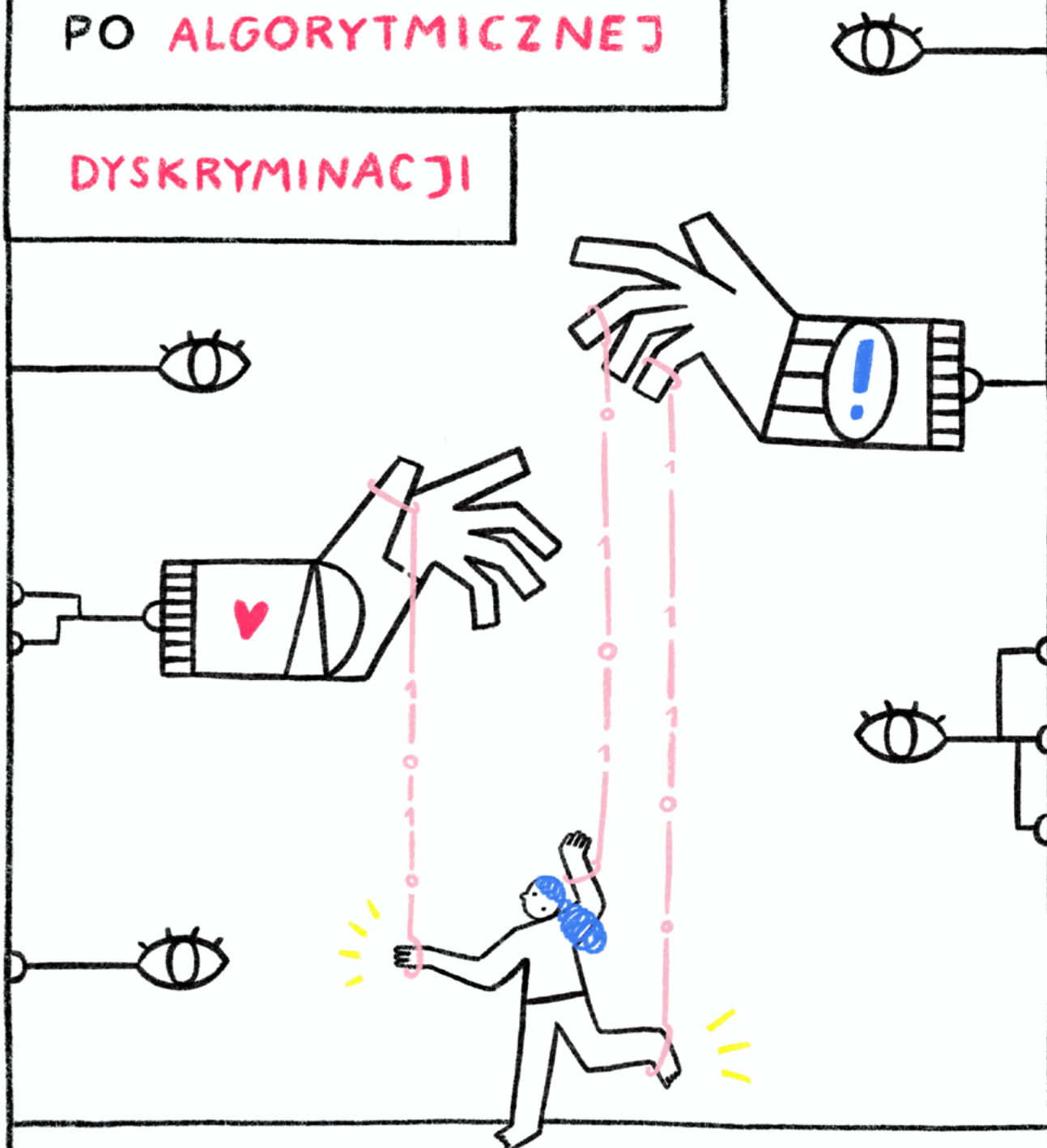


PRZEWODNIK

PO ALGORYTMICZNEJ

DYSKRYMINACJI



DLA ORGANIZACJI
POZARZĄDOWYCH,
OSÓB AKTYWISTYCZNYCH,
NAUCZYCIELI_NAUCZYCIELEK.

AMNESTY
INTERNATIONAL

BRONIMY PRAW CZŁOWIEKA

AUTOR:

Dariusz Grzemny

**REDAKCJA
MERYTORYCZNA:**

Anna Matras,
Katarzyna Salejko,
Dominika Sokołowska

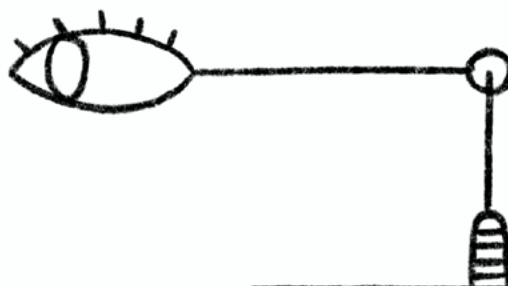
**REDAKCJA
JĘZYKOWA
I KOREKTA:**

Zuzanna Żółtowska

GRAFIKA I SKŁAD:

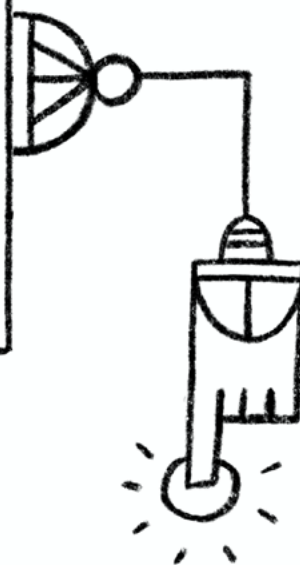
Magda Rysuje

Copyright by Stowarzyszenie
Amnesty International,
Warszawa 2026



Sfinansowane ze środków UE.
Wyrażone poglądy i opinie
są jedynie opiniami autora
lub autorów i niekoniecznie
odzwierciedlają poglądy i opinie
Unii Europejskiej lub Komisji
Europejskiej. Unia Europejska
ani organ przyznający nie ponoszą
za nie odpowiedzialności.

W całym tekście zastosowano
język równościowy, nazywany
też inkluzywnym. Jego celem
jest podkreślenie obecności
i włączenie perspektyw osób
z grup marginalizowanych,
a także tworzenie bardziej
sprawiedliwych i włączających
norm językowych, które
odzwierciedlają różnorodność
świata.



FUNDACJA
BATOREGO

**PROGRAM
RÓWNYCH
PRAW**



BRONIMY PRAW CZŁOWIEKA



**Finansowane przez
Unię Europejską**

Materiał powstał w projekcie *Algorytmy bez uprzedzeń – wzmacnianie praw mniejszości w sieci*, który jest finansowany przez Komisję Europejską w ramach Programu Równych Praw realizowanego ze środków programu CERV finansowanego ze środków Komisji Europejskiej w ramach programu *Obywatele, Równość, Prawa i Wartości* na lata 2021–2027.

SPIS TREŚCI

Internet jako przestrzeń
praw człowieka

> 4

Co to są algorytmy?

> 7

Algorytmiczna dyskryminacja
jako zagrożenie dla praw człowieka

> 10

Specyfika algorytmicznej
dyskryminacji

> 12

Algorytmy i przemoc
w sieci

> 15

Strategie ochrony
i reagowanie

> 18



INTERNET JAKO PRZESTRZEŃ PRAW CZŁOWIEKA

Internet nie stanowi odrębnej rzeczywistości, lecz jest integralną częścią życia społecznego, w której realizujemy nasze potrzeby, relacje i aktywności obywatelskie. Prawa człowieka obowiązują w sieci tak samo jak i poza nią, ponieważ nie są związane z konkretnym miejscem, tylko z samym faktem bycia człowiekiem.

Dlaczego to jest ważne?



UNIWERSALNOŚĆ PRAW CZŁOWIEKA

Prawa człowieka są powszechne i przysługują każdej osobie niezależnie od tego, w jakiej przestrzeni funkcjonuje – offline czy online. Na przykład prawo do wolności słowa obejmuje zarówno wypowiedzi w rozmowie twarzą w twarz, jak i w mediach społecznościowych.



INTERNET JAKO CZĘŚĆ ŻYCIA SPOŁECZNEGO

Dziś wiele kluczowych aktywności przeniosło się do sieci: edukacja, praca, debata publiczna. Ograniczanie praw człowieka w internecie oznaczałoby ograniczanie udziału ludzi w społeczeństwie.



OCHRONA PRZED NARUSZENIAMI

W sieci może dochodzić do takich samych nadużyć jak poza nią – np. do mowy nienawiści, dyskryminacji, pogwałcenia prywatności czy nękania. Dlatego w obu sferach potrzebne są te same standardy ochrony.



