

MARCIN SZULC

Uniwersytet Gdański, Instytut Psychologii

MANIPULOWANIE INFORMACJĄ W SIECI ZA POMOCĄ FAKE NEWSÓW JAKO ZAGROŻENIE DLA MŁODZIEŻY

Streszczenie: Oprócz niewątpliwych korzyści płynących z użytkowania mediów cyfrowych nie można przejść obojętnie wobec zagrożeń, jakie ze sobą niosą. Dotychczas uwaga badaczy najczęściej skupiała się na cyberprzemocy i uzależnieniu. Obecnie coraz częściej dostrzega się nowe niebezpieczeństwo, jakim stała się łatwość multiplikowania nieprawdziwych informacji w sieci. Młodsze pokolenia, dla których Internet jest naturalnym źródłem pozyskiwania wiedzy o świecie, mogą przejawiać problemy z krytyczną oceną informacji, tym bardziej, że metody oceny wiarygodności treści wymagają wiedzy, zaangażowania, czasu, a przede wszystkim

chęci odbiorcy, by podjął się ich weryfikacji. Problem wydaje się szerszy i nie dotyczy tylko nastolatków, coraz częściej bowiem osoby cechujące się dużą wiedzą i doświadczeniem życiowym wpadają w pułapkę fake newsów, zwłaszcza gdy informacje dotyczą sfer, w których nie są ekspertami. Celem artykułu jest scharakteryzowanie zjawiska fake news, psychologicznych побudek działania użytkowników świata cyfrowego nazywanych trollami oraz metod rozpoznawania nieprawdziwych informacji rozpowszechnianych w sieci.

Słowa kluczowe: manipulacja, fake news, trolling, młodzież.

WPROWADZENIE

Czy papież Franciszek poparł kandydaturę Donalda Trumpa na Prezydenta USA? Czy Hilary Clinton sprzedawała broń ISIS, wybory zaś przegrała dlatego, że zdobyła mniej głosów? Czy w Goeteborgu istnieją strefy no-go i nie można już swobodnie robić zakupów z powodu zagrożenia uzbrojonymi uchodźcami? Czy obecnie widmo zamachu terrorystycznego jest większe niż dawniej i czy głównymi ich sprawcami w Europie są jihadyści? Odpowiedzi na te nurtujące pytania można znaleźć w Internecie. Problemem jest to, czy są one prawdziwe. Manipulowanie informacją w mediach nie jest zjawiskiem nowym. W 1782 r. podczas amerykańskiej wojny o niepodległość, gdy Benjamin Franklin prowadził w Paryżu negocjacje pokojowe, opublikowano fałszywą wersję gazety *The Boston Independent*. Zamieszczono w niej artykuł na temat oskalpowania przez rdzenną ludność ponad 700 osób, w tym chłopów, kobiet, dzieci, a nawet niemowląt. Do tej zbrodni mieli rzekomo podżegać miejscowych Brytyjczycy. Ta fałszywa informacja była rozpowszechniona z inspiracji samego Franklina, aby podzielić opinię publiczną w Anglii i Europie oraz wzbudzić sprzeciw wobec wojny, a Bry-

tyczków podburzyć przeciwko Królowi Jerzemu III Hanowerskiemu (Berry, 2017; Hendricks, Vestergaard, 2019). Wynika stąd, że przedstawianie fałszywych informacji w mediach jako prawdziwych, aby osiągnąć określone, np. polityczne, cele ma długą historię, jednak digitalizacja życia społecznego znacznie ten proceder ułatwiła. Człowiek staje dziś nie tyle w obliczu braku dostępu do informacji, którego przyczyną jest wykluczenie cyfrowe, ile problemu jej jakości i rzetelności. Prawdziwość informacji zamieszczonych w sieci bardzo trudno bowiem poddać weryfikacji m.in. ze względu na jej nadmiar (Szpunar, 2007).

Liczba użytkowników mediów społecznościowych przekroczyła już 3 miliardy (obecnie to 3,48 mld) (Digital Trends by Hootsuite, 2019)¹⁾, czyli około połowę populacji ludności na świecie. Oznacza to globalną zmianę paradygmatu już nie tylko z „człowiek–świat” na „człowiek – media – świat”, ale na człowieka aktywnie tworzącego świat w mediach. Media odgrywają kluczową rolę w percepcji rzeczywistości (Maciołek, 2014). Internet stał się dla dzieci i młodzieży podstawowym źródłem wiedzy o świecie, naturalnym środowiskiem, w którym zdobywa doświadczenia i zaspokaja potrzeby społeczne, a także narzędziem efektywnej komunikacji i dostępu do informacji na niespotykaną dotychczas skalę (Laberschek, 2010). Przestrzeń cyfrowa to także nieodzowny element życia społecznego młodzieży, który wpływa na ich kształt dorastania i funkcjonowania. Można nawet twierdzić, że współczesna tożsamość dzieci i młodzieży jest tożsamością technologiczną (Pyżalski, 2012). Zmienił się kierunek przepływu wartości, świat jest bardziej zrozumiały dla dzieci i młodzieży, a młodsze pokolenia przekazują wiedzę techniczną pokoleniom starszym. Stosując terminologię Margaret Mead (1978), kultura prefiguratywna zastąpiła kulturę postfiguratywną.

NOWE NOWE MEDIA A TEORIA KORZYSTANIA I GRATYFIKACJI²⁾

Andrew Keen (2007) przekonuje, że rewolucja cyfrowa wpływa destrukcyjnie na kulturę, gospodarkę i wartości. Łatwość dostępu do narzędzi powszechnej komunikacji sprawia, że profesjonalistów zastępują amatorzy. Możliwość rozpowszechniania każdego rodzaju treści zagraża standardom kulturowym oraz wartościom (Keen, 2007). Używając języka Thomasa Henry’ego Huxleya technologia pozwala połączyć „wszystkie mały z wszystkimi maszynami do pisania”, co jednak nie daje gwarancji powstania arcydzieła, zamiast tego ukazuje „nieskończony cyfrowy zbiór mieronot” (Keen, 2007, s.26). Paul Levinson (2010), twórca pojęcia „nowe nowe media” wymienia ich najważniejsze cechy: możliwość jednoczesnego bycia zarówno konsumentem, jak i producentem, autentyczność wynikająca z braku profesjonalizmu, możliwość samodzielnego wyboru mediów, „nie płacisz, a dostajesz”, relacja katalityczna mimo rywalizacji, większe znaczenie mediów niż zwykła wyszukiwarka czy poczta elektroniczna.

Teoria korzystania i gratyfikacji wyrosła z funkcjonalistycznego ujmowania mediów i stawia w centrum zainteresowania nie medium a ich użytkownika. Zakłada, że ludzie, będąc świadomi swoich potrzeb, są ukierunkowani na realizację celów. Tym, co motywuje lub zniechęca odbiorcę do korzystania z mediów, jest jego doświadcze-

¹⁾ <https://wearesocial.com/global-digital-report-2019>.

²⁾ Media nowsze niż nowe media, powstałe w latach 60. XX w. Paulowi Levinsonowi chodzi o media społecznościowe, w których odbiorca stał się również nadawcą (Kuczer, 2010).

nie i oczekiwania (Polański, 2017). Człowiek jako aktywny twórca używa mediów, aby osiągnąć konkretne korzyści. Media, konkurując z innymi źródłami zaspokajania potrzeb człowieka, realizują je tylko w określonym zakresie. Odbiorcy zazwyczaj są świadomi zainteresowań i motywów, dla których korzystają z mediów, przekazom zaś przypisują własne znaczenie i emocje.

Denis McQuail wymienia cztery potrzeby, które zaspokajają media: informacyjną (zaspokajania ciekawości, poszukiwania rady), kształtowania tożsamości (potwierdzania własnego systemu wartości, empatii wobec innych, podwyższania samooceny), integracji i interakcji społecznej (przynależności), rozrywki (odwrócenia uwagi od problemów i realizacji potrzeb związanych z kulturą) (Iwanowska, 2016; Polański, 2017). Elihu Katz, Michael Gurevitch i Hadassah Haas zidentyfikowali aż 14 potrzeb, zaspokajanych poprzez media, które pogrupowali w pięć nadrzędnych kategorii: potrzeby poznawcze, afektywne, osobistej spójności, społecznej integracji, potrzeby eskapistycznej (za: Turska-Kawa, 2013).

Paul F. Lazarsfeld i Robert K. Merton wskazują także na negatywną funkcję mediów jako środka narkotyzującego, który pochłania odbiorcom coraz więcej czasu i zastępuje bezpośredni kontakt z rzeczywistością kontaktem sztucznym i fasadowym. Wspomniany wcześniej Elihu Katz zauważył, że nawet najbardziej rzetelne wiadomości nie mają wpływu na jednostkę, gdy ta nie znajduje dla nich zastosowania w społecznym i psychologicznym kontekście własnego życia (Turska-Kawa, 2013). Koncepcja ta wydaje się bardzo aktualna i ma wyraźne zastosowanie także do mediów społecznościowych, które ułatwiają gromadzenie osób o określonych poglądach w zamkniętych przestrzeniach echo chamber³). W tych właśnie przestrzeniach odbiorcy otrzymują gratyfikujące informacje, które potwierdzają ich przekonania.

