Jak wiele informacji o sobie udostępniamy w Internecie

Paulina Brzęcka 184701 Marek Borzyszkowski 184266

4 kwietnia 2025

SPIS TREŚCI

SPIS TREŚCI 2
1. WSTĘP I CEL RAPORTU 5
1.1. Cel pracy 5
2. ANALIZA PODSTAWOWYCH DANYCH I BUDOWA PROFILU 6
2.1. Zakres i charakter danych osobowych w kontekście budowy profilu 6
2.1.1. Elementy PII i podstawowe dane osobowe 6
2.1.2. Jakie dane najczęściej ujawniamy w Internecie? 6
2.1.3. Konsekwencje łączenia danych z różnych źródeł 7
2.2. Profilowanie: definicje, zastosowania i zagrożenia 7
2.2.1. Techniki profilowania i sposoby wykorzystania 7
2.2.2. Zagrożenia wynikające z profilowania 7
2.2.3. Implikacje dla budowania profilu 8
2.3. Wykorzystanie OSINT w analizie podstawowych danych 9
2.3.1. Kluczowe zasoby i narzędzia 9
2.3.2. Przykłady zastosowania OSINT w budowaniu profilu 9
2.3.3. Znaczenie ogólnodostępnych źródeł 9
2.3.4. Znaczenie OSINT na etapie „podstawowych danych” 9
2.3.5. Podsumowanie 10
2.4. Przykład re-identyfikacji - jak pozornie anonimowe dane mogą ujawniać tożsamość 10
2.4.1. Re-identyfikacja na podstawie danych podstawowych 10
2.4.2. Rola otwartych i powszechnie dostępnych źródeł 11
2.4.3. Konsekwencje dla prywatności i bezpieczeństwa 12
2.4.4. Podsumowanie: wzór dla pozostałych analiz 12
3. ŹRÓDŁA POZYSKIWANIA INFORMACJI 13
3.1. Wprowadzenie - co mówi prawo? 13
3.2. Identyfikacja i analiza publicznie dostępnych źródeł 14
3.2.1. Media społecznościowe 14
3.2.2. Fora, sekcje komentarzy, blogi 16
3.2.3. Rejestry publiczne i portale firmowe 16
3.2.4. Urządzenia IoT 16
3.2.5. Czat AI jako nieświadome źródło informacji 16
3.3. W jaki sposób dane trafiają w niepowołane ręce? 17
4. WYKORZYSTANIE DANYCH PRZEZ RÓŻNE PODMIOTY 19
4.1. Firmy reklamowe i sektor marketingu 19
4.1.1. Personalizacja i targetowanie reklam 19
4.1.2. Wykorzystanie Big Data w analizie zachowań konsumenckich 19
4.1.3. Nowe formaty reklamowe i ekosystem wymiany danych 20
4.2. Branża HR i rekruterzy 20
4.2.1. Analiza śladów w Internecie 20
4.2.2. Weryfikacja ścieżek kariery i referencji 20

4.3. Cyberprzestępcy i hakerzy ... 21
4.3.1. Phishing i spear phishing ... 21
4.3.2. Handel danymi i kradzież tożsamości ... 21
4.3.3. Cybergangi i ataki łańcuchowe ... 21
4.4. Brokerzy danych i inni pośrednicy ... 21
4.4.1. Koncepcja brokera danych ... 21
4.4.2. Wykorzystanie danych przez brokerów w analizie rynkowej ... 22
4.5. Inne podmioty i obszary wykorzystania ... 22
4.5.1. Partnerzy biznesowi i organizacje międzynarodowe ... 22
4.5.2. Sektor naukowo-badawczy ... 22
4.6. Przykłady synergii między podmiotami ... 23
4.6.1. Wspólne cele i nieoczywiste współprace ... 23
4.6.2. Rozwijające się platformy i nowe zastosowania ... 23
4.7. Podsumowanie ... 23
5. ZAGROŻENIA ... 25
5.1. Zagrożenia wynikające z udostępniania danych ... 25
5.1.1. Wprowadzenie ... 25
5.1.2. Kradzież tożsamości ... 25
5.1.3. Naruszenie prywatności i poufności danych ... 26
5.1.4. Cyberstalking i nękanie ... 29
5.1.5. Oszustwa socjotechniczne i internetowe ... 30
5.1.6. Konsekwencje reputacyjne ... 31
5.1.7. Odpowiedzialność prawna ... 32
6. ZALECENIA DOTYCZĄCE PRYWATNOŚCI ... 34
6.1. Czego nie udostępniać w internecie? ... 34
6.1.1. Adres e-mail i numer telefonu ... 34
6.1.2. Adres domowy i lokalizacja geograficzna ... 34
6.1.3. Zdjęcia nieletnich ... 34
6.1.4. Kompromitujące zdjęcia ... 34
6.1.5. Dokumenty osobiste ... 34
6.1.6. Opinia, skargi i kontrowersyjne komentarze ... 34
6.1.7. Prywatne rozmowy ... 34
6.2. Jak się chronić? ... 34
6.2.1. Aktualizuj posiadane oprogramowanie ... 35
6.2.2. Zabezpiecz domową sieć Wi-Fi ... 35
6.2.3. Skonfiguruj ustawienia prywatności ... 35
6.2.4. Używaj silnych, unikalnych haseł ... 35
6.2.5. Korzystaj z uwierzytelniania dwuskładnikowego (2FA) ... 35
6.2.6. Uważaj na podejrzane linki ... 35
6.2.7. Korzystaj z zaufanych źródeł cyberbezpieczeństwa ... 36
6.2.8. Nie publikuj wrażliwych informacji ... 36
6.2.9. Wyłącz geolokalizację ... 36
6.2.10. Zachowaj ostrożność w komentarzach ... 36
6.2.11. Rozważ korzystanie z usług typu VPN ... 36
6.2.12. Unikaj podawania numeru telefonu ... 36

6.2.13.Świadomie analizuj, z jakich usług korzystasz i jakimi danymi płacisz za „darmowy” dostęp 36
6.2.14.Zastrzeż swój numer PESEL 36
6.3. Co zrobić jeśli wycieknąą dane osobowe? 37
6.3.1. Przyczyny wycieków danych 37
6.3.2. Możliwe skutki wycieku danych 37
6.3.3. Jak sprawdzić, czy doszło do wycieku danych? 37
6.3.4. Co zrobić po wycieku danych? 37
7. PODSUMOWANIE 39
WYKAZ LITERATURY 41
WYKAZ RYSUNKÓW 42

1. WSTĘP I CEL RAPORTU

W dobie powszechnej cyfryzacji coraz więcej aspektów naszego życia przenosi się do świata wirtualnego. Codziennie korzystamy z mediów społecznościowych, wyszukiwarek internetowych, aplikacji mobilnych i wielu innych platform, nie zawsze zastanawiając się, jakie informacje o sobie udostępniamy i jakie mogą być tego konsekwencje. Dane, które pozostawiamy w sieci - świadomie lub nieświadomie - mogą obejmować zarówno podstawowe informacje, takie jak imię i nazwisko, wiek czy miejsce zamieszkania, jak i bardziej wrażliwe dane, takie jak zainteresowania, preferencje zakupowe czy historia przeglądania stron. Współczesny Internet sprawia, że budowanie profilu użytkownika na podstawie dostępnych informacji jest niezwykle łatwe, a same dane mogą być wykorzystywane na różne sposoby przez rozmaite podmioty.

Niniejszy raport ma na celu analizę skali i sposobów udostępniania danych w Internecie. W kolejnych rozdziałach omówiona zostanie budowa profilu użytkownika na podstawie podstawowych informacji, źródła pozyskiwania danych oraz sposoby ich wykorzystywania przez firmy, instytucje i inne organizacje. Szczególna uwaga zostanie poświęcona zagrożeniom wynikającym z nadmiernego udostępniania informacji, takim jak kradzież tożsamości, nieuprawnione śledzenie aktywności czy manipulacja preferencjami użytkowników. W końcowej części raportu przedstawione zostaną rekomendacje dotyczące ochrony prywatności i minimalizacji ryzyka związanego z publikowaniem danych w sieci.

1.1. Cel pracy

Celem niniejszego raportu jest zbadanie, w jakim stopniu i w jaki sposób użytkownicy Internetu udostępniają swoje dane oraz jakie niesie to za sobą konsekwencje. Szczególny nacisk zostanie położony na analizę różnych źródeł informacji, mechanizmów zbierania danych oraz podmiotów, które je przetwarzają i wykorzystują. Poprzez dokładne przeanalizowanie tego procesu możliwe będzie wskazanie zagrożeń związanych z nadmiernym udostępnianiem informacji oraz ocena, w jakim stopniu użytkownicy mają nad nim kontrolę.

Drugim, równie istotnym celem raportu, jest przedstawienie praktycznych zaleceń dotyczących ochrony prywatności w Internecie. Omówione zostaną sposoby minimalizowania ilości udostępnianych danych, techniki zabezpieczania informacji oraz narzędzia, które mogą pomóc w ochronie tożsamości cyfrowej. Ostatecznym efektem pracy będzie zwiększenie świadomości czytelników na temat zagrożeń i sposobów zabezpieczenia swoich danych w sieci, co pozwoli im podejmować bardziej świadome decyzje dotyczące własnej prywatności.

2. ANALIZA PODSTAWOWYCH DANYCH I BUDOWA PROFILU

2.1. Zakres i charakter danych osobowych w kontekście budowy profilu

Analizując proces gromadzenia i wykorzystywania informacji, warto najpierw zrozumieć, co w świetle różnorodnych definicji i regulacji uznaje się za *dane osobowe*. W literaturze i dokumentach anglojęzycznych często używa się określenia **Personally Identifiable Information (PII)**[1], które obejmuje wszelkie dane pozwalające na bezpośrednie lub pośrednie zidentyfikowanie konkretnej osoby. W kontekście ustawodawstwa europejskiego termin ten pokrywa się w znacznej mierze z terminem **dane osobowe** w rozumieniu *RODO*, a w praktyce oznacza wszelkie informacje związane z osobą fizyczną, które mogą posłużyć do jej zidentyfikowania.

2.1.1. Elementy PII i podstawowe dane osobowe

Według definicji *U.S. Department of Energy*[1], za PII uznaje się:

- Dane *jednoznacznie* identyfikujące daną osobę: między innymi imię, nazwisko, numer paszportu, numer PESEL (w kontekście polskim) lub inne unikalne identyfikatory.
- Dane, które *pozwalają* na ustalenie tożsamości osoby, nawet jeśli nie ma w nich bezpośrednich identyfikatorów. Może to być data urodzenia, kod pocztowy czy stan cywilny, jeśli te informacje odpowiednio skorelowane z innymi źródłami umożliwiają wskazanie konkretnej osoby.

W polskich realiach **imię i nazwisko** także stanowią dane osobowe, choć w niektórych sytuacjach *same w sobie* mogą nie być wystarczające do jednoznacznej identyfikacji - zależy to od kontekstu. Niemniej jednak, jak podkreślają eksperci[2], imię i nazwisko, w połączeniu z innymi informacjami (takimi jak miejsce zamieszkania czy data urodzenia), zdecydowanie umożliwia powiązanie takich danych z konkretną jednostką.

2.1.2. Jakie dane najczęściej ujawniamy w Internecie?

Według analizy przeprowadzonej na portalu *CyberDefence24*[3], większość internautów regularnie udostępnia:

- **Imię i nazwisko** - choćby na portalach społecznościowych lub forach, gdzie rejestracja wymaga konta opartego na rzeczywistych danych.
- **Adres e-mail** i **numer telefonu** - używane powszechnie do zakładania kont, newsletterów lub w procesach odzyskiwania haseł.
- **Zdjęcia i nagrania wideo** - w wielu przypadkach z metadanymi (np. lokalizacją GPS), które mogą ujawniać dodatkowe informacje.
- **Informacje o miejscu zamieszkania** - często w postach dotyczących codzienności lub jako część ustawień konta (w tym np. kod pocztowy).
- **Zainteresowania, poglądy, aktywności** - w formie polubień stron, członkostwa w grupach tematycznych czy interakcji w komentarzach.

Co istotne, *nie musimy* samodzielnie publikować wszystkich tych danych. Część informacji zostaje ujawniona poprzez aktywność naszych znajomych i bliskich, którzy np. *oznaczają* nas w lokalizacjach

lub na zdjęciach, komentują sytuacje z naszego życia prywatnego czy rodzinnego, a także publikują różnego rodzaju **listy kontaktów**.

2.1.3. Konsekwencje łączenia danych z różnych źródeł

Nawet jeśli pojedyncza dana (np. adres e-mail) nie wydaje się sama w sobie istotna, to można ją bez trudu *zmapować* z kolejnymi rejestrami, w których podawane są imię, nazwisko czy numer telefonu. Stąd już tylko krok do poznania rozbudowanej historii aktywności danej osoby w Internecie. Teoretycznie drobne fakty składają się na **profil cyfrowy**:

- Dokładniejszy wgląd w preferencje, poglądy i preferowane aktywności.
- Możliwość wyśledzenia relacji rodzinnych i towarzyskich.
- Analiza trendów w zachowaniu i przewidywanie przyszłych aktywności.

