

Krzysztof Bartosz Klimiuk, Łukasz Władysław Balwicki

WHAT IS INFODEMOLOGY? AN OVERVIEW AND ITS ROLE IN PUBLIC HEALTH

CZYM JEST INFODEMOLOGIA? PRZEGLĄD I ROLA W ZDROWIU PUBLICZNYM

Division of Public Health & Social Medicine, Medical University of Gdańsk, Poland
Zakład Zdrowia Publicznego i Medycyny Społecznej, Gdański Uniwersytet Medyczny, Polska

ABSTRACT

The exponential increase in internet use and the consequent surge in content generation present both opportunities and challenges for public health. Infodemiology, an emerging field at the intersection of information science and public health, seeks to harness the vast amounts of health-related data generated online for public health benefits. This paper provides a general overview of infodemiology, examining its development, methodologies, and potential to address public health challenges. We discuss the role of infodemiology in identifying and mitigating the spread of misinformation, especially in the context of the COVID-19 pandemic, which underscored the dangers of the “infodemic” – an overabundance of information, both accurate and not, that complicates public health responses. Through both demand and supply-based studies, infodemiology offers unique insights into health trends, misinformation dynamics, and the digital behaviors of health information seekers. Tools such as sentiment analysis are highlighted in navigating the vast digital landscape for real-time health data analysis. Despite the potential of infodemiology, challenges such as data overload or misinformation. The paper concludes by emphasizing the importance of interdisciplinary collaboration, the development of advanced analytical tools, and the need for guidelines to maximize the field’s impact on public health policy and practice.

Keywords: *health information management, health communication, infodemiology, infoveillance*

STRESZCZENIE

Wykładniczy wzrost używania internetu oraz wynikający z tego wzrost ilości generowanych treści stwarza zarówno szanse, jak i wyzwania dla zdrowia publicznego. Infodemiologia, niedawno powstała dziedzina, pojawia się na przecięciu nauk związanych z informacją i zdrowia publicznego. Dąży do wykorzystania ogromnych ilości danych zdrowotnych generowanych w internecie na korzyść zdrowia publicznego. Niniejszy artykuł przedstawia ogólny przegląd infodemiologii, pokazując jej rozwój, metodologie oraz potencjał w rozwiązywaniu wyzwań zdrowia publicznego. Omówiona zostaje rola infodemiologii w identyfikowaniu i zapobieganiu rozprzestrzenianiu się dezinformacji, szczególnie w kontekście pandemii COVID-19, która podkreśliła niebezpieczeństwa „infodemii” – nadmiaru informacji, zarówno prawdziwych, jak i fałszywych, utrudniających działania zdrowia publicznego. Poprzez badania oparte na popycie i podaży, infodemiologia oferuje unikalną perspektywę na trendy w zdrowiu, dynamikę dezinformacji oraz cyfrowe zachowania osób szukających informacji zdrowotnych. Narzędzia takie jak analiza sentymentu są wskazane jako istotne dla analizy danych zdrowotnych w czasie rzeczywistym. Pomimo potencjału infodemiologii, omówione są wyzwania takie jak przeciążenie danymi, czy dezinformacja. Artykuł kończy się podkreśleniem znaczenia interdyscyplinarnej współpracy, rozwoju zaawansowanych narzędzi analitycznych oraz potrzeby wytycznych, aby maksymalizować wpływ tej dziedziny na politykę i praktykę zdrowia publicznego.

Słowa kluczowe: *infodemiologia, komunikacja zdrowotna, zarządzanie informacją zdrowotną, nadzór informacyjny*

INTRODUCTION

In 2023, internet access was available in 93.3% of households in Poland, maintaining a trend of stability observed in previous years (1). The youngest adult age group in Poland (18-24 years) spends up to 6 hours per day using smartphones, and over 60% of the population actively uses social media, generating an unprecedented amount of data (2,3). The increasingly frequent use of internet resources is part of a global trend; in 2023, 92% of individuals in European Union aged 16 to 74 used the internet at least once in the three months prior to the survey date (4). Globally, the number of young internet users has risen in both developed and developing countries to 94% and 67% respectively as of 2017 (5). This widespread use of the internet, facilitated by smartphones, laptops, and other portable devices, allows people to access it from almost any location, leading to the creation of an extraordinarily large volume of data on all aspects of life, including health (6,7).

It is estimated that the data of a Polish internet user covering a year of online activity is valued at approximately 40 PLN (8). These data are collected, for example, by Google, where users consent to data collection, such as routes taken on Google Maps or data from applications logged in using a Google account (9). Later, these data are used to create personalized ads, which are said to provide better conversion rates, leading to almost a 10% increase in sales (10). However, these data need not be exclusively the domain of commercial marketing and can also be utilized by public health initiatives.

According to Eysenbach's research, nearly 4.5% of internet queries are focused on health, translating to over 6 million daily searches (11). A significant portion of the Polish population, with 62.9% seeking treatment options, 55.8% checking symptoms, and 54.5% researching medications online, reflects this trend (12). It's observed that online communities are preferred for health inquiries, as they often include individuals with firsthand experience or deeper knowledge about specific conditions (13). Particularly in mental health, where stigma and complexity play a role, a notable number of individuals turn to the internet for information, especially regarding depression (14). However, the rapid dissemination of information online can have detrimental effects. Misinformation can lead to serious consequences, as highlighted by a study in the *American Journal of Tropical Medicine and Hygiene*, which noted that false information on social media led to about 5,800 hospital admissions (15). Additionally, there were reported that misinformation related to COVID-19 resulted in at least 800 deaths globally in a three-month period (16).

