

Magdalena Filipiak  
Uniwersytet im. Adama Mickiewicza w Poznaniu

## **Sfera publiczna, inkluzja i algorytmy, czyli o racjonalności komunikacyjnej Jürgena Habermasa w kontekście wykluczenia algorytmicznego w mediach społecznościowych**

### **Wstęp**

**K**omunikacja stanowi podstawową ludzką aktywność. To właśnie ona w rozumieniu Habermasa zapewnia stabilność społeczną oraz jako przestrzeń językowa i symboliczna umożliwia trwanie społeczeństwa w czasie. Właśnie z tego powodu zaproponowana przez Habermasa teoria działania komunikacyjnego pretenduje do roli rekonstrukcji różnych form ludzkiej aktywności, z których szczególne znaczenie należy się właśnie działaniu komunikacyjnemu. Racjonalność komunikacyjna zgodnie z tym założeniem powinna przeciwdziałać społecznej polaryzacji oraz rozwiązywać na drodze dyskusywnej pojawiające się konflikty poprzez racjonalne procedury argumentacji dążące w pierwszej kolejności do zrozumienia i porozumienia. Nowoczesność, zgodnie z diagnozą Habermasa, zmierza w kierunku racjonalizacji umożliwionej poprzez relację komunikacyjną, która wyraża podstawowy stosunek społeczny i potencjał emancypacyjny. Potoczne wyobrażenia o tzw. mediach społecznościowych jako arenie nieskrępowanej wymiany poglądów mogłyby odpowiadać wizji ucieleśnienia sfery publicznej w rozumieniu Jürgena Habermasa. Warto zastanowić się, czy media społecznościowe urzeczywistniają habermasowski potencjał racjonalizacji nowoczesności oraz wynikający z tego ideał emancypacji?

Zgodnie z ujęciem Habermasa, szeroko rozumiane media elektroniczne, w tym w szczególności tzw. media społecznościowe, powinny być szczególnie predestynowane do bycia byc areną wszelkich deliberacji i debat politycznych jako medium służące sferze publicznej. Media społecznościowe mają potencjał przewyższenia błędów, które Habermas zarzuca mass mediom. Sam status mass mediów w koncepcji Habermasa

jest niejednoznaczny. Tadeusz Buksiński zauważa, że z jednej strony Habermas przypisuje im rolę tworzenia globalnej abstrakcyjnej sfery publicznej, z drugiej jednak obwinia je za krępowanie autentycznej komunikacji i zniekształcanie sfery publicznej.<sup>1</sup> Komunikacja realizowana za pośrednictwem platform internetowych mogłaby uniknąć zarzutów stawianych tradycyjnym mediom poprzez otwartość na wszystkich użytkowników, a nie pojedynczych aktorów, publicystów, którzy przedstawiają odbiorcom swoje racje. Media społecznościowe w potocznym rozumieniu charakteryzowane są poprzez inkluzywność, otwartość, dostępność, czy różnorodność. W ten sposób Internet mógłby stanowić urzeczywistnienie idei emancypacji umożliwionej przez racjonalność tkwiącą w samym języku i regułach komunikacji.

W rzeczywistości okazuje się, że media społecznościowe, mimo deklaracyjnej warstwy łączenia ludzi, raczej ich dzielą i kategoryzują. Odpowiedzialność za podsyćanie polaryzacji społecznej nie leży wyłącznie po stronie samych użytkowników mediów elektronicznych, ale technologii, która decyduje o rodzaju prezentowanej użytkownikom treści. Wykorzystywanie systemów sztucznej inteligencji w życiu ludzi, oprócz wielu korzyści, niesie ze sobą również szereg zagrożeń. W perspektywie funkcjonowania sfery publicznej jednym z poważniejszych problemów jest stronnicze czy wykluczające działanie algorytmów. Pojęcie sfery publicznej jest podstawowe dla Habermasowskiej teorii demokracji. Jest to tkanka życia społecznego, w którym toczą się wszelkie dyskusje na temat wartości, potrzeb, prawomocności władzy, podstaw porządku politycznego. Jest to przestrzeń, w której krytykuje się władzę, broni i uzasadnia swoje przekonania. Jako jeden z warunków demokracji sfera publiczna tworzy przestrzeń debaty politycznej, w której obywatele mogą prezentować swoje poglądy na temat prawomocności istniejących struktur władzy, w ramach której kształtuje się wola obywateli i opinie na temat funkcjonowania życia społecznego, władz i instytucji.<sup>2</sup> Zasadniczy problem znaczenia mediów społecznościowymi dla racjonalności komunikacyjnej i funkcjonowania sfery publicznej znajduje się zatem po stronie algorytmów i wyraża się w unieważnieniu zakładanej przez Habermasa inkluzywności.

## **Sfera publiczna w mediach społecznościowych**

Jeśli komunikacja według Habermasa niesie ze sobą potencjał emancypacji, obietnicę racjonalizacji, jest tkanką społeczeństwa, fundamentem racjonal-

<sup>1</sup> T. Buksiński, *Publiczne sfery i religie*, Wydawnictwo IFUAM, Poznań 2011, s. 31.