W poprzednich częściach raportu wskazywaliśmy przypadki re-identyfikacji, w których kluczowe były właśnie tzw. *quasi-identyfikatory* (data urodzenia, miejsce zamieszkania, płeć). Ich połączenie z innymi zasobami - często ogólnodostępnymi - prowadzi do *deanonimizacji* osób, które uważały się za bezpieczne za sprawą samego ograniczenia ilości ujawnianych danych.

2.2. Profilowanie: definicje, zastosowania i zagrożenia

Wraz ze wzrostem ilości informacji dostępnych w Internecie, coraz więcej podmiotów, zarówno komercyjnych, jak i publicznych, decyduje się na **profilowanie danych osobowych**. Jak wskazuje opracowanie PARP[4], profilowanie to zautomatyzowany proces przetwarzania danych osobowych, służący do oceny określonych czynników dotyczących osoby fizycznej, w szczególności do analizy lub prognozy aspektów związanych z wynikami w pracy, sytuacją ekonomiczną, zdrowiem, preferencjami osobistymi czy zainteresowaniami.

2.2.1. Techniki profilowania i sposoby wykorzystania

Profilowanie jest możliwe przede wszystkim dzięki:

- **Zaawansowanej analityce danych (Data Analytics)** - wykorzystuje się algorytmy uczenia maszynowego i sztucznej inteligencji, które na podstawie dużych zbiorów informacji są w stanie kategoryzować użytkowników i przewidywać ich zachowania.
- **Danym z różnych platform** - integracja danych z serwisów społecznościowych, systemów bankowych, firm ubezpieczeniowych czy sklepów internetowych pozwala na bogate wnioski dotyczące preferencji i ryzyka.
- **Plikom cookies i innym formom śledzenia** - znaczniki i identyfikatory internetowe (np. adres IP, identyfikatory urządzeń mobilnych) umożliwiają śledzenie aktywności na wielu stronach i aplikacjach, budując pełen obraz nawyków użytkowników.

W efekcie firmy mogą np. *personalizować* oferty handlowe, reklamy, rekomendacje produktowe czy decyzje kredytowe. Z jednej strony zwiększa to wygodę użytkownika (propozycja usług najlepiej pasujących do jego potrzeb), z drugiej - *rodzi wiele pytań* o kontrolę nad tym, jakie informacje i w jaki sposób są wykorzystywane.

2.2.2. Zagrożenia wynikające z profilowania

Choć profilowanie niesie potencjalne korzyści, wiąże się też z licznymi zagrożeniami:

1. **Brak transparentności** - użytkownik często nie ma świadomości, że jest profilowany ani w jakim zakresie dane na jego temat są gromadzone.
2. **Ryzyko dyskryminacji** - przy przyznawaniu kredytów, składek ubezpieczeniowych czy ofert pracy systemy oparte na algorytmach mogą *wykluczać* określone grupy osób.
3. **Nieodwracalna identyfikacja** - gdy dane raz zostaną połączone z tożsamością użytkownika, trudno przywrócić wcześniejszy stan anonimowości, zwłaszcza jeśli informacje zostały skopiowane do wielu niezależnych baz.
4. **Luka w ochronie prywatności** - różne systemy prawne (np. w USA, w Unii Europejskiej czy w innych regionach) w odmienny sposób regulują zasady przetwarzania danych. W rezultacie przechwycenie danych z jednego kraju i przetworzenie ich w innym może *osłabiać* gwarancje prywatności.

W kontekście przepisów RODO (Rozporządzenie Ogólne o Ochronie Danych Osobowych), każda osoba ma prawo do informacji o profilowaniu oraz, w pewnych przypadkach, do *wniesienia sprzeciwu* wobec takiego przetwarzania. Rzeczywiste zastosowanie tych uprawnień bywa jednak utrudnione, zwłaszcza gdy zakres i sposób pozyskiwania danych nie jest przejrzysty.

2.2.3. Implikacje dla budowania profilu

Mechanizmy profilowania znacząco przyspieszają proces *budowania zaawansowanego profilu*, o czym traktuje niniejszy rozdział raportu. Systemy oparte na uczeniu maszynowym są w stanie:

- **Powiązać różne źródła** danych (np. z mediów społecznościowych, rejestrów genealogicznych, analiz geolokalizacyjnych) w jeden spójny model.
- **Wnioskować** o cechach i zachowaniach, które nie zostały bezpośrednio udostępnione przez użytkownika (np. stan zdrowia, skłonność do podróży, zainteresowanie konkretną dziedziną).
- **Generować prognozy i klasyfikacje**, które nie zawsze są zgodne z rzeczywistością, ale w wystarczająco wielu przypadkach pozwalają na skuteczne podejmowanie decyzji handlowych, politycznych czy społecznych.

W połączeniu z **re-identyfikacją**, omawianą w poprzednich sekcjach, profilowanie tworzy środowisko, w którym pojedynczy użytkownik traci kontrolę nad własnym wizerunkiem i prywatnością w przestrzeni cyfrowej. Z tego względu rośnie znaczenie edukacji społecznej na temat świadomości cyfrowej oraz *podstawowych praktyk* związanych z ochroną danych.

Wnioski i perspektywy Omawiane w tej sekcji zjawiska ilustrują, jak budowanie profilu użytkownika *wykracza* daleko poza analizę podstawowych danych osobowych. Dodatkowe techniki i narzędzia (takie jak profile behawioralne, narzędzia Big Data czy rozpoznawanie wzorców w aktywnościach online) **poszerzają** zakres i dokładność takiej analizy. Jednocześnie zmienia się status prawny i społeczny użytkowników, którzy muszą mierzyć się z *możliwą dyskryminacją* lub *utrudnieniem* dostępu do usług, jeśli wynik profilowania okaże się dla nich niekorzystny.

W kolejnych rozdziałach przyjrzymy się dokładniej mechanizmom i strategiom minimalizowania ryzyka związanego z profilowaniem oraz re-identyfikacją. Zastanowimy się też, jak w praktyce *ograniczyć* nadmierne udostępnianie danych, aby nasza aktywność w Internecie nie stała się źródłem niekontrolowanego przepływu informacji o naszej tożsamości, zachowaniach i preferencjach.

2.3. Wykorzystanie OSINT w analizie podstawowych danych

Istotnym elementem pozyskiwania informacji, pozwalającym na *ulepszenie* profilu osoby już na wczesnym etapie analizy, jest **OSINT (Open Source Intelligence)**. Polega on na wykorzystywaniu otwartych i publicznie dostępnych źródeł danych, co w znacznym stopniu ułatwia skorelowanie nawet *podstawowych faktów* (np. daty urodzenia, miejsca zamieszkania, informacji genealogicznych) z dodatkowymi szczegółami na temat konkretnej jednostki.

2.3.1. Kluczowe zasoby i narzędzia

Istnieje szereg zróżnicowanych repozytoriów i narzędzi wspierających OSINT, z czego najpopularniejsze to:

- *OSINT Framework*[5] - interaktywna „mapa" serwisów i aplikacji służących do wyszukiwania osób, adresów e-mail, numerów telefonów czy aktywności w mediach społecznościowych.
- *IntelTechniques*[6] - obszerna kolekcja linków i skryptów, ułatwiających przeszukiwanie konkretnych platform i rejestrów pod kątem danych o osobie.

2.3.2. Przykłady zastosowania OSINT w budowaniu profilu

W kontekście **analizy podstawowych danych i tworzenia profilu** jednostki, OSINT wspomaga:

- **Wyszukiwanie danych kontaktowych i danych demograficznych** (m.in. imię, nazwisko, wiek, płeć, miejsce zamieszkania), które często można odnaleźć na publicznych profilach w mediach społecznościowych lub w rejestrach online.
- **Powiązywanie istniejących informacji** - np. data i miejsce urodzenia *zestawione* z rejestrami genealogicznymi, co pozwala szybciej zidentyfikować członków rodziny oraz kluczowe fakty biograficzne danej osoby.
- **Pogłębianie wiedzy** o aktywnościach internetowych (dawne posty, wzmianki w forach, archiwach stron), co umożliwia uzupełnienie profilu o dodatkowe szczegóły, takie jak zainteresowania lub przybliżony stan majątkowy.

2.3.3. Znaczenie ogólnodostępnych źródeł

Wspomniane narzędzia bazują głównie na **powszechnie dostępnych danych**, do których należą m.in.:

- **Serwisy genealogiczne** (np. FamilySearch, WikiTree), zawierające dane o dacie i miejscu urodzenia czy historii rodziny.
- **Rejestry i bazy urzędowe** (krajowe i międzynarodowe) - od podstawowych statystyk demograficznych po niektóre listy adresowe.
- **Platformy społecznościowe**, gdzie użytkownicy nierzadko sami dzielą się informacjami, *ułatwiając* tym samym ustalenie ich tożsamości oraz najważniejszych cech profilu.

2.3.4. Znaczenie OSINT na etapie „podstawowych danych"

W ramach **pierwszego etapu** analizy, gdy punktem wyjścia są *proste* informacje (wiek, płeć, region zamieszkania), OSINT pomaga w:

- **Weryfikacji autentyczności** - potwierdzeniu, czy dane elementy (np. data urodzenia) są spójne z innymi źródłami.

- **Szybszym zawężeniu obszaru poszukiwań** - przeszukiwanie publicznych zasobów z użyciem narzędzi OSINT pozwala od razu odrzucić dane niewłaściwe i skupić się na informacjach naprawdę użytecznych w budowie profilu.
- **Korelowaniu różnych identyfikatorów** - przykładowo, adres e-mail zarejestrowany w jednym serwisie może być powiązany z kontem w innej platformie, co otwiera dostęp do kolejnych danych.

2.3.5. Podsumowanie

OSINT stanowi **skuteczne wsparcie** w procesie analizy podstawowych danych i **budowania profilu** osoby, uwidaczniając, jak wiele szczegółów można ustalić *bez* zastosowania wyspecjalizowanych, zamkniętych baz czy usług komercyjnych. Otwarta struktura Internetu oraz rosnąca liczba publicznie dostępnych rejestrów sprawiają, że umiejętne korzystanie z OSINT stanowi nieodłączny element każdego przedsięwzięcia, w którym celem jest zebranie jak najbardziej szczegółowych i zweryfikowanych informacji na temat danej osoby.

2.4. Przykład re-identyfikacji - jak pozornie anonimowe dane mogą ujawniać tożsamość

W ramach rozdziału warto przyjrzeć się przykładowi opisującemu, jak dane uznawane za anonimowe mogą zostać użyte do ponownej identyfikacji konkretnych osób. Przypadek opisany przez Adama Tannera w magazynie „Forbes” [7]

odnosi się do sytuacji, w której badacz z Uniwersytetu Harvarda połączył próbki DNA rzekomo „anonimowych” ochotników z publicznie dostępnymi danymi genealogicznymi. Dzięki temu zdołał przypisać je do konkretnych osób.

2.4.1. Re-identyfikacja na podstawie danych podstawowych

W omawianym przykładzie kluczową rolę odgrywały wyłącznie *podstawowe informacje*, takie jak:

- Dane demograficzne (np. rok urodzenia, płeć, stan cywilny).
- Częściowe dane genealogiczne (np. nazwiska przodków, zarys drzewa rodowego).
- Ogólnodostępne rejestry i bazy, gdzie można zestawić powyższe informacje (np. rejestry ludności, dokumenty archiwalne).

Okazało się, że nawet przy anonimowych próbkach DNA można użyć **drobnych, publicznie dostępnych faktów** do precyzyjnego ustalenia tożsamości, jeśli tylko połączy się ze sobą odpowiednie elementy informacyjne pochodzące z wielu źródeł. Analiza ta nie wymagała specjalistycznych czy drogich rozwiązań; wystarczyły ogólnodostępne bazy genealogiczne oraz archiwa, które może przeszukiwać każdy internauta.

Znaczenie tzw. *quasi-identyfikatorów* Kluczowe w procesie re-identyfikacji okazują się dane, które *nie zawierają* jednoznacznych oznaczeń personalnych (np. imię i nazwisko wprost), ale w połączeniu z innymi źródłami tworzą wyraźny profil jednostki. Takie dane często nazywa się **quasi-identyfikatorami**. Należą do nich m.in.:

- data urodzenia (w szczególności dzień i miesiąc),
- kod pocztowy lub nazwa miejscowości,

- płeć i wiek,
- stan cywilny czy wykształcenie.

Choć z osobna nie stanowią kompletnego zestawu, to kiedy zestawi się je z innymi publicznymi rejestrami (np. spisy powszechne czy listy mieszkańców), zwykle szybko można zawęzić krąg osób do kilku lub wręcz jednej.

