



[www.aileaders-project.eu](http://www.aileaders-project.eu)

# Algorytmy rekomendacyjne i

# Bańki informacyjne



Co-funded by  
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Niniejszy OER jest **studium przypadku dotyczącym algorytmów rekomendacyjnych i baniek informacyjnych** oraz ich związku z **praktykami rekomendacji algorytmicznych w reklamie**. Studium przypadku dzieli się zasadniczo na dwie części. Po pierwsze, podkreśla rolę rekomendacji algorytmicznych w kształtowaniu dystrybucji informacji i związane z tym ryzyko dla ogółu społeczeństwa. W celu wprowadzenia pojęcia baniek informacyjnych analizuje przypadek systemu rekomendacji YouTube, wyjaśniając jego mechanizm poprzez praktyczne ćwiczenie polegające na porównaniu różnych zmiennych. Dyskusje w klasie koncentrują się na rekomendacjach treści YouTube i związanych z nimi wyzwaniach, a także możliwych rozwiązaniach.

W pierwszej części studenci są również zachęceni do zastanowienia się, w jaki sposób bańki informacyjne i problemy związane z systemami rekomendacji mogą odnosić się do praktyk reklamowych. Druga część przypadku koncentruje się na rekomendacjach algorytmicznych stosowanych w reklamie internetowej, wyjaśniając, jak działają tego typu algorytmy, na przykładzie serwisu Amazon. Następnie przypadek przechodzi do potencjalnych problemów związanych z reklamą algorytmiczną, w tym analizy przypadku reklam na Facebooku i tego, jak mogą one prowadzić do dyskryminacji. Studenci są następnie zachęceni do poszukiwania i proponowania rozwiązań.

## Cel/przeznaczenie

Celem niniejszego studium przypadku jest zilustrowanie, w jaki sposób algorytmy rekomendacyjne stały się istotną częścią sposobu, w jaki konsumujemy informacje i treści w naszych społeczeństwach, a także zwrócenie uwagi na problemy i wyzwania wynikające z powszechnego stosowania tej technologii. Następnie w studium przypadku przedstawiono powiązania między algorytmami rekomendacyjnymi w ujęciu ogólnym a algorytmami stosowanymi w reklamie internetowej, zwiększając świadomość na temat potencjalnie problematycznych zastosowań i zachęcając studentów do poszukiwania potencjalnych rozwiązań.



## Oczekiwane efekty kształcenia

- 01** Student będzie potrafił **zidentyfikować** wyzwania etyczne i kwestie regulacyjne związane z personalizacją opartą na sztucznej inteligencji dla ogółu społeczeństwa i **powiązać** je z praktykami związanymi z reklamą.
- 02** Student będzie potrafił **ocenić** wpływ praktyk personalizacji na ekosystemy informacyjne.
- 03** Student będzie potrafił **zaproponować** możliwe rozwiązania mające na celu złagodzenie negatywnego algorytmów rekomendacyjnych na nasze społeczeństwa, które będą priorytetowo traktować sprawiedliwość i inkluzję.

## Sugerowane podejście metodologiczne

Niniejszy przypadek najlepiej sprawdza się jako nauczanie oparte na problemach, w ramach którego instruktorzy powinni poprowadzić dyskusję ze studentami po zapoznaniu ich z koncepcjami i technologią. Podano tematy do dyskusji i potencjalne obawy, ale instruktorzy powinni zachęcać studentów do samodzielnego myślenia i identyfikowania innych potencjalnych obaw, które mogą mieć. Przypadek YouTube powinien służyć jako materiał wprowadzający problem, który można powiązać z konkretnymi wyzwaniami, a przypadek Tournesol przedstawia możliwe rozwiązanie. W tym drugim przypadku należy zachęcić studentów do dyskusji na temat skuteczności rozwiązania oraz do zaproponowania ulepszeń, poprawek, a nawet innych możliwych rozwiązań.



### Słowa kluczowe

Algorytmy rekomendacyjne, bańki informacyjne, Reklama algorytmiczna, umiejętność korzystania z informacji i danych, stronniczość

## Wprowadzenie

Algorytmy rekomendacyjne stały się kluczowym elementem nowoczesnych platform cyfrowych, wpływając na wybory konsumentów w takich sektorach jak handel elektroniczny, rozrywka czy konsumpcja wiadomości. Chociaż algorytmy rekomendacyjne **poprawiają komfort użytkowania i zwiększają zaangażowanie**, budzą one również **poważne obawy etyczne i prawne** związane z prywatnością, stronniczością, manipulacją, a także dezinformacją i fałszywymi informacjami<sup>1</sup>.

W odniesieniu do dezinformacji i fałszywych informacji algorytmy rekomendacyjne budzą obawy ze względu na to, jak mogą wpływać na zdrowie naszych **ekosystemów informacyjnych**, a co za tym idzie, na samą **demokrację**. Aby obywatele mogli dokonywać właściwych wyborów zarówno w życiu prywatnym, jak i publicznym, niezbędny jest dostęp do właściwych informacji, które umożliwią im podejmowanie decyzji na podstawie wiarygodnych danych, do których mają dostęp.