<sup>2</sup> A. Dupeyrix, *Zrozumieć Habermasa*, przeł. M.N. Wróblewska, Oficyna Naukowa, Warszawa 2013, s. 147-165.

ności, społecznej solidarności, podstawą demokracji, to czy media społecznościowe również niosą ze sobą taki potencjał? Początkowe analizy mediów społecznościowych podkreślały szczególnie mocno ten właśnie aspekt. Media społecznościowe mogłyby stanowić platformę społecznej deliberacji ze względu na swój potencjalnie inkluzyjny charakter. Trudno jest jednoznacznie scharakteryzować, czym są media społecznościowe. Grupa badaczy (Thomas Aichner, Matthias Grünfelder, Oswin Maurer, Deni Jegeni), posługując się metodologią systematycznego przeglądu anglojęzycznej literatury, przeanalizowała szereg definicji mediów społecznościowych pomiędzy rokiem 1994 a 2020. Stwierdzili, że według stanu na styczeń 2020 r. istnieje ponad 110 000 publikacji, które mają w tytule termin „media społecznościowe”. Wytypowali 88 artykułów, z których wyciągnęli 21 reprezentacyjnych definicji mediów społecznościowych. Pierwszym ciekawym wnioskiem było określenie gęstości leksykalnej, tj. procentowego udziału przymiotników, rzeczowników, przysłówków, które występowały łącznie. Okazało się, że w 57,5% najczęściej używanym słowem z 23 wystąpieniami jest „społeczny”, następnie „ludzie” z 12 wystąpieniami oraz „wirtualny”, „treść”, „użytkownik” i „sieć” z 8 wystąpieniami każdy. W zakresie fraz dwuwyrazowych, po „sieci społecznościowe” (8 wystąpień) następuje „społeczne” media” i „sieci społecznościowe” (po 5 wystąpień), a także „społeczności wirtualne” (4 wystąpienia). Badacze zauważyli również, że zasadnicza zmiana w definiowaniu tego, czym są media społecznościowe, zaszła po roku 2010. Wcześniejsze definicje koncentrowały się na łączności pomiędzy ludźmi o wspólnych zainteresowaniach. Po roku 2010 definicje podkreślały możliwość tworzenia i generowania treści przez samych użytkowników.<sup>3</sup> W definicję tę wpisuje się zaproponowane przez Andreasa M. Kaplana i Michaela Haenleina w 2010 roku określenie mediów społecznościowych jako zbioru internetowych aplikacji stworzonych na podstawie ideologii Web 2.0, które umożliwiają tworzenie i wymianę treści generowanych przez użytkowników.<sup>4</sup>

W przypadku sfery publicznej, jeśli przyjmiemy za Habermasem, że sfera publiczna jest przestrzenią swobodnego wyrażania opinii, debat, dyskusji zarówno w wymiarze formalnym, jak i nieformalnym, która niesie ze sobą potencjał przenoszenia problemów z peryferii do centrum działań politycznych, to media społecznościowe byłyby idealnym miejscem do tego typu działalności. Habermas zakładał, że zróżnicowane, zmieniające się w czasie techniki praktyk komunikacyjnych od druku, poprzez radio i telewizję

<sup>3</sup> T. Aichner, M. Grünfelder, O. Maurer, D. Jegeni, *Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019*, *Cyberpsychology, Behavior and Social Networking*, no. 24, no 4, 2021, s. 217-220.

<sup>4</sup> A.M. Kaplan, M. Haenlein, *Users of the world, unite! The challenges and opportunities of Social Media*, „Business Horizons” vol. 53, 2010, s. 59.

umożliwiają powszechny dostęp do uczestnictwa w różnorodnych płaszczyznach sfer publicznych. Umożliwiają one wypowiedanie się na zróżnicowane tematy w dowolnych kontekstach, co ułatwia utworzenie się bogatej i wielopłaszczyznowej siatki stref publicznych, np. lokalnych i globalnych, naukowych, literackich i politycznych itp.<sup>5</sup>

Sfera publiczna w rozumieniu Habermasa jest gwarancją demokracji, ochroną przed interesownością i manipulacją w dyskursie publicznym. Jej forma umożliwia przejrzystą komunikację w ramach debat społecznych i politycznych. Jest ona zatem miejscem łączącym ideały Habermasa, czyli emancypację umożliwiającą przez medium języka i komunikacji, potencjał krytyczny oraz poszukiwanie prawdy.<sup>6</sup> Funkcjonowanie sfery publicznej opiera się na normatywnym umocowaniu w postaci rozumu komunikacyjnego. Sfera publiczna mediuje pomiędzy systemem politycznym a światem życia. Wobec polityki pełni funkcje opiniujące, legitymizujące, krytyczne. W świecie życia upublicznia, racjonalizuje i przekazuje do rozwiązania pojawiające się problemy. Sfera publiczna nie jest organem decyzyjnym, ponieważ ma charakter wyłącznie debatujący, służący wytworzeniu opinii publicznej.<sup>7</sup>

Stosunek Habermasa do mass mediów nie jest przy tym jednoznaczny. Z jednej strony są one szansą na spopularyzowanie sfery publicznej do postaci globalnej, z drugiej jednak narażają ją na różnorodne patologie. Jednym ze skutków działania mass mediów na sferę publiczną jest tworzenie przez poszczególne przedsiębiorstwa medialne, np. wydawnictwa czy stacje telewizyjne, własnych form władzy w postaci skomercjalizowania publiczności, a nie kształtowania racjonalnej opinii publicznej. Łączy się to również z ewentualnym problemem manipulacji i propagandy, którymi mogą posługiwać się przedsiębiorstwa medialne w celu kreowania rzeczywistości zgodnie z interesami określonych korporacji gospodarczych czy politycznych. Mass media dezintegrują w opinii Habermasa tkankę społeczną, tworząc pseudopolityczne fora, w których występuje rozwarstwienie na aktorów czynnie uprawiających politykę, czyli publicystów, i biernych konsumentów. Mass media mogą również odpolitycznić sferę publiczną na inne sposoby, np. poprzez personalizowanie, fragmentaryzowanie czy skomercjalizowanie kwestii politycznych.<sup>8</sup>

---

<sup>5</sup> J. Habermas, *Filozoficzny dyskurs nowoczesności*, przeł. M. Łukasiewicz, Universitas, Kraków 2005, s. 404.