Koncepcja korzystania i gratyfikacji dobrze opisuje i identyfikuje motywacje użytkowników mediów, którzy aktywnie poszukują treści spełniających ich oczekiwania, co jednocześnie staje się także źródłem poważnego zagrożenia. Zwłaszcza, gdy zniekształcenia (deficyty) poznawcze, polegające na przypisywaniu swoim lub cudzym doświadczeniom znaczenia w sposób niedokładny lub skrzywiony, znajdą treści je potwierdzające.

DEZINFORMACJA ODBIORCÓW Z UŻYCIEM FAKE NEWSÓW

Coraz większym problemem w przestrzeni digitalnej stają się tak zwane *fake news*, czyli wiadomości medialne, które jednocześnie nie są ani prawdą, ani kłamstwem, opierają się na dezinformacji i mogą zawierać fragmenty prawdziwe, lecz przeinaczone lub wyrwane z kontekstu (Bąkowicz, 2019; Piotrowski, 2010). Celem ich jest wprowadzenie w błąd odbiorców, wywołanie chaosu, dezorientowanie i sianie strachu oraz podżeganie społeczeństwa przeciwko sobie. Słownik Języka Polskiego PWN definiuje fake newsy jako: „nieprawdziwe, fałszywe wiadomości, najczęściej rozpowszechniane przez tabloidy w celu wywołania sensacji bądź zniesławienia kogoś (najczęściej polityka)”⁴). Pojęcie *fake news* zastępowane jest czasami określeniem „fakty alternatywne”, które jest pozbawione

³) Bańka informacyjna – przestrzeń w mediach społecznościowych gromadząca osoby o podobnych poglądach, w której każda opinia z nimi zgodna potwierdza się i jest przyjmowana z uznaniem.

⁴) <https://sjp.pwn.pl/mlodziejowe-slowo-roku/haslo/fake-news;6368870.html>.

sensu logicznego, gdyż łamie Arystotelesowską zasadę niesprzeczności (Hendricks, Vestergaard, 2019). Jak zauważają Edson C. Tandoc Jr., Zheng Wei Lim i Richard Ling (2017), pojęcie „fałszywe wiadomości” to oksymoron, gdyż zadaniem dziennikarza jest dostarczanie obywatelom informacji, których potrzebują, aby pozostawać wolnymi i niezależnymi. Zatem oczywiste jest, że muszą być one prawdziwe, obiektywne i dokładne. Problemem jest to, że informacje podlegają subiektywnej ocenie samych dziennikarzy, którzy mogą je interpretować zgodnie z własnymi poglądami. Fakty są także „wrażliwe” na siły zewnętrzne, takie jak naciski polityków, preferencje odbiorców i oczekiwania reklamodawców. Można powiedzieć, że wiadomości są towarem wyjątkowym, ponieważ są sprzedawane odbiorcom, odbiorcy zaś sprzedawani są reklamodawcom, co czyni je szczególnie wrażliwymi na siły rynkowe (Tandoc i in., 2017).

Tygodnik *Polityka* podaje za BuzzFeed News, że w amerykańskiej sieci aż dwie trzecie dezinformacji rozpowszechniono w celu osiągnięcia zysku lub politycznej korzyści (Polityka.pl, 2017). Już na początku 2000 r. formułowano obawy, że nadmierna różnorodność poglądów w Internecie doprowadzi do tworzenia przez użytkowników o podobnych poglądach wspomnianych wcześniej „banków informacyjnych”, w których będą pozbawieni możliwości zapoznania się z innymi perspektywami (Allcott, Gentzkow, 2017). Nadreprezentacja określonego typu informacji w mediach może skłonić obserwatora bądź do przeceniania zagrożenia, np. spowodowanego falą uchodźców, niedoceniaenia zagrożenia, jakim jest promowanie dostępu do broni czy minimalizowanie zagrożenia, jakie niesie ze sobą używanie pewnych substancji psychoaktywnych.

Najbardziej spektakularne oddziaływanie *fake news* nastąpiło podczas wyborów prezydenckich w USA pomiędzy Hilary Clinton a Donaldem Trumpem. Fałszywe informacje rozpowszechniane w sieci, np. dotyczące rzekomej sprzedaży broni organizacji terrorystycznej ISIS przez Hilary Clinton, znacząco wpłynęły na decyzje wyborców, a w konsekwencji wygrał Trumpa (Allcott, Gentzkow, 2017). Strona Endingthefed.com, odpowiedzialna za cztery z dziesięciu najpopularniejszych fałszywych wiadomości na Facebooku waloryzujących Trumpa i deprecjonujących Clinton, była prowadzona przez 24-letniego obywatela Rumunii (Townsend, 2016). Fałszywe informacje stają się produktem do sprzedaży ze względu na popyt, jakim się cieszą. Amerykańska firma DisinfoMedia posiada wiele witryn oferujących nieprawdziwe wiadomości, w tym NationalReport.net, USAToday.com.co i WashingtonPost.com.co, a jej właściciel twierdzi, że zatrudnia od 20 do 25 twórców (Allcott, Gentzkow, 2017). Cechą charakterystyczną witryn dostarczających fałszywych wiadomości jest ich tendencja do krótkotrwałości (Allcott, Gentzkow, 2017).

Należy odróżniać celowe fałszowanie informacji od działania przypadkowego. Do zbliżonych fake newsom treści można zaliczyć:

- nieprawdziwe informacje (niepotwierdzone fakty) podane w sposób niezamierzony przez gazety lub portale, a następnie zdementowane;
- plotki;
- satyrę błędnie interpretowaną jako prawdę;
- fałszywe oświadczenia polityków powielane przez media;
- informacje zniekształcone, np. fałszywe wnioski z prawdziwych przesłanek;
- teorie spiskowe (część ludzi w nie wierzy, np. na temat kosmitów, tajnych organizacji, działań służb specjalnych).

To, co jednak budzi szczególnie niepokój, to technologia deepfake wykorzystująca sztuczną inteligencję, działająca na podstawie algorytmów głębokiego uczenia, która

wprawdzie dziś jeszcze wymaga dużej mocy obliczeniowej, lecz przy błyskawicznym postępie w tym obszarze będzie w niedługim czasie powszechnie dostępna. Deepfake pozwoli każdemu stworzyć materiał wideo, w którym znane osoby będą przedstawiać treści podburzające opinię publiczną. Krótko mówiąc, w usta celebrytów lub polityków zostaną włożone treści, których nigdy nie wypowiedzieli lub nie wypowiedzieliby (BBC News, 2017; CyberDefence24, 2018). Materiałem wyjściowym staną się obrazy filmowe wraz z zapisem głosowym oryginalnej postaci. Algorytm nauczy się, w jaki sposób naśladować i odtwarzać mimikę oraz tembr głosu i słownictwo osoby, co uniemożliwi ocenę rzetelności oglądanego materiału odbiorcom niedysponującym programami wykrywającymi manipulację (Grotto, 2018). Obecne systemy rozpoznawania twarzy oparte na sieciach neuronowych VGG i Facenet nie są w stanie wykryć filmów deepfake (Korshunov, Marcel, 2018).

Z międzynarodowego raportu GLOBSEC Policy Institute wynika, że w Polskim Internecie najaktywniej działają zewnętrzne podmioty dezinformujące, których celem jest szerzenie propagandy prorosyjskiej, antyukraińskiej i antyestablishmentowej. Z badań prowadzonych w polskiej przestrzeni cyfrowej wynika, że portale takie jak „Kresy.pl” i „Ukraińiec nie jest moim bratem” są jednymi z najbardziej wpływowych stron szerzących prorosyjską narrację na Facebooku. Pozostałe nierzetelne adresy to: zmianynaziemi.pl, dzienniknarodowy.pl, alexjones.pl, Wolna-polska.pl, Pch24.pl. (Wierzejski, 2017). Interesujące wydaje się to, że młodzi Polacy żywią przekonanie, iż potrafią zidentyfikować nieprawdziwą informację w Internecie znacznie skuteczniej niż ich rodzice, nie potrafią jednak osadzić jej w kontekście, tzn. odpowiedzieć na pytania dotyczące celu jej powstania (Klingová, Hajdu, 2018).

Badacze, próbując wytłumaczyć problem wiary w nieprawdziwe informacje, np. teorie spiskowe, odwołują się do zjawisk i czynników, takich jak: skłonność do radykalizmu politycznego, manifestowanie prawicowego autorytaryzmu, potrzeba unikatowości, brak poczucia kontroli nad sytuacją, potrzeba wyjaśnień, myślenie teleologiczne, skrzywienie epistemiczne, błąd konfirmacji czy osobowość schizotypowa (Damisch, Stoberock, Mussweiler, 2010; Darwin, Neave, Holmes, 2011; Grzesiak-Feldman, 2016; Imhoff, Lamberty, 2017; Kelemen, Rosset, 2009; Marchlewska, Cichocka, Kossowska, 2017; Van Prooijen, Van Dijk, 2014; Van Prooijen, Krouwel, Pollet, 2015).

