



Kibernetska varnost - napredna

Umetna inteligenca in kibernetske grožnje

Deepfake, glasovne prevare
in avtomatizirani napadi

Predavatelj: dr. Miha Dvojmoč
Naročnik: GRIFFIN d.o.o.
Izvajalec: MD svetovanje d.o.o.
Datum: 27. 3. 2025



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



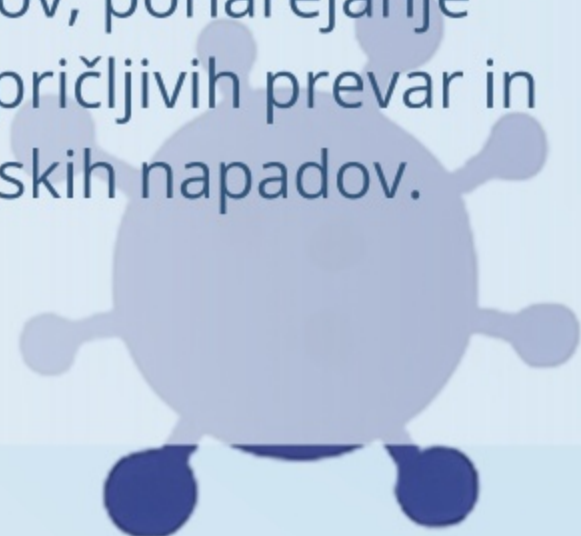
REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Umetna inteligenca in kibernetске grožnje

Umetna inteligenca prinaša številne koristi, hkrati pa ustvarja tudi nove varnostne izzive. Napadalci jo lahko uporabljajo za izdelavo lažnih posnetkov, ponarejanje glasu, pripravo bolj prepričljivih prevar in avtomatizacijo kibernetских napadov.



Kaj je umetna inteligenca?



Omogoča prepoznavanje vzorcev in obdelavo jezika.



Ustvarja besedila, slike in videoposnetke.



Izvaja prepoznavanje govora.



Odloča na podlagi podatkov in avtomatizira naloge.



Koristno orodje, ki pa ga lahko napadalci zlorabijo.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Zakaj je tema pomembna?



Lažni videoposnetki (deepfakes) so prepričljivi.



Ponarejen glas omogoča prevare.



Phishing prevare so bolj realistične.



Avtomatizirani napadi so hitrejši in obsežnejši.



Tveganja vključujejo lažne profile in manipulacijo informacij.

Kako napadalci uporabljajo umetno inteligenco?

- ➔ Pisanje lažnih e-sporočil in ustvarjanje lažnih slik.
- ➔ Ponarejanje glasu in izdelava deepfake videov.
- ➔ Iskanje tarč in avtomatizacija napadov.
- ➔ Prilagajanje prevar posamezniku.
- ➔ AI napadalcem omogoča hitrejše, cenejše in bolj prepričljivo delovanje.

Kaj je deepfake?

Deepfake je umetno ustvarjen ali spremenjen posnetek, pri katerem je lahko obraz, glas ali vedenje osebe ponarejeno. Lahko vključuje ponarejen obraz, glas, lažen govor, lažno gibanje ustnic, spremenjen videoposnetek in izmišljeno situacijo. Deepfake posnetek lahko ustvari vtis, da je oseba rekla ali naredila nekaj, česar v resnici ni.



Zakaj so deepfake posnetki nevarni?



Širjenje lažnih novic in zavajanje javnosti.



Škodovanje ugledu in izsiljevanje.



Politična manipulacija in poslovne prevare.



Lažno predstavljanje in kraja identitete.



Zaupanje v informacije je v krizi, ključen je preverjen vir.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Deepfake v poslovnem okolju



Lažni video direktorja ali
ponarejeno navodilo za plačilo.



Lažni sestanki in manipulacija
zaposlenih.

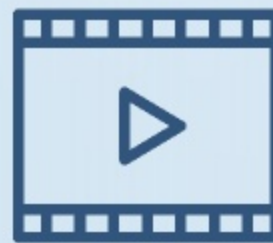


Izsiljevanje podjetja in škoda za
ugled organizacije.



Napadalec prepriča zaposlenega,
da izvede napačno dejanje.

Deepfake v osebnem življenju



Lažni posnetki posameznika in spletno nadlegovanje.



Izsiljevanje in lažne obtožbe.



Kraja identitete in škodovanje odnosom.



Širjenje sramotilnih vsebin povzroča psihološko in družbeno škodo.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Kako prepoznati deepfake?



Nenaravno gibanje obraza ali
čudno premikanje ustnic.



Nenavadna osvetlitev ali
nepravilne sence.



Neujemanje glasu in slike ali
nenavadno mežikanje.



Nejasen izvor posnetka. Tehnologija se
izboljšuje, zato je prepoznavanje težko.

Kaj storiti ob sumu na deepfake?

