

Wprowadzenie do deepfake

Czym jest deepfake?

Klasy IV–VIII (10–16 lat) · 45 minut · bez kont uczniów
michal.kids/blog/akademia/deepfake-wprowadzenie

Cele i zakres

Uczniowie poznają:

- Co to jest deepfake?
- Krótka historia deepfake
- Typy deepfake
- Deepfake a zwykła edycja
- Dlaczego deepfake jest niebezpieczny?

Przebieg (45 minut)

10 min – Wprowadzenie

Zadaj pierwsze pytanie do rozmowy i przeczytajcie pierwszą część lekcji.

- Jak rozpoznasz, że zdjęcie lub film jest sfalszowany?
- Co to jest deepfake?

15 min – Treść lekcji

Przeczytajcie kolejne części (na rzutniku albo z konspektu) i omówcie pytania po każdej z nich.

- Krótka historia deepfake
- Typy deepfake
- Deepfake a zwykła edycja
- Dlaczego deepfake jest niebezpieczny?

15 min – Karta pracy

Uczniowie samodzielnie odpowiadają na pytania, potem sprawdzacie odpowiedzi razem – klucz jest w konspekcie.

5 min – Podsumowanie

Uczniowie uzupełniają część „Co zrobię, gdy...” i zabierają kartę do domu – na dole jest krótka informacja dla rodzica.

- Co zrobisz, jeśli zobaczysz deepfake kogoś znajomego?

Pytania do rozmowy

- Jak rozpoznasz, że zdjęcie lub film jest sfalszowany?
- Dlaczego ważne jest, żeby nie ufać każdemu wideo w internecie?
- Co zrobisz, jeśli zobaczysz deepfake kogoś znajomego?

Karta pracy – klucz odpowiedzi

1. Słowo „deepfake” powstało z połączenia jakich dwóch angielskich słów?
A. deep learning + fake
B. deep web + fake news
C. deep state + fake photo
D. deep blue + face
Prawidłowa odpowiedź: A. deep learning + fake
2. Kiedy anonimowy użytkownik Reddita po raz pierwszy opublikował deepfake wideo?
A. W 2015 roku
B. Pod koniec 2017 roku
C. W 2019 roku
D. W 2020 roku
Prawidłowa odpowiedź: B. Pod koniec 2017 roku
3. Który z poniższych NIE jest typem deepfake?
A. Deepfake wideo (zamiana twarzy w nagraniu filmowym)
B. Deepfake audio (klonowanie głosu)
C. Deepfake hardware (fałszowanie procesorów)
D. Deepfake tekstu (generowanie treści przez AI)
Prawidłowa odpowiedź: C. Deepfake hardware (fałszowanie procesorów)
4. Czym deepfake różni się od tradycyjnej edycji zdjęć w Photoshopie?
A. Deepfake wymaga więcej czasu niż edycja ręczna
B. Photoshop tworzy lepszej jakości fałszywki, bo pracuje nad nimi człowiek
C. Deepfake działa tylko na filmach, a Photoshop tylko na zdjęciach
D. AI generuje treść automatycznie, ucząc się z tysięcy przykładów
Prawidłowa odpowiedź: D. AI generuje treść automatycznie, ucząc się z tysięcy przykładów
5. Kto współtworzył ostrzegawczy deepfake z Barackiem Obamą w 2018 roku?
A. BuzzFeed i Jordan Peele
B. Google i YouTube
C. Netflix i Steven Spielberg
D. Elon Musk i CNN
Prawidłowa odpowiedź: A. BuzzFeed i Jordan Peele
6. Jakie jest największe niebezpieczeństwo związane z deepfake dla społeczeństwa?
A. Ogromne zużycie prądu przez komputery, które przez całą dobę generują deepfake
B. Erozja zaufania – ludzie przestają ufać nawet prawdziwym materiałom
C. Wirusy ukryte w darmowych aplikacjach do tworzenia deepfake
D. Kradzież zdjęć i filmów z telefonów przez aplikacje deepfake
Prawidłowa odpowiedź: B. Erozja zaufania – ludzie przestają ufać nawet prawdziwym materiałom

7. Co dziś wystarczy do stworzenia prostego deepfake?

- A. Superkomputer i zespół naukowców
- B. Specjalistyczna kamera 3D
- C. Smartfon i darmowa aplikacja
- D. Wiedza z zakresu programowania i sieć neuronowa

Prawidłowa odpowiedź: C. Smartfon i darmowa aplikacja

Gdzie szukać pomocy

116 111 – telefon zaufania dla dzieci i młodzieży (bezpłatny),
800 12 12 12 – Dziecięcy Telefon Zaufania Rzecznika Praw Dziecka (bezpłatny),
112 – gdy zagrożone jest życie.