Przykłady łączenia danych Praktyka pokazuje, że re-identyfikacja bywa możliwa nawet w sytuacjach, gdy dane *pozornie* zostały zanonimizowane. Oprócz badań genetycznych, podobne zagrożenia występują w innych dziedzinach:

- **Informacje medyczne:** w przeszłości dowiedziono, że dane medyczne (np. rekordy ze szpitali) można powiązać z osobami, jeśli połączy się je z danymi demograficznymi z rejestrów urzędowych.
- **Logi aktywności internetowej:** w przypadku wycieków logów wyszukiwarki internetowej, nawet częściowo zanonimizowane identyfikatory internautów okazywały się wystarczające do wydobycia szczegółów o ich życiu prywatnym (np. miejsca zamieszkania, zainteresowań, problemów zdrowotnych).
- **Bazy genealogiczne:** gdy ktoś upublicznia fragment drzewa rodzinnego (nazwiska przodków, daty narodzin i zgonów), badacz może *dosztukować* brakujące elementy, korzystając z rejestrów lub innych źródeł, w efekcie dopasowując profil do konkretnej osoby.

Dynamika zmian w dostępnych źródłach Nie należy zapominać, że dostępność danych w Internecie **nieustannie się poszerza**. Dzięki digitalizacji i masowemu udostępnianiu zasobów (np. projektom bibliotek cyfrowych czy otwieraniu kolejnych baz urzędowych), dzisiejsza „bezpieczna" porcja informacji może jutro okazać się niewystarczająca do zachowania prywatności. Wraz z pojawianiem się nowych zbiorów - na przykład szczegółowych map geolokalizacyjnych, kolejnych edycji spisów powszechnych czy społecznościowych inicjatyw genealogicznych - *również historyczne dane zyskują nowy kontekst* i mogą prowadzić do ujawnienia wrażliwych szczegółów.

Podsumowanie zagadnienia re-identyfikacji Zjawisko re-identyfikacji opartej na podstawowych danych unaocznia, że proces anonimizacji powinien uwzględniać coraz to nowsze i łatwiej dostępne źródła informacji. Jeśli nawet kilka lat temu uznano jakiś zbiór za bezpieczny (bo odarte z niego dane uznano za „nieszkodliwe"), to w obecnej rzeczywistości, gdzie poszczególne rejestry można swobodnie łączyć, **tylko szczególnie rygorystyczne metody anonimizacji** dają realną gwarancję ochrony prywatności. W praktyce zaś większość danych „anonimowych" może zostać *deanonimizowana*, o ile dysponuje się wystarczającymi zasobami i wiedzą do ich korelowania z ogólnodostępnymi bazami.

2.4.2. Rola otwartych i powszechnie dostępnych źródeł

Zasadniczy problem ujawniony w tej historii polega na tym, że spora część naszych danych znajduje się w **otwartych i łatwo dostępnych** zasobach internetu. Dotyczy to m.in.:

1. **Publicznych baz rządowych i organizacyjnych**:
    - W Polsce są to choćby `https://dane.gov.pl/` czy `https://bdl.stat.gov.pl/` (Bank Danych Lokalnych GUS).
    - W USA bogatym źródłem statystyk jest `https://data.census.gov/`, a na poziomie europejskim - `https://data.europa.eu/`.

- Informacje o zameldowaniach, urodzeniach czy zgonach znajdują się w różnorodnych spisach i rejestrach regionalnych.

2. **Serwisów genealogicznych i historycznych**:
   - Popularne platformy (np. `https://www.familysearch.org/`, `https://www.wikitree.com/`) gromadzą nazwiska, daty, a czasem nawet dokumenty rodzinne, co umożliwia odtworzenie fragmentów drzew genealogicznych.
   - W niektórych krajach udostępniane są rejestry małżeństw, akt urodzeń i zgonów, które - choć ograniczone prawem - często dają się przeglądać bez większych formalności.
3. **Mediów społecznościowych i forów tematycznych**:
   - Nawet jeśli użytkownicy nie udostępniają publicznie wszystkich danych, to często dzielą się szczegółami (np. datą urodzenia, miejscem zamieszkania, relacjami rodzinnymi) w otwartych grupach czy w komentarzach.
   - Analiza np. zdjęć może zdradzić lokalizację (dzięki metadanym GPS lub charakterystycznym elementom w tle).

Dzięki wymienionym zasobom, osoba mająca wystarczającą motywację i podstawową wiedzę o metodach wyszukiwania jest w stanie zebrać rozmaite informacje o konkretnej jednostce. Początkowo są to drobne fakty (wiek, miejsce pochodzenia), jednak po złożeniu wszystkich elementów w całość otrzymujemy **szczegółowy profil** konkretnej osoby.

2.4.3. Konsekwencje dla prywatności i bezpieczeństwa

Opisany przykład wskazuje, że **anonimizacja** danych to często proces wyłącznie pozorny, zwłaszcza gdy nie uwzględnia się rosnącej liczby publicznie dostępnych zbiorów, z którymi można je zestawiać. Dla ochrony prywatności istotne jest, aby:

1. **Świadomie udostępniać informacje**: Zastanowić się, czy publikacja konkretnego faktu (np. dokładnej daty urodzenia) jest faktycznie niezbędna w danym miejscu (profil społecznościowy, forum internetowe, ankieta itp.).
2. **Rozpoznawać ryzyko łączenia danych**: Nawet dane wrażliwe (np. zdrowotne czy genetyczne) mogą zostać dopasowane do osoby, jeśli w innej bazie lub w mediach społecznościowych widnieje choć część informacji „uzupełniających” (data urodzenia, nazwisko panieńskie matki, nazwa miejscowości).
3. **Szukać mechanizmów ochronnych**: Sprawdzać, czy wybrana platforma oferuje wystarczającą kontrolę nad prywatnością i przetwarzaniem danych. Zgłębiać też przepisy takie jak RODO (w Unii Europejskiej) czy HIPAA (w Stanach Zjednoczonych) - regulacje te ustalają m.in. obowiązki podmiotów przetwarzających dane.

2.4.4. Podsumowanie: wzór dla pozostałych analiz

Omówiony przypadek z artykułu z „Forbes” ukazuje zjawisko, które dotyczy nie tylko badań genetycznych, ale praktycznie każdej sytuacji, w której istnieje potencjał do **połączenia rzekomo anonimowych informacji z innymi zbiorami**. To właśnie w taki sposób, krok po kroku, tworzy się i uszczegóławia profil danej osoby.

3. ŹRÓDŁA POZYSKIWANIA INFORMACJI

3.1. Wprowadzenie - co mówi prawo?

Informacje o osobach mogą być przechowywane w publicznych źródłach, takich jak Internet, media społecznościowe czy rejestry publiczne. Często sami użytkownicy publikują te dane. Jednak fakt, że dane są publiczne, nie oznacza, że można je wykorzystywać w dowolny sposób.

Umieszczenie danych osobowych w otwartych źródłach, np. na publicznych stronach internetowych, jest uznawane za przetwarzanie danych osobowych. Oznacza to, że zastosowanie ma Rozporządzenie o Ochronie Danych Osobowych (RODO), a więc konieczne jest posiadanie podstawy prawnej do takiego przetwarzania.

Wyjątkiem są cele osobiste i domowe. RODO nie obowiązuje, jeśli dane są przetwarzane wyłącznie w celach osobistych lub domowych. Można umieścić własne dane osobowe w otwartych źródłach - nie trzeba mieć ku temu podstawy prawnej. Każda osoba może dowolnie dysponować własnymi danymi osobowymi.

W przypadku zamiaru umieszczenia danych osobowych innych osób w otwartych źródłach:

- należy posiadać podstawę prawną przetwarzania danych,
- należy posiadać konkretny i uzasadniony cel,
- należy spełniać pozostałe wymagania RODO.

Obowiązki te należy spełnić **przed** opublikowaniem danych, aby uniknąć naruszenia przepisów RODO.

Dodatkowo obowiązuje zakaz przetwarzania danych szczególnej kategorii (np. dane zdrowotne, poglądy polityczne) oraz danych karnych. Można je przetwarzać jedynie wtedy, gdy zastosowanie znajduje wyjątek od tego zakazu. Publikując dane, należy samodzielnie ocenić, czy:

- dane należą do kategorii szczególnych lub karnych,
- można powołać się na wyjątek od zakazu ich przetwarzania.

Dane osobowe, które zgodnie z RODO zostały legalnie opublikowane w otwartym źródle, można wykorzystywać wyłącznie do użytku osobistego lub domowego (np. prywatny kalendarz urodzin, książka adresowa). W takich przypadkach RODO nie ma zastosowania.

Niedozwolone jest wykorzystywanie tych danych w celach zawodowych lub komercyjnych, jak również ich udostępnianie poza najbliższym gronem (np. rodziną, przyjaciółmi).

Dane mogą być legalnie dostępne w otwartym źródle, ale to nie oznacza, że wolno je automatycznie przetwarzać ponownie.

Zgoda osoby, której dane dotyczą, jest często niewystarczająca jako podstawa do ponownego przetwarzania — zwłaszcza w przypadku przetwarzania na większą skalę (np. gromadzenie danych do baz).

Każde nowe przetwarzanie danych wymaga sprawdzenia, czy konieczne jest przeprowadzenie **Oceny Skutków dla Ochrony Danych** (ang. DPIA - Data Protection Impact Assessment).

DPIA pozwala:

- zidentyfikować ryzyka dla prywatności,
- podjąć środki zaradcze jeszcze przed rozpoczęciem przetwarzania.

Nawet jeśli DPIA nie jest obowiązkowa, jej przeprowadzenie jest zalecane.

Organizacje, które automatycznie przeszukują Internet w celu gromadzenia danych osobowych, również są podmiotami przetwarzającymi dane i muszą przestrzegać przepisów RODO. Oznacza to, że muszą:

- wykazać istnienie podstawy prawnej,
- stosować zasadę minimalizacji danych,
- przetwarzać tylko dane niezbędne do realizacji określonego celu.

Dla danych szczególnych i karnych obowiązuje zakaz przetwarzania. Nawet jeśli dane te zostały legalnie opublikowane, nie można ich ponownie przetwarzać — **chyba, że** poza podstawą prawną istnieje również uzasadniony wyjątek od zakazu przetwarzania.

To, że dane są publicznie dostępne (lub zostały opublikowane przez samą osobę), nie oznacza, że można je bez ograniczeń ponownie wykorzystywać. Każde nowe przetwarzanie danych wymaga osobnej analizy zgodności z przepisami RODO [8].

Chociaż wszelkie źródła informacji muszą być prowadzone z uwzględnieniem praw osób, których dane dotyczą i w granicach wyznaczonych przez obowiązujące przepisy prawa, to w praktyce użytkownik jest i tak narażony na szereg zagrożeń.

3.2. Identyfikacja i analiza publicznie dostępnych źródeł

W dobie powszechnego dostępu do Internetu oraz ogromnej popularności mediów społecznościowych, zdobycie informacji na temat osoby prywatnej, przedsiębiorcy czy pracownika jest dziś łatwiejsze niż kiedykolwiek wcześniej. Użytkownicy często sami - świadomie lub nieświadomie - pozostawiają po sobie szereg danych, które można wykorzystać do stworzenia precyzyjnego profilu.

Według raportu „Digital 2021" (We Are Social i Hootsuite), z Internetu w Polsce korzysta 31,97 mln osób, a 25,9 mln aktywnie używa mediów społecznościowych. To oznacza, że większość społeczeństwa codziennie generuje cyfrowe ślady, które mogą być później analizowane.[9]

3.2.1. Media społecznościowe

Facebook, Instagram, X, LinkedIn czy Snapchat gromadzą dane zarówno świadomie podawane przez użytkowników (np. miejsce pracy), jak i te zbierane automatycznie - lokalizacja, adres IP, urządzenia. Dane mogą pochodzić także z aplikacji firm trzecich, programów lojalnościowych czy partnerów marketingowych.[10]

Szczegółowe zestawienie, jakie dane są gromadzone w mediach społecznościowych zostały przedstawione w poniższej tabeli. [11]

Zakres danych / Działanie	**Facebook**	**X**	**LinkedIn**
DANE OSOBOWE			
Imię i nazwisko	+	+	+
E-mail	+	+	+
Numer telefonu	+	+	+
Data urodzenia	+	+	+
Adres zamieszkania	+	+	+
Poprzednie miejsca zamieszkania	+		
Zdjęcia profilowe	+	+	+
Rodzina i związki	+		

Wykształcenie	+		+
Języki obce	+		+
Poglądy polityczne	+		
Przekonania religijne	+		
Wydarzenia z życia	+		
DANE ZAWODOWE I LOKALIZACYJNE			
Miejsce pracy	+		+
Kwalifikacje	+		+
Wynagrodzenie			+
Płatności	+		+
Lokalizacja	+	+	+
Urządzenia	+	+	+
Adres IP	+	+	+
Kalendarz			+
ŹRÓDŁA POZYSKIWANIA INFORMACJI			
Od użytkownika	+	+	+
Od innych użytkowników	+	+	+
Od partnerów zewnętrznych	+		+
Z aplikacji zewnętrznych	+	+	
Z aktywności (kliknięcia)	+	+	
ZASTOSOWANIA I UDOSTĘPNIANIE			
Reklamy	+	+	+
Personalizacja	+	+	+
Oznaczanie na zdjęciach	+		
Meldowanie w lokalizacjach	+	+	
Sugestie kontaktów	+		+
Udostępnianie firmom	+	+	+
Cele badawcze i analityczne	+	+	+

Jesli chodzi o Facebook, to gromadzi on najszerszy zakres danych - zarówno prywatnych, jak i zawodowych, w tym dane kontaktowe, lokalizacyjne, relacyjne i behawioralne. Dane są również pozyskiwane od partnerów i aplikacji zewnętrznych. Warto zaznaczyć, że domyślne ustawienia prywatności, których większość użytkowników nigdy dokładnie nie czyta przed akceptacją, bardzo często sprzyjają upublicznieniu informacji, a około 40% użytkowników nie zmienia nigdy ustawień prywatności, co czyni ich dane łatwo dostępnymi.