<sup>6</sup> A. Dupeyrix, *op. cit.*, s. 45.

<sup>7</sup> T. Buksiński, *op. cit.*, s. 25, 26.

<sup>8</sup> *Ibidem*, s. 31.

## Wykluczenie algorytmiczne

Rozwój technologii pozwala na gromadzenie i przetwarzanie coraz większej ilości danych określanych mianem big data ze względu na swoją objętość oraz rosnące znaczenie w wielu dziedzinach. Do analizy zbiorów big data wykorzystywane są technologie i narzędzia sztucznej inteligencji. Oczywiście niesie to ze sobą trudne do przecenienia korzyści, jak również rodzi szereg dylematów. W kontekście sfery publicznej w rozumieniu Habermasa warto zapytać, czy technologia służąca do organizowania wyników wyszukiwań lub wyświetlania postów w mediach społecznościowych jest stronnicza? Działania ludzi w sieci są w znacznej mierze sprofilowane przez sztuczną inteligencję. Na podstawie danych umieszczonych świadomie lub nieświadomie w sieci algorytmy wyznaczają treści, filmy, reklamy, które dla konkretnego użytkownika będą interesujące. Oznacza to, że wybory ludzi, np. co kupić, przeczytać, obejrzeć w znacznej mierze zdane są na łaskę algorytmów sztucznej inteligencji. Ogromny zasób materiałów w sieci w połączeniu z profilowaniem użytkowników prowadzi do prezentowania im treści określonego typu oraz ograniczenia pozostałych. Algorytm mediów społecznościowych zaprojektowany jest po to, aby wyświetlać jak najbardziej interesujące i atrakcyjne dla konkretnego użytkownika treści, ponieważ czas spędzany przed ekranem stanowi mierzalną wartość. Model ten pozostaje w zgodzie z definiowaniem współczesnego kapitalizmu przez pryzmat kapitalizmu nadzoru, inwigilacji czy komunikacji, w którym rosnącą wartość symboliczną oraz ekonomiczną mają dane behawioralne pozyskiwane na podstawie ludzkiej obecności w świecie cyfrowym.<sup>9</sup>

Algorytm w znaczeniu nieformalnym to ściśle określona procedura polegająca na przekształceniu danych wejściowych w dane wyjściowe, czyli wynik. Algorytm jest również listą wytycznych służących rozwiązaniu konkretnego problemu obliczeniowego. Wskazanie kwestii do rozwiązania polega na zdefiniowaniu ścisłych instrukcji względem relacji pomiędzy danymi wejściowymi a wyjściowymi. Algorytm jest procedurą obliczeniową gwarantującą poprawne zachowanie wskazanej relacji.<sup>10</sup> W życiu codziennym często posługujemy się różnego rodzaju algorytmami, np. wprowadzając w życie przepisy kulinarne czy podejmując decyzje. Procedury oparte na algorytmach wykorzystywane są powszechnie m.in. w sektorze usług medycznych, bankowych, administracji, przemyśle itd. Rosnące znaczenie przemysłu big

<sup>9</sup> J. Dean, *Communicative Capitalism and Revolutionary Form*, „Millennium: Journal of International Studies”, 3(47), 2019, *passim*.

<sup>10</sup> T.H. Cormen, Ch.E. Leiserson, R.L. Rivest, *Wprowadzenie do algorytmów*, wyd. 4, Wydawnictwa Naukowo-Techniczne, Warszawa 2001, s. 21, 22.

data, możliwość przetwarzania coraz większych i trudniejszych ilości danych, rozwój sztucznej inteligencji oraz uczenia maszynowego otwierają nowe perspektywy wykorzystania algorytmów. Świat technologii algorytmów fundamentalnie zmienił sposób funkcjonowania ludzi w każdej płaszczyźnie życia, od codzienności — jak poprzez rekomendację filmów, książek, dobór ubezpieczenia, wskazanie drogi czy produktów konsumenckich — po wykorzystanie ich w instytucjach społecznych, np. sądownictwie czy szkolnictwie. Algorytmizacja procesów decyzyjnych oparta na maszynowym uczeniu się może mieć swoją ciemną stronę określaną mianem stronniczości algorytmów, czy algorytmicznego wykluczenia (*algorithmic bias*).