ROLA TROLLI W ROZPOWSZECHNIANIU NIEPRAWDZIWYCH INFORMACJI

Witryny fałszujące informacje są wspierane przez użytkowników sieci nazywanych trollami. Nazwa „trolling” prawdopodobnie nawiązuje do terminologii stosowanej w wędkarstwie i oznacza technikę połowu ryb na haczyk lub błyskę (ang. *trolling for fish*). Jednym z najważniejszych kroków w rozpowszechnianiu nieprawdziwych treści jest zinfiltrowanie społeczności użytkowników, którzy już są zaangażowani w dyskusje na dany temat, aby użyć ich jako „rozsiewaczy” dezinformacji w ich podsieciach (Mustafarai, Metaxas, 2017). Jak zauważa Marta Juza (2015), jednym z powodów obniżenia standardów komunikacji cyfrowej była jej coraz większa powszechność, a zarazem różnorodność środowiska użytkowników. Coraz częściej stawali się nimi ludzie młodzi, gorzej wykształceni, z mniejszych ośrodków i niższych klas społecznych.

Trolling za pośrednictwem mediów społecznościowych (takich jak portale społecznościowe lub fora dyskusyjne) jest często próbą skłócania, denerwowania, ale

przede wszystkim podżegania przeciwko sobie użytkowników poprzez zamieszczenie złośliwych, prowokacyjnych komentarzy (Buckels, Trapnell, Paulhus, 2014; Hardaker, 2010, 2013). Hardaker (2010) przeprowadził analizę treści postów Usenetu⁵⁾ i zidentyfikował cztery podstawowe cechy trollingu: agresja, oszustwo, zakłócanie i sukces. Portal o nazwie Poltergeist wprost podaje wskazówki, w jaki sposób skutecznie trollować. Jest to swoisty septalog trolla:

1. Bądź bezczelny i niewychowany.
2. Przyczepiaj etykiety.
3. Uogólniaj, czepiaj się słówek.
4. Kieruj się zasadą „im gorzej, tym lepiej”.
5. Zadawaj miliony bezsensownych pytań.
6. Bełkocz i gadaj bez sensu.
7. Stosuj argumenty *ad personam* i *ad baculum* (Poltergeist, 2012).

W początkach istnienia Internetu ich motywacja ograniczała się do wywołania w uczestnikach dyskusji negatywnych reakcji emocjonalnych dla zabawy. Podobna motywacja przyświeca części współczesnych sprawców tej poznawczej przemocy. W lutym 2017 r. mieszkaniec Kostaryki James McDaniel założył stronę o żartobliwym charakterze, której celem było wykazanie naiwności odbiorców. Zamieszczał tam zmyślane informacje opracowane z myślą o wyborcach Donalda Trumpa. W ciągu krótkiego czasu te nieprawdopodobne treści były powielane przez tysiące Internautów, ich zaś sceptycyzm wcale nie wzrósł, gdy McDaniel poinformował, że padli ofiarą żartu z jego strony (Hendricks, Vestergaard, 2019).

Najczęściej jednak trollowanie jest wykorzystywane przez zorganizowane grupy, np. sympatyków określonych poglądów politycznych lub nawet z premedytacją opłacane przez partie w celu szkalowania oponentów politycznych (Bradshaw, Howard, 2017; Gorwa, 2017; Molga, 2017). Ich zadaniem jest udostępnianie na stworzonych przez siebie profilach, przypominających prywatne konta zwykłych użytkowników, nieprawdziwych czy tendencyjnych informacji. Podobnie jak wspomniane już manipulacje podczas wyborów prezydenckich w USA, trolle aktywizują się na forach i biorą czynny udział w dyskusjach na wrażliwe społecznie tematy, np. w USA stali się bardzo aktywni w temacie zmian zasad dostępu do broni po masowym zabójstwie, którego dokonał 14.02.2018 r. Nicolas Cruz.

Niezwykle silnie aktywizują także ruchy antyszczepionkowe. Broniatowski i wsp. (2018) odkryli, że tweety na temat szczepionek pochodzą z tych samych kont, które zostały wcześniej zidentyfikowane i obwiniane przez amerykański Departament Sprawiedliwości o ingerencję w wybory w Stanach Zjednoczonych. Rosyjscy trolle tweetują 22 razy częściej na temat szczepionek (#VaccinateUS) niż przeciętny użytkownik Twittera. Celem tych działań jest właśnie sianie zamętu i polaryzowanie społeczeństwa, wysyłając zarówno pro-, jak i antyszczepionkowe komunikaty, ale o charakterze skrajnym (pełne agresji lub utrzymywane w tonie prowokacyjnym) (Broniatowski i in., 2018). Sposób działania trolli jest zgodny z rosyjską doktryną wojny informacyjnej. Dezinformowanie, sianie chaosu, skłócanie społeczeństw stanowi jedno z kluczowych narzędzi wojny hybrydowej, w której – jak wskazują eksperci – uczestniczymy już od dłuższego czasu (Ćwieliuch, Różański, 2018; Giles, 2015; Hunter, Pernik, 2015; Łuczuk, 2017).

⁵⁾ Ogólnosiwiatowy system grup dyskusyjnych w Internecie, składający się z tysięcy grup tematycznych, ułożonych w strukturę hierarchiczną.

W dniu 17.01.2019 r. Facebook poinformował o usunięciu 364 stron (289 fan page'y i 75 profili) zaangażowanych w nieautentyczne skoordynowane zachowania, powiązane z częścią sieci, mającą swoje źródło w Rosji (Facebook Newsroom, 2019). Strony nieautentyczne cechuje brak informacji na temat prawdziwych celów i źródeł. Funkcjonowały w Kazachstanie, Kirgistanie, Tadżykistanie, Uzbekistanie, Armenii, Azerbejdżanie, Gruzji, Mołdawii, Rumunii, Białorusi, Estonii, Łotwie i Litwie. Poruszały różne tematy, część z nich cechował ostry ton antyzachodni i antynatowski (Biel, 2019). Facebook uznał, że są one powiązane z koncernem medialnym Rossija Siegodaya (Aleksejeva i in., 2017). Oprócz osób zawodowo zajmujących się trollingiem istnieje także ogromna liczba tych, którzy trollują z przyczyn ideologicznych (*evangelist marketing*), przyjmując strategię „bo jak nie my, to oni”, a także tych, których cechuje niski krytycyzm w ocenie rzetelności udostępnianych informacji. Gwoli ścisłości należy podkreślić, że pracę trolli wspierają boty, czyli zaprogramowane narzędzia służące rozpowszechnianiu treści w Internecie i tworzeniu sztucznego ruchu wokół jakiegoś tematu. Istnieją dowody na wzmożoną aktywność trolli i botów związanych z rosyjską firmą IRA (Internet Research Agency) podczas wyborów prezydenckich w USA w 2016 r. oraz separatystycznego aktu Katalończyków próbujących oderwać się od Hiszpanii w 2017 r. (*Mały leksykon postprawdy*, 2018).

Z trollingiem trudno jest walczyć, nawet prezes Twittera przyznał się do swojej nieskuteczności w radzeniu sobie z incydentami trollingowymi (Hern, 2015). W niektórych krajach takie antyspołeczne zachowanie w sieci jest surowo karane. Na przykład w Wielkiej Brytanii trwa debata nad przedłużeniem do dwóch lat dotychczasowego sześciomiesięcznego okresu pozbawienia wolności za trolling online (Maltby i in., 2016). Ze względu na poprawę skuteczności wykrywania fake newsów przez system, trolle cały czas zmieniają strategię działania i obecnie aktywnie angażują się w dyskusje pod postami.

Portret psychologiczny trolla

Warto zastanowić się, z jakich pobudek uruchamia się proces trollowania. Wielu uczestników kultury mediów społecznościowych bywa trochę trollami niezależnie od własnej woli. Zjadliwe, złośliwe komentowanie postu albo wypowiedzi innych w digitalnej rzeczywistości pozbawionej kontekstu, uszczuplonej emocjonalnie i dającej poczucie względnej anonimowości stwarza pokusę i ułatwia każdemu Internaucie taką aktywność. Z całą pewnością trollowanie należy do kategorii zachowań nieprzyzwoitych. Należy jednak pamiętać, że bycie nieprzyzwoitym jest naturalne i dość łatwe, sprzyja mu bowiem: egoizm i chciwość (bycie przyzwoitym wymaga namysłu, bycie zaś nieprzyzwoitym jest zautomatyzowane), nieaktywność norm (np. osoby, które chodzą do kościoła w USA, rzadziej ściągają filmy porno w niedzielę), niedostatek samokontroli (samokontrola jest zasobem wyczerpywalnym), bycie istotą społeczną (w grupie łatwiej się wspierać) oraz sprawność myślenia (sprawne produkowanie usprawiedliwień dla tego, co robimy zwłaszcza w sytuacji niejasnej) (Ariely, 2012; Strefa Psyche Uniwersytetu SWPS, 2017). Warto podkreślić, że ocena danego zachowania jako przyzwoitego bądź nie zależy od etosu wyznawanego przez grupę oceniającą. Zatem w mniemaniu jednych zachowanie trolla, który multiplikuje nieprawdziwe informacje, ale zgodne z ich poglądem będzie właściwe, w ocenie zaś oponentów będzie zachowaniem nagannym. Ta dychotomia postaw ułatwia rozprzestrzenianie się nie-

prawdziwych informacji, bo tak rozumując prawda przestaje być obiektywna, lecz jest w oku patrzącego. Ta niejednoznaczność pojęcia prawdy jest *nihili novi*, dostrzegł ją już przecież Poncjusz Piłat.