WPROWADZENIE

W 2023 roku, dostęp do Internetu miało 93,3% gospodarstw domowych w Polsce, co wpisuje się w stabilny trend wzrostowy obserwowany w poprzednich latach (1). Najmłodsza grupa dorosłych w Polsce (18-24 lata) spędza do 6 godzin dziennie na korzystaniu ze smartfonów, a ponad 60% populacji jest aktywnymi użytkownikami mediów społecznościowych, generując znaczącą ilość danych (2,3). Coraz powszechniejsze korzystanie z Internetu jest częścią globalnego trendu; w 2023 roku 92% osób w Unii Europejskiej w wieku od 16 do 74 lat korzystało z Internetu przynajmniej raz w ciągu trzech miesięcy (4). Na świecie, liczba młodych użytkowników Internetu wzrosła zarówno w krajach rozwiniętych, jak i rozwijających się do odpowiednio 94% i 67% w 2017 roku (5). Wszecobecne korzystanie z Internetu, ułatwione przez smartfony, laptopy i inne przenośne urządzenia, umożliwia ludziom dostęp do niego z prawie każdego miejsca, co prowadzi do tworzenia ogromnej ilości danych ze wszystkich aspektów życia, w tym zdrowia (6,7).

Szacuje się, że dane polskiego użytkownika Internetu obejmujące rok aktywności w sieci mają wartość około 40 PLN (8). Dane te są zbierane przykładowo przez Google, gdzie użytkownicy wyrażają zgodę w polityce prywatności na zbieranie danych takich jak trasy przebyte za pomocą Google Maps, czy dane z aplikacji, do których zalogowano się przy użyciu konta Google (9). Następnie te dane są wykorzystywane do tworzenia spersonalizowanych reklam, które szacunkowo prowadzą do wzrostu sprzedaży o prawie 10% (10). Jednakże, dane te nie muszą być wykorzystywane wyłącznie w dziedzinie marketingu komercyjnego i mogą również znaleźć zastosowanie w inicjatywach zdrowia publicznego.

Według badań Eysenbacha, blisko 4,5% zapytań w Internecie dotyczy zdrowia, co przekłada się na ponad 6 milionów wyszukiwań dziennie (11). Znacząca część polskiej populacji odzwierciedla ten trend, gdzie 62,9% szuka w Internecie opcji leczenia, 55,8% sprawdza objawy, a 54,5% sprawdza swoje leki online (12). Obserwuje się, że społeczności online są preferowaną przestrzenią dla zapytań związanych ze zdrowiem, ponieważ często obecne są tam osoby z bezpośrednim doświadczeniem lub głębszą wiedzą na temat konkretnych chorób (13). Szczególnie w dziedzinie zdrowia psychicznego, gdzie stygmatyzacja i złożoność osobistej sytuacji odgrywają istotną rolę, znacząca liczba osób szuka informacji w Internecie, zwłaszcza na temat depresji (14). Jednak, powszechne korzystanie z informacji z Internetu może mieć szkodliwe skutki. Znaczenie obecnej tam dezinformacji jest podkreślone przez badanie opublikowane w *American Journal of Tropical Medicine and Hygiene*, w którym zauważo-

INFODEMOLOGY AS A MANAGEMENT OF ONLINE HEALTH DATA

In this digital age, the widespread use of information and communication technologies significantly impacts every sector of society, particularly in the realm of knowledge sharing and decision-making facilitated by social media and online platforms (17). The vast amounts of user-generated data have opened doors to the exploration and analysis of patterns in information-seeking behavior, providing valuable insights into how societal behaviors evolve over time (18). It is within this context that the concept of ‘infodemiology’ emerges. This innovative field focuses on the study and analysis of health-related data and information created by users on the internet, offering unique perspectives on online health information behaviors (19). While traditional epidemiological methods are indispensable, infodemiology introduces a novel approach by offering real-time, unbiased insights into online health behaviors, thus complementing traditional epidemiological data (20).

Infodemiology is fundamentally based on the relationship between public health and the patterns of information and communication on the internet (20). It is defined as the ‘science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy’ (21). Positioned at the intersection of public health informatics, consumer health informatics, infometrics, and web analytics, infodemiology is a highly interdisciplinary and emerging field. It necessitates collaboration among information scientists, computer scientists, epidemiologists, medical experts, public health professionals, behavioural scientists, and statisticians (20,21). While Eysenbach initially described infodemiology as focusing on the misinformation, later it has shifted the main focus of the determinants and distribution of health information and misinformation (22). Furthermore, infodemiology is variously defined in relation to its scope, including techno-social predictive analysis (23), digital disease detection (24), web data or digital disease surveillance (25,26), and e-epidemiology (27).

HISTORY OF INFODEMOLOGY

Although the intersection of health information and the Internet predates the coining of the term ‘infodemiology’, and there were numerous internet-based publications for health and information before this field was formally named (28), the term ‘infodemiology’ itself was first introduced by Gunther Eysenbach in 2002 (29). Initially, its use diverged from

no, że fałszywe informacje w mediach społecznościowych doprowadziły do około 5800 hospitalizacji (15). Dodatkowo, zgłoszono, że dezinformacja związana z COVID-19 spowodowała co najmniej 800 zgonów na całym świecie w przeciągu tylko trzech miesięcy (16).