Pojęcie wykluczenia algorytmicznego dotyczy niezamierzonych, systematycznych i powtarzalnych błędów wynikających z zastosowania algorytmów, które prowadzą do uprzywilejowania określonych osób lub grup kosztem innych. Przyczynia się to do pogłębiania nierówności społecznych oraz niesprawiedliwego traktowania ludzi. Szczególnie jaskrawym przykładem są błędy w systemach identyfikacji twarzy, które wykazują lepszy stopień rozpoznawania twarzy o jaśniejszej karnacji niż ciemnej oraz męskiej w porównaniu do kobiecej. Częstym przykładem przestrzeni narażonej na algorytmiczne wykluczenie jest ocena ryzyka popełnienia przestępstwa lub recydywy przez określone grupy osób wykorzystywane przez systemy sądownictwa lub bezpieczeństwa np. w celu ustalenia wysokości kary, kaucji czy skierowania patrolu policji do określonej lokalizacji.<sup>11</sup>

O stronniczości w systemach komputerowych mówiły już w latach 90. XX wieku Batya Freidman i Helen Nissenbaum.<sup>12</sup> Wskazały one trzy przyczyny tego zjawiska, czyli stronniczość wynikającą z wcześniejszych uprzedzeń, zakorzenionych w instytucjach społecznych, postawach, praktykach, ograniczeń technicznych oraz stronniczość pojawiającą się już w trakcie użytkowania algorytmu. Co ważne, termin „stronniczość” posiada wydźwięk moralny. Odnoszony jest przez Friedman i Nissenbaum do systemów komputerowych, które systematycznie i niesprawiedliwie dyskryminują określone osoby lub grupy osób, stawiając je w gorszej pozycji względem innych osób lub grup. Stronnicze nie będzie więc przykładowo odmówienie udzielenia kredytu osobie, której dokumentacja wskazuje na niewypłacalność. Stronnicze natomiast będzie nieprzyznanie pożyczki na podstawie klasyfikacji przypisującej słabe ratingi kredytowe osobom z nazwiskami etnicznymi.<sup>13</sup>

<sup>11</sup> P.N.K. Schuetz, *Fly in the Face of Bias: Algorithmic Bias in Law Enforcement's Facial Recognition Technology and the Need for an Adaptive Legal Framework*, „Minnesota Journal of Law & Inequality”, 1(39), 2001, s. 223.

<sup>12</sup> B. Friedman, H. Nissenbaum, *Bias in Computer Systems*, „ACM Transactions on Information Systems”, Vol. 14, No. 3, 1996.

<sup>13</sup> *Ibidem*, s. 332.

Wykluczenie algorytmiczne to bardzo ważny aspekt badań nad sztuczną inteligencją, ponieważ dotyczy już niemal każdego aspektu ludzkiego życia i ma swoje bezpośrednie przełożenie na życie codzienne. Wiele różnorodnych organizacji opiera swoje systemy decyzyjne na predykcjach algorytmicznych, np. podczas zatrudniania pracowników, wypłacania premii, przyznania ubezpieczenia, ustalania pierwszeństwa do usług medycznych, przewidywania zakresu przestępczości w danym regionie czy w interesującym nas kontekście prezentowania wyników wyszukiwarek internetowych, ekspozycję treści w mediach społecznościowych. Algorytmy stosowane w różnego typu mediach społecznościowych, podlegające uczeniu maszynowemu, opierają się na wprowadzonych do algorytmu danych. Jeśli są one wadliwe lub niepełne, to mogą one utrwalać ewentualne błędy, np. stereotypy, uprzedzenia, rozpowszechniać tzw. fake newsy czy pomijać określone grupy ludzi, ich opinie oraz potrzeby.

## Racjonalność w mediach społecznościowych

Nowoczesność to zgodnie ze słynnym przemówieniem Habermasa z 1980 roku z okazji otrzymania nagrody Adorno, czyli „Modernizm — niedokończony projekt”, swoista obietnica racjonalności, której źródła możemy doszukiwać się w oświeceniu. Nowoczesność została przedstawiona jako niedokończony projekt racjonalizacji, projekt oświeceniowej emancypacji oparty na rozumie, który realizuje na polu nauki, moralności i sztuki. Potencjał racjonalizacji nowoczesności znajduje się w tym, co najbliższe człowiekowi, czyli w samym języku<sup>14</sup>. To właśnie racjonalność zawarta w języku może przeciwdziałać autonomizacji i profesjonalizacji wspomnianych dziedzin, czyli nauki, moralności i sztuki w kierunku racjonalności kognitywno-instrumentalnej, moralno-praktycznej i estetyczno-ekspresyjnej oraz umożliwić wykorzystanie ich potencjału w praktyce. Daje to sposobność do „rozumnego kształtowania stosunków życiowych”<sup>15</sup>. Język jednoczy zatem wskazane typy racjonalności oraz umożliwia realizację racjonalności w praktyce. Następnie doprecyzowanie tego, czym jest nowoczesność znajdujemy w „Filozoficznym dyskursie nowoczesności”, czyli zbiorze wykładów pomiedzy rokiem 1983 i 1984, w którym nowoczesność jawi się jako konsekwencja procesu modernizacji i racjonalizacji zapoczątkowanej przez rozum zarówno w wymiarze praktycznym, jak i technicznym. W tej perspektywie Habermas nawiązuje do koncepcji racjonalizacji Maxa Webera, który analizuje racjoni-

<sup>14</sup> J. Habermas, *Modernizm — niedokończony projekt*, tłum. M. Łukasiewicz, [w:] *Postmodernizm. Antologia przekładów*, red. R. Nycz, Wydaw. Baran i Suszczyński, Kraków 1997, s. 34-36.