W aplikacji X domyślnie wszystkie posty (tweety) są publiczne. Dodatkowo dane gromadzone również poprzez partnerów i aplikacje zewnętrzne, a polityka firmy jest dosyć liberalna i udostępnia dane badaczom.

Firma LinkedIn dodatkowo profiluje użytkowników głównie pod kątem danych zawodowych (miejsce pracy, kwalifikacje, języki), które potem są wykorzystywane komercyjnie i sprzedawane firmom rekrutacyjnym. Może być to szczególnie niebezpieczne przy atakach typu spear-phishing, gdzie precyzyjne informacje są używane do podszywania się i oszustw.

3.2.2. Fora, sekcje komentarzy, blogi

Uznawane za „anonimowe", w rzeczywistości łatwe do deanonimizacji. Powtarzające się pseudonimy, zdjęcia profilowe, adresy IP - wszystko to pozwala powiązać konto z konkretną osobą, po mniej lub bardziej wnikliwym poszukiwaniu.

3.2.3. Rejestry publiczne i portale firmowe

Z racji powszechnej cyfryzacji, przed udostępnieniem niektórych danych w internecie nie można się uchronić. Szczególnie osoby prowadzące firmy lub działalności gospodarcze, a nawet osoby posiadające nieruchomości. Nie wspominając o tym, że prawie każdy dorosły człowiek w dzisiejszych czasach pracuje w jakiejś firmie, gdzie często dane kontaktowe, adresy e-mail, a czasem także numery telefonów pracowników czy właścicieli firm są udostępniane na stronie firm, bardzo często nawet bez wiedzy samych zainteresowanych.[12]
Niektóre dane są dostępne w postaci rejestrów internetowych:

- **CEIDG** - zawiera dane o jednoosobowej działalności gospodarczej, w tym adres, który nierzadko jest miejscem zamieszkania.
- **KRS** - dane zarządu, często z numerem PESEL, umożliwia pozyskanie numeru PESEL i daty urodzenia członków zarządu spółek.
- **CRBR** - udostępnia dane beneficjentów rzeczywistych, w tym numery PESEL udziałowców spółek.
- **Elektroniczne Księgi Wieczyste** - umożliwiają dotarcie do danych właścicieli nieruchomości.
- **Portale ogłoszeniowe, firmowe** - e-maile, telefony, nazwiska

3.2.4. Urządzenia IoT

Inteligentne urządzenia domowe (głośniki, żarówki, lodówki) zbierają dane o rutynie, lokalizacji i obecności użytkownika. Coraz więcej urządzeń domowych (np. tostery, żarówki, głośniki) jest wyposażanych w moduły internetowe. Nawet jeśli nie oferują realnych funkcji online, dane z nich są zbierane i mogą być wykorzystywane marketingowo. Koszt wdrożenia łączności IoT dla producenta jest niewielki, a dane mają wysoką wartość.

> „Kupując sprzęt domowy, możemy nieświadomie nabyć urządzenie IoT. Kluczowym zasobem są dane - które mogą zostać sprzedane lub wykorzystane przez producenta." - Mikko Hypponen

3.2.5. Czat AI jako nieświadome źródło informacji

Coraz więcej osób korzysta z narzędzi opartych na sztucznej inteligencji (AI), takich jak ChatGPT czy Bing Chat, aby wspierać się w pracy zawodowej. Według badań aż 43% pracowników używa AI do realizacji swoich zadań.[13] Choć AI oferuje wygodne i szybkie odpowiedzi, istnieją zagrożenia dotyczące prywatności i bezpieczeństwa danych, o których warto pamiętać, że:

- dane mogą być przetwarzane lub analizowane przez ludzi,
- platformy zastrzegają prawo do przechowywania zapytań,
- AI modele, takie jak ChatGPT, mogą uczyć się na podstawie zadawanych pytań. Dlatego nie powinno się wprowadzać do nich żadnych danych osobowych ani poufnych.

Co zdarza się ludziom udostępnić, a nie powinno się tego robić:

- Imię i nazwisko,
- Data urodzenia,
- Adres zamieszkania,
- Numer dowodu osobistego,
- Telefon, adres e-mail,
- Kod źródłowy lub dane z projektów w fazie rozwoju,
- Poufne informacje o produktach i usługach,
- Niezaprezentowane jeszcze pomysły lub rozwiązania,
- Numery kart kredytowych,
- Numery kont bankowych,
- Hasła i dane logowania,
- Wyniki badań,
- Diagnozy lekarskie,
- Informacje dotyczące leczenia.

Aby korzystać z AI w sposób odpowiedzialny i bezpieczny, warto stosować się do kilku prostych zasad:

- Używanie jedynie informacji ogólnodostępnych i nieidentyfikujących,
- Podawanie dane fikcyjne lub zanonimizowane, jeśli konieczne jest podanie przykładu.
- Nie publikowanie szczegółowych planów podróży lub wakacji — może to narażać na kradzież lub inne ryzyko offline.
- Założenie drugiego, anonimowego konta, które nie jest powiązane z właściwą tożsamością.
- Korzystanie z bezpiecznego połączenia internetowego — unikanie otwartych sieci Wi-Fi w kawiarniach czy hotelach.
- Nie udostępnianie haseł ani nie proszenie AI o ich generowanie na potrzeby konkretnej usługi.
- Zachowanie ostrożności przy danych zdrowotnych, finansowych i prywatnych,
- Zapoznanie się z regulaminem i polityką prywatności chatu AI,
- Sprawdzenie, jakie dane są przechowywane, jak długo i komu mogą być udostępniane,
- Zwrócenie uwagi, czy narzędzie przewiduje możliwość przeglądu zapytań przez człowieka w celach jakościowych.

3.3. W jaki sposób dane trafiają w niepowołane ręce?

Chociaż skoro myślimy, że udostepniamy bardzo mało danych, bądź wcale, istnieją aktualnie techniki, które są na tyle zaawansowane, że na podstawie jednej, bądź kilku informacji można stworzyć czyiś profil osobowy. Do jednej z nich zalicza się biały wywiad. Open Source Intelligence, cyli biały wywiad, to wyspecjalizowana technika polegająca na analizie dostępnych publicznie informacji w celu identyfikacji i gromadzenia danych o wybranej osobie.

- przeszukiwanie stron internetowych pod kątem danych kontaktowych,
- wyszukiwanie kont społecznościowych na podstawie adresu e-mail lub loginu,

- analizę zdjęć w celu rozpoznania wizerunku i odnalezienia powiązanych treści,
- korelowanie fragmentarycznych danych z różnych źródeł (np. fora, blogi, komentarze).

Aktualne narzędzie umożliwiają bardzo wnikliwą analizę, między innymi:

- automatyczne przeszukiwanie domen internetowych w celu odnalezienia adresów e-mail,
- sprawdzanie, na jakich portalach społecznościowych zarejestrowano konkretne nazwy użytkowników,
- analizę zdjęć za pomocą sztucznej inteligencji w celu identyfikacji osoby.

Chociaż prawda jest taka, że nie potrzeba żadnych profesjonalnych narzędzi, aby komuś uprzykrzyć życie. Dostęp do danych wrażliwych może mieć każda osoba prywatna, wystarczy tylko odrobina negatywnej motywacji oraz chwila czasu spędzona na poszukiwaniu i wnioskowaniu, ale o tym wiecej w nastepnych rozdziałach.

4. WYKORZYSTANIE DANYCH PRZEZ RÓŻNE PODMIOTY

Współcześnie informacje gromadzone na temat użytkowników Internetu stały się **strategicznym zasobem** w wielu branżach. Dane te mogą być wykorzystywane zarówno w legalnych i powszechnie akceptowanych procesach biznesowych, jak i w działaniach sprzecznych z prawem oraz zasadami etyki. W niniejszym rozdziale omówiono główne kierunki wykorzystania danych przez różnorodne podmioty, ze szczególnym uwzględnieniem sektora reklamowego, firm rekrutacyjnych oraz środowiska cyberprzestępców.

4.1. Firmy reklamowe i sektor marketingu

4.1.1. Personalizacja i targetowanie reklam

Dynamiczny rozwój platform cyfrowych umożliwił firmom reklamowym gromadzenie i przetwarzanie ogromnych ilości danych, takich jak:

- *historia przeglądania stron internetowych*,
- *aktywność w mediach społecznościowych*,
- *dane transakcyjne* (np. preferencje zakupowe, częstotliwość dokonywania zakupów online).

Jak podkreśla raport *Behind the One-Way Mirror* opublikowany przez EFF[14], tak szeroki zakres informacji pozwala na **targetowanie behawioralne**, w którym reklamy są dopasowywane do konkretnych grup odbiorców, uwzględniając ich indywidualne cechy i preferencje.

W praktyce oznacza to, że użytkownikowi, który często wyszukuje informacje o zdrowym stylu życia, mogą się wyświetlać reklamy związane z dietą i aktywnością fizyczną. Natomiast osoby zainteresowane planowaniem podróży będą otrzymywać propozycje ofert turystycznych. Mechanizm ten, choć zwiększa *trafność* komunikatów marketingowych, jednocześnie - jak wskazują liczne analizy rynkowe - **podnosi** skuteczność kampanii reklamowych i skłania użytkowników do częstszych interakcji z treściami sponsorowanymi.

4.1.2. Wykorzystanie Big Data w analizie zachowań konsumenckich

Zgodnie z danymi statystycznymi przywoływanymi w raporcie *Digital 2025 Global Overview Report*[15], liczba użytkowników Internetu stale rośnie, a firmy reklamowe i agencje marketingowe mogą dzięki temu gromadzić coraz bardziej rozbudowane zasoby Big Data. Informacje o ścieżkach zakupowych czy upodobaniach pozwalają:

- **segmentować** odbiorców według wielu kryteriów (np. styl życia, wiek, dochód),
- **przewidywać** zachowania (np. kto najchętniej dokona zakupu w danej kategorii produktów),
- **monitorować** skuteczność kampanii w czasie rzeczywistym.

Takie dane służą również do tworzenia *dynamiki cenowej* (tzw. *price discrimination*), gdzie użytkownicy otrzymują zindywidualizowane oferty w zależności od historii wyszukiwań i wcześniejszych interakcji z reklamami. Dzięki temu możliwe jest zwiększanie zysków firm, ale pojawiają się też wątpliwości etyczne co do stopnia ingerencji w prywatność konsumentów.

4.1.3. Nowe formaty reklamowe i ekosystem wymiany danych

Raporty dotyczące rynku cyfrowego wskazują, że w obszarze marketingu zachodzi *przesunięcie* z tradycyjnych form reklamy (banery, wyskakujące okienka) w stronę treści natywnych oraz tzw. **mikrotargetowania**. W tym procesie kluczową rolę pełnią:

- **wielkie platformy internetowe** (np. Google, Facebook, TikTok), które gromadzą i scalają dane o aktywnościach użytkowników,
- **agencje analityczne** oferujące narzędzia do profilowania i śledzenia zachowań online,
- **marketerzy i działy komunikacji** w firmach, które poszukują maksymalnego zwrotu z inwestycji w kampanie reklamowe.

W efekcie cały ekosystem *nieustannie* się rozrasta, a dane o preferencjach konsumenckich krążą między różnymi platformami i stają się przedmiotem intensywnej wymiany.

4.2. Branża HR i rekruterzy

4.2.1. Analiza śladów w Internecie

W rekrutacji zawodowej dane osobowe oraz aktywność kandydatów w sieci mogą przesądzać o *ocenie* ich przydatności na określone stanowiska. Jak wynika z lokalnych analiz przywoływanych w *Digital 2025 Local Country Headlines*[16], wzrasta liczba przedsiębiorstw regularnie korzystających z portali zawodowych i mediów społecznościowych w procesach HR. Rekruterzy:

- sprawdzają **wiarygodność** profili kandydatów,
- weryfikują doświadczenie zawodowe na podstawie informacji publicznych,
- analizują **styl komunikacji** i **zainteresowania**, co może mieć znaczenie w ocenie dopasowania do danej kultury organizacyjnej.