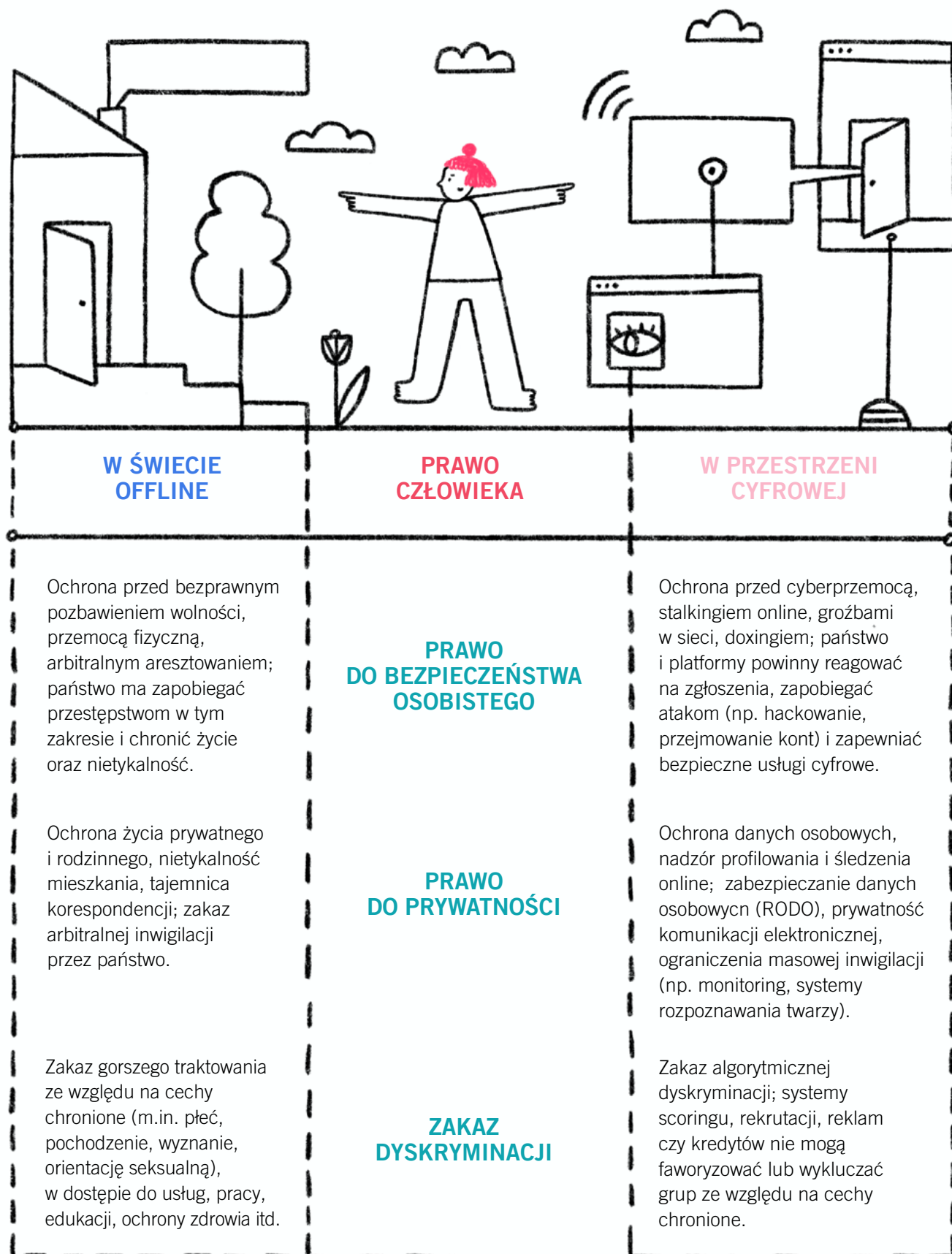
SPÓJNOŚĆ PRAW I ODPOWIEDZIALNOŚCI W SFERACH ONLINE I OFFLINE

Jeśli dane prawo obowiązuje offline, ale nie online, powstaje luka, w której łatwo o wykroczenia. Należy dążyć do sytuacji, w której zarówno użytkownicy_użytkowniczki, jak i instytucje czy firmy technologiczne będą ponosić odpowiedzialność za swoje działania.

CZY WIESZ, ŻE...

Europejska deklaracja praw i zasad cyfrowych w cyfrowej dekadzie to dokument przyjęty wspólnie przez Parlament Europejski, Radę Europejską i Komisję Europejską. Zapisano w nim, że transformacja cyfrowa w UE ma być „ukierunkowana na człowieka” i w pełni zgodna z prawami człowieka oraz Kartą praw podstawowych. Deklaracja mówi m.in. o prawie do bezpiecznego i inkluzywnego internetu, ochronie prywatności i danych osobowych, a także o tym, że żadna osoba nie powinna być wykluczona z życia społecznego z powodu braku dostępu do technologii. Dokument zawiera też zasady dotyczące odpowiedzialnego rozwoju sztucznej inteligencji i platform internetowych określone po to, aby nowe technologie wzmacniały demokrację, a nie ją osłabiały.

W dobie internetu coraz częściej mówi się o prawach cyfrowych. Nie są one czymś odrębnym od praw człowieka – stanowią raczej ich rozwinięcie i przełożenie na środowisko online. Technologia zmusza nas do reinterpretacji istniejących przepisów, np. w kwestii tego, jak stosować zakaz dyskryminacji w kontekście algorytmów albo jak chronić prawa człowieka w sytuacji masowej dezinformacji czy mikrotargetowania politycznego. A oto inne przykłady.



Absolutny zakaz zadawania cierpienia fizycznego i psychicznego przez funkcjonariuszy_ funkcjonariuszki w placówkach publicznych, m.in. w więzieniach, komendach policji, szpitalach psychiatrycznych itd.

Prawo do głoszenia poglądów, otrzymywania i przekazywania informacji, krytykowania władz; ograniczane tylko wyjątkowo (np. mowa nienawiści, bezpieczeństwo narodowe).

Prawo do rzetelnej informacji publicznej, pluralizm mediów; państwo ma działać w sposób przejrzysty.

Prawo do udziału w wyborach, tworzenia organizacji i działania w nich, do udziału w debacie publicznej; państwo ma gwarantować warunki dla demokracji.

Prawo do bezpłatnej i dostępnej edukacji podstawowej oraz równego dostępu do wyższych poziomów edukacji.

ZAKAZ TORTUR I NIELUDZKIEGO, PONIŻAJĄCEGO TRAKTOWANIA

WOLNOŚĆ SŁOWA/ WYRAŻANIA OPINII

PRAWO DO INFORMACJI

PRAWO DO UDZIAŁU W ŻYCIU PUBLICZNYM

PRAWO DO EDUKACJI

Zakaz stosowania technologii do niehumanitarnego traktowania (np. używanie systemów nadzoru do szykanowania, zastraszania; wykorzystywanie danych do wymuszania), ochrona przed cyberprzemocą prowadzącą do poważnej krzywdy psychicznej; zakaz używania AI do tortur (np. w połączeniu z urządzeniami).

Wolność wypowiedzi w mediach społecznościowych, na platformach i w komunikatorach; problem cenzurowania platform, nadużywania moderacji, ale też mowy nienawiści i dezinformacji; prawo do przejrzystych zasad moderacji treści i możliwości odwołania się od decyzji platform.

Prawo do dostępu do internetu i usług cyfrowych jako warunku korzystania z informacji; przejrzystość algorytmów rekomendacji, oznaczanie treści sponsorowanych, walka z dezinformacją przy poszanowaniu wolności słowa.

Udział cyfrowy: e-partycypacja, konsultacje online, e-głosowania, kampanie społeczne w sieci; prawo do niebycia manipulowanym poprzez mikrotargetowanie polityczne i profilowanie wyborców, ochrona integralności procesów demokratycznych w środowisku cyfrowym.

Prawo do rozwijania kompetencji cyfrowych i dostępu do edukacji online; państwo jest zobowiązane dbać, by szkoły i uczelnie zapewniały narzędzia cyfrowe i uczyły krytycznego myślenia, ochrony prywatności, bezpieczeństwa w sieci.

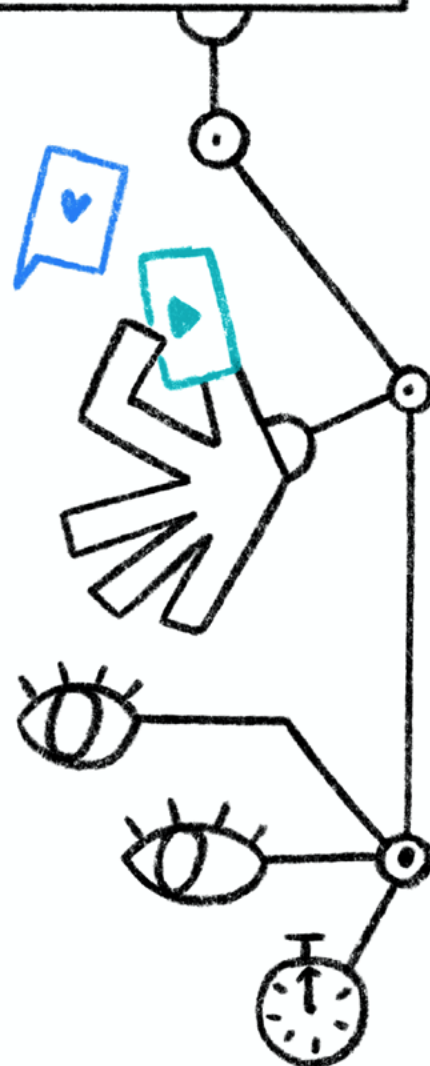
CO TO SĄ ALGORYTMY ?

Słowo „algorytm” kojarzy się wielu osobom ze skomplikowanym kodem albo sztuczną inteligencją. Tymczasem algorytm to po prostu zestaw kroków, który mówi komputerowi, co ma zrobić w danej sytuacji. To precyzyjny przepis, który na podstawie danych wejściowych generuje dane wyjściowe. Współczesne algorytmy działają bardzo szybko, przetwarzają ogromne ilości danych i często są ukryte przed wzrokiem użytkownika.

Gdy korzystamy z internetu, praktycznie zawsze spotykamy się z działaniem algorytmów. Kiedy używamy mediów społecznościowych, algorytm decyduje, jakie posty pokaże nam na górze, a jakie w ogóle się nie wyświetlą. Wyszukiwarka układa wyniki w takiej kolejności, jaka – zdaniem algorytmu – będzie dla nas najbardziej przydatna lub interesująca. Platformy z filmami i serialami podsuwają kolejne tytuły z dopiskiem „może ci się spodobać”, a sklepy internetowe wyświetlają reklamy produktów, które mogą przyciągnąć naszą uwagę. Wszystko to jest efektem szeregu obliczeń wykonanych na podstawie danych, które sami dostarczamy, korzystając z internetu.