<sup>15</sup> *Ibidem*, s. 35.

zają społeczeństwa w kontekście gospodarki kapitalistycznej i biurokratycznej administracji oraz racjonalizacji „form życia w szczególności w ramach sekularyzacji i desakralizacji wizji świata.<sup>16</sup>

Habermas pragnie zmienić zakres racjonalizacji i przekonuje, że racjonalność nie jest tożsama z odczarowaniem i instrumentalizacją, ale wyraża się w potencjale komunikacyjnym, czyli możliwości dyskursywnego osądu wszelkich problemów czy to społecznych, politycznych, czy ogólnie wszelkich problemów zaistniałych w świecie życia. Oczywiście jest to również główny problem *Teorii działania komunikacyjnego*, która wskazuje nowe możliwości społecznej teorii krytycznej zawarte właśnie w racjonalności komunikacyjnej.<sup>17</sup> Habermas już w przedmowie do *Teorii działania komunikacyjnego* nakreśla zasadniczy cel swojego dzieła właśnie jako sprzeciw względem zawężenia rozumu w kierunku kognitywno-instrumentalnym.<sup>18</sup> W ten sposób racjonalność odnoszona jest nie do kryteriów prawdziwości czy skuteczności wiedzy o świecie, ale dyspozycji podmiotów do intersubiektywnych interakcji zorientowanych na uznanie wzajemnie wysuwanych roszczeń do prawdziwości, słuszności, szczerości. Habermas proponuje tu, aby na racjonalność spojrzeć przez pryzmat rozumu komunikacyjnego w przeciwieństwie do rozumu ześrodkowanego w podmiocie.<sup>19</sup>

Komunikacja to codziennie tocząca się wymiana językowa zorientowana na wzajemne zrozumienie oraz porozumienie. Komunikacja zakłada, że wszyscy uczestnicy wyrażają swoje racje, które mogą zostać w każdej chwili podane w wątpliwość, a w razie konieczności uzasadnione oraz skorygowane. Racje te opierają się na roszczeniach, czyli gwarancjach prawomocności wypowiedzi do zrozumienia, prawdy, prawdomówności i moralnej słuszności. Racje i stojące za nimi roszczenia mogą w każdej chwili zostać podane w wątpliwość. Jeśli wysuwane racje zostaną zakwestionowane, to komunikacja przechodzi na poziom dyskursu. Uruchamia to proces argumentowania, czyli eksplikowania problematycznych roszczeń ważnościowych. W tej procedurze ujawniają się racjonalne mechanizmy komunikacji w ramach uzasadniania wysuwanych argumentów, ich uzasadniania oraz korygowania.<sup>20</sup> Normatywna instancja określająca ważność postulowanej procedury argumentacyjnej odnosi się do kilku przesłanek idealnej sytuacji

<sup>16</sup> A. Dupeyrix, *op. cit.*, s. 94, 95.

<sup>17</sup> J. Habermas, *Teoria działania komunikacyjnego. Racjonalność działania a racjonalność społeczna*, t. 1, przeł. A.M. Kaniowski, Wydawnictwo Naukowe PWN, Warszawa 1999, s. 631.

<sup>18</sup> *Ibidem*, s. 9.

<sup>19</sup> J. Habermas, *Filozoficzny dyskurs nowoczesności*, przeł. M. Łukasiewicz, Universitas, Kraków 2000, s. 334-362.

<sup>20</sup> J. Habermas, *Teoria działania...*, s. 45, 46.

mowy. Wśród nich Habermas wyróżnia cztery najważniejsze, czyli inkluzję, równouprawnienie, wykluczenie złudzeń i iluzji oraz swobodę. Oznacza to, że nikt nie może zostać wyłączony z dyskursu, wszyscy mają równe szanse wypowiedzenia się, inicjowania dyskursu oraz argumentowania, wyeliminowane są wszelkie ograniczenia oraz zapewnione zostaje zaspokojenie interesów każdego zainteresowanego.<sup>21</sup>

Wydaje się, że przynajmniej część z przesłanek argumentowania mogłaby być realizowana w przestrzeni Internetu, w której potencjalnie każdy zainteresowany użytkownik ma szansę wyrazić swoje zdanie. Media społecznościowe mogłyby stanowić przestrzeń dla spontanicznej sfery publicznej, będącej inkluzyjną siecią różnorodnych, wzajemnie ząbających się subkulturowych sfer publicznych, które nie mają ograniczonej granicy czasoprzestrzennej. Brak proceduralnego uregulowania tej sfery prowadzi do większej swobody, samorozumienia i wrażliwości na różnorodne problemy i ich ekspresje<sup>22</sup>.

Przebieg cyfrowej komunikacji regulowana jest jednak mechanizmami kapitalizmu inwigilacji<sup>23</sup> czy komunikacji<sup>24</sup>. Według Jodi Dean trzydzieści lat temu Internet ucieleśniał fantazję i obietnicę demokratyzacji przestrzeni społecznej, dostępnej i przejrzystej dla wszystkich, którzy pragną zabrać głos w debacie politycznej. Jednak wizja epoki kierowanej przez działania komunikacyjne w myśl intencji Habermasa w demokratycznej sferze publicznej okazała się fikcją zastąpioną przez nierówności i niesprawiedliwości funkcjonujące w ramach kapitalizmu cyfrowej komunikacji<sup>25</sup>. Określenie „kapitalizm inwigilacji” odnosi się do logiki, która stoi za technologią i nią zarządza. To eksploatacja jako cennego surowca danych pozyskiwanych z inwigilacji ludzkiego doświadczenia w szeroko rozumianym świecie cyfrowym a następnie wykorzystywania ich w celu predykcji zachowań konsumenckich.<sup>26</sup>

Rozwój Internetu w kierunku sieci Web 2.0 zapoczątkował rewolucję w sposobie funkcjonowania ludzi w przestrzeni wirtualnej i stanowił linię demarkacyjną pomiędzy Internetem treści a Internetem wymiany i interak-

<sup>21</sup> J. Habermas, *Działanie komunikacyjne i detranscendentalizacja rozumu*, przeł. W. Lipnik, Oficyna Wydawnicza, Warszawa 2004, s. 38.

