaping the victim's choices, information flow, or access to help. Keep outside counsel involved, preserve independent communication channels, and reject demands that limit what the target can read, say, verify, or disclose.
13. Pause all payment movement and use a second reviewer before sending funds. Never use gift cards, crypto, wire transfers, payment apps, or overpayment refunds for unexpected requests; call the bank or merchant through verified channels. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
14. End recurring access and repeated pressure systematically. Block, report, rotate credentials, revoke sessions and app permissions, preserve evidence, and tell trusted contacts what happened so recontact attempts are easier to recognize.
15. Escalate when safety, dignity, identity, reputation, devices, or family stability are at risk. Preserve evidence, involve trusted supporters, contact platforms, schools, banks, legal aid, or emergency services as appropriate, and prioritize the victim's physical and emotional safety over embarrassment.
16. Use a password manager, passkeys or MFA, and official account recovery pages only. Never share one-time codes, recovery links, passwords, session tokens, or screen access with someone who initiated contact. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
17. Refuse to act on the imposed timeline. Capture the message, independently contact the court, agency, school, employer, or company, and require written documentation before any payment, disclosure, or account change.
18. Confirm identity through an independent contact path and ask for a detail the real person or organization can verify outside the suspicious conversation. Treat refusal, urgency, or channel switching as a stop signal. Apply this rule consistently across channels so the decision is governed by verification, documented evidence, and trusted review instead of the adversary's framing.
19. Create a pause-and-verify rule for Shame or Fear Destabilization. Before acting, confirm the claim through an independent source, involve a trusted reviewer, preserve evidence, and refuse secrecy, payment, credentials, data, or device access until the request is verified.