Powiązana kwestia jest obawa, że **bańki informacyjne**<sup>2</sup> – czyli sytuacje, w których algorytmy eksponują użytkownikom tylko niektóre informacje, a inne pomijają – mogą nie tylko wpływać na jakość otrzymywanych przez nas informacji, ale także przyczyniać się do utrwalania naszych uprzedzeń, nie eksponując nam innych punktów widzenia i izolując nas w naszej własnej rzeczywistości. Istnieją obawy,

że bańki informacyjne mogą stać się tak ekstremalne, że mogą nawet doprowadzić niektórych ludzi do radykalizacji i ekstremizmu.<sup>3</sup>

W niniejszym studium przypadku analizujemy wyzwania etyczne związane z algorytmami rekomendacyjnymi w naszych społeczeństwach. Algorytmy rekomendacyjne wykorzystują **uczenie maszynowe i sztuczną inteligencję** do analizowania zachowań użytkowników i sugerowania spersonalizowanych treści. Firmy takie jak Amazon, Netflix, YouTube i Facebook stosują **zaawansowane systemy rekomendacyjne w celu zwiększenia zaangażowania użytkowników i generowania przychodów**. Systemy te opierają się na **danych** takich jak **interakcje użytkowników, historia przeglądania i preferencje**, aby tworzyć **spersonalizowane rekomendacje**.

1 W swojej książce „Filterworld” dziennikarz Kyle Chayka omawia znaczącą rolę algorytmów rekomendacyjnych w naszym społeczeństwie, zarówno tę pozytywną, jak i negatywną. Oto wywiad na temat pracy Chayki: <https://www.theverge.com/24094338/kyle-chayka-filterworld-algorithmic-recommendation-tiktok-instagram-culture-decoder-interview>

2 Oto wykład TED na temat baniek informacyjnych, znanych również jako bańki filtrujące: [https://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles?language=en](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=en)

3 Więcej informacji: W sprawie Gonzalez przeciwko Google, rozpatrywanej przez Sąd Najwyższy Stanów Zjednoczonych, omówiono możliwość, że algorytmy rekomendacyjne mogą prowadzić do radykalizacji terrorystycznej. Studenci mogą uzyskać dostęp do podcastu poświęconego tej sprawie: <https://www.techpolicy.press/a-deep-dive-into-gonzalez-v-google/>. Więcej informacji na temat tej sprawy: <https://www.oyez.org/cases/2022/21-1333>

4 Jak Netflix rekomenduje filmy do obejrzenia: <https://www.youtube.com/watch?v=nq2QtatuF7U>

5 Oto wyjaśnienie działania algorytmu rekomendacji Spotify: [https://www.youtube.com/watch?v=pGntmcy\\_HX8&t=37s](https://www.youtube.com/watch?v=pGntmcy_HX8&t=37s)

6 Jak Tik Tok Cię rozgryza: <https://www.youtube.com/watch?v=nfczi2cl6Cs>



## System rekomendacji YouTube – jak działa?

Jednym z najczęściej analizowanych systemów rekomendacji jest **algorytm YouTube**. Algorytm rekomendacji YouTube to **złożony system oparty na sztucznej inteligencji**, zaprojektowany w celu **maksymalizacji zaangażowania użytkowników** poprzez **sugerowanie filmów** dostosowanych do indywidualnych preferencji, często na podstawie **historii oglądania**. Wykorzystuje on połączenie **technik głębokiego uczenia się**, **analizy danych użytkowników** i **oceny treści**, aby zapewnić **spersonalizowane rekomendacje**.

## Co napędza algorytm rekomendacji YouTube? algorytm rekomendacji YouTube?<sup>7</sup> ..

### 01 Gromadzenie i przetwarzanie danych:

YouTube śledzi interakcje użytkowników, w tym historię oglądania, historię wyszukiwania, polubienia, niepolubienia, komentarze i udostępnienia, a także subskrypcje i preferencje dotyczące powiadomień, czas oglądania i czas trwania sesji.

### 02 Modele rankingowe i prognostyczne:

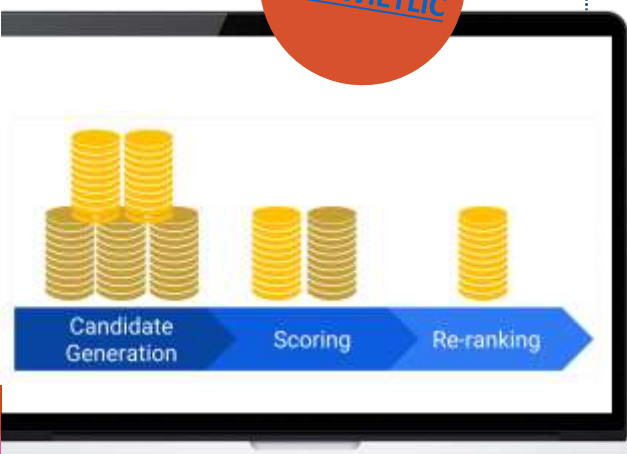
**Modele rankingowe i prognostyczne:** Algorytm wykorzystuje **modele głębokiego uczenia się** do **analizy przeszłych zachowań i przewidywania, co użytkownik prawdopodobnie obejrzy w następnej kolejności**. Ocenia różne czynniki, takie jak częstotliwość kliknięć użytkowników w polecane filmy, czas, przez jaki użytkownicy oglądają film (czas oglądania), oraz polubienia, udostępnienia i komentarze (wskaźniki zaangażowania).