<sup>22</sup> J. Habermas, *Faktyczność i obowiązki. Teoria dyskursu wobec zagadnień prawa i demokratycznego państwa prawnego*, przeł. A. Romaniuk, R. Marszałek, Wydaw. Naukowe Scholar, Warszawa 2005, s. 327.

<sup>23</sup> S. Zuboff, *Wiek kapitalizmu inwigilacji. Walka o przyszłość ludzkości na nowej granicy władzy*, przeł. A. Unterschuetz, Wydawnictwo Zysk i S-ka, Poznań 2019, *passim*.

<sup>24</sup> J. Dean, *op. cit.*, *passim*.

<sup>25</sup> J. Dean, *Preface*, [w:] *Affective Politics of Digital Media. Propaganda by Other Means*, ed. M. Boler, E. Davis, Routledge, New York, London 2021, s. 8, 9.

<sup>26</sup> S. Zuboff, *op. cit.*, s. 29-32.

cji. W myśl idei Web 2.0 użytkownicy Internetu nie są już tylko odbiorcami, ale użytkownikami, a Internet to platforma wspólnego działania i interakcji. Powstanie i ogromna popularność mediów społecznościowych jest naturalną konsekwencją rozwoju Internetu, który nie służy odąd tylko do konsumowania treści, jak w tradycyjnych mediach, ale ich bezpośredniego kreowania. Definicja odpowiada potocznemu rozumieniu mediów społecznościowych oraz lansowanej przez przedsiębiorstwa informatyczne wizji Internetu jako przestrzeni wspólnych działań, nieograniczonych możliwości komunikacji, forum nieskrępowanej wymiany poglądów i idei, przestrzeń równości i demokracji, płynącej mlekiem i winem krainę obfitości i możliwości uczestnictwa w życiu społecznym, z pominięciem podstawowej cechy funkcjonowania komercyjnego. Do podobnych wniosków doszedł Michael Fleischer, badając skojarzenia na temat Facebooka wśród grupy polskich studentów.<sup>27</sup> Najczęstsze odpowiedzi na pytanie o to, czym jest, z czym się kojarzy oraz jaki jest Facebook, odwoływały się do takich kategorii, jak portal społecznościowy, komunikator, zdjęcia, znajomi, informacje. Na pytanie o użyteczność Facebooka studenci najczęściej charakteryzowali go przez pryzmat kategorii związanych z informacjami i wiadomościami, np. czerpaniem informacji, dzieleniem się informacjami czy ich publikowaniem oraz komunikowaniem się.<sup>28</sup> Aspektem, którego się nie dostrzega, transparentną stroną Facebooka jest jednak jego model biznesowy i działalność ekonomiczna. Fleischer podkreśla, że celem Facebooka nie jest komunikacja, ale gromadzenie danych i zasobów zgodnie z jego przeznaczeniem komercyjnym<sup>29</sup>, co wpisuje się w wizję rzeczywistości jako kapitalizmu nadzoru czy inwigilacji.

O ile takie normatywne założenia idealnego dyskursu, jak swoboda, równouprawienie czy wykluczenie iluzji, mimo wielu problemów i kontrowersji, wydają się spełnialne w kontekście funkcjonowania tzw. mediów społecznościowych, to wysoce problematyczną cechą racjonalności komunikacyjnej jest inkluzja, czyli otwartość na wszelkie poglądy. Zgodnie z definicją inkluzji Habermasa „nikt kto mógłby wnieść jakiś istotny wkład ze względu na kontrowersyjne ważnościowe roszczenia, nie może zostać wykluczony”<sup>30</sup>. Inkluzja jest przeciwieństwem ekskluzji, czyli redukcji, wyłączenia, wykluczenia określonych stanowisk. Inkluzja to skłonność do przyjęcia innych punktów widzenia, które są potencjalnie istotne w dyskursie. Ważność inkluzji tkwi tym samym w możliwości integrowania różnych, równouprawionych w dys-

<sup>27</sup> M. Fleischer, *Facebook i okolice*, [w:] *idem*, *Konstrukcja rzeczywistości 3*, Wydaw. Libron — Filip Lohner, Kraków 2017, s. 63-83.

<sup>28</sup> *Ibidem*, s. 63-75.

<sup>29</sup> *Ibidem*, s. 17.