W mediach społecznościowych algorytmy obserwują m.in. to, co ktoś lajkuje, na co reaguje, jakie pisze komentarze, przy jakich treściach zatrzymuje się na dłużej, co udostępnia. Na tej podstawie tworzą obraz tej osoby: jej zainteresowań, poglądów, relacji. Decyzje algorytmu nie są obiektywną oceną w sensie moralnym czy społecznym, lecz wynikiem rachunku prawdopodobieństwa.

Algorytm nie wie, czy dana treść jest dobra dla zdrowia psychicznego młodego człowieka, czy wspiera jego rozwój; wie za to, że podobni użytkownicy spędzali przy takich treściach wiele czasu. Nie ocenia, czy reklama jest etyczna, tylko sprawdza, czy zwiększa szansę na kliknięcie i zakup. Z punktu widzenia systemu liczy się głównie skuteczność przewidywań.



ZASTANÓW SIĘ...

> Jakie konsekwencje dla młodych osób może mieć to, że internet jest filtrowany przez algorytmy (poczucie własnej wartości, relacje w klasie/grupie, dostęp do informacji)?

> Na ile ty sam_sama czujesz, że rozumiesz, jak działają algorytmy wykorzystywanych przez siebie narzędzi (np. dziennik elektroniczny, platforma edukacyjna, media społecznościowe danej organizacji)?

> W jaki sposób możesz włączać temat algorytmów do rozmów z młodzieżą – tak aby nie straszyć technologią, ale wspierać sprawczość i krytyczne myślenie?

Algorytmy same w sobie nie są ani dobre, ani złe – ich wpływ na prawa człowieka zależy od tego, jak zostały zaprojektowane, w jakim celu są wykorzystywane oraz czy ktoś kontroluje je z perspektywy praw i potrzeb użytkowników_użytkowniczek.

POZYTYWNY WPŁYW ALGORYTMÓW NA PRAWA CZŁOWIEKA

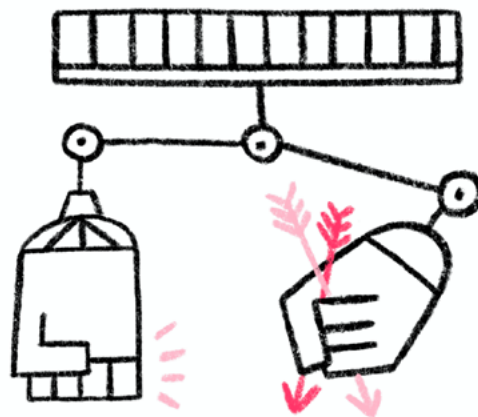
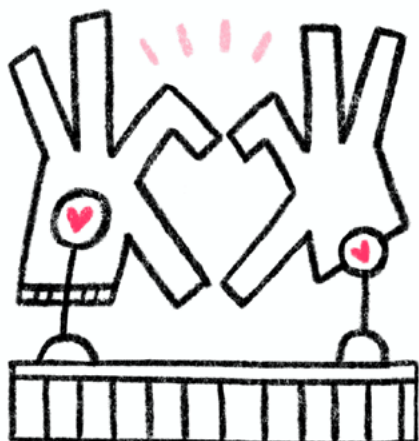
> Mogą ułatwiać dostęp do informacji i edukacji, podpowiadając materiały dopasowane do potrzeb młodych osób (np. treści w prostym języku, materiały wideo zamiast tekstu).

> Mogą wspierać włączanie osób z niepełnosprawnościami, np. dzięki automatycznemu tworzeniu napisów, tłumaczeniom materiałów na języki obce czy narzędziom do odczytu tekstu na głos.

> Mogą pomagać szybciej wykrywać treści nielegalne lub szkodliwe (np. przemocowe), co wzmacnia prawo do bezpieczeństwa i ochrony przed agresją w sieci.

> Mogą ułatwiać nagłaśnianie problemów społecznych i historii grup marginalizowanych, dzięki czemu ich głos staje się lepiej słyszalny w debacie publicznej.

> Mogą wspierać działanie organizacji społecznych, które wykorzystują analizę danych do identyfikowania obszarów nierówności, dyskryminacji czy wykluczenia i lepszego planowania działań.



ZAGROŻENIA DLA PRAW CZŁOWIEKA ZE STRONY ALGORYTMÓW

> Mogą utrwalać i wzmacniać istniejące uprzedzenia oraz nierówności, jeśli uczą się na stronniczych danych (np. częściej promować treści tworzone przez określone grupy, a pomijać inne).

> Mogą prowadzić do dyskryminacji, gdy podejmują decyzje (np. o zasięgach treści, dostępie do usług, widoczności kont) w sposób systematycznie krzywdzący określone osoby lub grupy.

> Mogą naruszać prywatność, gdy zbierają i analizują ogromne ilości danych bez kontroli i zrozumienia ze strony użytkowników_użytkowniczek, w tym dzieci i młodzieży.

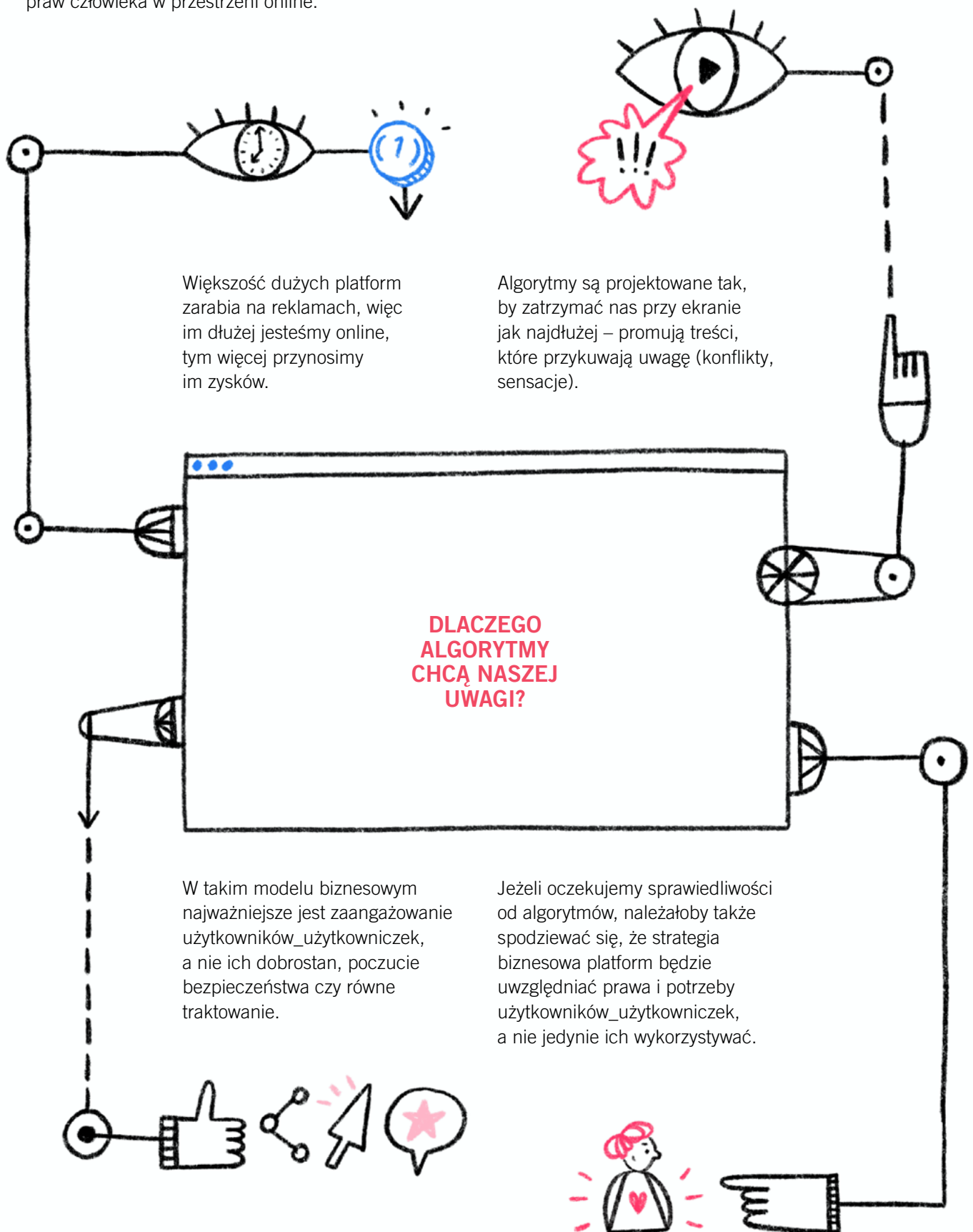
> Mogą ograniczać wolność wypowiedzi i nasz udział w życiu społecznym, jeśli błędnie oznaczają treści jako „szkodliwe” lub „nieodpowiednie” i usuwają je, nie dając uzasadnienia ani możliwości odwołania się.

> Mogą obniżać dobrostan psychiczny, promując treści emocjonujące, różnicujące lub porównawcze, co sprzyja lękowi, poczuciu gorszości czy presji bycia zawsze online.

> Mogą wzmacniać dezinformację i teorie spiskowe, jeśli okazują się skuteczne w przyciąganiu uwagi użytkowników_użytkowniczek i zatrzymywaniu ich przy ekranie.

> Mogą osłabiać poczucie sprawczości u młodych osób, które mają wrażenie, że system decyduje za nie, co zobaczą i jak zostaną ocenione, i nie rozumieją, jak to zmienić.