### 03 Proces rekomendacji:

YouTube stosuje proces rekomendacji, który można opisać w następujący sposób:

- **Generowanie kandydatów:** System najpierw filtruje miliony filmów, aby stworzyć mniejszy zestaw odpowiednich rekomendacji.
- **Ocena punktowa:** Inny model **ocenia i klasyfikuje kandydatów w celu wybrania zestawu elementów, które zostaną wyświetlone użytkownikowi** (od 1 do 10). Ponieważ model ten ocenia stosunkowo niewielki podzbiór elementów, system może wykorzystać bardziej precyzyjny model oparty na dodatkowych zapytaniach.
- **Ponowne rankingowanie:** System musi uwzględnić **dodatkowe ograniczenia dotyczące ostatecznego rankingu**. Na przykład system **usuwa pozycje, które użytkownik wyraźnie nie polubił, lub podnosi ocenę nowszych treści**.

KLIKNIJ, ABY  
WYŚWIETLIĆ



<sup>7</sup> Zobacz <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf>

```

graph LR
    UC[user history and context] --> CG[candidate generation]
    UC --> R[ranking]
    VC[(video corpus)] -- millions --> CG
    OCS[other candidate sources] --> CG
    VF[(video features)] --> R
    CG -- hundreds --> R
    R -- dozens --> Out[ ]
    
```

Figure 2: Overview of YouTube recommendation algorithm (source: Coxington et al., ItacSys'16).

Tak **mniej więcej** wygląda działanie systemu rekomendacji YouTube<sup>8</sup>, który decyduje o tym, co użytkownik widzi na tej platformie:

KLIKNIJ, ABY WYŚWIETLIĆ

## Zobacz, jak działa algorytm YouTube

W ramach projektu **DataSkop** realizowanego przez **Platform Dynamiken**<sup>9</sup> stworzono symulację systemu rekomendacji YouTube przy użyciu **danych przekazanych** przez ochotniczych użytkowników. Dzięki temu możemy **omówić i zrozumieć**, w jaki sposób **różne zmienne** wpływają na listę filmów wyświetlanych w serwisie YouTube. **Kliknij poniższy obrazek**, aby uzyskać dostęp do symulacji:



KLIKNIJ, ABY WYŚWIETLIĆ

<sup>8</sup> <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>  
<sup>9</sup> <https://dataskop.net/recommender-sim/?en>

## Co robią platformy – działania YouTube

Jako firma, YouTube (i jego właściciel, Google) wydaje się być **świadomy problemów** związanych z algorytmami rekomendacji. Firma podjęła działania mające na celu rozwiązanie tych problemów.

Na przykład YouTube wprowadził funkcje takie jak „Up Next” (Następne), aby **zwiększyć różnorodność treści** dostarczanych użytkownikom i **zachęcić** ich do **odkrywania nowych treści**. YouTube zachęca również użytkowników do robienia przerw, aby zapobiec nadmiernemu angażowaniu się w określone rodzaje treści. YouTube **zmodyfikował** również swój algorytm, aby **uwzględnić nowe i popularne treści** w celu **zwiększenia możliwości** odkrywania różnorodnych treści.

YouTube stosuje również weryfikację faktów, aby **złagodzić skutki dezinformacji** i fałszywych informacji na swojej platformie, ogranicza

monetyzację wprowadzających w błąd treści i wprowadził dalsze zmiany w swoich algorytmach, aby **ograniczyć rozpowszechnianie szkodliwych treści**. Użytkownicy otrzymali również większą kontrolę nad **ustawieniami rekomendacji**, co pozwala im usuwać historię oglądania lub oznaczać treści jako „nieinteresujące” w celu **udoskonalenia rekomendacji**.

YouTube stara się również **wyjaśniać, jak działa jego algorytm**, aby **zwiększyć świadomość** użytkowników na temat tego, w jaki sposób wybierane są treści, które widzą.<sup>10</sup>

## Interwencje stron trzecich

Aby rozwiązać problemy związane z algorytmami rekomendacji, badane są różne rozwiązania stron trzecich, z których wiele koncentruje się na „**miękkich interwencjach niezależnych od treści**”, takich jak „**wyłączniki wirusowości**”, które tymczasowo zatrzymują algorytmiczne wzmacnianie szybko rozprzestrzeniających się postów, lub wprowadzenie „**ukierunkowanego tarcia**” oraz stosowanie podpowiedzi i wyskakujących okienek, które zachęcają użytkowników do przeczytania artykułu przed udostępnieniem.

Ponadto przepisy takie jak unijna ustawa o usługach cyfrowych (DSA) mają również na celu zapewnienie użytkownikom większej autonomii poprzez zobowiązanie dużych platform do oferowania co najmniej jednej alternatywnej metody rekomendacji, która nie opiera się na profilowaniu użytkowników. **Zgodnie z prawem** duże platformy są zobowiązane do aktywnego monitorowania i podejmowania działań mających na celu ograniczenie tego, co DSA nazywa „ryzykiem systemowym” lub zagrożeniami, **jakie algorytmy lub ich stosowanie mogą stanowić między innymi dla praw człowieka lub samej demokracji**.

Oprócz domagania się zmian w platformach takich jak YouTube, istnieją projekty skupiające się na współpracy użytkowników w celu poprawy doświadczeń osób korzystających z systemów rekomendacji, mające na celu zapewnienie im pewnej kontroli nad tym, co widzą w Internecie. Jednym z przykładów jest **Project Tournesol**, platforma open source, która zapewnia narzędzie do wspólnego podejmowania decyzji.<sup>11</sup> **Głównym celem**

**projektu Tournesol jest wspólne identyfikowanie najpopularniejszych filmów będących przedmiotem zainteresowania publicznego poprzez uzyskiwanie opinii użytkowników na temat jakości treści w celu stworzenia dużej otwartej bazy danych zawierającej oceny jakości filmów.**

Bezpośrednim efektem tej bazy danych jest **poprawa jakości filmów rekomendowanych przez YouTube dzięki wykorzystaniu opinii tysięcy osób korzystających z platformy**. Tournesol udostępnia również użytkownikom **rozszerzenie przeglądarki**, które pozwala **wyświetlać filmy rekomendowane przez społeczność** bezpośrednio na stronie głównej YouTube.