<sup>30</sup> J. Habermas, *Działanie komunikacyjne...*, s. 38.

kursie punktów widzenia. Jest to zdolność do dyskursywnego rozwiązywania konfliktów wyrażająca się w założeniu, że porozumienie budowane jest na podstawie racjonalnej procedury argumentacji. Porządek społeczny nie jest ustalony na podstawie tradycji czy wierzeń, ale jest oparty na deliberacji, debatach, dyskusjach i dialogach wykorzystujących argumenty i racje, które stanowią ich podłoże. Nadzieje emancypacji oddane są do rąk ludzi działających w ramach komunikacyjnej praktyki społecznej i politycznej. Nie ma „czystego rozumu”, który wyraża się w języku, ponieważ rozum od początku jest ucieleśniony w relacjach komunikacyjnych oraz strukturach świata przeżywanego.<sup>31</sup>

W ujęciu Habermasa różnorodne siatki sfer publicznych mogą wzajemnie się przenikać, uzupełniać i upowszechniać. Na powszechne upublicznienie otwierają się wszelkiego typu dyskursy, nawet te lokalnie ograniczone. Rozwój mediów w kierunku tzw. mediów społecznościowych powinien zgodnie z tym założeniem pogłębić ten proces. Zasadność tej perspektywy podważał już ponad dekadę temu aktywista Eli Parieser, główny popularyzator określenia „bańka filtrująca” w odniesieniu do personalizacji materiałów internetowych. Dostosowanie zawartości do zainteresowań użytkowników prowadzi do zjawiska całkowicie przeciwnego demokracji i racjonalizacji dyskursu. Określenie „bańka filtrująca” odwołuje się do mechanizmów sortowania i filtrowania treści udostępnianej użytkownikom sieci zgodnie z ich spersonalizowanymi preferencjami, w wyniku czego mogą oni otrzymywać materiały określonego typu. W ten sposób personalizacja wyników wyszukiwań internetowych czy zawartości mediów społecznościowych fragmentaryzuje i sortuje sferę publiczną zamiast ją wspierać.<sup>32</sup>

Algorytmy doboru treści na naszej facebookowej tablicy działają upraszczająco, przez co nie uwzględniają niezwykle istotnego aspektu racjonalności komunikacyjnej Habermasa, a mianowicie inkluzyjności. Upraszczający charakter algorytmów wyraża się w pytaniu o naturę danych — w jakim stopniu odzwierciedlają one populację, jakiego typu błędami są obarczone, co i kogo pomijają, jakiego typu dane biorą pod uwagę itp.?<sup>33</sup> Optymalizacja, schematyzacja algorytmiczna prowadzi w konsekwencji do wykluczenia algorytmicznego oraz społecznej polaryzacji. Obecna rzeczywistość charakteryzuje się niespotykanym dotąd rozwojem technologii cyfrowych i wzrostem ilości danych. Safyia Umoja Noble podkreśla, że należy prześledzić technologie cyfrowe, ponieważ są one uwikłane w globalne stosunki opre-

<sup>31</sup> J. Habermas, *Filozoficzny dyskurs nowoczesności...*, s. 338.

<sup>32</sup> E. Pariser, *The Filter Bubble. What the Internet is Hiding from You*, The Penguin Press, New York 2011, s. 89-91.

<sup>33</sup> A. Badia, *The Information Manifold. Why Computers Can't Solve Algorithmic Bias and Fake News*, The MIT Press, Cambridge, London 2019, s. 231-233.

syjnej władzy rasowej. Represyjny charakter algorytmów prowadzi do wykluczenia grup szczególnie narażonych na marginalizację ze względu na płeć czy kolor skóry. Noble unaocznia, że dyskryminacja jest częścią architektury technologii cyfrowej i przejawia się ona już na poziomie wyszukiwarek internetowych. Świat algorytmów nie jest neutralny. Jest stworzony przez ludzi, którzy świadomie lub nieświadomie tworzą opresyjne, polaryzujące społecznie algorytmy.<sup>34</sup> Rzeczywistość algorytmów w pewnym sensie odzwierciedla nierówności, uprzedzenia i dyskryminację w świecie rzeczywistym, ponieważ wynika z ludzkiej natury lub odzwierciedla błędy poznawcze ludzi. Algorytmy są równaniami matematycznymi, formułami logicznymi i statystycznymi opartymi na danych w celu rozwiązania konkretnego problemu. Błędy w algorytmach mogą wynikać z wielu czynników, od niepełnych lub wadliwych danych po błędy ludzi, którzy je tworzą.<sup>35</sup>

## Zakończenie

Myśliciele XX wieku uczynili kwestię języka a później komunikacji problemem *par excellence* filozoficznym. Jürgen Habermas w komunikacji widział warunki ludzkiej racjonalności, możliwość emancypacji, demokratyzacji przestrzeni publicznej. Dwudziesty pierwszy wiek za sprawą Internetu stawia przed filozofią komunikacji nowe wyzwania i problemy. Nowe, elektroniczne przestrzenie realizacji ludzkiej aktywności komunikacyjnej rodzą nowe dylematy. Internet, który w początkach swojego istnienia ucieleśniał nadzieję na obiektywizację informacji, rozwój wiedzy, popularyzację nauki, przestrzeń debat i społecznych deliberacji, stał się obecnie areną konfliktów, nierówności, wykluczenia i polaryzacji społecznej. W tej perspektywie idea racjonalności komunikacyjnej realizującej się poprzez sferę publiczną okazuje się problematyczna, co dostrzega również Habermas<sup>36</sup>. Niewralgiczną kwestią w kontekście sfery publicznej funkcjonującej w mediach cyfrowych jest przy tym inkluzyjność oraz ekskluzywność. W mediach cyfrowych dostęp do argumentacji regulowany jest nie przez racjonalność podmiotów biorących udział w dyskursie, ale system algorytmów decydujących o rodzaju treści, do których mają dostęp użytkownicy mediów cyfrowych.

