

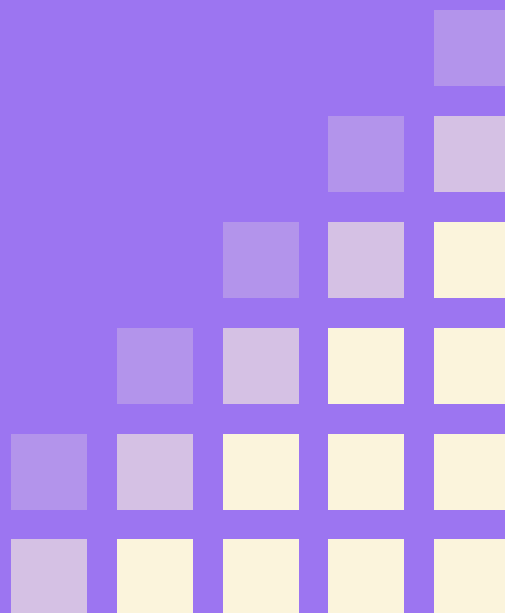
**nagranie, którego nigdy nie było.  
głos, którego nikt nie wypowiedział.  
twarz, która nie należy do nikogo.  
obraz, który zakrzywia rzeczywistość.**

# deep fake

**brzmi znajomo? żyjesz w tym codziennie.**

Poznaj od środka technologię, która produkuje fałsz na skalę przemysłową. Naucz się, jak ją rozpoznawać, po co powstaje i co możemy z nią zrobić.

0 s  90 s



# co to właściwie jest?

Deepfake to narzędzie modeli sztucznej inteligencji, które tworzy fałszywe zdjęcia, nagrania wideo lub dźwięki. AI potrafi „przykleić” czyjąś twarz do innego ciała. Sprawić, że ktoś na filmie mówi rzeczy, których nigdy nie wypowiedział. Wygenerować głos nie do odróżnienia od prawdziwego. Stworzyć obraz, który nigdy nie istniał.

To wszystko widzisz na TikToku, Instagramie, X czy grupach na komunikatorach. I staje się to coraz lepsze, profesjonalne, bezbłędne.

## skąd nazwa?

**deep learning** (głębokie uczenie - sposób, w jaki uczy się AI)

+ **fake** (coś fałszywego). Termin powstał w 2017 roku.

9 lat później AI doszło prawie do perfekcji.

photoshop - kiedyś

Wycinanie i sklejanie fragmentów istniejących zdjęć. Nienaturalne, „upiększone” obrazy - łatwo je zauważyć.

deepfake - dziś

Tworzy nowe treści od zera - piksel po pikselu, dźwięk po dźwięku, w kilkadziesiąt sekund. Uczy się, jak wyglądasz i brzmisz.

**cel? stworzyć treść, która oszuka widza, że jest prawdziwa.**

## 2 jak to działa?

Dlaczego deepfake działa na tak wysokim poziomie - i doszło do tego w zaledwie 9 lat? To wyścig **dwóch modeli AI**, które uczą się od siebie nawzajem:

### AI nr 1

twórca fake'ów

generuje twarz, głos,  
deepnude - jak najlepiej potrafi

VS

### AI nr 2

detektyw

wyłapuje fałsz i wytyka każdy  
błąd

♾️ powtarza się to miliony razy

Za każdym razem, gdy AI nr 2 wyłapie fałsz, AI nr 1 uczy się na błędzie i generuje lepiej - aż fake'a nie da się odróżnić od prawdy. Ta rywalizacja ma fachową nazwę: **GAN**. Ale nazwa jest nieważna. Ważny jest mechanizm: **AI trenuje samą siebie w oszukiwaniu**.

## czego AI potrzebuje, żeby zrobić deepfake?

Zdjęć, nagrań lub opisów. Kiedyś - setek. **Dziś wystarczy jedno zdjęcie lub kilka sekund głosu**. Im więcej Twoich zdjęć krąży w sieci, tym łatwiej stworzyć deepfake - AI uczy się na nich, jak przerobić Twój wizerunek tak, żeby wyglądał prawdziwie.

1 zdjęcie

tyle dziś wystarczy

## cała rodzina technik

Deepfake to nie jedna rzecz - to nazwa na całą rodzinę technik generowania obrazu i dźwięku. Część z nich widzisz codziennie:

### 1 zamiana twarzy

face swap - twarz jednej osoby nakładana na ciało innej.

### 2 synchronizacja ust

lip sync - istniejące nagranie zmienia się tak, że osoba „mówi” coś innego.

### 3 klonowanie głosu

AI mówi czyimś głosem dowolne zdanie. Wystarczy próbka z TikToka.

### 4 gadające zdjęcia

nieruchome zdjęcie zostaje „ożywione” i zaczyna mówić.

### 5 twarze, które nie istnieją

realistyczna twarz zmyślonego człowieka - m.in. pod fałszywe konta.