INFODEMOLOGIA JAKO ZARZĄDZANIE DANymi ZDROWOTNYMI ONLINE

W dobie cyfrowej rzeczywistości, wszechobecne użycie technologii informacyjnych i komunikacyjnych znacząco wpływa na każdy sektor społeczeństwa, zwłaszcza w dziedzinie dzielenia się wiedzą i podejmowania decyzji, co jest ułatwione przez media społecznościowe i inne platformy internetowe (17). Ogromne ilości danych generowanych przez użytkowników otworzyły drzwi do eksploracji i analizy wzorców wyszukiwania informacji, dostarczając wglądu w to, jak zachowania społeczne zmieniają się z biegiem czasu (18). To właśnie w tym kontekście pojawia się koncepcja „infodemiologii”. Ta innowacyjna dziedzina koncentruje się na badaniu i analizie danych i informacji zdrowotnych tworzonych przez użytkowników Internetu, oferując unikalną perspektywę na zachowania związane z informacjami zdrowotnymi online (19). Choć tradycyjne metody epidemiologiczne są niezastąpione, infodemiologia wprowadza nowe podejście, oferując obiektywny wgląd w internetowe zachowania zdrowotne w czasie rzeczywistym, uzupełniając tym samym tradycyjne dane epidemiologiczne (20).

Infodemiologia zasadniczo opiera się na związku między zdrowiem publicznym a wzorcami informacji i komunikacji w Internecie (20). Jest definiowana jako „nauka o dystrybucji i determinantach informacji w środowisku elektronicznym, głównie w Internecie, z ostatecznym celem informowania zdrowia publicznego i polityki publicznej” (21). Stojąca na skrzyżowaniu nauk o zdrowiu, badań nad zachowaniami konsumentów, metryki informacyjnej i analizy danych internetowych, infodemiologia jest wielodyscyplinarną, dynamicznie rozwijającą się dziedziną. Jej rozwój wymaga kooperacji między ekspertami z różnych dziedzin, w tym naukowcami zajmującymi się danymi, epidemiologami, lekarzami, specjalistami zdrowia publicznego, badaczami zachowań społecznych oraz statystykami (20,21). Choć początkowo Eysenbach opisał infodemiologię jako skupiającą się na dezinformacji, później przesunął główny nacisk na określanie determinantów i dystrybucji informacji (22). Co więcej, infodemiologia jest różnie definiowana w zależności od jej obszaru zastosowania, włączając w to analizę prognostyczną technospołeczną (23), cyfrowe wykrywanie chorób (24), dane internetowe lub cyfrowy nadzór nad chorobami (25,26), lub e-epidemiologię (27).

what we currently understand, as it was applied in the context of identifying misinformation. However, it soon became clear that Internet queries could attempt to predict influenza epidemics but also analyze the demand for published web content. This shift involved examining people's needs and monitoring their health information-seeking behavior. Although initially, the focus was on information prepared for users, it later emerged that studying the content users themselves posted could yield a wealth of data to better understand the population (29).

Professor Gunther Eysenbach is credited with coining 'infodemiology' in 2002 (29), defining it as *"The study of the determinants and distribution of health information and misinformation, which may be useful in guiding health professionals and patients to quality health information on the Internet. Information epidemiology, or infodemiology, identifies areas where there is a knowledge translation gap between best evidence (what some experts know) and practice (what most people do), as well as markers for 'high-quality' information."* (29).

In subsequent work, Eysenbach expanded on the scope of infodemiology, proposing a framework that includes basic metrics and differentiates between supply-side features (like publishing activity on the Internet) and demand-based features (such as search and navigation behavior). He argued that *"these metrics and methods could benefit public health practice and research and should be further developed and standardized"* (17). He later refined the definition of infodemiology to *"the science of distribution and determinants of information in an electronic medium, specifically the Internet, with the ultimate goal of informing public health and public policy"* (20).

Eysenbach also introduced the terms "infoveillance", referring to the systematic surveillance of information applications in public health, and "infodemic", defined as *"an excessive amount of unfiltered information about a problem, making the solution more difficult"* (20). The term "infodemic" was largely unnoticed until the COVID-19 era when Dr. Tedros Adhanom Ghebreyesus, Director-General of the World Health Organization (WHO), highlighted the challenge of fighting not only an epidemic but also an infodemic at the Munich Security Conference on February 15, 2020 (22).

RESEARCH IN THE FIELD OF INFODEMOLOGY

The area of infodemiology has generated studies based on its unique methodology, proven valid for identifying public health challenges (30).

HISTORIA INFODEMOLOGII

Chociaż analizy informacji zdrowotnych w Internecie z wykorzystaniem metod typowych dla infodemiologii były prowadzone jeszcze przed oficjalnym ukształtowaniem tej dyscypliny (28), to za jej formalny początek przyjmuje się moment opublikowania terminu „infodemiologia” przez Gunthera Eysenbacha w 2002 roku (29). Początkowe użycie tego terminu różniło się od współczesnego znaczenia, gdyż odnosiło się przede wszystkim do identyfikacji dezinformacji. Jednakże szybko okazało się, że zapytania internetowe mogą służyć do analizy popytu na opublikowane treści internetowe, prób przewidywania epidemii grypy, jak również mogą dostarczać cennych wglądów poprzez koncentrację na stronie podaży. Ta zmiana kierunku badań skupiała się na analizie potrzeb ludzi i monitorowaniu ich zachowań w poszukiwaniu informacji zdrowotnych. Mimo, że na początku skupiono się na dostarczanych użytkownikom treściach, to z czasem ujawniło się, że badanie materiałów publikowanych przez samych użytkowników może dostarczyć bogatego zbioru danych, umożliwiającego głębsze zrozumienie populacji (29).