- ➔ Preverimo vir in poiščemo dodatne informacije.
- ➔ Preverimo uradne kanale in ne delimo vsebine nepremišljeno.
- ➔ Opozorimo odgovorne osebe in shranimo dokaz, če gre za zlorabo.
- ➔ Pomembno je preverjanje, ne hitro širjenje sumljivih posnetkov.

Glasovne prevare

Glasovne prevare uporabljajo ponarejen ali posnet glas za zavarjanje žrtve. Napadalci lahko posnemajo direktorja, sodelavca, družinskega člana, prijatelja, bančnega uslužbenca ali predstavnika organizacije. Z umetno inteligenco je mogoče ustvariti glas, ki zveni zelo podobno resnični osebi.



Kako delujejo glasovne prevare?

- ➔ Napadalec uporabi kratek zvočni posnetek osebe za ustvarjanje ponarejenega glasu.
- ➔ Nato pokliče žrtev, pošlje glasovno sporočilo ali zahteva denar/podatke.
- ➔ Pogosto naroči nujno dejanje ali ustvari občutek panike.
- ➔ Napadalci izkoriščajo nujnost in čustveni pritisk.

Primer glasovne prevare v družini

Napadalec pokliče starša ali starega starša in se pretvarja, da je družinski član v stiski (nesreča, potreba po denarju, izgubljen telefon). Pogosto zahteva takojšnje ukrepanje in tajnost. Takšne prevare so nevarne, ker izkoriščajo strah, skrb in željo pomagati bližnjemu.



Primer glasovne prevare v podjetju

V podjetju napadalec posnema glas direktorja ali vodje in zahteva nujno plačilo, nakazilo na nov račun, pošiljanje dokumentov, razkritje gesel, spremembo podatkov dobavitelja ali odobritev transakcije. Glasovno navodilo samo po sebi ne sme biti dovolj za izvedbo občutljivega poslovnega dejanja.



Kako se zaščititi pred glasovnimi prevarami?



Ne ukrepamo samo na podlagi klica;
preverimo osebo prek drugega kanala.



Uporabimo dogovorjeno
varnostno vprašanje.



Ne razkrivamo gesel in ne
pošiljamo denarja pod pritiskom.



Sumljiv klic prijavimo.



Nenavadne, nujne ali denarne
zahteve vedno dodatno preverimo.

AI in phishing prevare

- ➔ AI ustvarja jezikovno pravilna, osebno prilagojena in realistična phishing sporočila.
- ➔ Sporočila so lahko brez očitnih napak, oblikovana kot uradna komunikacija.
- ➔ Prilagojena so lahko delovnemu mestu žrtve, kar poveča prepričljivost.
- ➔ Lažno sporočilo danes ni nujno več polno slovničnih napak, ampak je lahko zelo prepričljivo.



Personalizirane prevare

- ➔ AI pomaga napadalcem ustvarjati sporočila, prilagojena posamezniku.
- ➔ Uporabijo se ime, delovno mesto, javni podatki, informacije z družbenih omrežij.
- ➔ Tudi podatki o organizaciji in slog komunikacije se lahko izkoristijo.
- ➔ Bolj kot je sporočilo prilagojeno, večja je možnost, da mu bo žrtev verjela.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Lažni profili z umetno inteligenco



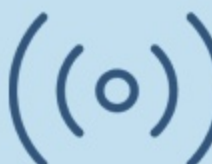
Lažni profili z AI fotografijami in besedili
za navezovanje stikov in prevare.



Uporabljajo se za širjenje
dezinformacij in zbiranje podatkov.



Služijo lažnemu predstavljanju
in vplivanju na mnenja.



Profilna fotografija in napisano besedilo
ne dokazujeta resnične identitete.

Avtomatizirani napadi

- ➔ AI omogoča hitrejša in obsežnejša izvajanje napadov.
- ➔ Vključujejo množično pošiljanje sporočil in iskanje ranljivosti.
- ➔ Omogočajo prilagajanje prevar, ustvarjanje lažnih vsebin in testiranje gesel.
- ➔ Avtomatizacija pomeni, da napadalci v kratkem času dosežejo veliko več tarč.

AI in zlonamerna programska oprema

- ➔ AI pomaga pri razvoju in prilagajanju zlonamerne programske opreme.
- ➔ To vodi do bolj prilagodljivih napadov in hitrejša ustvarjanje zlonamernih sporočil.
- ➔ Omogoča izogibanje zaznavi in bolj prepričljive prevare.
- ➔ AI povečuje učinkovitost napadalcev, zato obrambni ukrepi potrebujejo izboljšave.

AI kot orodje za obrambo



Zaznavanje nenavadnega vedenja in odkrivanje prevar.



Analiza varnostnih dogodkov in filtriranje neželene pošte.



Prepoznavanje zlonamerne programske opreme.