Dodatkowo rekruterzy często zwracają uwagę na tak zwany *e-wizerunek* kandydata: sposób, w jaki wyraża się on w sieci, jakie treści publikuje i z kim wchodzi w interakcje. Nawet informacje, które kandydat uznaje za nieistotne lub prywatne, mogą rzutować na finalną decyzję o zatrudnieniu.

4.2.2. Weryfikacja ścieżek kariery i referencji

Duże korporacje i agencje rekrutacyjne coraz częściej sięgają po zaawansowane narzędzia analityczne, aby automatycznie oceniać *historię zatrudnienia* kandydata. Dane z portali zawodowych (np. LinkedIn) konfrontuje się z rejestrami branżowymi, co pozwala:

- wykryć *nieścisłości* w CV,
- zweryfikować *rzetelność* deklarowanych umiejętności,
- ustalić ewentualne *powiązania* z kluczowymi organizacjami.

Choć taka praktyka bywa postrzegana jako kontrowersyjna w kontekście prywatności, wiele firm utrzymuje, że *szersza analiza* danych usprawnia proces rekrutacji oraz ogranicza ryzyko zatrudnienia osób niedopasowanych do zespołu bądź wprowadzających pracodawcę w błąd.

4.3. Cyberprzestępcy i hakerzy

4.3.1. Phishing i spear phishing

Według *ENISA Threat Landscape 2022*[17] cyberprzestępcy, korzystając z wykradzionych lub publicznie dostępnych informacji, przeprowadzają **phishing** na skalę masową, a także **spear phishing** - spersonalizowane kampanie wymierzone w konkretne osoby lub grupy. Mechanizm ataku obejmuje:

1. Zbieranie łatwo dostępnych danych o ofierze (stanowisko, profil zawodowy, kontakty).
2. Przygotowanie *wiarygodnej* wiadomości e-mail lub wiadomości w mediach społecznościowych, często stylizowanej na korespondencję od zaufanego nadawcy.
3. Skłonienie ofiary do **kliknięcia w złośliwy link** lub **otwarcia załącznika** zawierającego malware.

Jak zauważa *APWG Trends Report Q4 2022*[18], wzrastająca skuteczność inżynierii społecznej sprawia, że phishing jest wciąż jednym z **najefektywniejszych** sposobów na przejęcie wrażliwych danych.

4.3.2. Handel danymi i kradzież tożsamości

Z raportu *CERT Polska 2023*[19] wynika, że istotnym problemem jest **handel skradzionymi danymi**, obejmującymi loginy, hasła, dane osobowe oraz informacje finansowe. Cyberprzestępcy, wchodząc w posiadanie baz danych użytkowników:

- dokonują *kradzieży tożsamości* (np. zakładając kredyty w imieniu poszkodowanych),
- przejmują *konta w mediach społecznościowych* czy skrzynki e-mail,
- odsprzedają dane na *czarnym rynku* (tzw. **darknet markets**), co dalej napędza rozwój kolejnych ataków.

Przestępcy mogą również wykorzystywać dane w formule **double extortion**, żądając okupu nie tylko za *odszyfrowanie* zasobów, lecz także za *nieujawnienie* przechwyconych informacji.

4.3.3. Cybergangi i ataki łańcuchowe

W ostatnich latach coraz częściej obserwuje się zjawisko *cybergangów*, które profesjonalizują swoje działania, tworząc **wielopoziomowe łańcuchy** ataków. Jak podkreśla *ENISA Threat Landscape 2022*, procesy te obejmują m.in.:

- **rekonesans** (zbieranie danych o potencjalnych celach z publicznych źródeł),
- **fazę infiltracji** (phishing, spear phishing, przejęcie kont uprzywilejowanych),
- **rozprzestrzenienie** w infrastrukturze ofiary (instalacja złośliwych narzędzi, eskalacja uprawnień).

Dane wykradzione w tych kolejnych etapach mają *wartość* zarówno dla samej grupy przestępczej, jak i w szerszym obiegu handlu informacjami.

4.4. Brokerzy danych i inni pośrednicy

4.4.1. Koncepcja brokera danych

Poza bezpośrednimi użytkownikami informacji (firmami reklamowymi, rekruterami czy cyberprzestępcami), na rynku funkcjonuje **segment brokerów danych**. Są to podmioty specjalizujące się w:

- *gromadzeniu* danych z wielu źródeł (zarówno publicznych, jak i komercyjnych),

- *scalaniu* ich w rozbudowane profile zawierające szereg informacji demograficznych, behawioralnych czy finansowych,
- *odsprzedaży* tych baz firmom bądź organizacjom poszukującym kompleksowych zestawień na temat użytkowników bądź klientów.

Mechanizm ten opisują w swoich analizach zarówno *Behind the One-Way Mirror* (EFF), jak i *Digital 2025 Global Overview Report*, wskazując, że obrót danymi w ramach usług brokerskich jest często *niewidoczny* dla przeciętnego internauty.

4.4.2. Wykorzystanie danych przez brokerów w analizie rynkowej

Firmy zajmujące się **marketingiem**, **ubezpieczeniami** czy **rekrutacją** często kupują zbiory danych od brokerów, chcąc uzyskać:

- *rozszerzony zestaw informacji* (np. z różnych segmentów działalności użytkownika),
- *dane uzupełniające* pozwalające trafniej ocenić kandydata bądź klienta (finanse, historia zakupowa, styl życia),
- *opracowane rankingi* i predykcje (np. scoring kredytowy, analiza ryzyka).

Zyskiem dla brokerów jest rola pośrednika, który **konsoliduje** i **wzbogaca** dane, następnie oferując je zainteresowanym stronom. W rezultacie powstają wyjątkowo *szczegółowe* profile użytkowników, obejmujące dane z rozmaitych etapów ich życia w sieci.

4.5. Inne podmioty i obszary wykorzystania

4.5.1. Partnerzy biznesowi i organizacje międzynarodowe

Wiele międzynarodowych instytucji finansowych czy organizacji gospodarczych podejmuje starania w kierunku *standaryzacji* oraz *wymiany* danych o użytkownikach. Celem jest:

- **prognozowanie** trendów rynkowych w skali globalnej,
- **dostosowanie** usług finansowych do lokalnych potrzeb (np. mikrokredyty w regionach rozwijających się),
- **monitorowanie** zmian w zachowaniach konsumenckich w różnych państwach.

Tego rodzaju inicjatywy nie są wprost zorientowane na dane ściśle *personalne* - często chodzi o *zagregowane* statystyki. Niemniej jednak, skala i szczegółowość tych raportów nierzadko pozwala na bardzo wnikliwe wnioski o grupach użytkowników.

4.5.2. Sektor naukowo-badawczy

Część ośrodków naukowych i think-tanków wykorzystuje dane pochodzące z sieci w projektach badawczych, skupiając się na analizach statystycznych czy tworzeniu narzędzi służących zwiększeniu bezpieczeństwa. Jak wskazują analizy cytowane w *ENISA Threat Landscape 2022*, rosnące znaczenie cyberbezpieczeństwa motywuje do gromadzenia i badania *zagregowanych* informacji o incydentach, lukach w zabezpieczeniach czy metodach ataku. Choć dane te są *zazwyczaj* zanonimizowane, sam fakt, że można je precyzyjnie opisywać i analizować, świadczy o bogactwie informacji pojawiających się w przestrzeni publicznej.

4.6. Przykłady synergii między podmiotami

4.6.1. Wspólne cele i nieoczywiste współprace

Warto podkreślić, iż **różne** rodzaje organizacji (np. korporacje reklamowe, brokerzy danych, a nawet cyberprzestępcy) niejako *korzystają* z tych samych zasobów informacyjnych - choć ich cele są diametralnie różne. Przykładowo:

- **Agencje marketingowe** dążą do poprawy skuteczności kampanii, dokonując analizy zachowań konsumentów.
- **Cyberprzestępcy** wykorzystują zbliżone narzędzia (np. narzędzia do śledzenia ruchu) do przygotowania *spreparowanych* komunikatów.
- **Brokerzy danych** dostarczają materiały i firmom, i osobom wykorzystującym dane w mniej etyczny sposób (o ile nie istnieją przepisy lub mechanizmy wewnętrzne ograniczające zbywanie informacji do pewnych sektorów).

Wspólnym mianownikiem jest tu *wartość* płynąca z wiedzy o indywidualnych preferencjach i zachowaniach internautów.

4.6.2. Rozwijające się platformy i nowe zastosowania

Dodatkowym czynnikiem jest **rozszerzanie** platform interaktywnych, takich jak:

- *Media społecznościowe nowej generacji* (z elementami rozszerzonej rzeczywistości),
- *Serwisy streamingowe*, w których gromadzi się szczegółowe dane o gustach i aktywnościach użytkowników (np. co, kiedy i jak często oglądają),
- *Aplikacje do komunikacji* integrujące chatboty i analizę tekstu w czasie rzeczywistym.

Im więcej czasu internauci spędzają w takich środowiskach, tym bardziej *rosną* zbiory danych, a wraz z nimi możliwość jeszcze dokładniejszego profilowania przez przeróżne podmioty.

4.7. Podsumowanie

Przedstawione przykłady wyraźnie pokazują, że **dane** - zwłaszcza te generowane masowo i opatrzone metadanymi - *stanowią* kluczowy zasób, którego znaczenie stale rośnie. Firmy reklamowe oraz marketingowe używają ich do precyzyjnego **targetowania** treści i zwiększania sprzedaży, branża HR ocenia potencjalnych pracowników w oparciu o publicznie dostępne informacje o karierze i aktywności w sieci, zaś cyberprzestępcy nie ustają w wysiłkach, by wykradać dane i **monetyzować** je poprzez phishing czy handel skradzionymi tożsamościami.

Rola brokerów danych dodatkowo *zagęszcza* krajobraz, gdyż to oni scalają rozproszone fragmenty informacji, tworząc *obszerne* i *wszechstronne* profile użytkowników. W obliczu rosnącej liczby internautów - jak dokumentują raporty *DataReportal* - rośnie także dostępność i różnorodność danych, które mogą krążyć w obiegu. Niektóre z tych praktyk leżą na granicy zgodności z oczekiwaniami społecznymi, inne naruszają prawo, zaś jeszcze inne służą zupełnie legalnym i potrzebnym analizom.

W kontekście *wykorzystania danych* trudno dzisiaj mówić o jasnym podziale na „dobre” i „złe” cele, ponieważ te same informacje bywają używane zarówno do kreowania wygodniejszych usług (np. rekomendacje zakupowe), jak i do przeprowadzania skuteczniejszych ataków na użytkowników. Wartym podkreślenia pozostaje jednak fakt, że **coraz więcej** stron jest zainteresowanych pozyskiwaniem,

analizowaniem i *udostępnianiem* danych - co potwierdzają zarówno *ENISA Threat Landscape 2022*, *CERT Polska 2023*, *APWG Trends Report Q4 2022*, jak i liczne publikacje *EFF* oraz *DataReportal*.

5. ZAGROŻENIA

5.1. Zagrożenia wynikające z udostępniania danych

5.1.1. Wprowadzenie

Współczesne organizacje i użytkownicy indywidualni coraz częściej stają się ofiarami zagrożeń wynikających z niekontrolowanego udostępniania danych. Informacje osobiste mogą zostać wykorzystane zarówno przez zewnętrznych atakujących, jak i osoby z wewnątrz firmy — obecnych lub byłych pracowników, kontrahentów, a nawet przypadkowych użytkowników popełniających błędy.
Najczęstsze błędy popełniane przez użytkowników to:

- publikowanie selfie z oznaczeniem daty i lokalizacji,
- udostępnianie postów o planowanych lub trwających wyjazdach wakacyjnych,
- włączona na stałe geolokalizacja w telefonie
- stosowanie haseł łatwych do odgadnięcia, np. oparte na imionach członków rodziny lub zwierząt, dat urodzin i braku zmiany hasła co jakiś czas,
- hejtowanie, umieszczanie kontrowersyjnych treści w internecie.

Przestępcy mogą wykorzystać te dane do okradzenia mieszkania. Firmy windykacyjne mogą szybciej ustalić lokalizację użytkownika. Złodzieje mogą śledzić nawyki i rutynę użytkownika z pomocą danych GPS.
Z kolei stworzenie listy kombinacji haseł na podstawie udostępnianych w internecie informacji nie wymaga specjalistycznej wiedzy i pozwala na złamanie hasła w krótkim czasie. Ujawnianie danych osobistych i zwyczajów może prowadzić do łatwego odgadnięcia haseł do kont społecznościowych, bankowości elektronicznej, poczty, czy chmur obliczeniowych. Może to skutkować utratą danych osobistych i prywatnych zasobów, a także podszywanie się pod użytkownika, aby móc wyłudzić informacje lub pieniądze.
Kontrowersyjne treści publikowane w internecie mogą skutkować:

- Problemami zawodowymi - kontrowersyjne komentarze mogą dotrzeć do pracodawcy.
- Pozwem cywilnym - za naruszenie dóbr osobistych innych osób.
- Utratą reputacji - zarówno prywatnej, jak i zawodowej.