Większość platform internetowych zarabia przede wszystkim na reklamie i handlu uwagą użytkowników_użytkowniczek, co wynika z przyjętej przez nie strategii biznesowej. Im dłużej ktoś pozostaje na platformie, im częściej klika, przewija, ogląda i reaguje, tym więcej danych generuje i tym więcej reklam można mu wyświetlić. Nie pozostaje to bez wpływu na realizację praw człowieka w przestrzeni online.



ALGORYTMICZNA DYSKRYMINACJA JAKO ZAGROŻENIE DLA PRAW CZŁOWIEKA

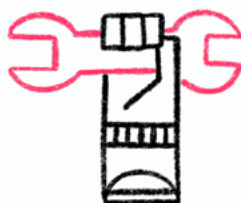
Uprzedzenia algorytmiczne, które mogą prowadzić do dyskryminacji, nie biorą się znikąd – najczęściej są odbiciem uprzedzeń i nierówności obecnych w naszym społeczeństwie.

Uprzedzenia algorytmiczne najłatwiej zrozumieć, gdy podzieli się ich źródła na trzy poziomy: dane, projektowanie systemu oraz sposób ich wykorzystywania. Każdy z nich może wprowadzać stroniczość, a razem mogą tworzyć środowisko sprzyjające dyskryminacji.



DANE

- > Algorytmy uczą się na danych historycznych: wynikach rekrutacji, ocenach, zachowaniach użytkowników_użytkowniczek, decyzjach urzędów itp.
- > Jeśli te dane zawierają informacje o wcześniejszej dyskryminacji (np. gdzieś rzadziej zatrudniano kobiety, częściej karano określone grupy), system uznaje takie sytuacje za normę.
- > Problemem jest też brak danych o pewnej części osób – grupy słabiej reprezentowane (np. mniejszości etniczne, narodowe, seksualne (LGBTI)) są dla algorytmu mniej rozpoznawalne, więc system częściej się na ich temat myli.



PROJEKTOWANIE SYSTEMU

- > Ktoś decyduje, co algorytm ma uznać za sukces (np. kliknięcia, czas spędzony na platformie, zgodność z wcześniejszymi decyzjami), jakie dane bierze pod uwagę i jak je waży.
- > Jeśli celem są głównie zysk i utrzymanie uwagi, a nie równe traktowanie i bezpieczeństwo użytkowników_użytkowniczek, system będzie premiował treści emocjonujące, dzielące, sensacyjne – nawet jeśli szkodzą one niektórym grupom.
- > Brak różnorodności w zespołach projektujących technologię sprawia, że łatwo przeoczyć perspektywę osób spoza dominującej grupy użytkowników_użytkowniczek i nie przewidzieć potencjalnej dyskryminacji.



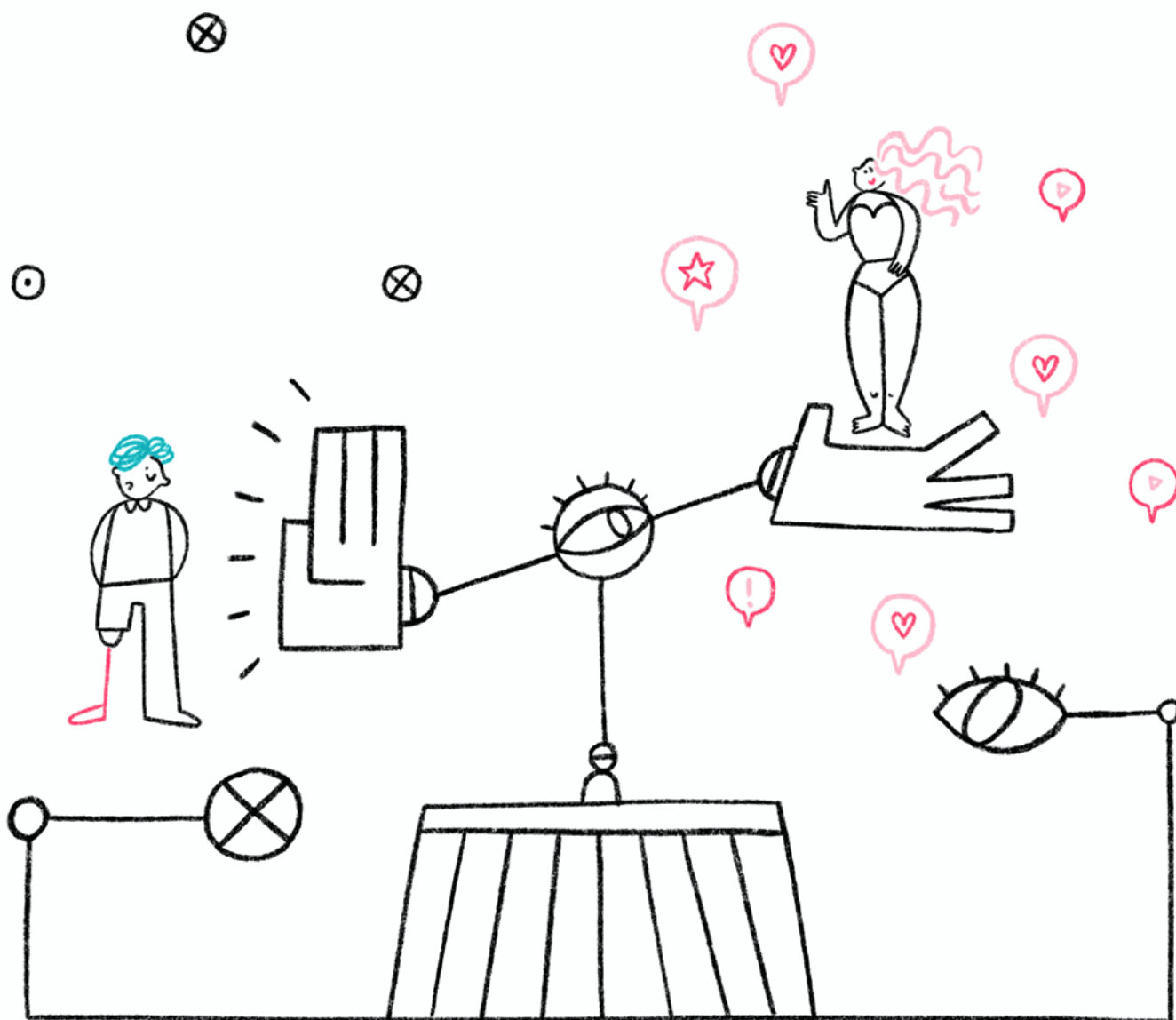
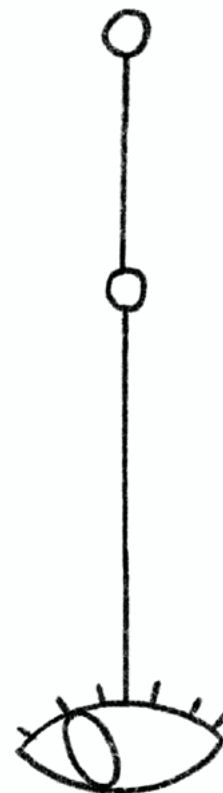
SPOSÓB WYKORZYSTYWANIA

- > Ten sam system może być pomocnym narzędziem wspierającym człowieka albo „czarną skrzynką”, której decyzji nikt nie kwestionuje.
- > Gdy szkoły, urzędy czy firmy traktują wyniki algorytmu jak obiektywną prawdę, bezkrytycznie i nie weryfikując kontekstu, uprzedzenia stają się trudniejsze do zauważenia i naprawienia.
- > Brak przejrzystości (nie wiadomo, jak działa system, gdzie zgłosić błąd, jak się odwołać) dodatkowo wzmacnia ryzyko, że algorytmiczne decyzje będą utrwały nierówności, zamiast im przeciwdziałać.

Uprzedzenia algorytmiczne nie są więc wyłącznie błędem technicznym, ale i skutkiem tego, jakie dane zbieramy, jak projektujemy systemy i jak bardzo im ufamy.

Problem staje się poważniejszy, gdy te czynniki zaczynają się przekładać na rzeczywiste gorsze traktowanie konkretnych osób. Wtedy pojawia się algorytmiczna dyskryminacja.

Dyskryminacja (poza internetem) i algorytmiczna dyskryminacja odnoszą się do tego samego zjawiska – niesprawiedliwego, gorszego traktowania pewnych osób ze względu na cechy chronione – ale dzieją się w innych środowiskach i stoją za nimi nieco inne mechanizmy. W obu kwestiach chodzi o naruszenie zasady równego traktowania i zakazu dyskryminacji – ktoś jest systematycznie gorzej traktowany ze względu na cechę, której nie powinno się używać jako kryterium różnicowania (np. płeć, pochodzenie, niepełnosprawność, wiek, orientacja seksualna, wyznanie). W obu przypadkach można mówić o dyskryminacji bezpośredniej (kogoś jawnie wykluczamy) i pośredniej (pozornie neutralna zasada w praktyce krzywdzi określoną grupę osób). Skutki dla ofiar są tu podobne: mniejszy dostęp do usług, ograniczenie szans edukacyjnych czy zawodowych, pogorszenie sytuacji materialnej, naruszenie godności osobistej, poczucie bycia mniej ważnym.



SPECYFIKA ALGORYTMICZNEJ DYSKRYMINACJI



> Zamiast konkretnej osoby (np. pracodawcy, urzędnika) decyzje podejmuje system bazujący na danych – np. algorytm rekrutacyjny, scoring kredytowy, system rekomendacji lub automatyczna moderacja treści.



> Dyskryminacja wynika najczęściej z pozornie neutralnych reguł i danych: system wykorzystuje cechy pośrednie (kod pocztowy, typ szkoły, historia aktywności), które zastępują cechy wrażliwe (np. pochodzenie, status społeczny).