Tournesol promuje **przejrzystość, dzielenie się wiedzą i umiejętność korzystania z mediów**, udostępniając swój algorytm i cały kod źródłowy jako **oprogramowanie open source i wolne oprogramowanie**. Udostępnia również swoją bazę danych na licencji Creative Commons, mając nadzieję, że pomoże to również w udoskonaleniu badań nad algorytmami rekomendacji.

<sup>10</sup> Aby zapoznać się z przeglądem tych działań, zobacz: <https://developers.google.com/machine-learning/recommendation/overview/types>

<sup>11</sup> Więcej informacji na temat Tournesol: <https://tournesol.app/about>



## Inne przykłady

[CaptainFact](#), narzędzie internetowe, które umożliwia wspólną weryfikację filmów wideo w Internecie poprzez nakładanie na nie wiarygodnych źródeł. Innym przykładem jest [Climate Feedback](#), które wykorzystuje sieć certyfikowanych naukowców do opatrywania artykułów internetowych adnotacjami dotyczącymi ich dokładności i klasyfikowania ich według wiarygodności.



KLIKNIJ, ABY  
WYŚWIETLIĆ

Ogólnie rzecz biorąc, inicjatywy takie jak te mają na celu **przeniesienie władzy na społeczności i ekspertów**, aby **wspólnie oceniać i kontekstualizować informacje**, stanowiąc **przeciwagę dla algorytmów rekomendacyjnych opartych wyłącznie na zaangażowaniu**.

## Dyskusja

### Skąd biorą się potencjalne problemy

Algorytm YouTube był krytykowany za tworzenie **baniek informacyjnych** lub „**króliczych nor**” treści, takich jak te omówione we wstępie do niniejszego przypadku.

Problem polega na tym, że wydaje się, iż algorytmy rekomendacyjne albo nieustannie dostarczają jednego rodzaju treści, albo w coraz większym stopniu polaryzują użytkowników, co prowadzi do potencjalnych dezinformacji i ideologicznych echokomór.

Ponieważ algorytm **nieustannie aktualizuje rekomendacje** w oparciu o aktywność użytkowników **w czasie rzeczywistym**, **priorytetowo traktuje tylko treści**, które mogą **utrzymać zaangażowanie użytkowników**, co czasami może prowadzić do sytuacji, w których tylko **podobne treści są wielokrotnie rekomendowane**.

## Omówienie problemów

Korzystając z doświadczeń związanych z algorytmem YouTube, możemy teraz omówić **potencjalne problemy społeczne** związane z wykorzystaniem algorytmów rekomendacyjnych, które określają, co ludzie oglądają w Internecie, a w szczególności z bankami informacyjnymi.

### Potencjalne problemy społeczne

01

#### Kwestie prywatności:

- **Naruszenie prywatności.** Algorytmy rekomendacyjne opierają się na ogromnych ilościach danych użytkowników, często gromadzonych bez wyraźnej zgody użytkowników.
- **Brak świadomości i brak przejrzystości.** Użytkownicy mogą nie być w pełni świadomi tego, w jaki sposób ich dane są wykorzystywane i udostępniane.

02

#### Stronniczość i dyskryminacja:

- **Wzmacnianie uprzedzeń.** Algorytmy mogą wzmacniać istniejące uprzedzenia, promując treści zgodne z dotychczasowymi zachowaniami, ograniczając dostęp do różnorodnych perspektyw.
- **Obawy dotyczące sprawiedliwości.** Historyczne uprzedzenia w danych szkoleniowych mogą skutkować niesprawiedliwym traktowaniem niektórych grup demograficznych.

03

#### Manipulacja i wykorzystywanie:

- **Algorytmy, które na pierwszym miejscu stawiają zysk, a dopiero potem zdrowe ekosystemy informacyjne.** Firmy mogą projektować algorytmy, które przedkładają treści nastawione na zysk nad dobro użytkowników, zachęcając do nadmiernej konsumpcji lub niezdrowych zachowań.
- **Manipulacja użytkownikami.** Niektóre platformy wykorzystują techniki perswazyjnego projektowania w celu maksymalizacji zaangażowania, często kosztem autonomii użytkowników.

04

#### Dezinformacja i radykalizacja:

- Algorytmy rekomendacyjne mogą **wzmacniać wprowadzające w błąd lub sensacyjne treści** ze względu na ich wysoki potencjał zaangażowania.
- Algorytmy rekomendacyjne **mogą również ukrywać** przed użytkownikami **ważne informacje**, obniżając jakość informacji, które otrzymują, ale także budząc obawy dotyczące możliwego **wykluczenia** lub **dyskryminacji**.
- Badania wykazały, że **takie algorytmy przyczyniają się do** rozpowszechniania teorii spiskowych i ideologii ekstremistycznych.



## Wpływ baniek informacyjnych na firmy

Dla firm media społecznościowe stanowią doskonały sposób na nawiązanie kontaktu z klientami online, dzielenie się informacjami o swoich produktach, angażowanie odbiorców i budowanie reputacji marki. Łatwość i szybkość działania mediów społecznościowych umożliwiły dotarcie do znacznie szerszego grona odbiorców niż było to możliwe wcześniej.

