



NEMZETI
KÖZSZOLGÁLATI
EGYETEM
LUDOVIKA

MI baj lehet? A Mesterséges Intelligencia kiberbiztonsági kockázatai

dataSTREAM 2023

Dr. Bányász Péter (banyasz.peter@uni-nke.hu)

NKE ÁNTK Kiberbiztonsági Tanszék

A kályha, avagy ne keverjük a szezont

- Általános MI:
 - *„a számítógép azon képességét jelenti, amellyel képes meghaladni az emberi eredményeket bármilyen intellektuális feladatban. Ilyen típusú AI-t láthatunk olyan filmekben, ahol a robotoknak tudatos gondolatai vannak, és saját céljaiknak megfelelően cselekszenek.”*



A kályha, avagy ne keverjük a szezont

- Szűk MI:

- *„A számítógépes rendszer azon képességét jelenti, hogy az embernél hatékonyabban el tud végezni egy szűken meghatározott feladatot. A szűk AI a legmagasabb AI-szint, amelyet az emberiség máig elért, és minden olyan AI, mellyel a való világban találkozunk, ebbe a kategóriába tartozik, ide értve az önvezető járműveket és a személyi digitális asszisztenseket is. Még ha úgy is tűnik, hogy az AI önállóan és valós időben gondolkodik, ilyenkor az valójában több, szűkre szabott folyamatot hangol össze, és előre meghatározott keretek között hoz döntéseket. A mesterséges intelligencia „gondolkodásából” hiányzik a tudatosság és az érzelem.”*

A kályha, avagy ne keverjük a szezont

- Mesterséges szuperintelligencia:
 - *„Egy szuperintelligenciával rendelkező számítógép képes lenne az embert csaknem minden területen túlszárnyalni, többek között a tudományos kreativitásban, az általános bölcsességben és a társadalmi készségekben is.”*



A kályha, avagy ne keverjük a szezont

- Gépi tanulás:

- *„A gépi tanulás az a folyamat, amelyet a számítógépes rendszerek használnak a mesterséges intelligencia eléréséhez. Algoritmusokkal azonosít mintákat az adatokban, amelyekkel ezután adatmodellt készít, és előrejelzéseket végez. A gépi tanulási modelleket adatok részhalmazain tanítják be. Ha a modell betanításához használt adatok pontosan képviselik az elemzésre váró teljes adatkészletet, akkor az algoritmus pontosabb eredményeket produkál. Ha a gépi tanulási modell elég jól van betanítva ahhoz, hogy gyorsan és pontosan elvégezze a feladatát, akkor elérte a szűk AI szintjét.”*



Let's see who you really are
machine learning

A kályha, avagy ne keverjük a szezont

- Mély tanulás:
 - *„A gépi tanulás egy fejlett típusa, amely az emberi agy szerkezete által ihletett algoritmus-hálózatokat, úgynevezett neurális hálózatokat használ. Egy mély neurális hálózat beágyazott neurális csomópontokból áll, és minden megválaszolt kérdés újabb kapcsolódó kérdésekhez vezet. A mély tanulás betanításához általában nagy méretű adathalmazok szükségesek: a mély tanuláshoz használt betanítási készletek általában több millió adatpontból állnak...*

A kályha, avagy ne keverjük a szezont

- Mély tanulás:

- *... Ha a mély neurális hálózatot betanították ezekkel a nagy méretű adathalmazokkal, akkor az a kisebb hálózatoknál jelentősen nagyobb kétértelműséget is képes kezelni. Ezért kiváltképp hasznos például képfelismeréshez, ahol a mesterséges intelligenciának meg kell találnia az alakzatok széleit ahhoz, hogy azonosítani tudja a kép tartalmát. Mély tanulóval tanítják be azt a típusú mesterséges intelligenciát is, amely képes meghaladni az emberi készségeket olyan összetett játékoknál, mint a sakk."*

A kályha, avagy ne keverjük a szezont

- Nagy nyelvi modell:
 - „... A nagy adatmennyiség segítségével pedig pontosan képes megjósolni, melyik szó következik a mondatban. Leggyakrabban ezt „következő szó-előrejelzés” formájában figyelhetjük meg. Leegyszerűsítve, a nyelvi modellt egy óriási szövegkorpuszon képzik ki, amelyből „megjegyzí”, mely szavak, mondatok és bekezdések állnak a leggyakrabban egymás mellett, és azok hogyan kapcsolódnak egymáshoz.”