> Jest trudna do zauważenia – osoba poszkodowana często nawet nie wie, że została gorzej potraktowana, bo nie widzi procesu decyzyjnego i nie ma porównania z innymi (nie wie, jakie reklamy, oferty czy decyzje zobaczyli inni użytkownicy_użytkowniczeki).



> Trudno ją zakwestionować prawnie – klasyczne prawo antydyskryminacyjne było pisane głównie z myślą o działaniach ludzi, a nie zautomatyzowanych systemów; udowodnienie złych intencji czy wskazanie konkretnej osoby odpowiedzialnej za naruszenie jest znacznie trudniejsze, gdy chodzi o sferę online.



> Szeroka skala zjawiska i szybkie tempo postępowania: jeden dyskryminujący model może w krótkim czasie zmotywować setki tysięcy decyzji, wzmacniając nierówności na masową skalę (np. w zatrudnieniu, reklamie, dostępie do kredytów itp.).



PRZYKŁAD

TARGETOWANIE REKLAM I NIEWIDZIALNOŚĆ W SIECI

Lokalna organizacja pozarządowa prowadzi kampanię informacyjną skierowaną do osób zagrożonych wykluczeniem społecznym. Decyduje się na płatne reklamy w mediach społecznościowych, bo chce dotrzeć do jak największej liczby osób potrzebujących wsparcia.

Algorytm reklamowy działa jednak inaczej, niż zakłada organizacja. Optymalizuje wyświetlanie pod kątem zaangażowania i klikalności na podstawie informacji, że młodsze, lepiej wykształcone i zamożniejsze grupy najczęściej klikają w treści pomocowe. W efekcie kampania trafia głównie do osób, które jej nie potrzebują. Osoby starsze, rzadziej aktywne w sieci, o niższym statusie ekonomicznym znikają z zasięgu. Algorytm nie dyskryminuje świadomie. Po prostu optymalizuje to, co mu kazano optymalizować – i przy okazji wzmacnia wykluczenie, które organizacja chciała zwalczać.

Algorytmiczna dyskryminacja zagraża prawom człowieka, ponieważ przenosi istniejące w społeczeństwie nierówności do sfery zautomatyzowanych decyzji, czyniąc je mniej widocznymi i działając na o wiele większą skalę.

Algorytmiczna dyskryminacja podważa podstawową zasadę praw człowieka, jaką jest równe traktowanie wszystkich ludzi, niezależnie od ich cech czy tożsamości. Systemy bazujące na danych mogą z reguły stawiać niektóre grupy osób w gorszej sytuacji – ograniczając ich dostęp do informacji, usług, edukacji czy możliwości udziału w życiu społecznym. Dzieje się to często po cichu, za sprawą nieprzejrzystych procesów decyzyjnych, przez co osoby poszkodowane nawet nie wiedzą, że zostały potraktowane niesprawiedliwie i trudno im dochodzić swoich praw. W efekcie naruszane są nie tylko prawo do równego traktowania i niedyskryminacji, ale także godność osobista, prawo do udziału w życiu publicznym oraz prawo do skutecznego środka odwoławczego. Dlatego temat algorytmów zawsze pociąga za sobą zagadnienie praw człowieka – nie chodzi tu wyłącznie o technologię.

CHECKLISTA: ZANIM ZAUFASZ ALGORYTMOWI

Zanim szkoła czy organizacja pozarządowa zdecyduje się skorzystać z jakiegoś cyfrowego narzędzia, warto by zatrzymała się na chwilę i odpowiedziała na kilka prostych, ale ważnych pytań. Algorytmy mogą wspierać naszą pracę, ale mogą też – niepostrzeżenie – utrwalac nierówności i wykluczać część osób z dostępu do szans. Poniższa checklista pomaga ocenić, czy dane narzędzie cyfrowe rzeczywiście wzmacnia prawa i dobrostan użytkowników_użytkowniczek, czy raczej niesie ryzyko algorytmicznej dyskryminacji.

CO WŁAŚCIWIE ROBI TO NARZĘDZIE?

- > Do czego ma służyć to narzędzie w naszej szkole/organizacji (ocena, selekcja, rekomendacje, moderacja, reklama)?
- > Jakie decyzje podejmie za nas, a co pozostanie w naszych rękach?

NA JAKICH DANYCH BAZUJE?

- > Skąd pochodzą dane, na których uczy się algorytm (kto je zbierał, kiedy, w jakim kontekście)?
- > Czy w tych danych dobrze widoczne są różne grupy naszych uczniów_uczennic, działaczy_działaczek (np. osoby z mniejszych miejscowości, z doświadczeniem migracji, z niepełnosprawnościami)?
- > Czy istnieje ryzyko, że dane odzwierciedlają wcześniejsze nierówności lub uprzedzenia?



JAKIE ZAŁOŻENIA WPISANO W SYSTEM?

- > Co algorytm uznaje za sukces (np. zaangażowanie, długość odpowiedzi, liczbę aktywności, liczbę kliknięć)?
- > Czy te kryteria nie faworyzują osób już uprzywilejowanych (np. z lepszym sprzętem, większym wsparciem w domu, wcześniejszym doświadczeniem)?
- > Czy ktoś sprawdził, jak narzędzie działa w przypadku różnych grup użytkowników_użytkowniczek?

KTO MA NAD NIM KONTROLĘ?

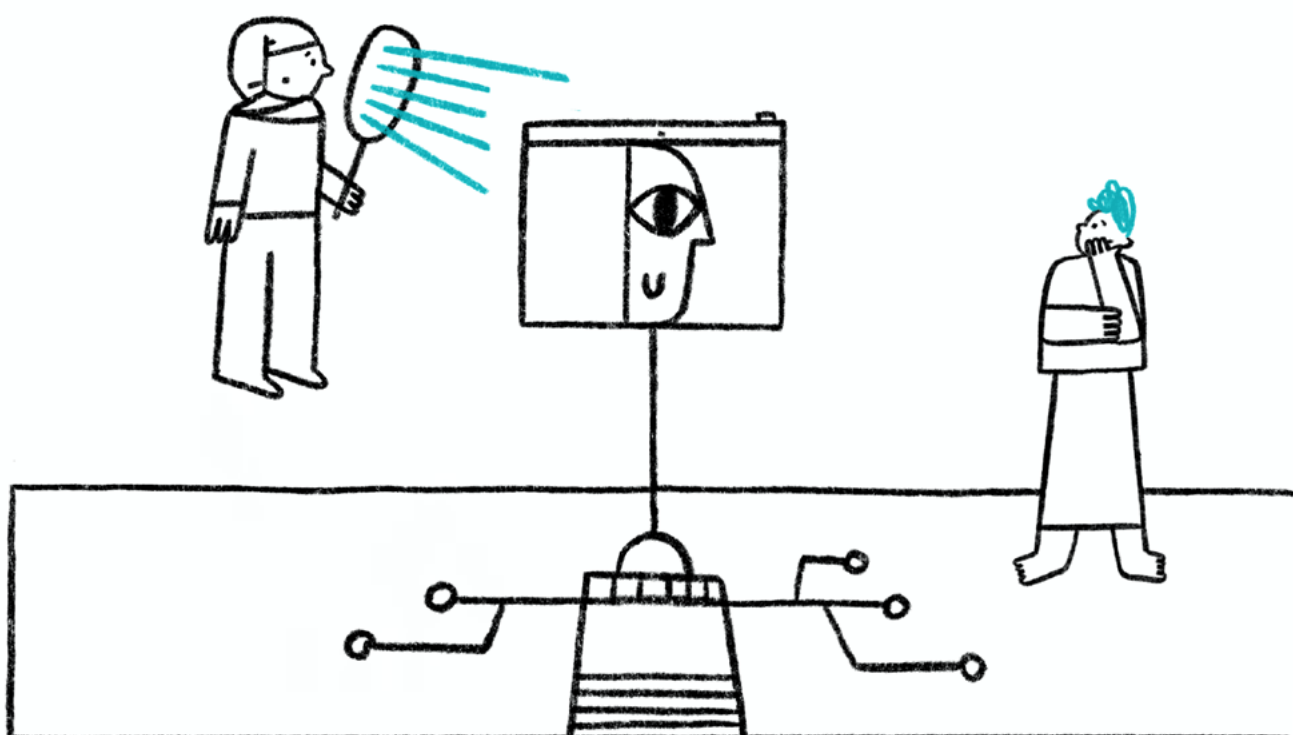
- > Czy decyzje algorytmu są traktowane jako pomoc dla człowieka, czy jako ostateczny werdykt?
- > Czy istnieje jasna procedura odwołania? Co może zrobić osoba, która uważa, że została potraktowana niesprawiedliwie?
- > Kto w szkole/organizacji odpowiada za monitorowanie skutków działania tego narzędzia?

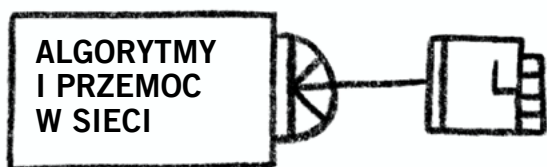
CZY MÓWIMY O TYM OTWARCIE Z UŻYTKOWNIKAMI_ UŻYTKOWNICZKAMI?

- > Czy uczniowie_uczennice, działacze_działaczki wiedzą, że w danym obszarze decyzje wspiera lub podejmuje algorytm?
- > Czy potrafimy im w prosty sposób wyjaśnić, jak działa narzędzie i co mogą zrobić, jeśli coś wydaje im się niesprawiedliwe?

JAKIE MOGĄ BYĆ SKUTKI DLA PRAW CZŁOWIEKA?