W następnej sekcji przedstawiono opis konkretnych zagrożeń wraz z przykładami z życia codziennego. [20, 21]

5.1.2. Kradzież tożsamości

Opis zagrożenia

Udostępnienie danych osobowych, takich jak imię, nazwisko, PESEL, numer dowodu czy adres e-mail, może skutkować:

- podszywaniem się pod ofiarę w celu zaciągnięcia pożyczki, otwarcia konta bankowego czy przeprowadzenia transakcji,
- wykorzystaniem danych w oszustwach ubezpieczeniowych, zdrowotnych lub podatkowych,

- długoterminowymi problemami finansowymi i prawnymi ofiary.

Przykłady z życia

W ataku na Marriott w 2020 r. wyciekło ponad 5 mln rekordów gości, zawierających dane kontaktowe, urodziny i numery kont lojalnościowych.

5 maja 2020 roku w internecie błędnie zidentyfikowano CEO firmy Propine, Tuhinę Singh, jako kobietę z Singapuru, która została aresztowana po tym, jak w viralowym nagraniu odmówiła założenia maseczki w miejscu publicznym. W wyniku pomyłki internauci zaczęli publikować jej zdjęcia, numer telefonu, prywatny adres e-mail oraz dane jej współpracowników, co doprowadziło do fali gróźb i rasistowskich obelg. Skala nagonki była na tyle duża, że firma Propine wydała oficjalne oświadczenie, w którym wyjaśniła, że doszło do pomyłki, a prawdziwa sprawczyni została już zidentyfikowana i odpowiednio ukarana przez władze.

5.1.3. Naruszenie prywatności i poufności danych

Opis zagrożenia

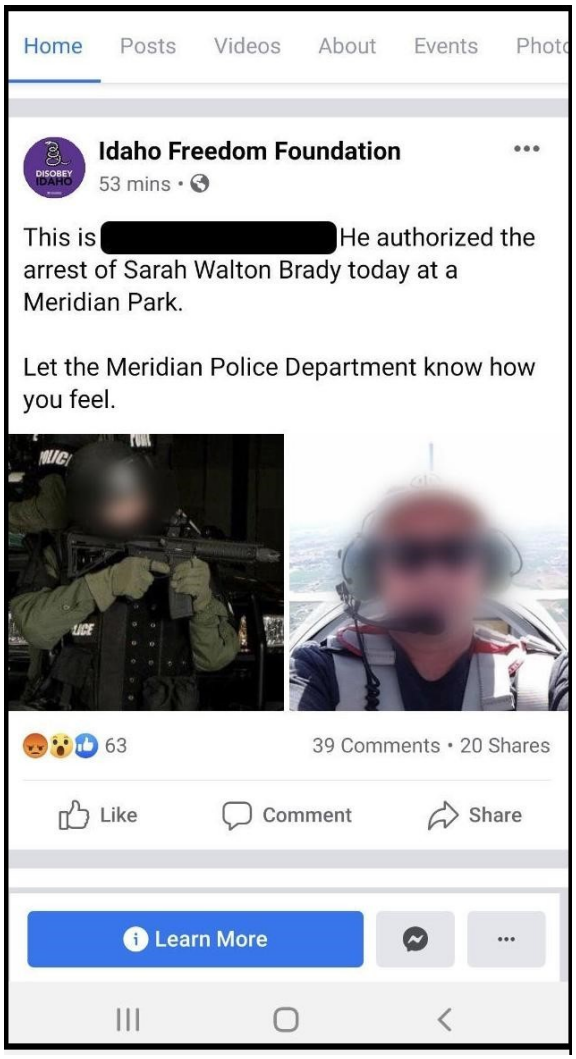
Dane, które zostaną ujawnione - nawet nieświadomie - mogą naruszyć prywatność pracowników, klientów czy partnerów biznesowych. Skutki to:

- utrata kontroli nad własnymi informacjami,
- możliwość ich dalszego rozpowszechniania,
- narażenie na nieautoryzowane monitorowanie.

Przykłady z życia

Zdeanonimizowanie konta dyrektora FBI na Twitterze na podstawie danych o obserwujących i członkach rodziny
Zidentyfikowanie tajnych baz wojskowych i osób w nich przebywających na podstawie map ciepła portalu Strava
Były pracownik Tesli w 2023 r. ujawnił dane ponad 75 000 pracowników, w tym numery SSN i informacje finansowe.

W kwietniu 2020 roku w Meridian (Idaho) policjant aresztował kobietę, która odmówiła opuszczenia zamkniętej części parku publicznego. Ograniczenie było związane z lockdownem COVID-19. Inne części parku pozostały otwarte. Po nagłośnieniu sprawy, libertariańska organizacja Idaho Freedom Foundation opublikowała na Facebooku imię, nazwisko i zdjęcie funkcjonariusza, wzywając ludzi, by „dali znać departamentowi policji w Meridian, co o tym sądzą". Dodatkowo, lider grupy Ammon Bundy rozpowszechnił adres domowy policjanta, m.in. wysyłając go na listę mailingową i zapisując na tablicy podczas spotkania grupy. Nagranie z widocznym adresem obejrzano ponad 1200 razy. Bundy wcześniej twierdził, że „jeśli prawa jakiejś osoby są łamane, tysiące ludzi powinno ją otoczyć, nagłośnić sprawę i pociągnąć sprawców do odpowiedzialności". Wypowiadał się też za obecnością broni na protestach, mówiąc: „Pierwsza poprawka jest chroniona przez drugą". W wyniku ujawnienia danych osobowych, przed domem policjanta zebrały się dziesiątki protestujących, żądając, by przyjął 13-stronicową skargę. Republikański poseł z Idaho, Greg Chaney, skomentował, że działania Bundy'ego naraziły policjanta i jego rodzinę na poważne zagrożenie i ryzyko.



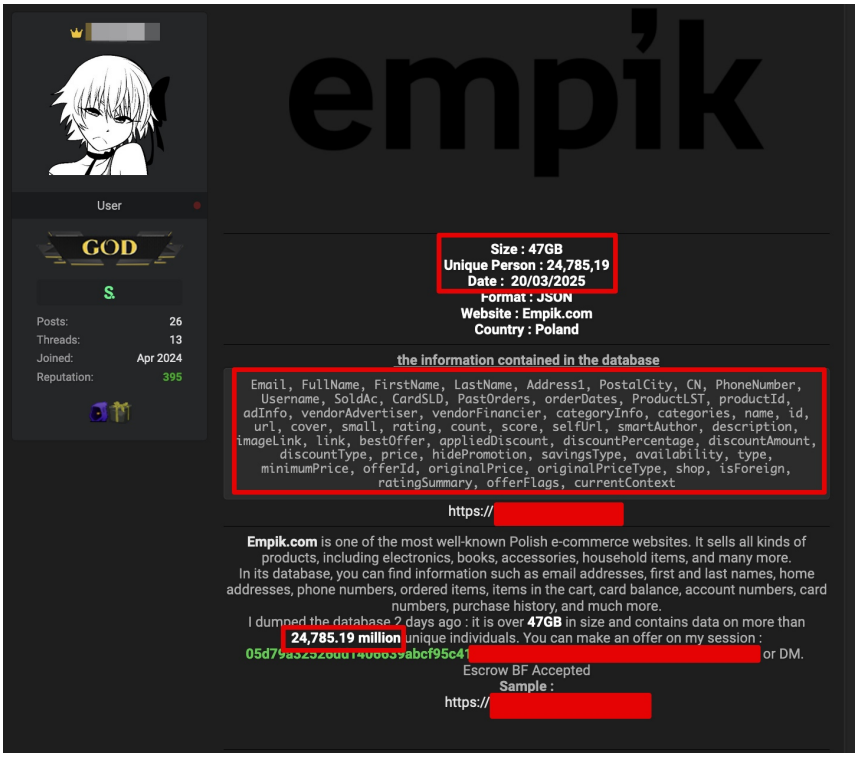
Rysunek 5.1: Post na Facebooku

W marcu 2025 roku w Niebezpieczniku[22] pojawiło się ogłoszenie dotyczące sprzedaży rzekomej bazy danych zawierającej informacje o 24 milionach klientów Empiku. Zgodnie z opisem, baza miała zawierać takie dane jak: imię i nazwisko, numer telefonu, adres (zamieszkania lub dostawy), adres e-mail oraz informacje o zamówieniach (ich liczba i daty). Co istotne, w ogłoszeniu oraz w udostępnionej próbce danych nie znalazły się żadne hasła ani ich hashe. Pojawiły się jednak wątpliwości co do autentyczności i pochodzenia danych - m.in. podejrzanie wysoka liczba rekordów (24 mln), która mogła wynikać z błędnego formatowania (sugerowano, że mogło chodzić o 2,4 mln). Sugerowano również, że źródłem danych może być nie Empik, lecz zewnętrzny system reklamowy lub trackingowy, który mógł generować duplikaty.
Empik zareagował na incydent, publikując oświadczenie, w którym poinformował, że:

- większość danych z próbki nie występuje w ich systemach,
- format danych różni się od stosowanego w systemach wewnętrznych firmy,
- część ujawnionych informacji to ogólnodostępne dane produktowe z katalogu Empik.com,
- nie doszło do wycieku haseł, historii zakupów ani danych płatniczych,
- zgodnie z normą PCI DSS firma nie przechowuje numerów kart płatniczych.

Zalecano klientom zachowanie ostrożności i zmianę hasła na wypadek, gdyby dane jednak pocho-

dziły z realnego wycieku i przypomniano, że na podstawie danych takich jak nazwisko, e-mail, adres czy numer telefonu można wzajemnie ustalić inne informacje o osobie. Aktualizacja z 23 marca 2025 roku przyniosła kolejne, kluczowe oświadczenie Empiku: po szczegółowej analizie stwierdzono jednoznacznie, że rzekoma baza danych była fałszywa i została spreparowana na podstawie danych pochodzących z historycznych wycieków z innych firm, najprawdopodobniej w celu oszukania potencjalnych kupujących. Empik zapewnił, że nie doszło do żadnego incydentu bezpieczeństwa po ich stronie, dane klientów są bezpieczne, a firma natychmiast po pojawieniu się sygnałów o możliwym wycieku uruchomiła proces weryfikacji z pomocą zespołu CERT. Firma również przestrzegła przed powielaniem niezweryfikowanych informacji dotyczących bezpieczeństwa danych. W tym przypadku było to akurat fałszywe ogłoszenie, ale bardzo ważne jest by firmy reagowały właściwymi działaniami na tego typu incydenty, tak jak zrobiła to firma Empik.



Rysunek 5.2: Ogłoszenie sprzedaży bazy danych klientów empiku

Rysunek 5.3: Próbka ogłoszenia

5.1.4. Cyberstalking i nękanie

Opis zagrożenia

Gdy dane kontaktowe lub lokalizacyjne stają się publiczne, rośnie ryzyko:

- śledzenia i nękania w sieci lub w świecie rzeczywistym,
- stosowania szantażu lub gróźb,
- prób kontaktu ze strony osób niepożądanych, np. stalkerów.

Przykłady z życia

Angela Dunn to epidemiolog nękana przez protestujących, po tym jak opublikowano ulotki z jej adresem na Facebooku, protestujący pojawili się pod jej domem, oskarżając ją o zniszczenie gospodarki przez lockdown.

Rysunek 5.4: Protesty pod domem Angeli Dunn

Laurie Jones, czyli urzędniczka zdrowia publicznego znalazła się pod ostrzałem gróźb. Po kontakcie z osobą zakażoną COVID-19, gdzie kobieta zadzwoniła do osoby zakażonej, aby przypomnieć jej o tym by została w domu, chociaż zakażona osoba tego nie przestrzegała. Pojawiły się później w internecie pod jej adresem fałszywe oskarżenia o szpiegostwo, szerzenie paniki, co wywołało falę hejtu i groźby śmierci. Kobieta doświadczyła traumy, musiała zainstalować zabezpieczenia w domu. Do dziś ma obawy przed wychodzeniem z domu i stała się chorobliwie zachowawcza.

W styczniu 2021 roku senator Mitch McConnell spotkał się z falą krytyki po tym, jak zablokował propozycję zwiększenia kwoty wypłat w ramach rządowej pomocy COVID-19. W odpowiedzi, niektórzy oburzeni obywatele odnaleźli jego domowy adres i dokonali aktu wandalizmu na jego drzwiach pojawił się napis „Gdzie są moje pieniądze?”, a na ganku umieszczono wulgarne hasła. Protest został zorganizowany przez grupę DC Under Siege, która opublikowała wydarzenie na Facebooku. Rankiem przed domem senatora zebrała się grupa demonstrantów wyposażonych w megafony i transparenty. McConnell odniósł się do zajścia, podkreślając, że docenia prawo obywateli do udziału w procesie demokratycznym, również tych, którzy się z nim nie zgadzają. Jednocześnie stanowczo potępił akty

wandalizmu i zastraszania, stwierdzając, że „polityka strachu nie ma miejsca w naszym społeczeństwie”.