Powiązanie z podstawą programową

Informatyka, klasy IV–VI – wymagania:

- 4.3) rozpoznaje podejrzane treści on-line i podejmuje podstawowe działania obronne – moduł medialny;
- 5.3) przestrzega zasad bezpiecznego i higienicznego korzystania z technologii i mediów cyfrowych – moduł medialny;
- 5.4) ocenia konsekwencje różnych zagrożeń cyfrowych oraz wskazuje sposoby reagowania na nie i możliwości uzyskania wsparcia;
- Godzina wychowawcza: zadania wychowawczo-profilaktyczne szkoły (załącznik nr 2, część IV) – realizowane w programie wychowawczo-profilaktycznym szkoły.

Źródło: rozporządzenie Ministra Edukacji z dnia 11 marca 2026 r. w sprawie podstawy programowej wychowania przedszkolnego oraz podstawy programowej kształcenia ogólnego dla szkoły podstawowej (Dz. U. z 2026 r. poz. 378, ze zm.), załącznik nr 2: informatyka, klasy IV–VI, oraz część IV „Zadania wychowawczo-profilaktyczne szkoły”. W roku szkolnym 2026/2027 ta podstawa obowiązuje w klasie IV; w klasach V–VIII obowiązuje dotychczasowa podstawa z 2017 r. (Dz. U. z 2017 r. poz. 356 z późn. zm.).

Treść lekcji

Co to jest deepfake?

Słowo **deepfake** powstało z połączenia dwóch angielskich słów: **deep learning** (głębokie uczenie – rodzaj sztucznej inteligencji) oraz **fake** (fałszywy). Deepfake to **syntetyczne media** – filmy, nagrania audio lub zdjęcia – stworzone przez sztuczną inteligencję tak, aby wyglądały lub brzmiały jak prawdziwe osoby.

Co odróżnia deepfake od tradycyjnej edycji? Zwykła obróbka zdjęć czy filmów wymaga ręcznej pracy człowieka – grafik musi piksel po pikselu modyfikować obraz. Przy deepfake to **AI generuje treść automatycznie**, ucząc się na tysiącach przykładów prawdziwych twarzy, głosów czy ruchów ciała.

Przykłady deepfake to między innymi: wideo, na którym ktoś wypowiada słowa, których nigdy nie powiedział; realistyczne zdjęcia osób, które w rzeczywistości nie istnieją; nagrania głosu brzmiące identycznie jak głos prawdziwej osoby.

Dlaczego to ważne? Jeszcze kilka lat temu stworzenie przekonującego deepfake wymagało zaawansowanej wiedzy technicznej i potężnych komputerów. Dziś wystarczy **smartfon i darmowa aplikacja**. To oznacza, że każdy – dosłownie każdy – może tworzyć fałszywe materiały. Dlatego umiejętność rozpoznawania deepfake staje się jedną z najważniejszych kompetencji cyfrowych naszych czasów.

Czas na quiz!

1. Z jakich dwóch angielskich słów powstało słowo „deepfake”?

- A. deep learning + fake
- B. deep web + fake news
- C. deep state + fake photo
- D. deep blue + fake video

Prawidłowa odpowiedź: A. deep learning + fake

2. Co odróżnia deepfake od tradycyjnej edycji zdjęć?

- A. Deepfake jest zawsze lepszej jakości
- B. Deepfake wymaga korzystania z drogiego komputera ze specjalistycznym sprzętem przeznaczonym do przetwarzania obrazu
- C. AI generuje treść automatycznie, zamiast ręcznej pracy człowieka
- D. Deepfake działa tylko na filmach, nie na zdjęciach

Prawidłowa odpowiedź: C. AI generuje treść automatycznie, zamiast ręcznej pracy człowieka

3. Czym są deepfake?

- A. Programy do ręcznej obróbki zdjęć
- B. Rodzaj wirusa komputerowego w filmach
- C. Specjalne filtry na Instagramie
- D. Syntetyczne media stworzone przez AI

Prawidłowa odpowiedź: D. Syntetyczne media stworzone przez AI

4. Co jest potrzebne dziś do stworzenia prostego deepfake?

- A. Superkomputer i zespół naukowców
- B. Specjalna kamera i mikrofon
- C. Smartfon i darmowa aplikacja
- D. Licencja od rządu

Prawidłowa odpowiedź: C. Smartfon i darmowa aplikacja

Krótką historia deepfake

Historia deepfake jest stosunkowo krótka, ale bardzo dynamiczna. Wszystko zaczęło się pod koniec **2017 roku**, kiedy anonimowy użytkownik serwisu Reddit o nazwie „deepfakes” zaczął publikować filmy z podmienionymi twarzami znanych osób. Użył do tego ogólnodostępnych narzędzi uczenia maszynowego. To od jego pseudonimu pochodzi nazwa całego zjawiska.