Profesor Gunther Eysenbach, wprowadzając w 2002 roku pojęcie „infodemiologii” (29), zdefiniował je jako: *„Badanie determinant i dystrybucji informacji zdrowotnych oraz misinformacji, które mogą przyczyniać się do ukierunkowania zarówno profesjonalistów ochrony zdrowia, jak i pacjentów ku wiarygodnym źródłom informacji zdrowotnych w sieci. Infodemiologia, rozumiana jako epidemiologia informacyjna, pozwala identyfikować obszary, w których występuje brak spójności między najlepiej udokumentowanymi dowodami (tym, co jest znane niektórym ekspertom), a rzeczywistą praktyką (tym, co stosuje większość ludzi), a także wskazuje na markery informacji o 'wysokiej jakości'."* (29).

W swoich kolejnych opracowaniach Gunther Eysenbach rozwinął i doprecyzował zakres infodemiologii, uwzględniając analizy zarówno z perspektywy podaży (np. aktywność publikacyjna użytkowników w internecie) jak i popytu (czyli treści, z którymi użytkownicy się stykają i których aktywnie poszukują). Argumentował, że *„te metody badawcze mogą przynieść istotne korzyści dla praktyki i nauki w dziedzinie zdrowia publicznego, wymagają jednak dalszego rozwoju i standaryzacji"* (17). Później doprecyzował definicję infodemiologii jako *„nauki o dystrybucji i determinantach informacji w środowisku elektronicznym, specyficznie w Internecie, z ostatecznym celem dostarczania informacji dla potrzeb zdrowia publicznego i polityki publicznej"* (20).

Eysenbach wprowadził również terminy takie jak „nadzór informacyjny” (ang. *infoveillance*), który od-

Infodemiology is characterized by its adaptability to the type of data used, which determines the design of the studies conducted. This includes demand-based studies, such as those analyzing advertisements, Google searches, or surveillance of articles and news containing medical data (19). These studies operate on the premise that individuals gain knowledge from the static and prepared part of the internet, where they are treated as consumers (31).

Demand-based research in infodemiology primarily focuses on the initial stage of the World Wide Web, characterized by static web pages and content primarily created by a small number of contributors, with the vast majority of users acting as consumers (32,33). An example of such research could involve identifying disinformation and disinformation regarding vaccinations (34) or the COVID-19 pandemic (35). However, this type of research faces several challenges. Firstly, the filtration and interpretation of the vast amounts of data pose a significant challenge. Secondly, distinguishing between high-quality information and false information requires robust analytical methods. Lastly, the majority of these studies are cross-sectional, meaning they only provide a snapshot of the situation at a specific moment, limiting their ability to track changes over time or establish causality (19,32).

On the other hand, there are studies based on content posted by internet users, known as "Supply based studies". Contrary to demand-based studies, where content is more static and prepared, supply-based research focuses on the dynamic and interactive nature of content. Dynamic development of social media communities have opened up broader possibilities for supply-based infodemiology research (31). These studies often examine posts on Twitter (X) and Facebook, as well as forums or blogs. Their goals include supporting the design of more effective health interventions and campaigns that meet the needs of online communities (19).

Supply-based studies in infodemiology delve into the content provided by users on online platforms to comprehend the health information being shared and its impact on public discourse (19). Such research includes analyses of comments from vaccine opponents (36) or efforts to identify suicide risk factors by tracking people's posts on Twitter (37). These studies focus on analyzing and verifying the credibility of user-generated content, understanding the context and intentions behind the shared information, and finding a balance between freedom of speech and the necessity of providing accurate and reliable health information. Similar to demand-based research, supply-based studies also present a snapshot in time,

nosi się do systematycznego monitorowania aplikacji informacyjnych w zdrowiu publicznym, oraz „infodemię” (ang. *infodemic*), definiowaną jako „*nadmiar niefiltrowanych informacji dotyczących określonego problemu, co utrudnia odnalezienie skutecznych rozwiązań*”. Termin „infodemia” pozostał w dużej mierze niezauważony aż do czasów pandemii COVID-19, kiedy to dr Tedros Adhanom Ghebreyesus, Dyrektor Generalny Światowej Organizacji Zdrowia (WHO), zwrócił uwagę na problem walcząc nie tylko z samą epidemią, ale również z infodemią. Podkreślił to wyzwanie podczas Monachijskiej Konferencji Bezpieczeństwa, która miała miejsce 15 lutego 2020 roku (22).

BADANIA W DZIEDZINIE INFODEMOLOGII

W dziedzinie infodemiologii rozwijane są badania oparte na specyficznej metodologii, która została oceniona jako efektywna w rozpoznawaniu wyzwań zdrowia publicznego (30). Infodemiologia charakteryzuje się adaptacyjnością do rodzaju wykorzystywanych danych, co decyduje o kształcie prowadzonych badań. Obejmuje to analizy oparte na popycie, takie jak badanie reklam, zapytań w Google czy monitorowanie publikacji i wiadomości zawierających dane medyczne (19). Prace te zakładają, że jednostki korzystają ze statycznej i przygotowanej części Internetu, gdzie pełnią rolę konsumentów (31).