### 6 sceny z tekstu

z samego opisu powstaje realistyczne wideo, które nigdy się nie odbyło.

wszystkim zajmują się modele sztucznej inteligencji.

## jak fake leci w świat?

Sam deepfake to dopiero połowa problemu. Druga połowa? To, jak błyskawicznie krąży po internecie. **Algorytmy kochają deepfake.**

1

**Ktoś tworzy treść** - apką, stroną, modelem AI. Dziś dostępne dla każdego.

2

**Treść trafia na platformę** - TikTok, Instagram, X, YouTube, Discord, komunikatory.

3

**Algorytmy robią swoje** - treści budzące silne emocje są podbijane, bo generują reakcje i wyświetlenia.

4

**Ludzie udostępniają dalej** - często bez sprawdzania, bo ktoś inny zrobił to samo.

5

**Kopie zostają w sieci** - nawet po usunięciu oryginału: screeny, reuploady.

### **dłaczego fałsz rozchodzi się szybciej niż prawda?**

Bo fałszywe treści mają jedno zadanie: wywołać emocje. Szok, złość, śmiech - cokolwiek, co sprawi, że klikniesz i udostępnisz. Sprostowania? Są nudne - dla algorytmów i dla ludzi. Filmik z „politykiem” widzą miliony; sprostowanie - garstka.

**algorytmy nie nagradzają prawdy - nagradzają zaangażowanie.**

**to, co szkodzi najbardziej, najlepiej się klika. i to jest cały haczyk.**





## skala problemu

# 96-98%

**99-100% ofiar to kobiety.**

70%

# 1 na 4

Deepfake'i osób publicznych napędzają fałszywe reklamy i dezinformację - dzielą społeczeństwo i „rozwadniają” przekaz tak, żeby nie było wiadomo, co jest prawdą.

■ ■ ■ ■ ■ ■ ■ ■ ■ ■ s. 7/10

## jak rozpoznać deepfake?

Technologia każdego dnia staje się lepsza - ale warto zwracać uwagę na:

### ■ oczy i mruganie

nienaturalny rytm, „szklane” spojrzenie

### ■ brzegi twarzy i włosy

rozmycia, migotanie przy tle

### ■ zęby, uszy, dłonie

za dużo palców to klasyk

### ■ usta i dźwięk

usta nie nadążają za słowami

### ■ światło i cienie

padają z różnych stron

### ■ głos

płaski, bez emocji, dziwny oddech albo jego brak

**najważniejsze jest jednak nie to, co widzi oko:**

→ **kto to opublikował i po co?**

→ **czy inne, wiarygodne źródła też o tym mówią?**

→ **czy to nagranie ma wywołać we mnie silną emocję?**

Silna emocja to częsty sygnał manipulacji.

**weryfikacja źródła bije wypatrywanie glitchy.**

## zasada stop

Zanim udostępnisz - cztery kroki, które trwają kilkanaście sekund:

**S****sprawdź źródło**

kto to wrzucił? wiarygodne konto czy świeży profil?

**t****trzy razy pomyśl**

zanim wyślesz dalej - po co ktoś to udostępnił?

**O****oceń emocje**

czy to gra na złości, strachu, sensacji?

**P****potwierdź gdzie indziej**

niezależne źródła mówią to samo?

### pamiętaj o dwóch rzeczach:

**Twój wizerunek należy do Ciebie.** Nikt nie ma prawa robić z Twoją twarzą i głosem tego, co chce.

**Udostępnianie też jest działaniem.** Jeśli podajesz dalej krzywdzący deepfake - stajesz się częścią problemu, nawet jeśli nie Ty go stworzyłeś.

## co możesz zrobić już dziś?

Nie uciekniemy od internetu ani od AI. Możemy jednak zapobiegać i minimalizować szkody:

- 1 Przemyśl, co publikujesz** - im mniej publicznych zdjęć i nagrań, tym mniej „paliwa” dla AI.
- 2 Ustaw konta na prywatne** i kontroluj, kto widzi Twoje treści.
- 3 Umów się z bliskimi na „hasło rodzinne”** - jeśli ktoś dzwoni „głosem” bliskiej osoby i prosi o pieniądze, zapytaj o hasło, którego AI nie zna.
- 4 Nie ufaj samemu nagraniu** - sensacyjne wideo potwierdzaj w innych źródłach.
- 5 Reaguj, nie udostępniaj** - widzisz krzywdzący deepfake? Zgłoś go, zamiast podawać dalej.

**im mniej paliwa dajesz, tym słabszy fake.**

Twoje zdjęcia, Twój głos, Twój klik „udostępnij” - to wszystko czyjś materiał do pracy. Decyduj świadomie.



# nie wszystko, co widzisz, wydarzyło się naprawdę.

**silniejsze od deepfake może być tylko myślenie  
krytyczne.**

116 111 - telefon zaufania, 24/7

dyzurnet.pl - zgłoś nielegalne treści



Ministerstwo  
Edukacji Narodowej



półtorej  
minuty

poltorejminuty.pl · kontakt@poltorejminuty.pl