W **kwietniu 2018 roku** serwis BuzzFeed we współpracy z reżyserem **Jordanem Peele** stworzył film ostrzegawczy – deepfake przedstawiający byłego prezydenta USA **Baracka Obamę**, który mówi rzeczy, których nigdy nie powiedział. Film miał na celu uświadomienie ludziom, jak łatwo można stworzyć przekonujące fałszywe wideo.

W **2019 roku** tworzenie deepfake trafiło na smartfony. Chińska aplikacja **Zao** (wrzesień 2019) pozwalała w kilka sekund wstawić własną twarz do znanych scen filmowych i z dnia na dzień stała się wiralem. Popularne były też inne aplikacje do zabawy twarzą: działający od 2017 roku FaceApp (postarzanie i odmładzanie twarzy), a od 2020 roku Reface (podmiana twarzy w klipach i gifach) – miliony ludzi bawiły się podmianą twarzy.

Od **2020 roku** do dziś technologia deepfake rozwija się w zawrotnym tempie. Jakość fałszywych materiałów rośnie z roku na rok, a jednocześnie narzędzia stają się coraz bardziej dostępne. To wyścig między twórcami deepfake a badaczami próbującymi je wykrywać.

Czas na quiz!

1. Kiedy użytkownik Reddita o nazwie „deepfakes” zaczął publikować podmienione filmy?
A. W 2015 roku
B. Pod koniec 2017 roku
C. W 2019 roku
D. W 2020 roku

Prawidłowa odpowiedź: B. Pod koniec 2017 roku

2. Kto współtworzył ostrzegawczy deepfake z Barackiem Obamą w 2018 roku?
A. BuzzFeed i Jordan Peele
B. Elon Musk i CNN
C. Netflix i Steven Spielberg
D. Google i Barack Obama osobiście

Prawidłowa odpowiedź: A. BuzzFeed i Jordan Peele

3. Co stało się w 2019 roku w kontekście deepfake?
A. Deepfake zostały zakazane na całym świecie
B. Aplikacje na smartfony (na przykład Zao) po raz pierwszy pozwoliły masowo tworzyć deepfake
C. AI przestała być używana do tworzenia deepfake
D. Reddit usunął wszystkie deepfake

Prawidłowa odpowiedź: B. Aplikacje na smartfony (na przykład Zao) po raz pierwszy pozwoliły masowo tworzyć deepfake

4. Skąd pochodzi nazwa „deepfake”?
A. Od nazwy firmy, która wynalazła tę technologię
B. Od nazwy pierwszego programu do deepfake
C. Od nazwy artykułu naukowego
D. Od pseudonimu użytkownika Reddita

Prawidłowa odpowiedź: D. Od pseudonimu użytkownika Reddita

Typy deepfake

Deepfake to nie tylko podmienione twarze w filmach. Istnieje **pięć głównych kategorii** syntetycznych mediów tworzonych przez sztuczną inteligencję.

1. Zamiana twarzy w wideo (face swap) – to najbardziej znany typ. AI wykrywa twarz jednej osoby w filmie i zastępuje ją twarzą kogoś innego, klatka po klatce. Efekt może być niezwykle realistyczny.

2. Klonowanie głosu (voice cloning) – AI uczy się brzmienia czyjegoś głosu na podstawie nagrań, a następnie generuje nową wypowiedź w tym głosie. Sklonowany głos może powiedzieć dowolny tekst.

3. Obrazy generowane przez AI – tworzenie realistycznych zdjęć osób, które nigdy nie istniały. Technologia jest tak zaawansowana, że większość ludzi nie odróżnia tych zdjęć od prawdziwych fotografii.

4. Tekst generowany przez AI – fałszywe artykuły, recenzje, wiadomości i posty w mediach społecznościowych tworzone automatycznie przez duże modele językowe. Mogą naśladować styl pisanie konkretnych osób lub mediów.

5. Synteza całego ciała – najbardziej zaawansowana forma deepfake. AI generuje nie tylko twarz, ale całą postać w ruchu – gesty, chód, mowę ciała. Można stworzyć film z osobą, która wykonuje czynności, których nigdy nie robiła.