Badania oparte na popycie w infodemiologii powstawały głównie na początkowym etapie rozwoju Internetu, który charakteryzował się statycznymi stronami internetowymi oraz treściami tworzonymi przez ograniczoną liczbę autorów, przy dominującej roli użytkowników jako konsumentów tychże informacji (32,33). Obecnie, przykładem takich badań może być identyfikacja dezinformacji i fałszywych informacji na temat szczepień (34) lub pandemii COVID-19 (35). Jednak ten typ badań napotyka szereg wyzwań. Po pierwsze, filtracja i interpretacja ogromnych ilości danych stanowią znaczącą przeszkodę. Po drugie, rozróżnienie między informacjami wysokiej jakości a fałszywymi wymaga solidnych metod analitycznych i zasobów merytorycznych. Wreszcie, większość tych badań ma charakter przekrojowy, co oznacza, że analizują one jedynie wycinek sytuacji w konkretnym momencie, przy ograniczonej zdolności do śledzenia zmian w czasie lub ustalania przyczynowości (19,32).

Z drugiej strony, rozwijane są badania oparte na treściach publikowanych przez użytkowników Internetu, znane jako badania oparte na podaży. W przeciwieństwie do badań opartych na popycie, gdzie treść jest bardziej statyczna i przygotowana, badania te skupiają się na treści o dynamicznej i interaktywnej naturze. Dynamiczny rozwój mediów społecznościowych oraz

highlighting the temporal limitations inherent in this type of research (19,32).

The sheer volume of online content, especially during health crises, means that researchers in infodemiology face significant challenges in manually analyzing and verifying the entirety of internet data (6,22). To address this, infodemiology leverages tools like Google Trends, which offers real-time access to data, allowing for the analysis of emerging trends. With its extensive database, Google, the most popular search engine, provides comprehensive geographic and demographic insights, enabling comparisons of search queries across different regions and times. Comparative analysis and the discovery and verification of research topics further enhance the utility of such tools in infodemiology (24). Additionally, sentiment analysis, or the assessment of the emotional tone of online text, whether positive, negative, or neutral, serves as a critical mechanism. Known also as emotion AI or opinion mining, sentiment analysis helps evaluate the author's attitude towards specific topics, proving invaluable in analyzing user-generated content (38,39).

FUTURE OF INFODEMOLOGY

As people increasingly turn to the internet, its role in our lives is set to expand significantly. This digital engagement offers a unique window into human life through the data made available online. Infodemiology, as a field, is poised for substantial growth, facing numerous opportunities and challenges ahead. Enhanced data quality and sophisticated algorithms are expected to improve the prediction of health threats and enable real-time monitoring of epidemiological events. This progress will also facilitate more effective implementation of social campaigns and more informed allocation of budgets. However, challenges loom large. With the internet becoming more crowded with data, it's becoming increasingly difficult for decision-makers, healthcare professionals, and patients to find high-quality and trustworthy information. The spread of misinformation, disinformation and malinformation online is rampant. Additionally, the digital landscape is inundated with content generated by language models, making interactions with non-human entities more common. These issues are indicative of the internet's transition into the Web 3.0 era, presenting both risks and opportunities for infodemiology in navigating this evolving digital ecosystem (40).

CONCLUSION

The unprecedented surge of content on the internet, notably during health crises, has substantially challenged the processes of manual verification and

społeczności online aktywnie z nich korzystającej rozszerzyły możliwości badań infodemiologicznych opartych na podaży (31). Ten rodzaj badań często obejmuje analizę wpisów na Twitterze (X), Facebooku, forach i blogach. Ich główne cele to identyfikacja bieżących trendów oraz luk w dostępnych online informacjach zdrowotnych. Dążą również do zrozumienia wpływu treści generowanych przez użytkowników na percepcję problemów zdrowotnych. Badania strony podażowej mają również na celu wsparcie w projektowaniu skuteczniejszych interwencji i kampanii zdrowotnych, odpowiadających na rzeczywiste potrzeby społeczności internetowych (19).

Badania oparte na podaży w infodemiologii zagłębiają się w treści dostarczane przez użytkowników na platformach online, aby zrozumieć udostępniane przez nich informacje zdrowotne i ich wpływ na dyskurs publiczny (19). Przykładem takich badań są analizy komentarzy od przeciwników szczepień (36) czy próby identyfikacji czynników ryzyka samobójstwa poprzez śledzenie postów na Twitterze (37). Skupiają się one na analizie i weryfikacji wiarygodności treści generowanych przez użytkowników, rozumieniu kontekstu i intencji stojących za udostępnianymi informacjami oraz szukaniu równowagi między wolnością słowa a koniecznością dostarczania dokładnych i wiarygodnych informacji zdrowotnych. Podobnie jak badania oparte na popycie, badania oparte na podaży również prezentują obraz sytuacji w danym momencie, co decyduje o czasowych ograniczeniach tego typu badań (19,32).