Muriel J. Bebeau i wsp. (1999) podkreślają, że „wrażliwość moralna” lub „świadomość tego, w jaki sposób nasze działania wpływają na innych ludzi” są równie istotne jak zdolność rozumowania i wydawania sądów. Cheolho Yoon (2011) wskazuje, że w przestrzeni cyfrowej cel uświęca środki, dominuje przede wszystkim relatywizm moralny i utylityzm. Andrea Flores i Carrie James (2013) zauważyli, że w Internecie kwestie moralne rozpatruje się w oparciu o myślenie indywidualistyczne, a nie kolektywistyczne, właściwe jest to, co jest dobre dla mnie, nie dla innych. Anthony McCosker (2013) uważa, że główną motywacją trolli jest prowokacja, badania zaś Buckels i in. (2014) wskazują na istnienie związków pomiędzy skłonnością do trollowania a występowaniem Mrocznej Triady: narcyzmu, makiawelizmu i psychopatii, a także niskiej ugodowości i skłonności do sadyzmu, ale manifestowanego tylko w przestrzeni cyfrowej. Badania ujawniły także silny pozytywny związek między częstotliwością komentowania online a zachwianiami trollingowymi, co może potwierdzać wcześniejsze podejrzenia o związkach zachodzących pomiędzy nadmiernym wykorzystywaniem technologii a antyspołecznością (Carr, 2011; Juvonen, Gross, 2008; Phillips, Butt, 2006; Rosen i in., 2013).

Nie wszyscy są zgodni co do kierunku tej zależności, nadal nierozstrzygnięte zostaje pytanie, czy osoby o cechach antyspołecznych chętniej wykorzystują Internet do własnych nikczemnych celów, czy raczej długotrwałe uczestnictwo w sieci sprzyja rozwojowi cech antisocjalnych (Buckels i in., 2014; Piotrowski, 2019). Natalie Sest i Evita March (2017) z Federation University w Australii potwierdzają w dużym stopniu przytoczone wcześniej wyniki badań nad występowaniem sadyzmu u sprawców internetowego trollingu. Przebadały online 415 uczestników (36% mężczyzn, 63% kobiet) i zauważyły, że jego autorami częściej są mężczyźni. Cechował ich wyższy poziom psychopatii i sadyzmu, byli też pozbawieni empatii afektywnej, ale nie empatii poznawczej, którą stosowali jako użyteczne narzędzie do rozumienia emocji innych (ofiar), tylko po to, aby skuteczniej ich ranić. W tym miejscu warto podkreślić, że w literaturze łączy się problematykę trollingu z cyberprzemocą.

Niezależnie od motywów trolli przestrzeń digitalna jest miejscem, w którym formalna kontrola zachowania jest w dużym stopniu ograniczona, a zachowanie innych może hamować tylko sam uczestnik cyfrowej rzeczywistości poprzez blokowanie, kasowanie, nierozpowszechnianie fałszywych informacji, lecz przede wszystkim unikanie wdawania się w dyskusje z prowokującym oponentem, czyli „nie karmienie trolla” (Szulc, 2018, 2019). Osoby, które wcielają się w rolę trolli na zlecenie rozpowszechniających fałszywe i krzywdzące informacje, z dużym prawdopodobieństwem mogą swoją aktywność traktować jako zwykłą pracę i oddzielać „ja realne” od „ja wirtualne”. Podobne postawy są znane z przekazów historycznych, np. Rudolf Höss komendant obozu w Auschwitz Birkenau bez kłopotu potrafił oddzielać „pracę” od życia prywatnego. Było to możliwe dzięki temu, że nie dopuszczał do przenikania się obu sfer pomimo paradoksu, jakim było wykorzystywanie przez niego więźniów do wykonywania przydomowych prac (Wysińska, 2013). Z dużym prawdopodobieństwem świat wirtualny ułatwia stworzenie podobnej naturalnej bariery, tym bardziej, że myślenie moralne nie musi pokrywać się z moralnym postępowaniem (Colby, Damon, 1992). Na przykład sprawcy cyberprzemocy spostrzegają swoją aktywność w Internecie jako mniej „prawdziwą” niż w realnym

świecie, a wyrządzone krzywdy jawią się im jako „dzieło” innych ludzi (Rzepa, Przybylska-Duda, 2015).

MŁODZIEŻ W OBLICZU FAKE NEWSÓW

Okresem dorastania umownie określa się czas pomiędzy 11. a 18. rokiem życia. Wyróżnia się wczesną fazę (11/12–14/15 lat) oraz późną dorastania (14/15–19/20 lat) (Piotrowski, Wojciechowska, Ziółkowska, 2014a, 2014b). Jest to złożony i niezwykle dynamiczny czas w życiu człowieka obfitujący w wielorakie transformacje i liczne kryzysy (Kielar-Turska, 2000). Cechą charakterystyczną są trudności wychowawcze związane z poszukiwaniem własnej tożsamości i odpowiedzi na pytania egzystencjalne. Oprócz rozwoju tożsamości kluczową rolę w tym okresie odgrywa kształtowanie się światopoglądu na podstawie własnej wiedzy, doświadczenia, procesów psychicznych, a także stosunków społeczno-politycznych. Na tworzenie się umysłowej reprezentacji świata nastolatek ma wpływ szereg czynników: czas historyczny (zdarzenia, atmosfera), czynniki osobowościowe (styl umysłowości, poczucie kompetencji życiowej), indywidualne doświadczenia życiowe oraz płeć (Kielar-Turska, 2000). W tym okresie pojawia się idealizm młodzieńczy, który przejawia się w optymistycznej wizji przyszłości połączonej z wiarą w realizację marzeń. Rozczarowania płynące ze zderzenia z rzeczywistością prowadzą do kontestacji, buntu i krytykowania innych. Dorastanie jest także czasem zwiększonej podatności na oddziaływanie różnego rodzaju grup, idei, mód i mediów (Appelt, 2004).

Dla 99% Polaków w wieku 16–24 lata Internet jest głównym źródłem informacji, a najchętniej odwiedzanymi stronami są Facebook i YouTube (Klingová, Hajdu, 2018; Raport Wave 9, 2018). Badania brytyjskie pokazują, że uczniowie szkół podstawowych znacznie gorzej niż uczniowie szkół średnich zdają sobie sprawę ze zjawiska fake newsów. A nawet ci, którzy mają jakąś wiedzę na ten temat, nie zawsze są w stanie ich skutecznie rozpoznawać (Common Sense Media, 2017; Wineburg, 2016). Nastolatki i młodzi ludzie częściej niż inne grupy wiekowe korzystają z Internetu, a zatem naturalnie mogą przyczynić się do rozsyłania tzw. alternatywnych faktów (Łódzki, 2017). Brak umiejętności rozpoznawania fake newsów stanowi zagrożenie dla procesów demokratycznych i podważa zaufanie do instytucji. Konsekwencji tego doświadczyli Amerykanie podczas wspomnianych wyborów Trump – Clinton oraz Brytyjczycy podczas referendum w sprawie opuszczenia Unii Europejskiej. Zwykły krytycyzm oceny tego, co się czyta i ogląda, może tu nie wystarczyć, konieczne wydaje się pogłębianie wiedzy nastolatków na temat mechanizmów sterowania ich poglądami za pomocą fake newsów. Młodzież ma słabe wyczucie tego, co stanowi dylemat moralny, natomiast kłamstwa, oszukiwanie i zastraszanie są w ich mniemaniu zjawiskiem powszechnym (Flores, James, 2012; Smith i in., 2011). Wykazuje większą akceptację dla nieetycznego postępowania w Internecie niż w świecie rzeczywistym, a do tego przejawia trudności w ocenie tego, czy ich komentarz może być dla odbiorcy krzywdzący (Moor, Heuvelman, Verleur, 2010; Poole, 2007). Patologie na portalach społecznościowych cechuje brutalna bezwzględność, brak współczucia i schadenfreude, szerzenie kłamstw i oszczerstw, brak zdolności do niuansowania rzeczywistości oraz przyjmowania w dialogu perspektywy odbiorcy, jak również zachęcanie użytkowników (także ze strony nieodpowiedzialnych polityków, artystów czy ogólnie działaczy z każdej strony sceny politycznej) do przyjmowania skrajnych poglądów i odrzucania

punktów widzenia innych uczestników dyskusji. Sprzyja temu słaba kontrola przez adminów stron tematycznych w serwisach społecznościowych. Bywa, że administratorzy celowo tworzą kontrowersyjny funpage, aby poprzez skandal zyskać większą popularność. Ułatwiają tym samym manifestowanie postaw mizogicznych (poniżania i przedmiotowego traktowania kobiet), rasistowskich (nienawiści wobec uchodźców i gasterbeiterów) oraz klasistowskich (utożsamiania beneficjentów programu 500 plus z patologią społeczną), albo zachęcają do krytykanctwa (ośmieszania innych).