Rysunek 5.5: Akt wandalizmu na drzwiach senatora

W lipcu 2020 roku fani Taylor Swift zaczęli atakować krytyczkę muzyczną Jillian Mapes po tym, jak opublikowała recenzję albumu Folklore, która ich zdaniem nie była wystarczająco pochlebna. Wkrótce po publikacji recenzji w internecie zaczęły krążyć prywatne dane Mapes, w tym jej adres domowy, numery telefonów oraz zdjęcia jej i jej domu. Już godzinę po pojawieniu się recenzji zaczęła otrzymywać telefony m.in. o drugiej w nocy oraz groźby przez Twittera, w tym nawoływania do „spalenia jej domu”. Jillian odpowiedziała na ataki we wpisie na Twitterze, pisząc, że dostaje wiadomości pełne nienawiści, ale mimo strachu, jaki wywołała sytuacja, jest bezpieczna i trzyma się dobrze.

5.1.5. Oszustwa socjotechniczne i internetowe

Opis zagrożenia

Ujawnione dane są wykorzystywane do:

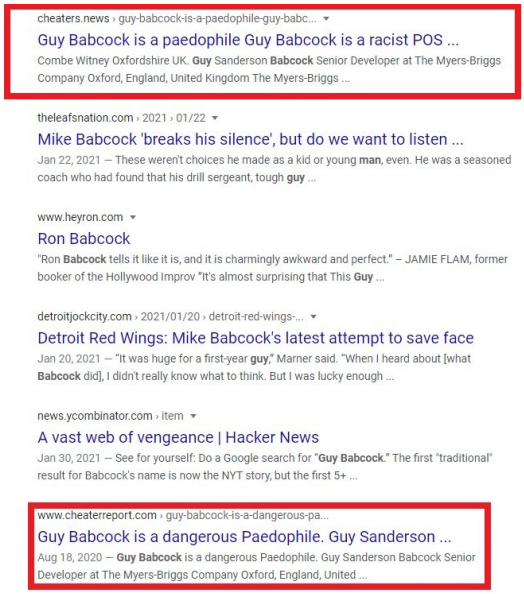
- phishingu (fałszywe e-maile i strony),

- spear-phishingu (ukierunkowane ataki na konkretną osobę),
- vishingu (oszustwa telefoniczne),
- przekonywania ofiar do ujawnienia loginów i haseł.

Przykłady z życia

W 2020 r. grupa atakujących uzyskała dostęp do systemów Twittera, podszywając się pod pracowników i kradnąc dane do kont znanych osób.

Guy Babcock został niesłusznie oskarżony w internecie o pedofilię i złodziejstwo. Fałszywe wpisy na różnych forach zarzucały Babcockowi i jego rodzinie przestępstwa seksualne i kradzieże. Jego udostepnionych w internecie dane, czyli zdjęcia, miejsce pracy, dane kontaktowe z LinkedIna i Facebooka umożliwiły oszustom wiarygodne przedstawienie jego osoby i jego rodziny jako pedofila oraz złodziei, w postaci wpisów a forum. Wkrótce jego zdjęcia zaczęły krążyć w internecie z podpisem pedofila. Mężczyzna i jego rodzina zaczęli bać się o swoje kariery, reputację i o przyszłość swoich dzieci. I chociaż powstał nawet prostujący te informacje artykuł w magazynie „New York Times", to rodzina wciąż boi się o swoje bezpieczeństwo i dobre mniemanie.



Rysunek 5.6: Fałszywe wpisy oskarżające o pedofilię

5.1.6. Konsekwencje reputacyjne

Opis zagrożenia

Dla osób publicznych, pracowników wysokiego szczebla lub właścicieli firm, ujawnienie danych może prowadzić do:

- utracenia zaufania klientów i partnerów,
- negatywnego rozgłosu medialnego
- trudności w znalezieniu nowej pracy lub rozwijaniu kariery.

Przykłady z życia

Wyciek danych w firmach takich jak Proofpoint, Apple czy Google nadszarpnął ich wizerunek jako liderów w dziedzinie innowacji i bezpieczeństwa.

Facebook był w 2011 roku cytowany w 33% spraw rozwodowych. Wskazuje to na to, że treści udostepniane w internecie mogą być wykorzystywane przeciwko nam, jako dowody w sprawach sądowych.

W wyniku mylnej identyfikacji mężczyzny z nagrania z zamieszek na Kapitolu w 2021 roku, internauci uznali, że to David Quintavalle zabił policjanta. Pan David miał udostępniony w internecie swoje imię, nazwisko, stare zdjęcia, adres domowy. W efekcie doprowadziło to do gróźb, nękań, media przed domem, policyjna ochrona, trwałe powiązanie nazwiska z fałszywym oskarżeniem. Musiał wynająć prawnika, który musiał mu pomóc odbudować jego dobre imię, chociaż bardzo trudno było w pełni pozbyć się nękań.

Rysunek 5.7: Fragment nagrania zamieszek na Kapitolu

5.1.7. Odpowiedzialność prawna

Opis zagrożenia

Udostępnianie danych bez zgody lub ich niewłaściwe zabezpieczenie może skutkować:

- pozwami cywilnymi ze strony poszkodowanych osób,
- karami finansowymi za naruszenie przepisów o ochronie danych (np. RODO/GDPR),
- odpowiedzialnością dyscyplinarną lub karną.

Przykłady z życia

Microsoft mógłby zostać ukarany grzywną do 20 mln euro za wyciek danych z systemów GitHub, gdyby doszło do naruszenia danych klientów z UE.

Japończyk Hibiki Sato użył zdjęć z mediów społecznościowych i na podstawie refleksów w oczach i Google Street View, by odnaleźć kobietę i ją zaatakować. Po tym wydarzeniu kobieta musiała się długi czas zmagać z traumą, natomiast stalker poniósł odpowiedzialność prawną - skazanie na 30 miesięcy pozbawienia wolności w 2020 roku.

Rysunek 5.8: Hibiki Sato

Yue Chen, pacjent z rakiem w IV stadium, który oskarżył swoich lekarzy o traktowanie go jak „małpę laboratoryjną”, odnalazł ich domowe adresy w rejonie zatoki San Francisco i planował ich zamordowanie. 31 maja 2017 roku członkowie rodziny Chen’a zgłosili jego zaginięcie. Gdy policja przybyła na miejsce, znalazła notatkę, w której mężczyzna napisał, że „musi dziś zabić tych lekarzy, ponieważ są źli”. Na szczęście Chen zgubił się i nie odnalazł żadnego z domów lekarzy. Policja zatrzymała go, gdy próbował wrócić do siebie. W jego samochodzie znaleziono dwie naładowane półautomatyczne pistolety, notatnik z ręcznie przepisanymi z Google Maps wskazówkami dojazdu do domów lekarzy oraz białą gumową maskę. Znaleziono również notatkę zatytułowaną „dlaczego zabijam”, w której napisał: „to jest możliwe, jeśli traktujecie ludzi jak zwierzę”.

6. ZALECENIA DOTYCZĄCE PRYWATNOŚCI

Poniżej przedstawione zostały aspekty, których przestrzeganie może nas uchronić przed zagrożeniami wynikającymi z udostepniania danych w internecie. Już na wstępie należy dodać, że eksperci jednogłośnie podkreślają, że nasza prywatność w sieci zależy przede wszystkim od naszej świadomości i ostrożności[23].

6.1. Czego nie udostępniać w internecie?

6.1.1. Adres e-mail i numer telefonu

Aby uniknąć spamu, phishingu, ataków socjotechnicznych, przejęcia kont, używaj oddzielnego e-maila do rejestracji i nie udostępniaj numeru telefonu publicznie.

6.1.2. Adres domowy i lokalizacja geograficzna

Aby zmniejszyć ryzyko włamania, śledzenia, identyfikacji miejsca zamieszkania, wyłącz geolokalizację w aplikacjach i unikaj publikowania informacji o podróżach czy codziennej rutynie.

6.1.3. Zdjęcia nieletnich

Dokładnie przemyśl publikację zdjęć dzieci i ustaw ścisłe ograniczenia prywatności, aby uniknąć wykorzystanie zdjęć przez osoby nieuprawnione, ryzyko ich rozpowszechnienia bez zgody.

6.1.4. Kompromitujące zdjęcia

Nigdy nie udostępniaj intymnych zdjęć, nawet zaufanym osobom — tracisz nad nimi kontrolę. Może to doprowadzić do szantażu zdjęciami i cyberprzemocy.

6.1.5. Dokumenty osobiste

Nie publikuj dokumentów tożsamości, umów czy informacji bankowych; jeśli musisz przechowywać je online — szyfruj dane, aby uniknąć kradzieży tożsamości, oszustw finansowych.

6.1.6. Opinia, skargi i kontrowersyjne komentarze

Przemyśl każdą publikację; internet nie zapomina, a wypowiedzi mogą być błędnie zinterpretowane i spowodować naruszenie reputacji, konflikty, nękanie.

6.1.7. Prywatne rozmowy

Korzystaj z szyfrowanych komunikatorów; nie przesyłaj wrażliwych danych w czatach, aby uniknąć ujawnienia poufnych informacji, ryzyka naruszenia prywatności innych osób.

6.2. Jak się chronić?

W wielu przypadkach nie ma jednak wyjścia. Wymagane jest udostepnienie w swoich danych w internecie. Poniżej znajduje się podsumowanie tego co warto zrobić, aby ochronić dane już przez nas

udostępnione [24].

6.2.1. Aktualizuj posiadane oprogramowanie

- Przestępcy wykorzystują luki w oprogramowaniu, zanim zostaną one załatane przez producentów.
- Zaleca się jak najszybsze instalowanie dostępnych aktualizacji.
- Najlepiej włączyć automatyczne aktualizacje dla:
  - oprogramowania antywirusowego,
  - przeglądarki internetowej,
  - systemu operacyjnego,
  - aplikacji mobilnych.

6.2.2. Zabezpiecz domową sieć Wi-Fi

- Router jest punktem dostępowym do Internetu dla wszystkich urządzeń domowych.
- Zainfekowane urządzenie może rozprzestrzenić złośliwe oprogramowanie na inne w sieci.
- Zaleca się zabezpieczenie routera hasłem i aktualizacją oprogramowania.

6.2.3. Skonfiguruj ustawienia prywatności

Na swoich kontach - ogranicz widoczność treści tylko do osób zaufanych.

6.2.4. Używaj silnych, unikalnych haseł

Nieopartych na łatwo dostępnych danych osobowych. Regularnie je zmieniaj i nie używaj ich wielokrotnie.

- Hasło powinno być długie (co najmniej 15 znaków) i zawierać duże i małe litery, cyfry oraz znaki specjalne.
- Można stosować losowe hasła lub frazy z przypadkowych słów.
- Alternatywy:
  - Automatycznie generowane hasła przez przeglądarkę.
  - Menedżery haseł - umożliwiają tworzenie i bezpieczne przechowywanie haseł. Dostępne w większości przeglądarek. Dostępne są również dedykowane aplikacje przechowujące hasła.

6.2.5. Korzystaj z uwierzytelniania dwuskładnikowego (2FA)

- Nawet silne hasło można złamać — 2FA zwiększa bezpieczeństwo.
- Najczęściej spotykane formy:
  - jednorazowe kody przez SMS lub e-mail,
  - aplikacje autoryzujące,
  - fizyczne klucze bezpieczeństwa.

6.2.6. Uważaj na podejrzane linki

Sprawdzaj źródło wiadomości i nie klikaj w nieznane odnośniki. Przestępcy stosują wiadomości e-mail lub SMS-y z fałszywymi linkami i załącznikami.

6.2.7. Korzystaj z zaufanych źródeł cyberbezpieczeństwa

Np. Incibe, OSI.

6.2.8. Nie publikuj wrażliwych informacji

Szczególnie tych dotyczących lokalizacji, planów i życia prywatnego.

6.2.9. Wyłącz geolokalizację

Jeśli nie jest potrzebna.

6.2.10. Zachowaj ostrożność w komentarzach

Nawet pozorna anonimowość może być złudna.

6.2.11. Rozważ korzystanie z usług typu VPN

Szczególnie w publicznych sieciach Wi-Fi (na laptopie, smartfonie, tablecie).

6.2.12. Unikaj podawania numeru telefonu

Tam, gdzie nie jest to absolutnie konieczne, ewentualnie do celów bezpieczeństwa (np. 2FA).