Ogromna ilość treści online, szczególnie podczas kryzysów zdrowotnych, sprawia, że badacze w dziedzinie infodemiologii napotykają znaczne wyzwania w ręcznej analizie i weryfikacji całokształtu danych internetowych (6,22). W odpowiedzi na to, infodemiologia wykorzystuje narzędzia takie jak Google Trends, które oferuje dostęp do danych w czasie rzeczywistym oraz analizę pojawiających się trendów. Dzięki obszernej bazie danych, Google, najpopularniejsza wyszukiwarka, dostarcza wszechstronnych analiz geograficznych i demograficznych, umożliwiając porównania zapytań wyszukiwania w różnych regionach i czasach. Porównawcza analiza oraz odkrywanie i weryfikacja tematów badań dodatkowo zwiększają użyteczność takich narzędzi w infodemiologii (24). Dodatkowo, analiza sentymentu, czyli ocena emocjonalnego tonu tekstu online, niezależnie od tego, czy jest pozytywny, negatywny, czy neutralny, pełni kluczową rolę. Zna również jako AI emocji czy głęboka analiza opinii, analiza sentymentu pomaga ocenić stosunek autora do konkretnych tematów, okazując się nieocenioną w analizie treści generowanych przez użytkowników (38,39).

analysis. This scenario has been encapsulated by the World Health Organization (WHO) as an “infodemic,” defined by an overabundance of information – both accurate and inaccurate – that complicates responses to health emergencies (22). The COVID-19 pandemic exemplified this phenomenon, with the information on the virus expanding exponentially; by November 2020, the global publication of over 75,000 scientific papers related to COVID-19, with a new paper appearing every three minutes, rendered manual verification by researchers an impractical endeavor (41).

In addressing the complexities of internet use and its capacity to generate vast data volumes, this article underscores the essential role of infodemiology in the analysis and management of such information for public health optimization. Amidst the swift digitalization of society and the ensuing challenges brought forth by the infodemic – particularly pronounced during the COVID-19 pandemic – infodemiology has proven to be an indispensable interdisciplinary field (22). Infodemiology gives an opportunity to efficiently processes and interprets the surfeit of data, providing critical insights that can profoundly influence public health policies and interventions (21). Through meticulous examination of both demand and supply-based studies, infodemiology not only endeavors to narrow the information-misinformation divide but also aims to utilize the expansive digital realm to enhance global health outcomes.

REFERENCES

1. Główny Urząd Statystyczny. Społeczeństwo informacyjne w Polsce w 2023 r. 2023.
2. Social Media. Ilu Polaków korzysta z mediów społecznościowych? [cited 2024 Jan 25]. Available from: <https://www.ciekawostatystyki.pl/2021/08/ilu-polakow-korzysta-z-social-media.html>
3. Młodzi Polacy spędzają każdego dnia kilka godzin przed smartfonem - TELEPOLIS.PL. [cited 2024 Jan 25]. Available from: <https://www.telepolis.pl/wiadomosci/prawo-finanse-statystyki/mlodzi-polacy-korzystanie-ze-smartfonow-kilka-godzin-dziennie>
4. Eurostat. Digital economy and society statistics – households and individuals – Statistics Explained. 2023 [cited 2024 Jan 25]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals_#Internet_access_of_individuals.2C_2010_and_2023
5. The International Telecommunication Union. ICT facts and figures. 2017 [cited 2024 Jan 25]. Available from: <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf>

PRZYSZŁOŚĆ INFODEMIOLOGII

W miarę jak ludzie coraz częściej korzystają z Internetu, jego rola w naszym życiu znacznie się rozszerza. To cyfrowe zaangażowanie użytkowników oferuje unikalną perspektywę postrzegania ludzkiego życia poprzez dane dostępne online. Infodemiologia, jako dziedzina, jest gotowa do znacznego rozwoju, stojąc przed licznymi możliwościami i wyzwaniem. Oczekuje się, że lepsza jakość danych i zaawansowane algorytmy poprawią przewidywanie zagrożeń zdrowotnych oraz umożliwią monitorowanie wydarzeń epidemiologicznych w czasie rzeczywistym. Postęp ten będzie również sprzyjać skuteczniejszemu wdrażaniu kampanii społecznych oraz bardziej świadomemu alokowaniu budżetów. Jednak obecne są także poważne wyzwania. W miarę, jak Internet staje się coraz bardziej zatłoczony danymi, decydentom, profesjonalistom ochrony zdrowia i pacjentom coraz trudniej jest znaleźć informacje wysokiej jakości i godne zaufania. Rozpowszechnianie się informacji nieprawdziwych/błędnych, dezinformacji i informacji złośliwych online jest coraz powszechniejsze. Dodatkowo, krajobraz cyfrowy jest zalewany treściami generowanymi przez modele językowe AI, co sprawia, że interakcje z botami stają się coraz częstsze. Kwestie te wskazują na przejście Internetu w erę Web 3.0, co oznacza zarówno ryzyka, jak i możliwości dla infodemiologii w nawigowaniu po tym ewoluującym ekosystemie cyfrowym (40).