<sup>34</sup> S.U. Noble, *Algorithms of Oppression. How Search Engines Reinforce Racism*, New York University Press, New York 2018, s. 171, 172.

<sup>35</sup> T. Baer, *Understand, Manage and Prevent Algorithmic Bias. A Guide form Business Users and Data Scientists*, Apress, Kaufbeuren 2019. s. 4-7.

<sup>36</sup> J. Habermas, *Uniwersalizm moralny w czasach politycznego regresu. Jürgen Habermas o współczesności i dziele życia, rozmowa z Jürgenem Habermasem, Claudia Czingon, Aletta Diefenbach, Victor Kempf*, przeł. P. Pluciński, J. Balisz-Schmelz, „Władza sądzona” 19(2020), dostęp on-line [www.wladzasadzenia.pl/nr/19-2020/](http://www.wladzasadzenia.pl/nr/19-2020/), 15.08.2022, s. 16.

Czy zatem tzw. media społecznościowe stanowią arenę dla sfery publicznej w rozumieniu Habermasa? Byłoby to możliwe, gdyby nie fakt, że media społecznościowe działają komercyjnie. Model biznesowej inwigilacji, co podkreślała Zuboff, wskazując na przekształcenia współczesnego kapitalizmu w kierunku kapitalizmu cyfrowej inwigilacji, nastawiony jest na spieniężenie ludzkiego doświadczenia i codziennej rutyny.<sup>37</sup> Formuła algorytmów wynika z komercyjnego charakteru „mediów społecznościowych”. Algorytmiczne systemy rekomendujące i filtrujące określają rodzaj treści prezentowanych poszczególnym użytkownikom. Mechanizm ten daje niepełny zakres informacji, stanowisk, poglądów na daną kwestię lub preferuje określony światopogląd, podsyłając użytkownikom treści określonego typu.

Algorytmy poprzez dobierające treści w mediach społecznościowych poprzez systemy filtrujące ograniczają użytkownikom prezentowane treści, zamykając ich w określonym światopoglądzie i wykluczając inne punkty widzenia. Tym samym uniemożliwiają swobodną komunikację, podsycają konflikt, polaryzują społeczeństwo, czy wprowadzają chaos informacyjny. W ten sposób media społecznościowe, zamiast inkluzyjne, stają się ekskluzyjne. Zasadniczy problem nie leży jednak po stronie użytkowników i ich zdolności czy kompetencji komunikacyjnych, ale samym medium, czyli w przypadku mediów społecznościowych — w algorytmach, które decydują o tym, co, kiedy i w jakiej kolejności wyświetla się na naszej tablicy czy inaczej news feedzie. Stosunek ten wyraża znana formuła determinizmu technologicznego Marshalla McLuhana, według której przekąźnik jest przekazem i to właśnie on (przekąźnik, medium) określa charakter przekazu, a nie treść.

Racjonalna komunikacja to zgodnie z intencją Jürgena Habermasa językowe mechanizmy koordynacji dyskursu zmierzające do intersubiektywnego uznania wnoszonych przez partnerów roszczeń do ważności oraz możliwości osiągnięcia przez nich porozumienia. Potoczne wyobrażenia o mediach społecznościowych jako arenie racjonalnego dyskursu nie odpowiadają jednak ich rzeczywistemu działaniu. Media społecznościowe nie sprzyjają łagodzeniu nieporozumień, ponieważ ich funkcjonowanie nie odwołuje się do zasad rozumu, ale emocji, co wykorzystują algorytmiczne mechanizmy filtrowania treści. Algorytmy te stanowią narzędzie prognozowania zorientowanego komercyjnie zgodnie z prawami rynku, co wpisuje się we współczesne reguły kapitalizmu inwigilacji. Cyfrowy świat mediów elektronicznych jest fragmentaryczny. Informacje kierowane do publiczności internetowej są spersonalizowane i zindywidualizowane analogicznie jak w modelu targetowanych reklam. Celem jest utrzymanie uwagi odbiorców, ponieważ czas spędzany na przeglądaniu określonej platformy internetowej ma realną wartość. Intensy-

---

<sup>37</sup> S. Zuboff, *op. cit.*, s. 337.

fikacja komunikacji w kapitalizmie inwigilacji czy komunikacji nie niesie ze sobą potencjału racjonalizacji i emancypacji, czego chciałby Habermas.

*Magdalena Filipiak*

**The Public Sphere, Inclusion and Algorithms, or the Communicative Rationality of Jürgen Habermas in the Context of Algorithmic Exclusion in Social Media**

*Abstract*

The article aims to analyze the public sphere and the rationality of communication according to Jürgen Habermas in the context of the functioning of social media. The content filtering mechanisms available to social media users raise doubts as regards the fulfilment of Habermas's inclusiveness. Communicative rationality should counteract social polarization and discursively resolve emerging conflicts. From the popular point of view, social media is a place of free and unrestricted communication, inclusion, access to diverse opinions and pluralism. These features of social media express hope for the realization of the public sphere and emancipation, which is enabled by the rationality inherent in the very language and principles of communication. However, the main barrier to implementing the public sphere is the algorithms and recommendation systems that filter the content presented to users. A hazardous phenomenon is an algorithmic bias, which has a real impact on people's lives.

*Keywords:* communicative rationality, public sphere, algorithmic exclusion, social media.