Nem csak a ~~20~~-éveseké-Chat GPT-é a világ...

- Bár sok mindent megválaszol, python kódot ír stb, de...
- Csupán egy nyelvi modell, így ezek nem többek, mint a szavak gyakori kombinációi és kollokációi.
- Nem található mögötte tényleges tudás vagy logika.

**When ChatGPT
Gives You
Better
Response**

Than Your Professor

Nem csak a ~~20~~-éveseké-Chat GPT-é a világ...

- Amitől eltér:
 - képes valós idejű reagálásra,
 - megérti a kontextust,
 - valós időben megfelelő válaszokat generál,
 - személyre szabottan,
 - témák széles skáláját képes kezelni,
 - emberhez hasonló nyelvezetet alkalmaz,
 - nincs élő összeköttetésben az internettel, 2021-esek az utolsó adatok.



Nem csak a ~~20~~-éveseké-Chat GPT-é a világ...

- Szöveg alapú platformok:
 - Jasper,
 - Copy.ai,
 - Jenni,
 - Fake Name Generator;
- kép alapú platformok:
 - Dall-E,
 - Midjourney,
 - Hexo AI,
 - Hypar;
- hang alapú platformok:
 - ElevenLab,
 - D-ID,
 - AD Auris,
 - Vocal Clarity,
 - Mubert;



Nem csak a ~~20-éveseké~~-Chat GPT-é a világ...

- programozás alapú platformok:
 - Metabob,
 - Enzyme,
 - Replit, Warp,
 - Mutable AI;
- videó alapú platformok:
 - Pictory,
 - Peech, Veed.io,
 - Fathom,
 - Embly,
 - Yepic,
- „This person/city/stb does not exist” platformok, amelyek különböző nem létező dolgokat (arcok, vállalkozások profiljai, városok műholdképjei stb.) generálnak.