Jednocześnie niesie to ze sobą nowe zagrożenia dla firm, ponieważ mogą one stać się celem dezinformacji, która jest chętnie konsumowana przez znaczną liczbę osób w Internecie.<sup>12</sup> Inne rodzaje firm, bardzo różniące się od YouTube, wykorzystują algorytmy do rekomendowania użytkownikom treści lub produktów. Doskonałym przykładem jest **Amazon**, który wykorzystuje własny algorytm do rekomendowania użytkownikom produktów, które mogą chcieć kupić.

Algorytm Amazona opiera rekomendacje produktów na **korelacjach między produktami**, a nie na podobieństwach między profilami użytkowników i ich zachowaniami zakupowymi. Dzięki temu Amazon nie musi analizować historii zakupów w całej swojej bazie danych klientów. Zamiast tego badacze Amazona wykorzystali **miarę powiązania opartą na różnicach prawdopodobieństwa**: produkt B jest powiązany z produktem A, jeśli nabywcy produktu A są bardziej skłonni do zakupu produktu B niż przeciętny klient Amazona. Im większa różnica w prawdopodobieństwie, tym większe powiązanie między produktami.<sup>13</sup>

## W jaki sposób algorytmy rekomendacyjne wyświetlają reklamy

Algorytmy rekomendacyjne służą reklamie poprzez **analizę danych użytkowników w celu dostarczania spersonalizowanych reklam**. Algorytmy te gromadzą dane z historii przeglądania, zachowań zakupowych, zaangażowania w treści oraz informacje demograficzne w celu tworzenia **profilu użytkowników**. Na podstawie tych profili reklamy są kierowane do użytkowników, którzy najprawdopodobniej uznają je za istotne.

Aby dopasować reklamy do użytkowników, platformy wykorzystują techniki takie jak **filtrowanie oparte na treści, które rekomenduje reklamy podobne do wcześniej oglądanych treści, oraz filtrowanie oparte na współpracy**. Wiele systemów wykorzystuje również modele hybrydowe i głębokie uczenie się, aby jeszcze bardziej udoskonalić rekomendacje reklam. Google, Amazon lub Meta wykorzystują własne algorytmy do dostarczania reklam.

Po wyświetleniu reklamy algorytmy stale śledzą interakcje, takie jak kliknięcia lub konwersje, aby ulepszyć przyszłe rekomendacje. Tworzy to pętlę informacji zwrotnej, która z czasem optymalizuje dostarczanie reklam. Platformy takie jak Google<sup>14</sup>, Facebook<sup>15</sup> i Amazon<sup>16</sup> wykorzystują te metody, aby zmaksymalizować skuteczność reklam, jednocześnie utrzymując zaangażowanie użytkowników.

<sup>12</sup> [https://www.pssi.cz/download/docs/8209\\_764-blog-private-companies-and-disinformation.pdf](https://www.pssi.cz/download/docs/8209_764-blog-private-companies-and-disinformation.pdf)

<sup>13</sup> Zobacz: <https://www.amazon.science/the-history-of-amazons-recommendation-algorithm>

<sup>14</sup> Reklamy Google wyjaśnione przez Google: <https://business.google.com/es/google-ads/>

<sup>15</sup> Możesz posłuchać, jak to działa, od naukowców z Amazon: <https://www.youtube.com/watch?v=GSQj27ps854>

<sup>16</sup> Tak działa system Meta: <https://www.facebook.com/business/news/good-questions-real-answers-how-does-facebook-use-machine-learning-to-deliver-ads#:~:text=How%20does%20Facebook%20decide%20which,results%20of%20our%20ad%20auction>

## Potencjalne problemy: reklama i dyskryminacja

Problem z algorytmami rekomendującymi reklamy polega na tym, że algorytmy rekomendacyjne mogą **aktywnie dyskryminować ludzi**, ponieważ **nie są one wcale neutralne**. Są to **systemy społeczno-techniczne**, które mogą utrwalać społeczne uprzedzenia, ponieważ **algorytmy są projektowane przez ludzi i szkolone na podstawie danych historycznych**, które mogą być niedokładne, niekompletne lub tendencyjne, co może **prowadzić do „uprzedzeń automatyzacji”** wobec grup marginalizowanych ze względu na rasę, płeć lub orientację seksualną.<sup>17</sup>

Może to być **szczególnie szkodliwe w przypadku reklamy ukierunkowanej**, ponieważ zainteresowania reklamowe w połączeniu z technikami optymalizacji mogą być wykorzystywane do **wykluczania określonych grup demograficznych z możliwości uzyskania mieszkania, zatrudnienia lub kredytu**, nawet bez takiego zamiaru. Nawet jeśli sposób, w jaki technologia „decyduje” o tym, które reklamy są wyświetlane, a które nie, wydaje się neutralny, w rzeczywistości jest daleki od neutralności i może

prowadzić do utrwalania historycznej dyskryminacji osób znajdujących się w trudnej sytuacji, co jest znane jako „nierówny wpływ”.<sup>18</sup>

Nierówny wpływ wynikający z praktyk reklamowych w Internecie był wielokrotnie dokumentowany. Poniżej przedstawiono jeden z najbardziej znanych udokumentowanych przypadków związanych z praktykami reklamowymi Facebooka.

## Badania ProPublica dotyczące reklamy internetowej i wykluczenia

W 2016 r. organizacja dziennikarska ProPublica udokumentowała, w jaki sposób system reklamowy Facebooka wykluczał użytkowników z niektórych reklam ze względu na ich rasę, co jest praktyką niezgodną z prawem.