WNIOSKI

Bezprecedensowy wzrost ilości treści internetowych, zwłaszcza w czasie kryzysów zdrowotnych, stał się ogromnym wyzwaniem dla manualnej weryfikacji i analizy danych. Światowa Organizacja Zdrowia (WHO) opisała ten problem jako „infodemię”, zdefiniowaną przez nadmiar informacji – zarówno prawdziwych, jak i fałszywych – które utrudniają reagowanie na zagrożenia zdrowotne (22). Pandemia COVID-19 stanowiła przykład tego zjawiska – informacje o wirusie rozprzestrzeniły się w tempie wykładniczym, do listopada 2020 roku globalna liczba publikacji naukowych dotyczących COVID-19 przekroczyła 75 000, z nową pracą pojawiającą się co trzy minuty, co uczyniło manualną weryfikację przez badaczy praktycznie niemożliwą (41).

W odpowiedzi na złożoność wykorzystania Internetu i jego zdolność do generowania ogromnych ilości danych, niniejszy artykuł przedstawia kluczową rolę infodemiologii w analizie i zarządzaniu informacjami pojawiającymi się w Internecie w celu optymalizacji zdrowia publicznego. W obliczu szybkiego postępu cyfryzacji społeczeństwa i wynikających z tego

6. Ren C, Deng Z, Hong Z, Zhang W. Health information in the digital age: an empirical study of the perceived benefits and costs of seeking and using health information from online sources. *Health Info Libr J.* 2019 Jun 1 [cited 2024 Jan 25];36(2):153–67. Available from: <https://pubmed.ncbi.nlm.nih.gov/30737878/>
 7. Kim SM, Oh JY. Health information acquisition online and its influence on intention to visit a medical institution offline. *Information Research.* 2011[cited 2024 Jan 25];16(4). Available from: <https://khu.elsevierpure.com/en/publications/health-information-acquisition-online-and-its-influence-on-intent>
 8. OnAudience.com. Global Data Market Size 2017-2019. 2018.
 9. Polityka prywatności – Prywatność i warunki – Google. [cited 2024 Mar 4]. Available from: <https://policies.google.com/privacy?hl=pl>
 10. 2020 Trends in Personalization.
 11. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the Internet. *AMIA Annual Symposium Proceedings.* 2003 [cited 2024 Jan 25];2003:225. Available from: [/pmc/articles/PMC1480194/](https://pubmed.ncbi.nlm.nih.gov/1480194/)
 12. Prawie 63% Polaków szuka sposobu leczenia w Internecie. [cited 2024 Jan 25]. Available from: <https://www.biostat.com.pl/prawie-63-procent-polakow-szuka-sposobu-leczenia-w-internecie.php>
 13. Brady E, Segar J, Sanders C. “You get to know the people and whether they’re talking sense or not”: Negotiating trust on health-related forums. *Soc Sci Med.* 2016 Aug 1 [cited 2024 Jan 25];162:151. Available from: [/pmc/articles/PMC4962769/](https://pubmed.ncbi.nlm.nih.gov/2769769/)
 14. Chan JKY, Farrer LM, Gulliver A, Bennett K, Griffiths KM. University Students’ Views on the Perceived Benefits and Drawbacks of Seeking Help for Mental Health Problems on the Internet: A Qualitative Study. *JMIR Hum Factors.* 2016 Jan 1 [cited 2024 Jan 25];3(1). Available from: [/pmc/articles/PMC4797689/](https://pubmed.ncbi.nlm.nih.gov/27689/)
 15. Islam MS, Sarkar T, Khan SH, Kamal AHM, Murshid Hasan SM, Kabir A, et al. COVID-19-Related infodemic and its impact on public health: A global social media analysis. *American Journal of Tropical Medicine and Hygiene.* 2020 Oct 1 [cited 2024 Jan 25];103(4):1621–9. Available from: <https://www.bbc.com/news/world-53755067>
 16. Chater N. Facing up to the uncertainties of COVID-19. *Nat Hum Behav.* 2020 May 1 [cited 2024 Jan 25];4(5):439. Available from: <https://pubmed.ncbi.nlm.nih.gov/32221513/>
 17. Eysenbach G. Infodemiology: Tracking Flu-Related Searches on the Web for Syndromic Surveillance. *wyzwań, zwłaszcza widocznych podczas pandemii COVID-19, infodemiologia okazała się niezbędną dziedziną interdyscyplinarną (22). Umożliwia ona efektywne przetwarzanie i interpretowanie nadmiaru danych, dostarczając pogłębionego zrozumienia, które może zasadniczo wpłynąć na politykę zdrowotną i interwencje (21). Poprzez szczegółowe badania zarówno oparte na popycie, jak i podaży, infodemiologia nie tylko dąży do zniwelowania przepaści między informacją a dezinformacją, ale także ma na celu wykorzystanie rozległego cyfrowego świata w celu poprawy globalnych wyników zdrowotnych.*
-
- AMIA Annual Symposium Proceedings. 2006 [cited 2024 Jan 31];2006:244. Available from: [/pmc/articles/PMC1839505/](https://pubmed.ncbi.nlm.nih.gov/1839505/)
 18. Wang FY, Zeng D, Carley KM, Mao W. Social Computing. *IEEE Intell Syst.* 2007 Mar 1 [cited 2024 Jan 31];22(2):79–83. Available from: <https://dl.acm.org/doi/10.1109/MIS.2007.41>
 19. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. *Health Info Libr J.* 2018 Jun 1 [cited 2024 Jan 31];35(2):91–120. Available from: <https://pubmed.ncbi.nlm.nih.gov/29729073/>
 20. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res.* 2009 Jan 1 [cited 2024 Jan 31];11(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/19329408/>
 21. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med.* 2011 [cited 2024 Jan 31];40(5 Suppl 2). Available from: <https://pubmed.ncbi.nlm.nih.gov/21521589/>
 22. World Health Organization. 1st WHO Infodemiology Conference. 2020.
 23. Kamel Boulos MN, Sanfilippo AP, Corley CD, Wheeler S. Social Web mining and exploitation for serious applications: Technosocial Predictive Analytics and related technologies for public health, environmental and national security surveillance. *Comput Methods Programs Biomed.* 2010 Oct [cited 2024 Jan 31];100(1):16–23. Available from: <https://pubmed.ncbi.nlm.nih.gov/20236725/>
 24. Santillana M, Zhang DW, Althouse BM, Ayers JW. What can digital disease detection learn from (an external revision to) Google Flu Trends? *Am J Prev Med.* 2014 [cited 2024 Jan 31];47(3):341–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/24997572/>
 25. Guy S, Ratzki-Leewing A, Bahati R, Gwady-Sridhar F. Social media: A systematic review to understand the evidence and application in infodemiology. *Lecture Notes of the Institute*