Każdy z tych typów stanowi inne zagrożenie i wymaga innych metod wykrywania. W kolejnych lekcjach omówimy szczegółowo każdą kategorię.

Czas na quiz!

1. Ile głównych kategorii deepfake wyróżniamy?

- A. Trzy
- B. Cztery
- C. Siedem
- D. Pięć

Prawidłowa odpowiedź: D. Pięć

2. Który typ deepfake jest najbardziej zaawansowany?

- A. Zamiana twarzy w filmie
- B. Klonowanie głosu
- C. Synteza całego ciała
- D. Tekst generowany przez AI

Prawidłowa odpowiedź: C. Synteza całego ciała

3. Co potrafi klonowanie głosu?

- A. Tylko nagrywać czyjś głos
- B. Generować nową wypowiedź brzmiącą jak głos danej osoby
- C. Tylko podwyższać ton głosu bez uczenia się jego barwy
- D. Tylko przestawiać kolejność słów już wypowiedzianych w nagraniu, aby odtwarzać je w nowym układzie

Prawidłowa odpowiedź: B. Generować nową wypowiedź brzmiącą jak głos danej osoby

4. Czym są obrazy generowane przez AI w kontekście deepfake?

- A. Realistyczne zdjęcia osób, które nigdy nie istniały
- B. Zdjęcia prawdziwych osób edytowane ręcznie w Photoshopie
- C. Zdjęcia zrobione aparatem z filtrem AI
- D. Zdjęcia znanych osób pobrane z internetu i lekko zmienione

Prawidłowa odpowiedź: A. Realistyczne zdjęcia osób, które nigdy nie istniały

Deepfake a zwykła edycja

Ludzie edytują zdjęcia i filmy od dziesięcioleci. Programy takie jak **Adobe Photoshop** (wydany w 1990 roku) pozwalają na ręczne retuszowanie zdjęć – usuwanie niedoskonałości, zmianę kolorów, a nawet wklejanie elementów z innych zdjęć. Ale to wszystko wymaga **ludzkiej pracy, czasu i umiejętności**.

Deepfake działa zupełnie inaczej. Oto kluczowe różnice:

Szybkość – tradycyjna edycja profesjonalnego wideo może zająć godziny lub dni. AI potrafi stworzyć przekonujący deepfake w ciągu minut, a czasem sekund.

Jakość – AI uczy się na milionach przykładów prawdziwych twarzy, głosów i ruchów. Dzięki temu generowane treści mogą wyglądać niezwykle realistycznie, nawet bez udziału profesjonalnego grafika.

Skala – jedna osoba z komputerem może przy pomocy AI wygenerować setki lub tysiące fałszywych materiałów. Ręczna edycja w takiej skali byłaby niemożliwa.

Dostępność – kiedyś edycja wideo wymagała lat nauki i drogiego oprogramowania. Dziś wiele narzędzi deepfake jest darmowych i nie wymaga żadnych umiejętności technicznych.

Ważna uwaga: samo narzędzie nie jest złe ani dobre. Photoshop można użyć do poprawienia zdjęcia wakacyjnego albo do oszustwa. Tak samo AI do generowania treści – może służyć rozrywce, edukacji lub dezinformacji. **Kluczowe jest to, w jakim celu ktoś używa technologii.**

Czas na quiz!

1. Kiedy został wydany program Adobe Photoshop?

- A. W 1990 roku
- B. W 1985 roku
- C. W 1995 roku
- D. W 2000 roku

Prawidłowa odpowiedź: A. W 1990 roku

2. Ile czasu AI potrzebuje na stworzenie deepfake?

- A. Zawsze co najmniej kilka dni
- B. Minimum tydzień
- C. Tyle samo czasu, co ręczna edycja w Photoshopie
- D. Minuty, a czasem sekundy

Prawidłowa odpowiedź: D. Minuty, a czasem sekundy

3. Która cecha NIE jest przewagą deepfake nad tradycyjną edycją?

- A. Szybkość tworzenia
- B. Gwarancja, że wynik będzie etyczny
- C. Możliwość masowej produkcji
- D. Dostępność dla osób bez umiejętności technicznych

Prawidłowa odpowiedź: B. Gwarancja, że wynik będzie etyczny

4. Co decyduje o tym, czy deepfake jest szkodliwy?

- A. Wyłącznie to, jak realistycznie odtworzono wygląd osoby na obrazie lub brzmienie jej głosu
- B. Cel, w jakim ktoś używa technologii
- C. Wyłącznie to, czy materiał jest wideo
- D. To, czy użyta aplikacja jest płatna

Prawidłowa odpowiedź: B. Cel, w jakim ktoś używa technologii

Dlaczego deepfake jest niebezpieczny?