<sup>17</sup> Zobacz: Noble, S. U. (2018). Algorytmy ucisku: jak wyszukiwarki wzmacniają rasizm. New York University Press.

<sup>18</sup> [https://www.law.cornell.edu/wex/disparate\\_impact#:~:text=A%20disparate%20impact%20policy%20or,by%20the%20Wex%20Definitions%20Team%5D](https://www.law.cornell.edu/wex/disparate_impact#:~:text=A%20disparate%20impact%20policy%20or,by%20the%20Wex%20Definitions%20Team%5D)

## Rozmowa na temat problemów

- 01** Czy potrafisz wskazać podobne wyzwania do tych **zidentyfikowanych w algorytmie YouTube?**
- 02** Jakie potencjalne **problemy** mogą Twoim zdaniem wynikać z **personalizowanych reklam lub rekomendacji produktów?**

## Co możemy zrobić?

Jak widać, **algorytmy rekomendacyjne w reklamie** mają na celu personalizację treści poprzez analizę danych użytkowników i przewidywanie ich preferencji. Chociaż poprawia to trafność reklam i zaangażowanie użytkowników, może również prowadzić do **niezamierzonych konsekwencji, takich jak** wykluczenie użytkowników, jak pokazuje **przypadek ProPublica**.

Algorytmy te **mogą wzmacniać uprzedzenia**, ograniczając dostęp do różnorodnych produktów, usług lub ofert pracy, co ma nieproporcjonalny wpływ na niektóre grupy. Ponadto użytkownicy, którzy nie pasują do typowych wzorców zachowań, mogą otrzymywać mniej trafnych rekomendacji, co w praktyce ogranicza ich dostęp do możliwości.

W jaki sposób reklamodawcy i platformy mogą **zrównoważyć personalizację z inkluzywnością**, aby zapewnić, że systemy rekomendacji służą szerszej i bardziej zróżnicowanej grupie odbiorców?

*Jakie rozwiązania zaproponować byś?*



## O studium przypadku

Studium przypadku ma na celu dostarczenie instruktorom konkretnych przykładów algorytmicznych rekomendacji stosowanych przez odpowiednie firmy technologiczne w celu zbadania, jaką rolę odgrywają one w społeczeństwie i w kształtowaniu naszego ekosystemu informacyjnego.

Aby zrozumieć, jak działają algorytmiczne rekomendacje i jakie są związane z nimi problemy, przeanalizowano praktyki znanych firm, takich jak YouTube, Amazon czy Meta. Studium przypadku podzielono zasadniczo na dwie części. Pierwsza część poświęcona jest ogólnie algorytmom rekomendacyjnym, badając ich rolę w określaniu, jakie informacje i treści konsumują ludzie, a także związane z tym kwestie, takie jak bańki informacyjne. Oprócz analizy problemów i możliwych rozwiązań, poprzez przykłady i praktyczne demonstracje podczas studium przypadku, instruktor powinien upewnić się, że studenci dostrzegają związek między ogólnym wykorzystaniem algorytmów do rekomendacji treści w społeczeństwie a ich zastosowaniem we współczesnych praktykach reklamowych, na czym skupia się druga część.

Część druga jest krótsza, ponieważ część pierwsza powinna zapewnić studentom wystarczający kontekst i wiedzę, aby wypełnić luki, a część druga może zostać zastąpiona innymi istotnymi zastosowaniami algorytmów rekomendacyjnych w społeczeństwie. Jeśli instruktorzy zdecydują się dostosować studium przypadku, część druga może być otwartym polem do działania, gdzie instruktor może zdecydować się skupić na algorytmach rekomendacyjnych wykorzystywanych do rozpowszechniania wiadomości, platformach

podobnych do YouTube lub innych mediach społecznościowych, takich jak Instagram lub Tik Tok, lub na modelach biznesowych, takich jak te stosowane przez Netflix, Spotify lub eCommerce w przypadku Amazon. Instruktorzy mogą również zdecydować się na skupienie się na jego zastosowaniu w finansach i bankowości, opiece zdrowotnej, transporcie (Google Maps lub Uber) lub innych obszarach poza reklamą.

Studium przypadku szczególnie dobrze nadaje się do zbadania potencjalnych wyzwań i zagrożeń związanych z wykorzystaniem algorytmów rekomendacyjnych w społeczeństwie i jest najbardziej pomocne, gdy służy podnoszeniu świadomości na temat tych kwestii, aby **skłonić studentów do krytycznego myślenia o technologiach, które stały się istotną częścią naszej gospodarki i sposobu prowadzenia działalności online, a także, w idealnym przypadku, aby skłonić ich do unikania pułapek i proponowania rozwiązań.** Dlatego też niniejsze studium przypadku powinno być szczególnie przydatne zawsze, gdy chcesz przedstawić **ogólny kontekst tego, jak technologia zmienia społeczeństwo, obowiązki lidera biznesowego zajmującego się transformacją cyfrową i postępem technologicznym, lub w każdej klasie, w której chcesz wprowadzić materiał pomagający krytycznie myśleć o technologii, z której korzystamy dzisiaj.**

## W sekcji „Zasoby

W całym materiale wykładowcy i studenci znajdą zasoby, które pomogą im zapoznać się z tematyką omawianą w studium przypadku. Kilka filmów wideo pomaga wyjaśnić, jak działają algorytmy rekomendacyjne, oraz zidentyfikować znane problemy lub zagrożenia związane z ich stosowaniem.