Tworzenie w mediach społecznościowych pozorowanej symetrii pomiędzy nauką i antynauką prowadzi do trudnych w oszacowaniu negatywnych konsekwencji wśród dorastającej młodzieży. Rozprzestrzenianie informacji na temat spisku koncernów farmaceutycznych i rzekomych zagrożeniach, jakie niosą szczepienia, wywołuje wątpliwości co do ich stosowania nie tylko u osób cechujących się niskim krytycyzmem. Takie wątpliwości coraz częściej manifestują osoby dobrze wyedukowane, a nawet lekarze, co przyczynia się do pogłębienia chaosu, utwierdza w przekonaniu, że nie istnieje na ten temat naukowy konsensus. W konsekwencji prowadzi to do podważenia zaufania do sprawdzonych procedur medycznych oraz metod leczenia i profilaktyki (Mieśnik, 2018). Istnieją dowody, że osoby mające wątpliwości coraz częściej poszukują odpowiedzi na pytania dotyczące zdrowia w Internecie zamiast u specjalistów z zakresu zdrowia publicznego (Broniatowski i in., 2018; Dubé, Vivion, MacDonald, 2015).

Pomimo faktu, że młodzież przesuwą swoją aktywność z Facebooka na Instagram, który jeszcze nie jest portalem opanowanym przez rodziców jak pierwszy ze wspomnianych, można zauważyć pewne prawidłowości:

- dobór znajomych na Facebooku jest zgodny z preferencjami ideologicznymi;
- u osób o podobnych poglądach mediana udziału przyjaciół z przeciwną ideologią wynosi tylko 20% dla liberałów i 18% dla konserwatystów;
- użytkownicy udostępniają artykuły, które są zgodne z ich poglądami;
- użytkownicy, którzy czerpią wiadomości z Facebooka (lub innych mediów społecznościowych), rzadziej otrzymują dowody na temat prawdziwego stanu świata (Bakshy, Messing, Adamic, 2015);
- wzajemne utwierdzanie się w postawach ideologicznych powoduje zniekształcenie obrazu świata.

Najczęstszym sposobem weryfikacji wiarygodności danych jest polecenie strony przez znajomych, co samo w sobie może prowadzić do uwiarygodnienia informacji poprzez regułę społecznego dowodu słuszności (skoro tak mówią inni, to znaczy, że to prawda) oraz samospełniające się proroctwo (poszukiwanie informacji zgodnych z przekonaniami i ignorowanie tych, które do nich nie pasują) (Cialdini, 2016; Szpunar, 2007). Reasumując, główny problem pojawia się wtedy, gdy celem fałszywych doniesień staje się szerzenie nienawiści. Niebezpieczeństwo polega na tym, że osoba poruszająca się w wąskim kręgu znajomych traktuje otrzymany od niej fake news jako informację prawdziwą. Odkrycie prawdy utrudnia mechanizm „bańki filtrującej” (*filter bubble*), polegający na tym, że odbiorca nie widzi wszystkich informacji, lecz otrzymuje je jako spersonalizowane. Algorytm dobiera informacje na podstawie dotychczasowej historii przeglądania w taki sposób, aby odizolować odbiorców od informacji niezgodnych z ich punktem widzenia (*Mały leksykon postprawdy*, 2018). Konsekwencją takiego filtrowania jest kształtowanie się fałszywego obrazu świata (Rack, Wolde-michael, Kimmel, 2017).

SPOSOBY RADZENIA SOBIE Z FAKE NEWS

Konsekwencje fake newsów mogą wykroczyć poza świat wirtualny, stanowiąc jak najbardziej realne, fizyczne zagrożenie. Dnia 4.12.2016 r. do pizzerii w Waszyngtonie wkroczył mężczyzna uzbrojony w karabin szturmowy i wystrzelił kilka serii w sufit, na szczęście nikogo nie raniąc. Jak później zeznał policji, chciał osobiście przekonać się, czy w restauracji elity polityczne powiązane z Hilary Clinton rzeczywiście zajmowały się handlem nieletnimi w celach seksualnych. Motywowały go historie, które czytał na prawicowych blogach i mediach społecznościowych na Facebooku, Reddite i Twitterze (Tandor, Wei Lim, Ling, 2017).

Walka z fake newsami nie jest łatwa, gdyż zmyślone treści często odpowiadają psychologicznym lub politycznym potrzebom odbiorców (*Mały leksykon postprawdy*, 2018). Według badań przeprowadzonych przez YouGov na 1125 internautach ze Stanów Zjednoczonych aż 28% dorosłych Amerykanów korzystających z zasobów sieci przyznało, że zdarzało im się celowo podejmować „szkodliwą działalność online”, czyli działania wymierzone w urażenie innego internauty. Portal WP tech (2016) wymienia następujące cechy kont trolli:

- generują znacznie więcej treści (nawet do 12 razy więcej niż typowy internauta);
- częściej sięgają po wyrazy potoczne i wulgaryzmy;
- mają tendencję do odpowiadania niemal na każdy komentarz, w którym ktoś inny odnosi się do ich słów;
- na profilach trolli nie ma żadnych informacji na ich temat lub udostępniają artykuły o treściach ideologicznych nieopatrzone własnym komentarzem;
- najczęściej stosują awatary lub korzystają ze zdjęć innych osób, np. zdjęć stockowych lub osób zmarłych.

Warto także wiedzieć, w jaki sposób ocenić wiarygodność informacji, opierając się na czterech podstawowych elementach: ocenie źródła, faktu, zdjęcia i aktualności. Ocena źródła jest związana z identyfikacją autora, stopki portalu, składu redakcji. Brak danych adresowych albo skrytka pocztowa za granicą powinny wzbudzać nieufność. Ocena faktu polega na porównaniu informacji na innych portalach. Jeśli są tworzone metodą „kopiuj – wklej” albo treść jest uboga, ogranicza się do nagłówka, jest wysycona emocjami, jej celem jest wywołanie wzburzenia, pochodzi z niewiarygodnego źródła (podejrzanych blogów), informacje należy traktować z rezerwą. Właśnie szokujący tytuł, którego celem jest zwrócenie uwagi odbiorców, powinien być jednym z ważniejszych sygnałów alarmowych dla czytelnika. Krytyczna analiza zdjęcia polega na zadaniu sobie pytania, czy należy ono do tekstu i nie jest przetworzone. Warto zastosować funkcję Google „wyszukiwanie obrazem”, aby zidentyfikować źródło zdjęcia. Ocena aktualności polega na zidentyfikowaniu okresu, w jakim powstała wiadomość za pomocą filtra „szukanie zaawansowane Google”. Warto także zwrócić uwagę na niewiarygodnie dużą liczbę „lajków” pod postem (Rack i in., 2017). Ważnym elementem jest zidentyfikowanie wiarygodności użytkownika, który udostępnił artykuł, kiedy powstało jego konto i w jaki sposób wygląda aktywność na jego profilu, np. jakie treści najczęściej udostępnił.

Kluczowe znaczenie ma tutaj podjęcie pracy edukacyjnej z młodzieżą, której podczas zajęć warsztatowych podaje się specyficzną „mentalną szczepionkę”, tworzącą inteligentnego konsumenta, a więc takiego, który posiadał zdolność krytycznego myślenia. Jako pierwszy koncepcję „zaszczepiania” zapoczątkował William McGuire (1964), próbując wzbudzić postawę oporności wobec perswazji i propa-

gandy za pomocą metody analogicznej do immunizacji biologicznej. Zastrzyki zawierające osłabioną dawkę wirusa mogą zwiększyć odporność na przyszłą infekcję poprzez aktywację produkcji przeciwciał. Teoria zaszczepiania postuluje, że to samo można osiągnąć za pomocą „przeciwciał informacyjnych”. Polega ona na tym, że gdy osobie przedstawimy jakiś fałszywy argument, który będzie potrafiła odeprzeć, to w ten sposób uodparniamy ją przeciwko podobnej argumentacji. Wystawiając ludzi na osłabioną wersję nieprawdziwych treści, można zbudować skuteczną linię oporu wobec przyszłych prób perswazji (McGuire, Papageorgis, 1961). Eksperyment przeprowadzony przez Jona Roozenbeeka i Sandrę van der Linden (2018) dostarcza obiecujących wyników. Stworzyli oni grę (*fake news game*), w której uczestnicy mieli za zadanie stworzyć artykuł informacyjny na temat silnie upolitycznionego problemu (np. europejskiego kryzysu uchodźczego), stosując wprowadzające w błąd taktyki, wykorzystywane zazwyczaj przez producentów nieprawdziwych wiadomości. Randomizowane badanie terenowe na próbie 95 uczniów szkół średnich dostarczyło wstępnych dowodów na to, że zastosowanie gry zwiększało sceptycyzm i rozbudziło krytyczne myślenie. Zatem gry edukacyjne mogą być obiecującym narzędziem do „zaszczepienia” młodzieży przed fałszywymi wiadomościami.