Pomoč pri odzivu na incidente. Enaka tehnologija se lahko uporabi tudi za zaščito.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Kako preveriti sumljivo sporočilo?



Ali sem sporočilo pričakoval? Ali pošiljatelj zahteva nujno ukrepanje?



Ali gre za denar, gesla ali nenavadno zahtevo?



Ali je povezava uradna? Ali lahko osebo preverim po drugem kanalu?



Pred klikom, plačilom ali razkritjem podatkov si vedno vzamemo čas za preverjanje.

Preverjanje identitete



Povratni klic na znano številko ali preverjanje prek druge aplikacije.



Uporaba dogovorjenega varnostnega gesla.



Preverjanje pri sodelavcu ali uporaba uradnih kontaktnih podatkov.



Dvojna odobritev pri plačilih. Ne zaupajmo kontaktnim podatkom iz sumljivega sporočila.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Zaščita osebnih podatkov



Pazimo na javne fotografije,
posnetke glasu in podatke o službi.



Izogibajmo se objavam družinskih
informacij, lokacije in potovanj.



Bodimo previdni pri objavah
otrok in osebnih dogodkov.



Javno objavljeni podatki se lahko
uporabijo za prepričljive prevare.





USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Zaščita organizacij



Preverjanje plačilnih zahtev in dvojno potrjevanje transakcij.



Izobraževanje zaposlenih o varnostnih postopkih.



Omejevanje javnih informacij in načrt odziva na incidente.



Spremljanje zlorab identitete. Potrebni so jasni postopki.



Vloga izobraževanja

- ➔ Izobraževanje je ključno za prepoznavanje deepfake posnetkov in glasovnih prevar.
- ➔ Uporabniki morajo poznati phishing, socialni inženiring in varno preverjanje identitete.
- ➔ Pomembno je tudi poznavanje zaščite osebnih podatkov in prijave sumljivih dogodkov.
- ➔ Varnostno zavedanje uporabnikov se mora spreminjati s tehnologijo.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI

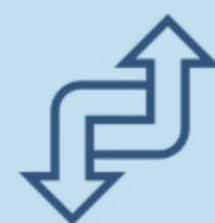


Financira
Evropska unija
NextGenerationEU

Kaj storiti, če smo žrtev?



Prekinemo komunikacijo z napadalcem in shranimo dokaze.



Spremenimo gesla in obvestimo banko, če gre za plačilo.



Obvestimo organizacijo ali odgovorne osebe ter prijavimo incident.



Opozorimo druge morebitne tarče. Hitro ukrepanje zmanjša škodo.

Etična uporaba umetne intelligence

- ➔ Ne ustvarjamo lažnih posnetkov za škodo in ne zlorablamo tuje podobe/glasu.
- ➔ Jasno označimo umetno ustvarjene vsebine in spoštujemo zasebnost.
- ➔ Preverjamo informacije in ne širimo nepreverjenih vsebin.
- ➔ AI ni problem sama po sebi, ampak njena zloraba za zavajanje in manipulacijo.



USPOSABLJAMO
IN PRILAGAJAMO
TRAINING AND ADJUSTMENT



NAČRT ZA
OKREVANJE
IN ODPORNOST



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA DELO, DRUŽINO,
SOCIALNE ZADEVE IN ENAKE MOŽNOSTI



Financira
Evropska unija
NextGenerationEU

Najpomembnejša pravila zaščite



Preverimo vir informacij in ne
zaupamo samo glasu ali videu.



Nenavadne zahteve preverimo
po drugem kanalu.



Ne klikamo sumljivih povezav, ne
razkrivamo gesel, pazimo, kaj objavljamo.



Uporabljamo večfaktorsko avtentikacijo
in prijavimo sumljive dogodke.



Kombinacija previdnosti, tehničnih
ukrepov in postopkov preverjanja.

Zaključek

- ➔ Deepfake posnetki, glasovne prevare in lažni profili so največja tveganja.
- ➔ Prepričljive phishing prevare in avtomatizirani napadi so vse pogostejši.
- ➔ Zloraba osebnih podatkov in manipulacija informacij sta pomembna izziva.
- ➔ V digitalnem svetu je ključno preverjanje informacij, ne le videnje ali slišanje.



Umetna inteligenca in kibernetske grožnje

Umetna inteligenca prinaša številne koristi, hkrati pa ustvarja tudi nove varnostne izzive. Napadalcji jo lahko uporabljajo za izdelavo lažnih posnetkov, ponarejanje glasu, pripravo bolj prepričljivih prevar in avtomatizacijo kibernetskih napadov.

Kibernetska varnost - napredna Umetna inteligenca in kibernetske grožnje

Deepfake, glasovne prevare
in avtomatizirani napadi

Predavatelj: dr. Miha Dvojmoč
Naročnik: GRIFFIN d.o.o.
Izvajalec: MD svetovanje d.o.o.
Datum: 27. 3. 2025