Znajdują się tu również linki do dłuższych tekstów, które pozwalają pogłębić wiedzę na omawiane tematy. Niektóre zasoby, zwłaszcza te przedstawione w postaci klikalnego obrazka w studium przypadku, są niezbędnymi elementami, ponieważ stanowią praktyczne ilustracje koncepcji omawianych w studium przypadku. W takich przypadkach pojawia się wyraźna prośba, aby wykładowca i studenci **kliknęli na obrazek**, który przeniesie ich do danego zasobu.

Materiały wideo i teksty są krótkie, co oznacza, że instruktor może zdecydować się na przydzielenie ich jako zadanie przygotowawcze przed przyjściem na zajęcia, aby pracować nad studium przypadku, lub, jeśli pozwala na to struktura kursu, na obejrzenie i przeczytanie podczas pracy nad studium przypadku w klasie.

W studium przypadku użyto sporo **terminów technicznych**. Przypisy często zawierają linki do filmów lub dostępnych tekstów, które wyjaśniają te terminy. Zaleca się, aby **instruktor zapoznał się z tymi terminami przed rozpoczęciem pracy z studium przypadku.**

### Wprowadzenie

Instruktor powinien przedstawić następujące ćwiczenie wstępne, aby zachęcić do refleksji nad omawianym tematem i upewnić się, że wszyscy uczniowie są na tym samym poziomie.

### Ćwiczenie wstępne:

01

Poproś uczniów, aby w grupach przejrzyli swoje profile na **Netflix<sup>19</sup>**, **Spotify<sup>20</sup>**, **TikTok<sup>21</sup>** (lub podobnych serwisach) i zapisali dziesięć najpopularniejszych programów telewizyjnych/piosenek, które widzą w aplikacji.

02

Poproś członków grupy, aby porównali to, co widzą, i omówili podobieństwa i różnice, które dostrzegają. **Dlaczego, ich zdaniem, pokazywane są im właśnie te treści?**

03

Poproś członków grupy, aby omówili, **jak ich zdaniem ma się to do reklam, które widzą w swoich feedach** w mediach społecznościowych, takich jak **Instagram** lub **YouTube**.

### Część pierwsza

Korzystając z doświadczeń związanych z algorytmem YouTube, instruktor powinien omówić **potencjalne problemy społeczne** związane z wykorzystaniem algorytmów rekomendacyjnych, które określają, co ludzie oglądają w Internecie, a w

szczegółności z bankami informacyjnymi. Instruktor powinien poprowadzić **dyskusję w klasie na następujące ważne tematy**, ale powinien również zachęcić uczniów do **przedstawienia i omówienia własnych obaw**.

### Potencjalne problemy społeczne

01

#### Kwestie prywatności:

- **Naruszenie prywatności.** Algorytmy rekomendacyjne opierają się na ogromnych ilościach danych użytkowników, często gromadzonych bez wyraźnej zgody użytkowników.
- **Brak świadomości i brak przejrzystości.** Użytkownicy mogą nie być w pełni świadomi tego, w jaki sposób ich dane są wykorzystywane i udostępniane.

02

#### Stronniczość i dyskryminacja:

- **Wzmacnianie uprzedzeń.** Algorytmy mogą wzmacniać istniejące uprzedzenia, promując treści zgodne z dotychczasowymi zachowaniami, ograniczając dostęp do różnorodnych perspektyw.
- **Obawy dotyczące sprawiedliwości.** Historyczne uprzedzenia w danych szkoleniowych mogą skutkować niesprawiedliwym traktowaniem niektórych grup demograficznych.

<sup>19</sup> Jak Netflix rekomenduje Ci rzeczy do obejrzenia: <https://www.youtube.com/watch?v=nq2QtatUF7U>

<sup>20</sup> Oto wyjaśnienie działania algorytmu rekomendacji Spotify: [https://www.youtube.com/watch?v=pGntmcy\\_HX8&t=37s](https://www.youtube.com/watch?v=pGntmcy_HX8&t=37s)

<sup>21</sup> Jak Tik Tok Cię rozgryza: <https://www.youtube.com/watch?v=nfczi2cl6Cs>

## • Potencjalne problemy społeczne

### 03 Manipulacja i wyzysk:

- Algorytmy, które na pierwszym miejscu stawiają zysk, a dopiero potem zdrowe ekosystemy informacyjne. Firmy mogą projektować algorytmy, które przedkładają treści nastawione na zysk nad dobro użytkowników, zachęcając do nadmiernej konsumpcji lub niezdrowych zachowań.
- **Manipulacja użytkownikami.** Niektóre platformy wykorzystują techniki perswazyjnego projektowania w celu maksymalizacji zaangażowania, często kosztem autonomii użytkowników.

### 04 Dezinformacja i radykalizacja:

- Algorytmy rekomendacyjne mogą **wzmacniać wprowadzające w błąd lub sensacyjne treści** ze względu na ich wysoki potencjał zaangażowania.
- Badania wykazały, że **takie algorytmy przyczyniają się do rozpowszechniania teorii spiskowych i ideologii ekstremistycznych.**

W ramach dyskusji na temat części pierwszej instruktorzy powinni zbadać „działania platform” oraz „interwencje stron trzecich” (patrz strona 5). Instruktorzy mogą sformułować dyskusję wokół możliwych rozwiązań, korzystając z następujących pytań do dyskusji:

## Kto to naprawi i w jaki sposób?

### • Pytania:

**01** Czy uważasz, że **YouTube** podejmuje wystarczające wysiłki, aby rozwiązać **problem baniek informacyjnych** na swojej platformie?