PODSUMOWANIE

Skuteczność manipulacji wynika z cech samych odbiorców. Zdaniem Marcina Iłowieckiego (2003) najchętniej przyjmują i akceptują oni te informacje, które potwierdzają ich wcześniej ugruntowane przekonania, odrzucają zaś opinie z nimi niezgodne. Nie są w stanie rozpoznać, czy informacja została zmodyfikowana, czemu sprzyja niechęć lub brak krytycyzmu. Bywa, że zachodzi reakcja odwrotna, wiadomości z innych źródeł są postrzegane jako całkowicie niewiarygodne. Cechuje ich skłonność do uogólniania i stosowania uproszczonych ocen rzeczywistości, tzw. klisz (stereotypów), np. żywią przekonanie, że wszyscy muzułmanie to potencjalni terroryści, każdy polityk to złodziej, światem zaś rządzą Żydzi. Często mają problemy ze zrozumieniem i właściwym zdefiniowaniem pojęć lub rzadziej stosowanych zwrotów przez dziennikarzy (suweren, koncyliacja, feminizm, legitymizacja, prawica, lewica, LGBT) oraz poddają się instynktowi stadnemu, odwołując się do zdania większości na jakiś temat. Odbiorcy przejawiają skłonność do dychotomizacji, widzenia rzeczy jako dobrych lub złych. Biało-czarna rzeczywistość zapewnia wprawdzie poczucie bezpieczeństwa, gdyż łatwiej ją pojąć i zaakceptować emocjonalnie w sytuacji zgeneralizowanego poczucia zagrożenia społecznego, jednak przyczynia się do pogłębienia konfliktów międzyludzkich (Pospiszyl, 2008).

Elias Aboujaoude (2012) przestrzega, że obserwowane w przestrzeni cyfrowej uproszczenia leksykalno-gramatyczne (niedbałość o ortografię, gramatykę i semantykę) mogą skutkować uszczupleniem kompetencji społecznych. Wydaje się też, że brak niuansowania rzeczywistości staje się poważną przeszkodą w dyskusji, zwłaszcza wtedy, gdy jedna ze stron stara się zachować wyważoną postawę, druga zaś świadomie wykorzystuje emocje, wiedząc, że w publicznej debacie za pośrednictwem mediów to głównie one dotrą do odbiorcy i go przekonają. Zatem przedstawiciel świata nauki opierający się na racjonalnych argumentach ma niewielkie szanse powodzenia w starciu medialnym ze zwolennikami tzw. medycyny alternatywnej. Stąd tak ważną rolę

w edukacji odgrywają popularyzatorzy nauki. Pojawiają się coraz częściej autorzy fanpejdżów uzbrojeni w rzetelną wiedzę poprzedzoną solidną edukacją i podejmują walkę z osobami naruszającymi dobre imię, np. zawodu lekarza. Na portalu „Lekarz Dawid Ciemiega” autor bezwzględnie rozprawia się z mitami na temat szczepień i informuje o kolejnych krokach prawnych podejmowanych wobec osób szkalujących i obrażających lekarzy. Na profilu „Patolodzy na klatce” autorka – lekarz patolog – w przystępny sposób przybliży problematykę chorób.

Tak ceniona przez młodzież wartość mediów społecznościowych, jaką jest interaktywność, niesie zagrożenie polegające na uleganiu pokusie dzielenia się swoimi refleksjami na każdy temat. Niczym nieskrępowany dostęp do tego swoistego „placu publicznego”, jakim są media społecznościowe, umożliwia nieograniczoną własną ekspresję nie zawsze przemyślaną (Tapscott, Williams, 2010). Łatwo też stać się pośmiewiskiem, zwłaszcza że nie jest przecież możliwe, by być ekspertem w każdej dziedzinie. Niektóre wypowiedzi zyskują ogromną popularność i zaczynają żyć własnym życiem, np. na portalach znajdujących się na wspomnianym już Facebooku: „Beka z antyszczepionkowców” albo „Dej, mam horom curke”^{6) 7)}.

Przykładem klasycznej manipulacji jest popularna także w mediach społecznościowych seria krótkich filmów pt. „Matura to bzdura”, w których pytana młodzież podczas sondy ulicznej daje spektakularny popis swojej ignorancji i niewiedzy głównie z zakresu historii, geografii i biologii oraz ogólnej wiedzy o świecie⁸⁾. Czytając niepochlebne, stereotypowe komentarze pod materiałem, widz pada ofiarą manipulacji i reaguje bezpośrednio na podany bodziec, nie zadając sobie trudu postawienia pytania, jak wiele zapytanych osób znało prawidłową odpowiedź na pytania, lecz nie zostały włączone do materiału, gdyż nie pasowały do założeń żartobliwej konwencji, której celem było wykazanie czegoś zgoła innego.

Raport przygotowany przez GLOBSEC Policy Institute pokazuje, że młodzież zdaje sobie sprawę, że nie wszystkie informacje, które pojawiają się w Internecie, są prawdziwe i rzetelne. Przyznaje też, że taka sytuacja budzi ich irytację i zgeneralizowane poczucie obrzydzenia oraz nieufności do źródeł informacji w ogóle. W rezultacie wzrasta nieufność i sceptycyzm wobec wszystkiego i wszystkich (Klingová, Hajdu, 2018).

Niezależnie od naszych postaw i poglądów nic nie zastąpi zdrowego rozsądku i świadomości, że przestała obowiązywać opracowana przez agencję prasową „Associated Press” zasada 5 (kto, co powiedział, kiedy, gdzie, dlaczego). Dlatego nieodzowna jest dziś edukacja medialna młodzieży w formie warsztatowej pracy nad rozpoznawaniem fałszywych informacji oraz konieczność zachowania ostrożności w upowszechnianiu niesprawdzonych informacji, a także dywersyfikacja jej źródeł.

Wypada także pamiętać, że edukacja w omawianym zakresie jest ważna, ale nie wystarczająca, przecież wiedza nie musi zmieniać postaw np. wobec uchodźców. Konieczna jest praca od najwcześniejszych poziomów edukacji nad empatią i zdolnością do przyjmowania innych punktów widzenia niż własne.

⁶⁾ <https://www.facebook.com/dejmamhoromcurke/>.

⁷⁾ <https://www.facebook.com/bekazantyszczepionkowcow/>.

⁸⁾ <http://maturatobzdura.tv/>.

BIBLIOGRAFIA

- Aboujaoude, E. (2012). *Wirtualna osobowość naszych czasów. Mroczna strona e-osobowości*. Kraków: Wydawnictwo UJ.
- Aleksejeva, N., Karan, K., Barojan, D., Nimmo, B. (2017). Facebook's Sputnik Take-down – In Depth. Inauthentic pages linked to Sputnik promoted its content. *Digital Forensic Research Lab*. <https://medium.com/dfirlab/facebook-sputnik-takedown-in-depth-f417bed5b2f8>.
- Allcott, H., Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.
- Appelt, K. (2004). Trudności wychowawcze w okresie dorastania. *Psychologia w Szkole*, 4, 19–25.
- Ariely, D. (2012). *The (honest) truth about dishonesty: How we lie to everyone---especially ourselves*. New York: HarperCollins.
- Bakshy, E., Messing, S., Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132, <http://dx.doi.org/10.1126/science.aaa1160>.
- Bąkiewicz, A. (2019). Wprowadzenie do definicji i klasyfikacji zjawiska fake newsa. *Studia Medioznawcze*, t. 20, 3(78), 280–289.
- BBC, News, (2017). *Fake Obama created using AI video tool*. <https://www.youtube.com/watch?v=AmUC4m6w1wo>.
- Bebeau, M.J., Rest, J.R., Narvaez, D. (1999). Beyond the promise: A perspective on research in moral education. *Educational Researcher*, 28(4), 18–26.
- Berry, J. (2017). The economic efficiency of fake news, *Oxford University Press's Academic Insights for the Thinking World, January 17th 2017. Verified November 28th 2017*: <https://blog.oup.com/2017/01/economic-efficiency-fake-news/>.
- Biel, B. (2019). *Tak działa rosyjska dezinformacja*. Konkret24. https://konkret24.tvn24.pl/swiat,109/tak-dziala-rosyjska-dezinformacja,901530.html?fbclid=IwAR2aPPXqCFFBVstTdb31DTgGoQm7YmruKPSnoa6vsvd_BqtенаQaWnmg-DvM4.
- Bradshaw, S., Howard, P.N. (2017). Troops, trolls and troublemakers: A global inventory of organized social media manipulation. *Working paper no. 2017.12, Computational Propaganda Project, University of Oxford*. <https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/07/Troops-Trolls-and-Troublemakers.pdf>.
- Broniatowski, D.A., Jamison, A.M., Qi, SiHua, AlKulaib, L., Chen, Tao, Benton, A., Quinn, S.C., Dredze, M. (2018). Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health*, 108(10), 1378–1384.
- Buckels, E.E., Trapnell, P.D., Paulhus, D.L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, 67, 97–102. doi: 10.1016/j.paid.2014.01.
- Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. New York, NY: W.W. Norton & Company.
- Cialdini, R. (2016). *Wywieranie wpływu na ludzi. Teoria i praktyka*. Gdańsk: GWP.
- Colby, A., Damon, W. (1992). *Some do care: Contemporary lives of moral commitment*. New York: Free Press.
- Common Sense Media (2017, March 8). News and America's Kids: How Young People Perceive and Are Impacted by the News. Common Sense Media. <https://www.commonsensemedia.org/research/news-and-americas-kids>.
- CyberDefence24 (2018). *Nowa broń w wojnie informacyjnej*. <https://www.cyberdefence24.pl/nowa-bron-w-wojnie-informacyjnej>.
- Ćwieluch, J., Różański, M. (2018). *Dlaczego przegramy wojnę z Rosją*. Wydawnictwo Znak Horyzont.
- Damisch, L., Stoerber, B., Mussweiler, T. (2010). Keep your fingers crossed! How superstition improves performance. *Psychological Science*, 21, 1014–1020. <https://doi.org/10.1177/0956797610372631>.
- Darwin, H., Neave, N., Holmes, J. (2011). Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, 50(8), 1289–1293. <https://doi.org/10.1016/j.paid.2011.02.027>.