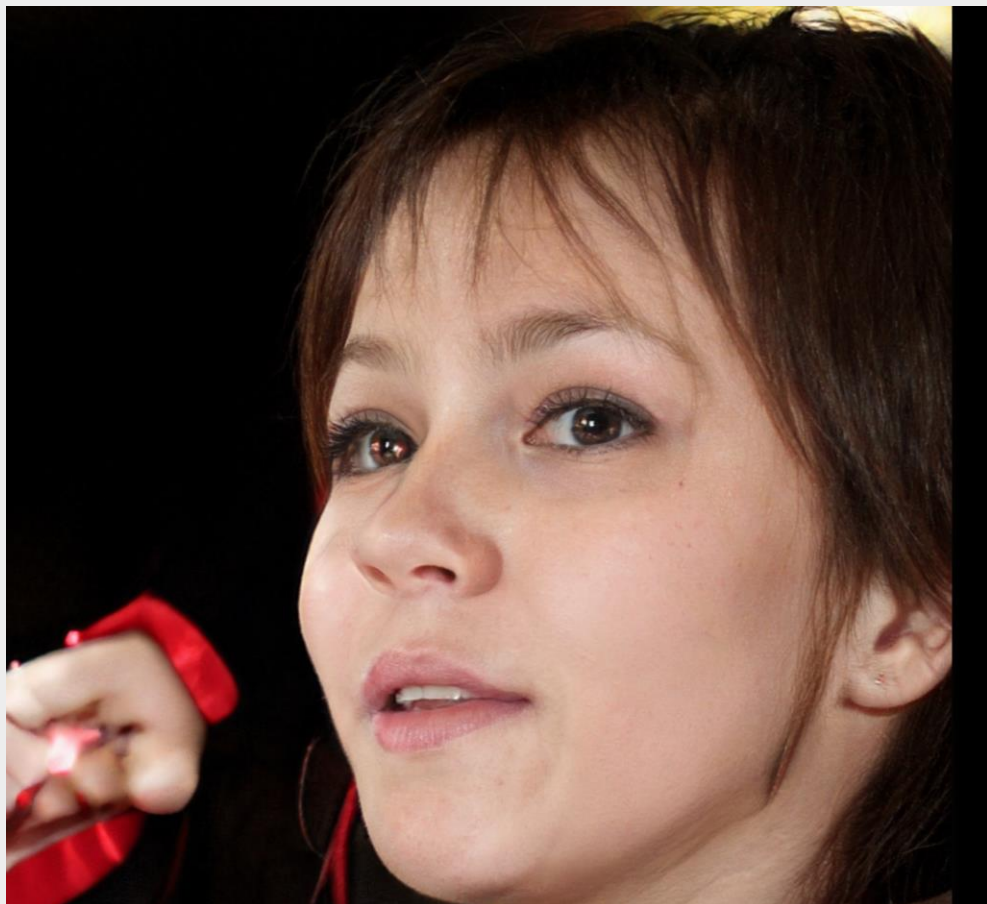




NEMZETI
KÖZSZOLGÁLATI
EGYETEM
LUDOVIKA

A kockázatok és mellékhatások tekintetében

Az illúzióink



Az illúzióink

Your Randomly Generated Identity

Gender

Name set

Country

[Generate](#) [Advanced Options](#)

These name sets apply to this count
Hungarian



Logged in users can view full social security numbers and can save their fake names to use later.



Györfi Csombor

3132 Nógrádmegyer
Apáczai Csere János u. 63.

Mother's maiden name Horváth

Geo coordinates **48.019535, 19.649203**

PHONE

Phone (32) 538-672

Country code 36

BIRTHDAY

Birthday June 22, 1993

Age 29 years old

Tropical zodiac Cancer

ONLINE

Email Address GyorfiCsombor@rhyta.com

This is a real email address. [Click here to activate it!](#)