- > Czy użytkowanie tego narzędzia może prowadzić do gorszego traktowania określonych grup osób (np. przez niższe zasięgi, częstsze błędne oznaczanie, mniejszy dostęp do informacji lub projektów)?
- > Jak może wpływać na godność, poczucie sprawczości i dobrostan psychiczny młodych osób?
- > Czy korzyści z używania tego narzędzia są większe niż ryzyko z nim związane?





Przemoc to nie tylko fizyczna agresja, ale też słowa i symbole, które ranią, wykluczają, normalizują nierówności i mogą prowadzić do realnej brutalności. Uderza w to, co mamy najcenniejszego: w naszą godność. W świecie online szacunek i bezpieczeństwo są fundamentalnym prawem użytkownika_użytkowniczki, niezależnie od pochodzenia, orientacji seksualnej, wyznawanej religii, płci czy wyglądu.

Bywa, że przemoc w sieci (hejt, nękanie, upokarzające treści, groźby, przemocowe nagrania) jest wzmacniana przez mechanizmy działania platform – rekomendacje, sekcje „polecane dla ciebie”, trendy.

Algorytmy, które decydują o tym, co widzimy w mediach społecznościowych, zostały zaprojektowane po to, by jak najdłużej skupiać naszą uwagę. Niestety treści upokarzające, dotyczące konfliktów czy przemocy często wywołują silne emocje, przez co stają się dla systemu szczególnie atrakcyjne. W efekcie algorytmy mogą nie tylko nie hamować przemocy, ale wręcz ją wzmacniać, zwiększając zasięg szkodliwych materiałów i ułatwiając ich rozprzestrzenianie.



ALGORYTMY PREMIUJĄ ZAANGAŻOWANIE, NIE BEZPIECZEŃSTWO

Algorytmy platform oceniają treści pod kątem tego, jak bardzo angażują one użytkowników_użytkowniczki – liczy się czas oglądania, komentarze, udostępnienia. Treści dotyczące konfliktów i przemocy często wywołują silne emocje, więc system chętnie je promuje. W efekcie materiały szkodliwe mogą zyskiwać większy zasięg niż treści spokojne, edukacyjne.



SPIRALA REKOMENDACJI

Jedno obejrzenie lub polubienie przemocowego materiału może uruchomić spiralę rekomendacji. Algorytm zakłada, że użytkownik_użytkowniczka lubi ten typ treści, i zaczyna podsuwać kolejne, coraz bardziej podobne. Użytkownik_użytkowniczka stopniowo widzi więcej brutalnych i toksycznych materiałów, nawet jeśli trafił_trafiła na nie tylko z ciekawości.



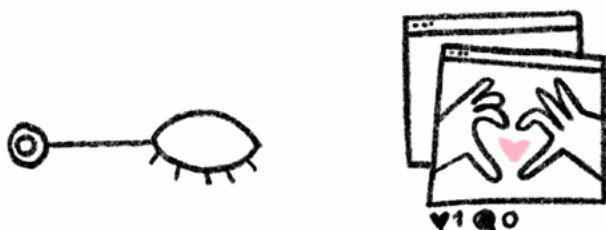
NORMALIZACJA PRZEMOCY

Częsty kontakt z przemocą w sieci prowadzi do jej osławiania i znieczulenia. To, co wcześniej szokowało, zaczyna być postrzegane jako zwykła rozrywka. Rośnie przyzwolenie na bycie biernym świadkiem_świadkinią, a czasem także gotowość, by samemu stosować przemoc dla popularności.



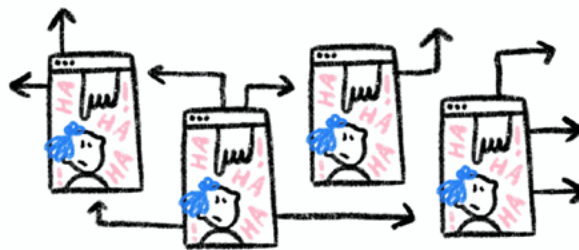
HEJT, MOWA NIENAWIŚCI I ALGORYTMY

Wątki pełne hejtu i mowy nienawiści generują wiele reakcji, więc algorytmy oceniają je jako interesujące. To może prowadzić do wzmacniania treści uderzających w osoby z grup mniejszościowych czy narażonych na wykluczenie. Osoby te widzą w sieci jeszcze więcej przekazów wymierzonych w ich tożsamość.



NIEWIDOCZNOŚĆ TREŚCI WSPIERAJĄCYCH

Spokojne, wyważone materiały edukacyjne zwykle generują mniej szybkich reakcji. Algorytmy promują je rzadziej niż kontrowersyjne filmiki. W praktyce oznacza to, że przekazy działaczy_działaczek i organizacji pozarządowych – nawet jeśli tworzą one wartościowe materiały o bezpieczeństwie – są słabiej widoczne niż treści sensacyjne.



ROZPRZESTRZANIE KRZYWDZĄCYCH TREŚCI

W pierwszej kolejności algorytmy wspierają materiały, które szybko zyskują dużą liczbę wyświetleń i reakcji. Nagranie przedstawiające upokorzenie lub przemoc może w krótkim czasie dotrzeć do bardzo szerokiego grona odbiorców_odbiorczyń. Dla osoby pokrzywdzonej oznacza to długotrwałe doświadczanie krzywdy – materiał jest wielokrotnie oglądany, komentowany i przerabiany, a możliwości sprawowania nad tym kontroli są bardzo ograniczone.



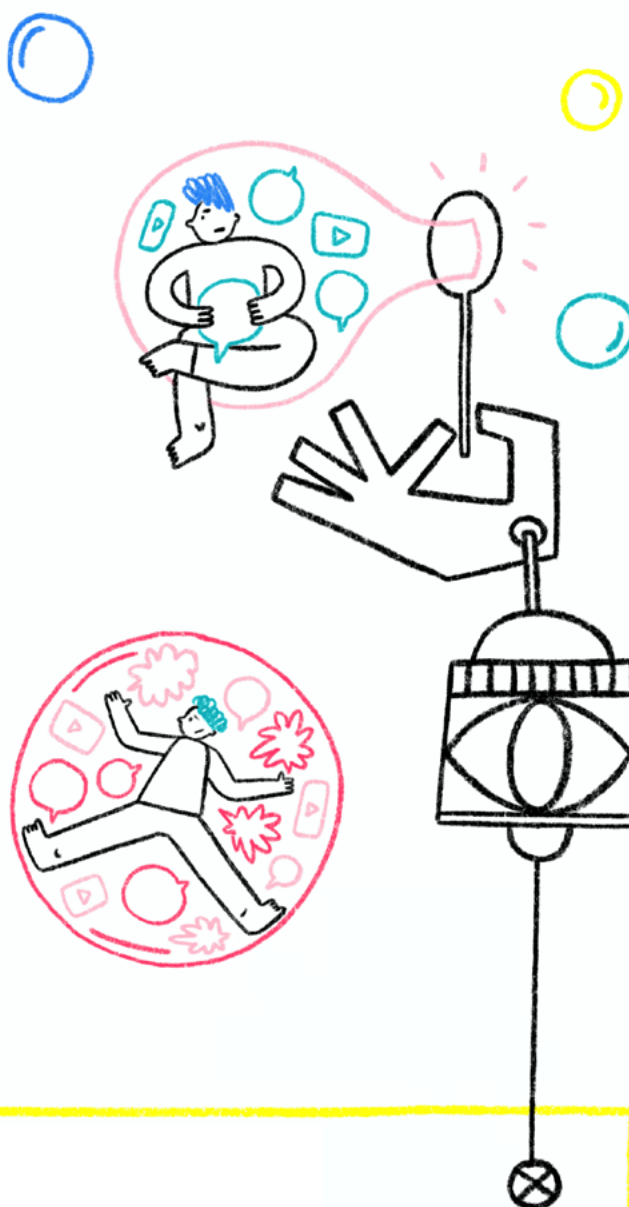
TREŚCI AUTODESTRUKCYJNE A OSOBY W KRYZYSIE

Jeśli młoda osoba szuka informacji o depresji czy samookaleczeniach, algorytm próbuje dopasować podobne treści. Obok materiałów pomocowych mogą pojawić się treści romantyzujące samobójstwo czy pokazujące samookaleczanie się jako sposób radzenia sobie z problemami. Zamiast wsparcia użytkownik_użytkowniczka dostaje coraz więcej niebezpiecznych przekazów.

BAŃKI INFORMACYJNE I KOMORY POGŁOSOWE

Bańka informacyjna to sytuacja, w której do użytkownika_użytkowniczki internetu docierają głównie takie treści, które są spójne z jego_jej dotychczasowymi poglądami, zainteresowaniami i schematami myślenia. Dzieje się tak dlatego, że algorytmy platform uczą się, co najczęściej oglądamy, lajkujemy i komentujemy, a następnie ograniczają widoczność treści odmiennych lub sprzecznych. W efekcie użytkownik_użytkowniczka ma wrażenie, że wszyscy myślą podobnie do niego_niej, a inne perspektywy są rzadkie lub niewiarygodne.

Komora pogłosowa to z kolei środowisko (np. grupa, kanał, społeczność online), w którym użytkownicy_użytkowniczki wyrażają podobne opinie, nie konfrontując ich z innymi punktami widzenia. W takim otoczeniu treści agresywne, dyskryminujące lub nienawistne mogą swobodnie krążyć, być aprobowane i stopniowo radykalizować użytkowników_użytkowniczki. Algorytmy sprzyjają zarówno bańkom informacyjnym, jak i komórkom pogłosowym, łącząc osoby podobnie reagujące na pewne treści. Rodzi to ryzyko, że młodzi użytkownicy_młode użytkowniczki uznają przemoc słowną i upokarzające materiały za normę, a nie za zachowanie przekraczające granice.