- Dubé, E., Vivion, M., MacDonald, N.E. (2015). Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. *Expert Rev Vaccines*, 14(1), 99–117.
- Facebook Newsroom (2019). *Removing Coordinated Inauthentic Behavior from Russia*. https://newsroom.fb.com/news/2019/01/removing-cib-from-russia/amp/?__twitter_impression=true.
- Flores, A., James, C. (2012). Morality and ethics behind the screen: Young people's perspectives on digital life. *New media & Society*, 15(6), 834–852.
- Giles, K. (2015). *Russia's hybrid warfare: A success in propaganda*. Arbeitspapier Sicherheitspolitik, 1, 1–5. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwi2w9iR2sXjAhXC-i0KHQ6EATwQFjAAegQIAxAB&url=https%3A%2F%2Fwww.researchgate.net%2Fpublication%2F280922184_Russia%2527s_Hybrid_Warfare_A_Success_in_Propaganda&usg=AOvVaw28qIraImIhCcyXhPK3RMB5.
- Gorwa, M. (2017). Computational Propaganda in Poland: False Amplifiers and the Digital Public Sphere, Working Paper 4, University of Oxford. <https://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comp-prop-Poland.pdf>.
- Grotto, A.J. (2018). *Andrew Grotto Testimony on "Cybersecurity and California Elections"*. Center for International Security and Cooperation. <https://cisac.fsi.stanford.edu/docs/andrew-grotto-testimony-cybersecurity-and-california-elections>.
- Grzesiak-Feldman, M. (2016). *Psychologia myślenia spiskowego*. Warszawa: Wydawnictwa UW.
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication. From user discussions to academic definitions. *Journal of Politeness Research*, 6(2), 215–242. doi: 10.1515/JPLR.2010.011.
- Hardaker, C. (2013). "Uh.... not to be nitpicky,,,,,but...the past tense of drag is dragged, not drug.". An overview of trolling strategies. *Journal of Language Aggression and Conflict*, 1(1), 58–86. doi: 10.1075/jlac.1.1.04har.
- Hendricks, V.F., Vestergaard, M. (2019). *Reality lost: Markets of attention, misinformation and manipulation*. Springer Open, Cham, <http://dx.doi.org/10.1007/978-3-030-00813-0>.
- Hern, A. (2015). *Twitter CEO: We suck at dealing with trolls and abuse*. *The Guardian*. <http://www.theguardian.com/technology/2015/feb/05/twitterceo-we-suck-dealing-with-trolls-abuse>.
- Hunter, E., Pernik, P. (2015). *The Challenges of Hybrid Warfare*. Hybrid Warfare and Cyber Security. International Centre for Defence and Security.
- Howiecki, M., (2003). *Krzywe zwierciadło. O manipulacji w mediach*. Lublin: Gaudium.
- Imhoff, R., Lamberty, P.K. (2017). Too special to be duped: Need for uniqueness motivates conspiracy beliefs. *European Journal of Social Psychology*, 47(6), 724–734. <https://doi.org/10.1002/ejsp.2265>
- Iwanowska, M. (2016). *Zrozumieć media (społecznościowe). Perspektywa psychologiczna*. W: M. Łosiewicz, A. Ryłko-Kurpiewska, *Media biznes kultura. Rzeczywistość medialna. Formy, problemy, aspiracje*. Gdańsk: Novae Res.
- Juvonen, J., Gross, E.F. (2008). Extending the school grounds? Bullying experiences in cyberspace. *Journal of School Health*, 78, 496–505. <http://dx.doi.org/10.1111/j.1746-1561.2008.00335.x>.
- Juza, M. (2015). Hejterstwo w komunikacji internetowej: charakterystyka zjawiska, przyczyny i sposoby przeciwdziałania. *Profilaktyka Społeczna i Resocjalizacja*, 25, 27–50.
- Keen, A. (2007). *Kult amatora. Jak internet niszczy kulturę*. Warszawa: WAiP.
- Kelemen, D., Rosset, E. (2009). The Human Function Compunction: Teleological Explanation in Adults. *Cognition*, 111, 138–143.
- Kielar-Turska, M. (2000). Rozwój człowieka w pełnym cyklu życia. W: J. Strelau (red.), *Psychologia, t. 1* (s. 314–315). Gdańsk: GWP.
- Klingová, K., Hajdu, D. (2018). *From online battlefield to loss of trust? Perceptions and*

- habits of youth in eight European countries*. GLOBSEC Policy Institute, Bratislava, Slovakia.
- Korshunov, P., Marcel, S. (2018). DeepFakes: a New Threat to Face Recognition? Assessment and Detection. CoRR, arXiv,abs/1812.08685, 2019. URL <https://arxiv.org/pdf/1812.08685.pdf>.
- Kuczer, R. (2010). „Nowe nowe media”, Paul Levinson, Kraków 2010: [recenzja]. Resovia Sacra. *Studia teologiczno-filozoficzne Diecezji Rzeszowskiej*, 17, 333–338.
- Laberschek, M. (2010). Treści sensacyjne w internetowych serwisach informacyjnych i ich wpływ na wzrost poziomu czytelności. *Zarządzanie w Kulturze*, 11, 81–98.
- Levinson, P. (2010). *Nowe nowe media*. Kraków: Wydawnictwo WAM.
- Łódzki, B. (2017). Fake news – dezinformacja w mediach internetowych i formy jej zwalczania w przestrzeni międzynarodowej. *Polityka i Społeczeństwo*, 4(15), 19–30. doi: 10.15584/polispol.2017.4.2.
- Łuczuk, P. (2017). *Cyberwojna. Wojna bez amunicji?* Wydawnictwo Biały Kruk.
- Maciołek, D. (2014). Wpływ manipulacji medialnej na kształtowanie się tożsamości współczesnej młodzieży. *Kultura – Media – Teologia*, 19, 45–59.
- Maltby, J., Day, L., Hatcher, R., Tazzyman, S., Flowe, H.D, Palmer, E.J, Frosch, C.A, O'Reilly, M., Jones, C., Buckley, Ch., Knieps, M., Cutts, K. (2016). Implicit theories of online trolling: evidence that attention-seeking conceptions are associated with increased psychological resilience. *British Journal of Psychology*, 107(3), 448–466.
- Mały leksykon postprawdy (2018). Fundacja Wolność i Demokracja. <https://wid.org.pl/maly-leksykon-postprawdy/>.
- Marchlewska, M. Cichocka, A., Kossowska, M. (2017). Addicted to answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European Journal of Social Psychology*, 48(2), 109–117.
- McCosker, A. (2013). Trolling as provocation: YouTube's agonistic publics. *The International Journal of Research into New Media Technologies*, 20(2), 201–217.
- McGuire, W.J. (1964). Some contemporary approaches. *Advances in Experimental Social Psychology*, 1, 191–229. doi:10.1016/S0065-2601(08)60052-0.
- McGuire, W.J., Papageorgis, D. (1961). The Relative Efficacy of Various Types of Prior Belief-Defense in Producing Immunity against Persuasion. *Journal of Abnormal and Social Psychology*, 62, 327–337.
- Mead, M. (1978). *Kultura i tożsamość Studium dystansu międzypokoleniowego*. Warszawa: PWN.
- Mieśnik, M. (2018). *Pediatra antyszczepionkowiec rezygnuje z pracy w państwowej przychodni. Wszczęto postępowanie*. WP Wiadomości. <https://wiadomosci.wp.pl/pediatra-antyszczepionkowiec-rezygnuje-z-pracy-w-panstwowej-przychodni-wszczeto-postepowanie-6284286099965569a>.
- Molga, T. (2017). *Bunt prawicowych trolli. Nawet Paweł Szefernaker stracił nad nimi kontrolę*. WP Wiadomości. <https://wiadomosci.wp.pl/bunt-prawicowych-trolli-nawet-pawel-szefernaker-stracil-nad-nimi-kontrola-6153688357341313a>.
- Moor, P.J., Heuvelman, A., Verleur, R. (2010). Flaming on YouTube. *Computers in Human Behavior*, 6, 1536–1546.
- Mustafarai, E., Metaxas, P.T. (2017). The fake news spreading plague: was it preventable? In Proceedings of ACM Web Science Conference, Troy, NY, USA, June 2017 (WebSci'17): 235–239 <https://doi.org/10.1145/3091478.3091523>.
- Phillips, J.G., Butt, S. (2006). Personality and self-reported use of mobile phones for games. *CyberPsychology & Behavior*, 9, 753–758. <http://dx.doi.org/10.1089/cpb.2006.9.753>.
- Piotrowski, A. (2010). Wizerunek medialny Służby Więziennej. *Przegląd Więziennictwa Polskiego*, 67–68, 85–93.
- Piotrowski, A. (2019). Relation of inmates' entitlement to identification with criminal groups, narcissism and readiness for self-improvement. *Current Issues in Personality Psychology*, 7(3), 252–263. <https://doi.org/10.5114/cipp.2019.89169>.
- Piotrowski, K., Wojciechowska, J., Ziółkowska, B. (2014a). *Niezbędnik Dobrego*