- for Computer Sciences, Social-Informatics and Telecommunications Engineering. 2012 [cited 2024 Jan 31];91 LNICST:1–8. Available from: https://link.springer.com/chapter/10.1007/978-3-642-29262-0_1
26. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? *JAMA*. 2014 Apr 9 [cited 2024 Jan 31];311(14):1399–400. Available from: <https://pubmed.ncbi.nlm.nih.gov/24577162/>
 27. Bexelius C, Larserics Digital Print). *E-epidemiology : adapting epidemiological methods for the 21st century*. 2009 Sep 18 [cited 2024 Jan 31]; Available from: <http://openarchive.ki.se/xmlui/handle/10616/39273>
 28. Zielinski C. Infodemics and infodemiology: a short history, a long future. *Rev Panam Salud Publica*. 2021 [cited 2024 Feb 1];45. Available from: <https://pubmed.ncbi.nlm.nih.gov/33995517/>
 29. Eysenbach G. Infodemiology: The epidemiology of (mis)information. *American Journal of Medicine*. 2002 Dec 15 [cited 2024 Feb 1];113(9):763–5. Available from: <http://www.amjmed.com/article/S0002934302014730/fulltext>
 30. Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. *J Health Commun*. 2010 [cited 2024 Mar 4];15 Suppl 3(SUPPL. 3):279–93. Available from: <https://pubmed.ncbi.nlm.nih.gov/21154099/>
 31. Atkinson RK, Sabo K, Conley Q. THE PARTICIPATORY WEB. *Inf Commun Soc*. 2012 Aug [cited 2024 Mar 4];91–120. Available from: <https://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.665935>
 32. MAvragani A. Infodemiology and Infoveillance: Scoping Review. *J Med Internet Res*. 2020 Apr 1 [cited 2024 Mar 5];22(4). Available from: <https://pubmed.ncbi.nlm.nih.gov/33995517/>
 33. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. Vol. 35, *Health Information and Libraries Journal*. Blackwell Publishing Ltd; 2018. p. 91–120.
 34. Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. *Vaccine*. 2010 Feb 17 [cited 2024 Mar 4];28(7):1709–16. Available from: <https://pubmed.ncbi.nlm.nih.gov/20045099/>
 35. Jarynowski A, Wójta-Kempa M, Belik V. Trends in interest of COVID-19 on Polish Internet. *Przegl Epidemiol*. 2020;74(2):258–75.
 36. Klimiuk K, Czoska A, Biernacka K, Balwicki Ł. Vaccine misinformation on social media - topic-based content and sentiment analysis of Polish vaccine-deniers' comments on Facebook. *Hum Vaccin Immunother*. 2021 [cited 2022 Feb 1];17(7):2026–35. Available from: <https://pubmed.ncbi.nlm.nih.gov/33517844/>
 37. Jashinsky J, Burton SH. Tracking suicide risk factors through Twitter in the US. *Crisis*. 2013 [cited 2020 Aug 28];35(1):51–9. Available from: <https://www.researchgate.net/publication/257754018>
 38. Analiza sentymentu - dlaczego Sentimenti jest najlepsze? | Sentimenti. [cited 2020 Aug 28]. Available from: <https://sentimenti.pl/blog/analiza-sentymentu-sentimenti-jakie-emocje/>
 39. Singh M, Dhillon HK, Ichhpujani P, Iyengar S, Kaur R. Twitter sentiment analysis for COVID-19 associated mucormycosis. *Indian J Ophthalmol*. 2022 May 1 [cited 2024 Mar 5];70(5):1773–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/35502071/>
 40. Rudman R, Bruwer R. Defining Web 3.0: Opportunities and challenges. *Electronic Library*. 2016 Feb 1;34(1):132–54.
 41. Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. *The BMJ*. 2020 May 12;369.

Received: 16.03.2024

Accepted for publication: 29.04.2024

Otrzymano: 16.03.2024 r.

Zaakceptowano do publikacji: 29.04.2024 r.

Address for correspondence:

Adres do korespondencji:

Krzysztof Bartosz Klimiuk

Zakład Zdrowia Publicznego i Medycyny Społecznej

Gdański Uniwersytet Medyczny

ul. Marii Skłodowskiej-Curie 3a,

80-210, Gdańsk, Polska;

email: krzysztof_klimiuk@gumed.edu.pl