ŚMIERĆ PRZEZ TYSIĄC CIĘĆ

Raport Amnesty International *Śmierć przez tysiąc cięć* opisuje, w jaki sposób model biznesowy platformy X i zaniedbania w moderowaniu treści umożliwiły szerzenie technologicznie wspieranej przemocy ze względu na płeć (ang. *technology-facilitated gender-based violence*, TFGBV).

System rekomendacji X, czyli algorytm stojący za treściami „dla ciebie”, został stworzony w celu maksymalizacji zaangażowania użytkowników_użytkowniczek tego serwisu. Nawet przy zastosowaniu zabezpieczeń system rekomendacji umożliwia szerzenie niebezpiecznych treści, które generują silne reakcje – po to, aby utrzymać poziom zaangażowania. Model ten bazuje na inwigilacji i polega na inwazyjnym gromadzeniu danych o użytkownikach_użytkowniczkach w celu kierowania do nich odpowiednich reklam. Podobne modele stosują także inne platformy społecznościowe. Taka strategia biznesowa w połączeniu z niewystarczającym moderowaniem treści sprawia, że społeczności marginalizowane, jak społeczność LGBTI w Polsce, stają się celem wspomaganej technologicznie przemocy ze względu na płeć.

Nieskuteczność X w ograniczaniu rozpowszechniania TFGBV na tej platformie od 2020 roku sugeruje brak chęci lub niezdolność firmy do wprowadzenia niezbędnych ulepszeń w celu złagodzenia spowodowanych szkód. X powinien niezwłocznie zmienić swoją strategię i dążyć do poprawy funkcjonowania, aby wywiązać się z obowiązku poszanowania praw człowieka, m.in. przeznaczając odpowiednie zasoby na moderowanie treści oraz nawiązując bardziej owocne stosunki z organizacjami społeczeństwa obywatelskiego w Polsce i we wszystkich krajach, w których prowadzi działalność.

STRATEGIE OCHRONY I REAGOWANIE

ZNACZENIE AKTU O USŁUGACH CYFROWYCH DLA SZKÓŁ I ORGANIZACJI POZARZĄDOWYCH

Akt o usługach cyfrowych (DSA) to unijne rozporządzenie, w którym określono jednolite zasady działania platform i innych pośredników internetowych w całej Unii Europejskiej. Jego celem jest stworzenie bezpiecznego, przejrzystego i godnego zaufania środowiska online, w którym prawa użytkowników_ użytkowniczek, a zwłaszcza dzieci i młodzieży, są skutecznie chronione.

DSA opisuje przede wszystkim sposób, w jaki mają działać duże platformy i pośrednicy internetowi, co jest szczególnie istotne dla szkół oraz organizacji pozarządowych.



SILNIEJSZA OCHRONA DZIECI I MŁODZIEŻY

Platformy muszą identyfikować i ograniczać zagrożenia dla małoletnich oraz projektować usługi w sposób bezpieczny dla dzieci. Daje to szkołom i NGO dodatkową możliwość reagowania na cyberprzemoc, nękanie czy wykorzystywanie w sieci.



REKLAMY I PROFILOWANIE

Reklamy w sieci muszą być wyraźnie oznaczone oraz opatrzone informacją, kto za nie płaci i dlaczego są wyświetlane danej osobie. Profelowanie reklam kierowanych do dzieci ma być istotnie ograniczone, co warto wykorzystać jako punkt wyjścia do edukacji na temat reklamy online.



ŁATWIEJSZE ZGŁASZANIE TREŚCI NIELEGALNYCH

Serwisy muszą udostępniać prosty mechanizm zgłaszania treści nielegalnych (np. groźby, pornografia dziecięca, mowa nienawiści). Szkoły i NGO mogą z niego korzystać jako z formalnego narzędzia interwencji, obok własnych procedur wychowawczych.



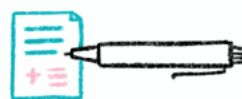
WIĘKSZE ZNACZENIE NGO JAKO STRAŻNIKÓW PRAW UŻYTKOWNIKÓW_ UŻYTKOWNICZEK

Organizacje pozarządowe mogą nie tylko zgłaszać nielegalne treści, ale też – w niektórych przypadkach – starać się o specjalny status zaufanego sygnalisty. Wzmacnia to ich rolę w zabieganiu o prawa użytkowników_ użytkowniczek oraz we współpracy ze szkołami na rzecz bezpieczeństwa w internecie.



PRZEJRZYSTOŚĆ FUNKCJONOWANIA PLATFORM I PRAWO DO ODWOŁANIA

Gdy platforma usuwa jakąś treść lub blokuje konto, musi wyjaśnić, na jakiej podstawie podjęła decyzję. Użytkownicy_ użytkowniczki – w tym szkoły i organizacje – zyskują prawo do odwołania, co ma znaczenie np. przy niesłusznym blokowaniu materiałów edukacyjnych czy kampanii społecznych.



WNIOSKI DLA SZKÓŁ I NGO

DSA daje szkołom i organizacjom pozarządowym konkretne narzędzia: możliwość skuteczniejszego zgłaszania szkodliwych treści i odwoływania się od decyzji platform, jasne standardy ochrony dzieci oraz przejrzystości reklam. Warto uwzględnić te elementy w programach wychowawczo-profilaktycznych i działaniach edukacyjnych, tak by młodzi ludzie znali nie tylko zagrożenia obecne w sieci, ale i obowiązujące w niej prawa.

ODPORNOŚĆ CYFROWA

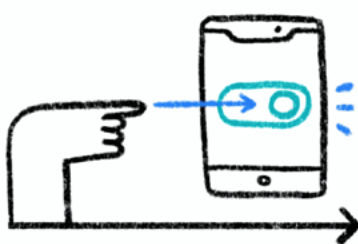
Odporność cyfrowa to umiejętność zachowania spokoju, przytomności i poczucia sprawczości w świecie, w którym to algorytmy decydują, jakie treści widzimy, a nasze dane są cenną walutą.

Kształtowanie odporności cyfrowej w przypadku dzieci i młodzieży jest kluczowe, ponieważ technologie przenikają wszystkie obszary ich życia – od nauki, przez relacje rówieśnicze, po uczestnictwo w debacie publicznej. Bez tej zdolności każde wejście do internetu może przynieść szkodę dla zdrowia psychicznego, poczucia własnej wartości czy bezpieczeństwa, zamiast wspierać rozwój. Budując cyfrową odporność, uczymy młodych rozpoznawania zagrożeń, radzenia sobie z trudnymi doświadczeniami online i szukania pomocy, gdy jej potrzebują. Takie podejście jest dalekosiężne: nie tylko chronimy młodzież przed zagrożeniami, ale też wzmacniamy u niej sprawczość, krytyczne myślenie i zdolność do radzenia sobie w sytuacjach kryzysowych. Dzięki temu szkoły i organizacje pozarządowe przygotowują młodzież do aktywnego, bezpiecznego i odpowiedzialnego korzystania z technologii.

JAK SZKOŁY I ORGANIZACJE POZARZĄDOWE MOGĄ WSPIERAĆ ODPORNOŚĆ CYFROWĄ U DZIECI I MŁODZIEŻY?



Włączać tematykę życia online do wychowania i profilaktyki: rozmawiać z młodymi ludźmi o ich doświadczeniach w sieci, normalizować proszenie o pomoc, ćwiczyć scenariusze reagowania (np. na hejt, wykluczenie, niepokojące treści).



Uczyć umiejętności samoregulacji cyfrowej: planowania czasu ekranowego, świadomego korzystania z powiadomień, rozpoznawania sygnałów przeciążenia i dbania o równowagę między sferą online i offline.



Rozwijać krytyczne myślenie i kompetencje medialne: analizować mechanizmy działania platform (algorytmy, rekomendacje), uczyć rozpoznawania manipulacji, dezinformacji i komercyjnych pułapek w aplikacjach.



Tworzyć wewnętrzną kulturę opartą na zaufaniu, a nie tylko kontroli: określać jasno zasady, ale też dawać przestrzeń do rozmowy o błędach bez zawstydzania, z naciskiem na naprawę sytuacji i wsparcie.



Włączać rodziców i opiekunów_ opiekunki: organizować spotkania, przygotowywać materiały edukacyjne i konsultacje, tak aby dorośli rozumieli, czym jest odporność cyfrowa i jak ją budować na co dzień.

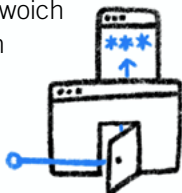


Angażować młodych: tworzyć zespoły młodych ekspertów lub wyznaczać liderów rówieśniczych, którzy będą planować i koordynować działania, współtworzyć zasady i kampanie dotyczące odpowiedzialnego korzystania z sieci.

AUDYT MOJEJ ODPORNOŚCI CYFROWEJ

Odpowiedz sobie na następujące pytania.

> Czy masz włączone 2FA (dwuetapowe logowanie) na swoich najważniejszych kontach?



> Czy w ciągu ostatniego tygodnia sprawdziłeś_sprawdziłaś źródło informacji, która cię zdenerwowała?



> Czy potrafisz spędzić 2 godziny bez sprawdzania powiadomień bez uczucia niepokoju?



> Czy masz wyłączone powiadomienia z aplikacji, która najbardziej cię rozprasza?



> Czy zdarza się, że po przeglądaniu wiadomości lub social mediów czujesz lęk, złość albo bezsilność?



> Czy masz choć jedną osobę, do której możesz się zwrócić, gdy coś trudnego wydarzy się online?