Username Helfinch

Az illúzióink



Az illúzióink



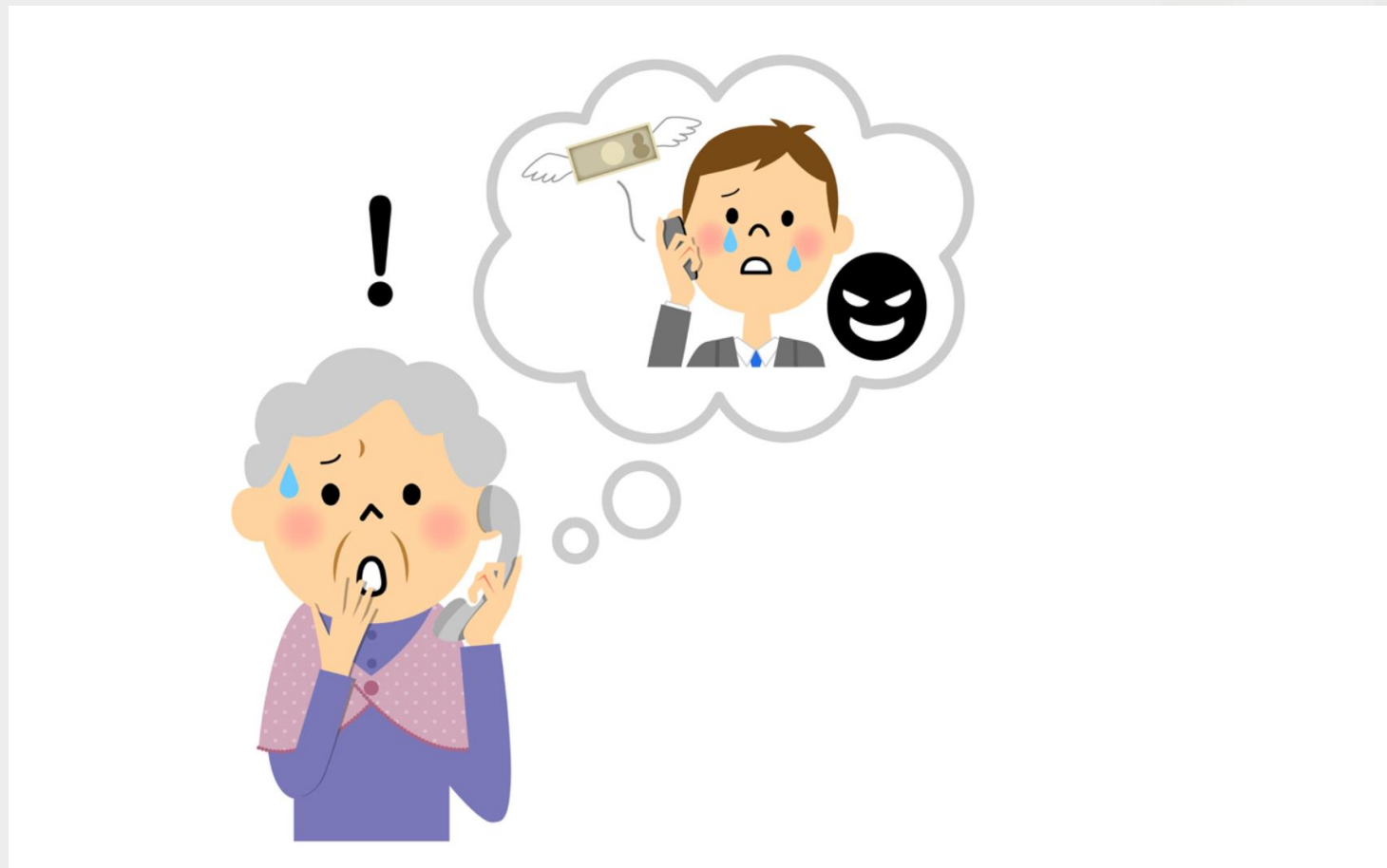
Az illúzióink



Az illúzióink



Az illúzióink



NKI szerint kockázatok

- Álhír terjesztés,
- SPAM,
- Adathalászat,
- Hamis weboldalak,
- Malware készítés.



can you decompile the following code to C?

```
func0(int*, int*):
    push rbp
    mov rbp, rsp
    mov QWORD PTR [rbp-24], rd
    mov QWORD PTR [rbp-32], rsi
    mov rax, QWORD PTR [rbp-24]
    mov eax, DWORD PTR [rax]
```

can you rewrite the C code above with more meaningful variable names?

Here is a version of the code with more descriptive variable names:

```
void swap(int *a, int *b) {
    int temp = *a;
    *a = *b;
    *b = temp;
}

void bubble_sort(int *arr, int n) {
    for (int i = 0; i < n - 1; i++) {
        for (int j = 0; j < n - i - 1; j++) {
            if (arr[j] > arr[j + 1]) {
                swap(&arr[j], &arr[j + 1]);
            }
        }
    }
}

void print_array(int *arr, int n) {
    for (int i = 0; i < n; i++) {
        printf("%d ", arr[i]);
    }
}
```

OpenAI ChatGPT FOR HACKERS 🤖

Ezen kívül

- Adatszivárgás,
- Szerzői jogok,
- Üzleti titkok,
- Plágium,
- Reputációs támadások,
- Kódolt előítélet,
- ...



NKI szerint pozitív felhasználás

- Rosszindulatú tevékenységek azonosítása,
- Sérülékenységek megértése (+ keresése),
- Kiberbiztonsági tanácsadás.





KÖSZÖNÖM A FIGYELMET!

uni-nke.hu