**02** Jakie **bariery** lub **konflikty interesów** dostrzegasz w działaniach **YouTube**?

**03** Co sądzisz o rozwiązaniach **Tournesol**? Czy dostrzegasz **potencjał pozytywnego wpływu**? Jakie są jego **ograniczenia**?

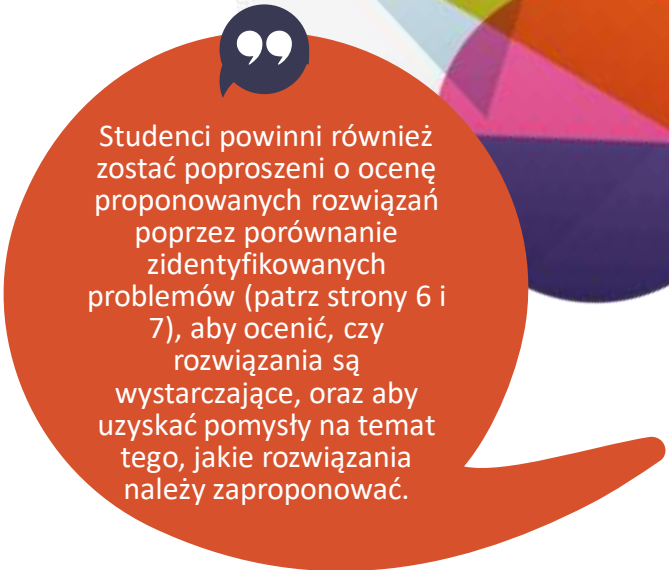
**04** Jakie strategie mogą wdrożyć platformy, decydenci i technolodzy, aby **systemy rekomendacji były bardziej zrównoważone, przejrzyste i odporne na stronniczość**, a jednocześnie nadal **angażowały użytkowników**?

**05** Jakie rozwiązania **zaproponowałbyś**?



## • Część pierwsza

Instruktor musi **upewnić się**, że studenci **dostrzegają powiązania między bardziej ogólnymi algorytmami rekomendacyjnymi, kwestiami związanymi z bańkami informacyjnymi, problemami opisanymi powyżej oraz wykorzystaniem algorytmów rekomendacyjnych w reklamie.**



Studenci powinni również zostać poproszeni o ocenę proponowanych rozwiązań poprzez porównanie zidentyfikowanych problemów (patrz strony 6 i 7), aby ocenić, czy rozwiązania są wystarczające, oraz aby uzyskać pomysły na temat tego, jakie rozwiązania należy zaproponować.

## • Aby wprowadzić dyskusję, instruktorzy mogą zadać następujące pytania:

- 01** W jaki sposób algorytmiczne systemy reklamowe **przyczyniają się do dyskryminacji**, szczególnie w takich obszarach jak rekrutacja pracowników, mieszkalnictwo i usługi finansowe? Czy potrafisz podać przykłady z życia wzięte, które ilustrują te zagrożenia?
- 02** Jaką rolę **odgrywają algorytmy rekomendacyjne** w kształtowaniu **sposobu kierowania reklam** do użytkowników?
- 03** W jaki sposób algorytmy rekomendacyjne mogą prowadzić **do wykluczenia niektórych grup demograficznych z dostępu do możliwości**?
- 04** W jaki sposób personalizacja reklam za pomocą algorytmów rekomendacyjnych tworzy „**bańki filtrujące reklamy**” i w jaki sposób może to **wzmacniać nierówności ekonomiczne lub społeczne**?
- 05** W jakim **stopniu firmy powinny ponosić odpowiedzialność za dyskryminację algorytmiczną** w kierowaniu reklam i jakie polityki lub regulacje mogłyby promować sprawiedliwość w reklamach algorytmicznych?
- 06** Jakie są **etyczne kompromisy między maksymalizacją trafności reklam** w celu zwiększenia zaangażowania i zysków a zapewnieniem inkluzywności w reklamach internetowych?
- 07** W jaki sposób firmy mogą **zrównoważyć te konkurujące ze sobą priorytety**? W jaki sposób mogą **zaradzić potencjalnym uprzedzeniom**?



## Więcej informacji

- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). Recommender systems: An introduction. Cambridge University Press. <https://www.cambridge.org/core/books/recommender-systems/C6471B59388D8A9F684C49C198691B53>
- Chayka, K. (2024). Filterworld: How Algorithms Flattened Culture. Doubleday. <https://www.penguinrandomhouse.com/books/695902/filterworld-by-kyle-chayka/>
- Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press. <https://nyupress.org/9781479837243/algorithms-of-oppression/>
- Selbst, A., boyd, d., Friedler, S., Venkatasubramanian, S. & Vertesi, J. (2018). Fairness and Abstraction in Sociotechnical Systems (August 23, 2018). 2019 ACM Conference on Fairness, Accountability, and Transparency (FAT\*), 59-68, Available at SSRN: <https://ssrn.com/abstract=3265913>
- Kourabi, F., & Tapp, A. (2021). The ethical challenges of recommender systems. arXiv. <https://arxiv.org/abs/2105.05265>
- Sweeney, L. (2013). Discrimination in online ad delivery. Communications of the ACM, 56(5), 44-54. <https://doi.org/10.1145/2447976.2447990>



Śledź naszą podróż

[www.aileaders-project.eu](http://www.aileaders-project.eu)



Co-funded by  
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.