- Nauczyciela. Diagnoza zasobów nastolatka i środowiska rozwoju. Wczesna faza dorastania.* Warszawa: Wydawnictwo IBE.
- Piotrowski, K., Wojciechowska, J., Ziolkowska, B. (2014b). *Niezbędnik Dobrego Nauczyciela. Diagnoza zasobów nastolatka i środowiska rozwoju. Późna faza dorastania.* Warszawa: Wydawnictwo IBE.
- Poole, D. (2007). A study of beliefs and behaviors regarding digital technology. *New Media & Society*, 9(5), 771–793.
- Polański, G. (2017). Gratyfikacje z korzystania z mediów. *Edukacja – Technika – Informatyka*, 3(21), 250–255.
- Polityka.pl (2017). *Konferencja: Czym są fake news i jak się przed nimi bronić.* <https://www.polityka.pl/tygodnikpolityka/kraj/1703036,1,konferencja-czym-sa-fake-news-i-jak-sie-przed-nimi-bronic.read>.
- Poltergeist (2012). *Jak skutecznie trollować?* <http://polter.pl/Jak-skutecznie-trollowac-b15536>.
- Pospiszyl, I. (2008). *Patologie społeczne.* Warszawa: WN PWN.
- Pyżalski, J. (2012). *Agresja elektroniczna i cyberbullying jako nowe ryzykowne zachowania młodzieży.* Kraków: Impuls.
- Rack, S., Woldemichael, D., Kimmel, B. (2017). *Fakt czy fake? Jak można demaskować fałszywe doniesienia pojawiające się w internecie?* <https://www.safe-rinternet.pl/pliki/publikacje/Fakt%20czy%20fake.pdf>; www.klicksafe.de.
- Raport Wave 9 (2018) *Polak zapatrzonny w Facebooka.* <http://www.universalmc-cann.pl/news/68/114/Wave-9-Polak-zapatrzonny-w-Facebooka>.
- Rosen, L.D., Whaling, K., Rab, S., Carrier, L.M., Cheever, N.A. (2013). Is Facebook creating “iDisorders?” The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. *Computers in Human Behavior*, 29, 1243–1254. <http://dx.doi.org/10.1016/j.chb.2012.11.012>.
- Roozenbeek, J., Van der Linden, S. (2018). The fake news game: actively inoculating against the risk of misinformation. *Journal of Risk Research*, 22(5), 570–580. doi: 10.1080/13669877.2018.1443491.
- Rzepa, T.N., Przybylska-Duda, M. (2015). Cyberbullying z perspektywy nastolatków i ich rodziców. *Annales Universitatis Mariae Curie-Skłodowska Lublin – Polonia*, 28(2) sectio J, 19–34.
- Sest, N., March, E., (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences*, 119, 69–72.
- Smith, C., Christoffersen, K., Davidson, H., Herzog, P.S. (2011). *Lost in Transition: The Dark Side of Emerging Adulthood.* New York: Oxford University Press.
- Strefa Psyche Uniwersytetu SWPS (2017). *Dlaczego tak trudno być przyzwoitym człowiekiem – debata z okazji benefisu prof. Bogdana Wojciszke.* <https://youtubersi.pl/film/261393>.
- Szpunar, M. (2007). Internet – medium informacji versus dezinformacji. *e-mentor*, GH, 2, 46–51.
- Szulc, M. (2018). Sposoby radzenia sobie ze stresem i poczucie kontroli u ofiar cyberprzemocy. *Pomeranian J. Life Sci.*, 64(3), 81–87.
- Szulc, M. (2019). Psychologiczne uwarunkowania, konsekwencje i przeciwdziałanie zjawisku cyberprzemocy. *WNCD*, 5–6 (272–273), 3–9.
- Tandoc, E.C. Jr., Lim, Z.W., Ling, R. (2017). Defining “fake news”. A typology of scholarly definitions. *Digital Journalism*, 6(2), 137–153. doi: 10.1080/21670811.2017.1360143.
- Tapscott, D., Williams, A.D. (2010). *Macrowikinomics: Rebooting business and the world.* Canada: Penguin Group.
- Townsend, Tess. (2016). *Meet the Romanian Trump Fan behind a Major Fake News Site.* Inc. <http://www.inc.com/tess-townsend/ending-fed-trump-facebook.html>.
- Turska-Kawa, A. (2013). Media jako źródło zaspokajania fundamentalnych potrzeb: Rozważania w kontekście teorii „użytkowania gratyfikacji”. *Studia Politicae Universitatis Silesiensis*, 10, 147–170.
- Van Prooijen, J.-W., Van Dijk, E. (2014). When consequence size predicts belief in conspiracy theories: The moderating role of perspective taking. *Journal of Experimental Social Psychology*, 55, 63–73.
- Van Prooijen, J.W., Krouwel, A.P., Pollet, T.V. (2015). Political extremism predic-

- ts belief in conspiracy theories. *Social Psychological and Personality Science*, 6, 570–578.
- Wierzejski, A. (2017). *Wojna informacyjna w Internecie. Przeciwdziałanie prokremlowskiej dezinformacji w państwach Europy Środkowej i Wschodniej*. Centrum Stosunków Międzynarodowych. http://informationwarfare.csm.org.pl/wp-content/uploads/2017/07/REPORT_Information-warfare_Polish.pdf.
- Wineburg, S., McGrew, S., Breakstone, J., Ortega, T. (2016). *Evaluating Information: The Cornerstone of Civic Online Reasoning*. *Stanford Digital Repository*. <http://purl.stanford.edu/fv751yt5934>.
- WP Tech (2016). *Sposób na internetowych trolli i hejterów*. <https://tech.wp.pl/sposob-na-internetowych-trolli-i-hejterow-6034811830030977a>.
- Wysińska, S. (2013). Wizerunek Rudolfa Hössa jako osoby prywatnej: na podstawie akt procesowych i relacji więźniów KL Auschwitz-Birkenau. *Wieki Stare i Nowe*, 5(10), 137–155.
- Yoon, C. (2011). Theory of planned behavior and ethics theory in digital piracy: An integrated model. *Journal of Business Ethics*, 100(3), 405–417.

MANIPULATING INFORMATION ONLINE USING FAKE NEWS AS A THREAT TO YOUNG PEOPLE

Abstract: Besides unquestionable benefits which we can gain from using media, the relevant threats cannot be ignored. So far, the researchers' attention mostly focused on cyberbullying and addiction. These days, more and more people recognise the danger, which is an ease with which fake news can be multiplied on the Internet. Younger generations, for whom Internet is a natural source of knowledge about the world, can have some troubles with critical rating of information. Furthermore, the assessment of reliability requires knowledge, commitment, time, and most importantly – a desire to verify of the viewer. The prob-

lem seems wider, and it doesn't concern only young people. More and more often people who are characterized by knowledge and life experience can be fooled by fake news, especially when they are not experts in some fields. The main aim of the article is to characterize the phenomenon of fake news, and also psychological reasons shared by the users of digital world, referred to as 'trolls'. Also, the methods used to identify fake news on the Internet are described.

Keywords: manipulation, fake news, trolling, youth.