Warsztaty 7: Deepfake i oszustwa AI
Checklista czerwonych flag

Uzywaj tej listy, gdy widzisz obraz, film, reklame, wiadomosc albo slyszysz glos, ktory prosi o szybkie dzialanie.

1. PRESJA

Uwaga, jesli pojawia sie:

[ ] zrob to teraz
[ ] ostatnia szansa
[ ] nie rozlaczaj sie
[ ] nikomu nie mow
[ ] usun, zanim zniknie
[ ] wyslij szybko
[ ] masz tylko kilka minut
[ ] strach, wstyd, gniew albo wzruszenie

Zasada:
Presja czasu nie jest dowodem prawdziwosci.
Presja czasu jest powodem, zeby zwolnic.

2. KOD, LINK, DANE, PIENIADZE

Zatrzymaj sie, jesli ktos prosi o:

[ ] kod BLIK
[ ] przelew
[ ] dane karty
[ ] haslo albo login
[ ] zainstalowanie aplikacji
[ ] klikniecie w link
[ ] zdjecie dokumentu
[ ] dane osobowe
[ ] szybka wplate

Zasada:
Glos, film, reklama ani wiadomosc nie wystarczaja do podania kodu, danych albo pieniedzy.

3. PODSZYCIE

Sprawdz, czy ktos nie udaje:

[ ] osoby bliskiej
[ ] znajomego
[ ] pracownika banku
[ ] urzednika
[ ] policjanta
[ ] lekarza albo eksperta
[ ] osoby publicznej
[ ] znanej organizacji

Zasada:
Tozsamosc sprawdzam drugim kanalem, najlepiej takim, ktory znam z wczesniej.

4. BRAK ZRODLA

Czerwona flaga:

[ ] brak oficjalnej strony
[ ] dziwny albo podobny adres strony
[ ] brak daty
[ ] brak miejsca
[ ] brak pelnego nagrania
[ ] brak regulaminu
[ ] brak danych organizatora
[ ] brak niezaleznego potwierdzenia

Zasada:
Jesli nie moge sprawdzic zrodla, nie klikam, nie place i nie przesylam dalej.

5. OBRAZ, FILM, GLOS

Mozliwe sygnaly:

[ ] ruch ust nie pasuje do glosu
[ ] glos jest plaski albo zbyt rowny
[ ] brakuje naturalnych pauz
[ ] akcent albo rytm jest dziwny
[ ] gesty nie pasuja do tresci
[ ] tlo wyglada nienaturalnie
[ ] obraz jest zbyt emocjonalny i bez kontekstu
[ ] znana osoba mowi cos, co nie pasuje do jej zwyklego sposobu wypowiedzi

Zasada:
Sygnal techniczny pomaga, ale nie wystarcza.
Najwazniejsze jest: czego ta tresc ode mnie chce?

6. RYTUAL STOP

1. STOP
   Nie dzialam od razu.

2. ZRODLO
   Sprawdzam, skad to przyszlo.

3. DRUGI KANAL
   Oddzwaniam na znany numer albo wchodze samodzielnie na oficjalna strone.

4. KOD/LINK/PIENIADZE
   Nie podaje pod presja.

5. DECYZJA
   Decyduje po sprawdzeniu.

7. GDY KTOS JUZ KLIKNAL ALBO PODAL DANE

Nie zaczynam od wstydu. Zatrzymuje szkody.

[ ] zapisuje link, numer, godzine i zrzut ekranu
[ ] nie kasuje wiadomosci ani historii polaczenia
[ ] kontaktuje bank, urzad albo organizacje samodzielnie znalezionym kanalem
[ ] blokuje karte albo konto, jesli pojawily sie pieniadze lub dane
[ ] zmieniam haslo, jesli moglo zostac ujawnione

[ ] zglaszam podejrzana tresc

Zasada:
Szybkie zgloszenie jest lepsze niz idealny opis sytuacji.

8. GDZIE SPRAWDZAC I ZGLASZAC

Wazne adresy i kanaly:

- CERT Polska: https://incydent.cert.pl/
- Podejrzany SMS z linkiem: przeslij na numer 8080
- Kontakt CERT Polska: https://cert.pl/kontakt/
- NASK o rozpoznawaniu deepfake: https://nask.pl/
- Policja o oszustwach BLIK i AI: https://policja.pl/
- CEBRF KNF o falszywych inwestycjach: https://cebrf.knf.gov.pl/

Zasada:
Adres wpisuje samodzielnie albo wybieram ze znanej zakladki.
Nie korzystam z linku przeslanego w podejrzanej wiadomosci.

ZDANIE NA KONIEC

Nie musze wiedziec od razu, czy to deepfake.
Mam zatrzymac akcje i sprawdzic zrodlo.