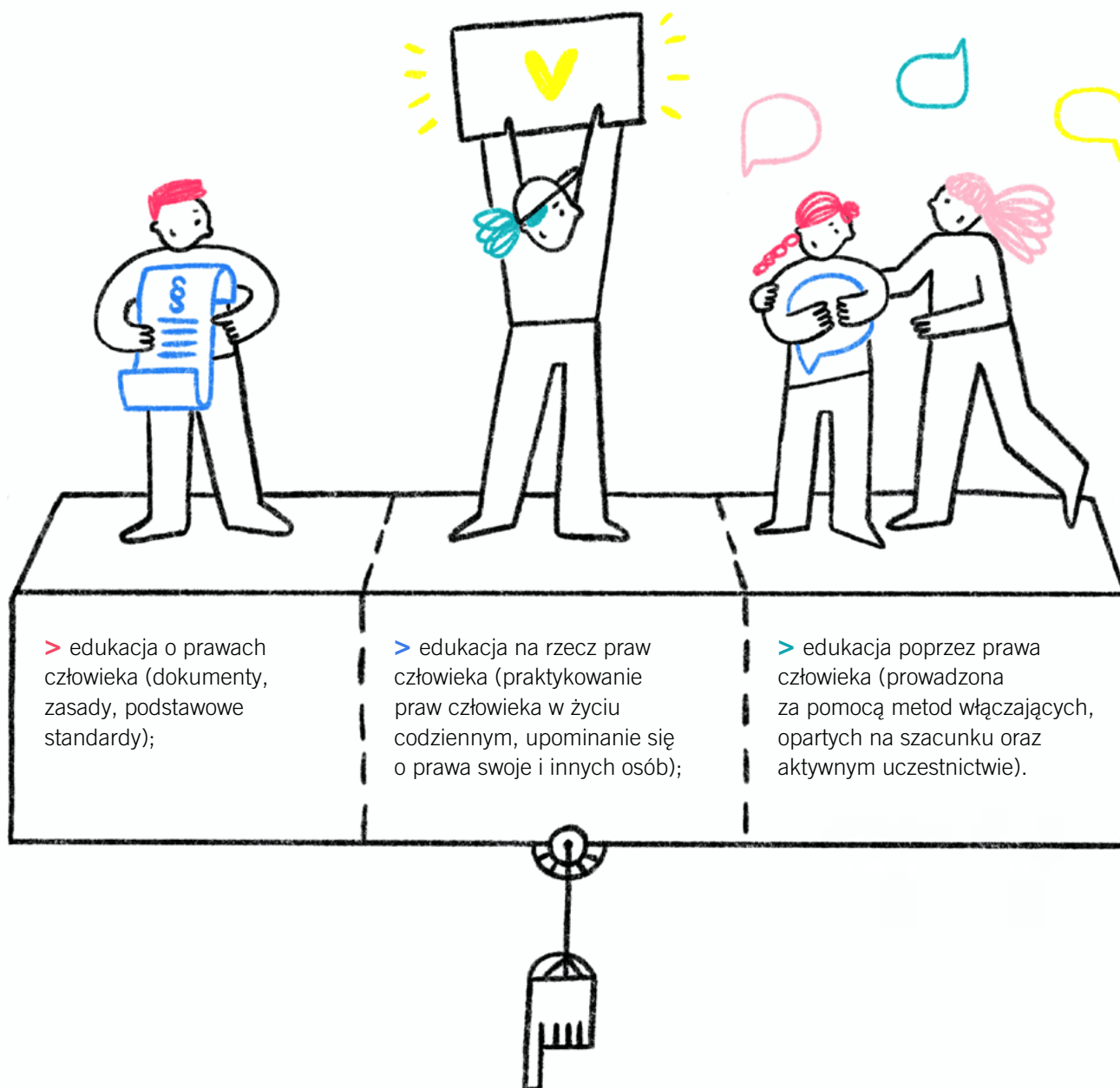


Odporność cyfrowa silnie zależy od relacji z innymi oraz poczucia, że w trudnych sytuacjach online nie zostajemy sami. Relacje i wsparcie oznaczają zarówno to, że mamy wokół siebie osoby, którym ufamy i do których możemy się zwrócić, gdy w sieci dzieje się coś niepokojącego (hejt, nękanie, wyciek danych, szokujące treści), jak i to, że potrafimy reagować, gdy to inni potrzebują pomocy. W praktyce chodzi o budowanie atmosfery, w której mówienie o trudnych doświadczeniach cyfrowych jest normalne, a nie wstydliwe, oraz o jasne komunikaty: jeśli coś cię zaniepokoi w internecie, masz prawo to przerwać, zablokować, zgłosić i poprosić o wsparcie. Dobrze działające sieci wsparcia – rodzina, grupa rówieśnicza, zespół w pracy czy społeczność szkolna – wzmacniają odporność cyfrową każdego z nas bardziej niż jakakolwiek aplikacja czy technologia.



ROLA EDUKACJI PRAW CZŁOWIEKA

Według Amnesty International edukacja praw człowieka jest zamierzonym, partycypacyjnym działaniem, którego celem jest wspieranie jednostek, grup i społeczności w zdobywaniu wiedzy, rozwijaniu umiejętności i kształtowaniu postaw zgodnych z międzynarodowymi zasadami praw człowieka. Edukacja taka rozumiana jest jako:



Problematyka zagrożeń w sieci, uprzedzeń algorytmicznych czy przemocy online jest ściśle związana z prawami człowieka, ponieważ dotyczy fundamentalnych wartości, takich jak godność, prywatność, wolność wypowiedzi i równość. W przestrzeni cyfrowej naruszane mogą być te same prawa, które chroni prawo międzynarodowe w świecie offline – np. poprzez nienawiść i dyskryminację w mediach społecznościowych, manipulację informacją czy nieujawnione formy nadzoru technologicznego. Odpowiedzialne korzystanie z technologii, przejrzystość algorytmów i przeciwdziałanie cyberprzemocy stają się więc kluczowymi elementami ochrony praw człowieka we współczesnym, cyfrowym społeczeństwie.

JAK EDUKACJA PRAW CZŁOWIEKA WSPIERA ODPORNOŚĆ CYFROWĄ?

- > Pokazuje, że w internecie obowiązują te same prawa co offline: prawo do prywatności, godności, ochrony przed przemocą i dyskryminacją, dostępu do rzetelnej informacji i udziału w życiu publicznym.
- > Wzmacnia poczucie uprawnienia i sprawczości: pokazuje, że użytkownicy nie są towarem, ale osobami chronionymi prawem, które mogą domagać się zmian, składać skargi i wspierać chroniące ich regulacje.
- > Buduje empatię i odpowiedzialność – pokazuje, jak nasze działania online (np. udostępnianie czyjegoś wizerunku, żarty z czyjejś tożsamości) mogą naruszać prawa innych; zachęca do reagowania na przemoc cyfrową i wspierania osób krzywdzonych.
- > Daje język i narzędzia – uczy, jak nazywać naruszenia (hejt, mowa nienawiści, pogwałcenie prywatności, dyskryminacja algorytmiczna) i gdzie szukać pomocy lub procedur zgłaszania.
- > Rozwija krytyczne myślenie i kompetencje medialne – uczy rozpoznawania dezinformacji, manipulacji emocjonalnej, stronniczych algorytmów, a także świadomego dobierania źródeł informacji.
- > Przypomina, że odporność cyfrowa to nie tylko indywidualne nawyki, ale też kwestia warunków systemowych: regulaminów platform, prawa, praktyk instytucji – pomaga je rozumieć i współkształtować.

JAK MOŻE REAGOWAĆ OSOBA DOŚWIADCZAJĄCA PRZEMOCY, DYSKRYMINACJI CZY UPRZEDZEŃ W SIECI?



PO PIERWSZE BEZPIECZEŃSTWO

Zadbaj o swoje granice: blokowanie, wyciszenie, ograniczanie możliwości komentowania.

Zbieraj dowody: zrzuty ekranu, linki, daty – mogą być potrzebne przy zgłoszeniu. Nie wchodź w samotne „wojny” z osobami, które wyraźnie chcą tylko prowokować.



PO DRUGIE WSPARCIE

Powiedz o sytuacji zaufanej osobie (kolega, koleżanka, nauczyciel, rodzic, pedagog, psycholog). Korzystaj z dostępnych form pomocy (telefon zaufania, lokalne organizacje, szkolny psycholog).

Ustal plan: co dalej, kto może pomóc w formalnych zgłoszeniach.



PO TRZECIE ZGŁOSZENIE

Funkcje report, „zgłoś” – szczególnie przy mowie nienawiści, nękanii, groźbach. W poważniejszych przypadkach – zgłoszenie na policję lub do innych właściwych instytucji (np. przy groźbach, przemocy seksualnej, uporczywym stalkingu).

Zaangażowanie szkoły: poinformowanie wychowawcy, pedagoga, dyrektora – zwłaszcza gdy sprawcą lub ofiarą jest uczeń_uczennica.

Algorytmy nie są ani neutralne, ani „naturalne” – odzwierciedlają decyzje ludzi, którzy je projektują, dane, na których się uczą, oraz modele biznesowe platform, które chcą jak najdłużej utrzymać naszą uwagę, często kosztem równego traktowania i dobrostanu użytkowników. To oznacza, że algorytmiczna dyskryminacja nie jest nieuchronnym skutkiem postępu technologicznego, ale czymś, czemu można i powinno się przeciwstawiać – w szkołach, organizacjach, samorządach, instytucjach publicznych. Każda osoba może zadawać trudne pytania dostawcom narzędzi, domagać się

przejrzystości, korzystać z checklist i standardów praw człowieka oraz wspierać młodych ludzi w krytycznym myśleniu o technologiach. Reagowanie na przejawy algorytmicznej niesprawiedliwości – zgłaszanie problemów, nagłaśnianie historii osób wykluczanych przez systemy, włączanie tych tematów do edukacji – to konkretna forma obrony praw człowieka w środowisku cyfrowym. Od naszej uważności, solidarności i gotowości do działania zależy, czy algorytmy będą wzmacniać istniejące nierówności, czy staną się narzędziem bardziej sprawiedliwego i inkluzywnego internetu.