Deepfake niesie ze sobą poważne zagrożenia, które dotyczą zarówno pojedyncze osoby, jak i całe społeczeństwa.

Erozja zaufania – to być może największe niebezpieczeństwo. Kiedy każdy film, zdjęcie czy nagranie głosu może być sfalszowane, ludzie przestają ufać nawet prawdziwym materiałom. Powstaje świat, w którym nikt nie jest pewien, co jest prawdą.

Dezinformacja – deepfake może być użyty do rozpowszechniania fałszywych informacji na masową skalę. Fałszywe wypowiedzi polityków, sfabrykowane relacje z wydarzeń, wymyślone skandale – wszystko to może wpływać na opinię publiczną i wyniki wyborów.

Naruszenie prywatności – wykorzystanie czyjegoś wizerunku bez zgody to poważne naruszenie praw osobistych. Deepfake umożliwia tworzenie kompromitujących materiałów z twarzą dowolnej osoby.

Oszustwa finansowe – przestępcy używają sklonowanych głosów, aby podszywać się pod szefów firm lub członków rodziny i wyłudzać pieniądze. Takie ataki stają się coraz częstsze.

Cyberprzemoc – to zagrożenie szczególnie dotyka młodych ludzi. Ktoś może stworzyć fałszywe, ośmieszające lub krzywdzące materiały z twarzą ofiary i rozpowszechnić je w internecie. Takie materiały mogą bardzo zranić.

Jeśli spotka to ciebie albo kogoś, kogo znasz: nie udostępniaj tego dalej, zgłoś materiał na platformie i powiedz dorosłemu, któremu ufasz. Razem zanotujcie link, nazwę konta i datę (intymnych zdjęć nie zapisuj). To nie twoja wina. Możesz też bezpłatnie zadzwonić pod 116 111 (telefon zaufania dla dzieci i młodzieży).

Dlatego właśnie umiejętność rozpoznawania deepfake nie jest tylko ciekawostką – to **niezbędna umiejętność obronna** w cyfrowym świecie.

Czas na quiz!

1. Co jest uznawane za być może największe zagrożenie ze strony deepfake?

- A. Wysokie koszty produkcji
- B. Erozja zaufania – ludzie przestają ufać nawet prawdziwym materiałom
- C. Zbyt wolne działanie technologii
- D. Brak dostępu do internetu

Prawidłowa odpowiedź: B. Erozja zaufania – ludzie przestają ufać nawet prawdziwym materiałom

2. Jak sklonowany głos może zostać wykorzystany w oszustwie finansowym?

- A. Do podszycia się pod bliską osobę proszącą o pieniądze.
- B. Do zmiany istniejącego nagrania bez możliwości stworzenia nowej prośby o przelew.
- C. Do tłumaczenia prośby o przelew na inny język bez naśladowania głosu rozmówcy.
- D. Do potwierdzenia, że rozmówca naprawdę jest bliską osobą, bo głos brzmi znajomo.

Prawidłowa odpowiedź: A. Do podszycia się pod bliską osobę proszącą o pieniądze.

3. Ktoś rozsyła ośmieszający deepfake z twarzą osoby z twojej klasy. Co warto zrobić?

- A. Przestać go dalej do całej klasy, żeby wszyscy zobaczyli, że to fałszywka
- B. Nic nie robić i nikomu nie mówić, bo to nie twoja sprawa i samo minie
- C. Odpowiedzieć sprawcy własnym, równie ośmieszającym deepfake'em
- D. Nie udostępniać, zgłosić na platformie i powiedzieć zaufanemu dorosłemu

Prawidłowa odpowiedź: D. Nie udostępniać, zgłosić na platformie i powiedzieć zaufanemu dorosłemu

4. Dlaczego rozpoznawanie deepfake jest ważną umiejętnością?

- A. Bo to niezbędna umiejętność obronna w cyfrowym świecie
- B. Bo jest wymagane na egzaminie szkolnym
- C. Bo pozwala tworzyć własne deepfake lepiej i szybciej niż inni
- D. Bo bez tego nie wolno korzystać z mediów społecznościowych

Prawidłowa odpowiedź: A. Bo to niezbędna umiejętność obronna w cyfrowym świecie

Materiały możesz bezpłatnie drukować i kopiować na potrzeby zajęć.