6.2.13. Świadomie analizuj, z jakich usług korzystasz i jakimi danymi płacisz za „darmowy" dostęp

6.2.14. Zastrzeż swój numer PESEL

Od niedawna w Polsce można zastrzec swój numer PESEL, aby nikt nie wykorzystał go bez wiedzy osoby, do której ten numer należy.

Z zastrzeżonym numerem PESEL można:

- zarejestrować się do lekarza,
- zrealizować receptę,
- wypłacić środki w bankomacie,
- zlecić przelew bankowy,
- wyjechać za granicę lub
- załatwić sprawę urzędową.

Natomiast z zastrzeżonym numerem PESEL:

- nie można zaciągnąć kredytu, pożyczki, leasingu,
- nie można otworzyć nowego rachunku bankowego,
- nie można zmienić umowy kredytu lub pożyczki,
- nie można wypłacić w placówce banku więcej pieniędzy niż trzykrotność minimalnego wynagrodzenia,
- nie można załatwić niektórych spraw notarialnych,
- nie można otrzymać duplikatu karty SIM.

Należy cofnąć zastrzeżenie numeru PESEL w momencie, gdy zajdzie potrzeba wykonania powyższych czynności.

Od 1 czerwca 2024 r. instytucje finansowe (np. banki) mają obowiązek weryfikować, czy numer PESEL jest zastrzeżony przy zawieraniu np. umowy kredytu lub pożyczki.

Funkcjonalność ta oferowana przez urząd państwowy pozwala uniknąć wielu przykrych i niebezpiecznych sytuacji. Zastrzeżenie PESELu to sprawa, którą można załatwić bardzo szybko - przez internet, bądź osobiście w Urzędzie [25].

6.3. Co zrobić jeśli wycieką dane osobowe?

Wyciek danych osobowych to incydent, w wyniku którego dochodzi do niezamierzonego lub celowego ujawnienia danych identyfikujących daną osobę. Może on mieć miejsce zarówno w przestrzeni cyfrowej (np. Internet), jak i offline, np. przez zgubienie lub kradzież dokumentów tożsamości [26].

6.3.1. Przyczyny wycieków danych

- ataki hakerskie lub działania przestępcze,
- błędy użytkowników,
- awarie systemów informatycznych,
- niesprawny sprzęt komputerowy.

6.3.2. Możliwe skutki wycieku danych

- kradzież tożsamości (wyłudzenia kredytów, zawieranie umów na cudze dane),
- utrata kontroli nad swoimi informacjami,
- straty finansowe i problemy z prawem,
- naruszenie reputacji i dobrego imienia,
- szkody gospodarcze i społeczne,
- negatywny wpływ na środowisko zawodowe.

6.3.3. Jak sprawdzić, czy doszło do wycieku danych?

Nie zawsze otrzymujemy bezpośrednie powiadomienie o naruszeniu danych, dlatego warto monitorować swoją aktywność w sieci. Pomocne w tym są następujące narzędzia:

- **haveibeenpwned.com** - pozwala sprawdzić, czy dane logowania (adres e-mail, hasło) znajdują się w bazie znanych wycieków.
- **Biuro Informacji Kredytowej (BIK)** - umożliwia monitorowanie zapytań kredytowych powiązanych z Twoimi danymi. W razie próby ich użycia, system wyśle powiadomienie.
- **Lokalizator wycieku danych** - pomocne narzędzie weryfikujące, czy PESEL lub inne dane zostały ujawnione.

6.3.4. Co zrobić po wycieku danych?

W przypadku potwierdzenia lub podejrzenia wycieku danych osobowych, należy działać natychmiast. Poniżej przedstawiono najważniejsze kroki:

Wyloguj się i zmień hasła

- Wyloguj się ze wszystkich kont, na których mogły zostać wykorzystane wyciekłe dane.
- Zmień hasła na silne i unikalne - co najmniej 12 znaków, zawierające litery, cyfry i znaki specjalne.
- Stosuj zasadę: jedno konto = jedno hasło.

Zastrzeż dokumenty w banku

- W przypadku wycieku danych dowodu osobistego, karty płatniczej itp. natychmiast je zastrzeż.
- Można to zrobić w dowolnym banku uczestniczącym w programie „Dokumenty Zastrzeżone”.
- W przypadku kart płatniczych można skorzystać z **Systemu Zastrzegania Kart**.

Zgłoś sprawę na policję

- Jeżeli uważasz, że doszło do przestępstwa, zgłoś to na policję.
- Zachowaj potwierdzenie zgłoszenia - może okazać się dowodem w postępowaniu sądowym lub administracyjnym.

Bezpieczeństwo w sieci dobrze podsumowuje poniższy cytat:

Twoje dane to Twoja odpowiedzialność.
W świecie cyfrowym prywatność nie jest dana raz na zawsze - trzeba ją stale chronić.

7. PODSUMOWANIE

Analiza zaprezentowana w niniejszym raporcie ukazuje skalę i złożoność problemu, jakim jest udostępnianie informacji o sobie w Internecie. Przedstawione przykłady dowodzą, że nawet *podstawowe* dane – takie jak wiek, płeć, miejsce zamieszkania czy data urodzenia – często wystarczają do utworzenia **szczegółowych profili** użytkowników. Proces ten jest dodatkowo usprawniany przez narzędzia OSINT oraz wzajemne powiązania między różnymi źródłami informacji (np. rejestrami publicznymi, platformami społecznościowymi czy rejestrami genealogicznymi).

Kluczowe wnioski z raportu:

- **Łączenie danych z rozmaitych źródeł**: Nawet drobne, pozornie nieszkodliwe elementy (tzw. quasi-identyfikatory) mogą prowadzić do *re-identyfikacji* osoby, jeśli są zestawiane z publicznie dostępnymi rejestrami bądź innymi źródłami.
- **Wielostronne wykorzystanie informacji**: Dane osobowe znajdują zastosowanie w sektorze reklamowym (personalizacja i targetowanie), branży HR (weryfikacja kandydatów), a także w działaniach *cyberprzestępczych* (phishing, kradzież tożsamości, handel bazami danych).
- **Rosnąca rola brokerów danych**: Funkcjonuje wyodrębniony rynek, na którym *skonsolidowane* informacje o użytkownikach stają się towarem. Pozwala to firmom na tworzenie wyjątkowo szczegółowych profilów, ale też *zwiększa* ryzyko nadużyć.
- **Konsekwencje dla prywatności i bezpieczeństwa**: Brak kontroli nad udostępnianymi informacjami może prowadzić do *kradzieży tożsamości*, szantażu, strat finansowych czy kłopotów reputacyjnych – zarówno w życiu prywatnym, jak i zawodowym.

Zalecenia dotyczące ochrony prywatności:

- **Minimalizacja ujawnianych danych**: Warto świadomie decydować, które informacje faktycznie muszą być publiczne (np. data urodzenia, lokalizacja, szczegóły dotyczące życia prywatnego).
- **Konfiguracja ustawień profili**: Regularne sprawdzanie, kto ma dostęp do publikowanych treści, i ograniczanie widoczności danych wyłącznie dla zaufanych znajomych bądź grup.
- **Używanie silnych haseł i szyfrowania**: Systematyczna zmiana haseł, unikanie tych samych kombinacji w wielu serwisach oraz korzystanie z weryfikacji dwuetapowej.
- **Świadomość technik socjotechnicznych**: Zachowanie ostrożności wobec nieoczekiwanych wiadomości (phishing, vishing), a także *weryfikowanie* nadawców i treści przed kliknięciem w odnośniki.
- **Zastrzeżenie numeru PESEL**: Korzystanie z procedury, która utrudnia zaciąganie kredytów lub pożyczek na skradzioną tożsamość.

Z perspektywy zarówno użytkowników indywidualnych, jak i organizacji, **kluczowe** jest zrozumienie, że w Internecie nie ma całkowitej gwarancji anonimowości. Wzrastająca liczba usług, platform i urządzeń połączonych z siecią sprawia, iż obieg informacji przyspiesza, a ewentualne błędy czy zaniedbania w zabezpieczeniu danych stają się tym łatwiejsze do wykorzystania na niekorzyść użytkowników. Raport potwierdza, że rozważne zarządzanie danymi oraz stała edukacja w zakresie mechanizmów ich przetwarzania powinny być *stałym elementem* kultury cyfrowej. Tylko wówczas można skutecznie minimalizować zagrożenia płynące z nadmiernego (lub nieświadomego) udostępniania informacji w sieci.

WYKAZ LITERATURY

[1] "Personally identifiable information (pii)," (Dostęp 2025-04-02). [Online]. Available: https://www.directives.doe.gov/terms_definitions/personally-identifiable-information-pii.

[2] "Czy imię i nazwisko to dane osobowe?" (Dostęp 2025-04-02). [Online]. Available: https://lexdigital.pl/czy-imie-i-nazwisko-to-dane-osobowe.

[3] "Jakie dane najczęściej ujawniamy w internecie?" (Dostęp 2025-04-02). [Online]. Available: https://cyberdefence24.pl/cyberbezpieczenstwo/jakie-dane-najczesciej-ujawniamy-w-internecie.

[4] "Profilowanie danych osobowych - na czym polega i z czym się wiąże?" (Dostęp 2025-04-02). [Online]. Available: https://poir.parp.gov.pl/component/content/article/87360%3Aprofilowanie-danych-osobowych-na-czym-polega-i-z-czym-sie-wiaze.

[5] "Osint framework," (Dostęp 2025-04-02). [Online]. Available: https://osintframework.com/.

[6] "Inteltechniques," (Dostęp 2025-04-02). [Online]. Available: https://inteltechniques.com/links.html.

[7] "Harvard professor re-identifies anonymous volunteers in dna study," (Dostęp 2025-04-02). [Online]. Available: https://www.forbes.com/sites/adamtanner/2013/04/25/harvard-professor-re-identifies-anonymous-volunteers-in-dna-study/.

[8] "Personal data in public sources," (Dostęp 2025-04-03). [Online]. Available: https://www.autoriteitpersoonsgegevens.nl/en/themes/internet-and-smart-devices/personal-data-on-the-internet/personal-data-in-public-sources.

[9] "Jakie ślady zostawiamy w internecie?" (Dostęp 2025-03-26). [Online]. Available: https://serwisy.gazetaprawna.pl/nowe-technologie/artykuly/1019082,prywatnosc-w-internecie.html.

[10] "Twitter, snapchat, internet rzeczy. dane konsumenta na wyciągnięcie ręki," (Dostęp 2025-03-26). [Online]. Available: https://serwisy.gazetaprawna.pl/nowe-technologie/artykuly/1017004,twitter-snapchat-iot-czyli-dane-konsumenta-na-wyciagniecie-reki-2000.html.

[11] Wiesław Wolny. "Bezpieczeństwo i prywatność danych w badaniach mediów społecznościowych," (Dostęp 2025-03-26). [Online]. Available: https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-06278ad0-d165-4b1c-8b2b-5817b3982416/c/07.pdf&ved=2ahUKEwjqnLHL35SMAxUwMRAIHewwFJAQFnoECBEQAQ&usg=AOvVaw3qYoSnGsBD_Z3EekwfDAoJ.

[12] "Co wie o nas internet?" (Dostęp 2025-03-26). [Online]. Available: https://www.infor.pl/prawo/prawa-konsumenta/konsument-w-sieci/5297247,Co-wie-o-nas-Internet.html.

[13] "Things you shouldn't share with chatgpt," (Dostęp 2025-03-26). [Online]. Available: https://telefonicatech.com/en/blog/things-you-shouldnt-share-with-chatgpt.

[14] "Behind the one-way mirror: A deep dive into the technology of corporate surveillance," (Dostęp 2025-04-02). [Online]. Available: https://www.eff.org/wp/behind-the-one-way-mirror.

[15] "Digital 2025: Global overview report," (Dostęp 2025-04-02). [Online]. Available: https://datareportal.com/reports/digital-2025-global-overview-report.

[16] "Digital 2025: Country headlines report," (Dostęp 2025-04-02). [Online]. Available: https://datareportal.com/reports/digital-2025-local-country-headlines.

[17] "Enisa threat landscape 2022," (Dostęp 2025-04-02). [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022.

[18] “Phishing activity trends report,” (Dostęp 2025-04-02). [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q4_2022.pdf.

[19] “Raport roczny z działalności cert polska w 2023 roku,” (Dostęp 2025-04-02). [Online]. Available: https://cert.pl/posts/2024/04/raport-roczny-2023/.

[20] “11 real-life insider threat examples,” (Dostęp 2025-03-26). [Online]. Available: https://www.mimecast.com/blog/insider-threat-examples/.

[21] “10 times someone’s online information made them unsafe in real life,” (Dostęp 2025-03-26). [Online]. Available: https://www.linkedin.com/pulse/10-times-someones-online-information-made-them-unsafe-bridges.

[22] “Wyciek danych klientów empiku,” (Dostęp 2025-03-30). [Online]. Available: https://niebezpiecznik.pl/post/wyciek-danych-klientow-empiku/.

[23] “Protect yourself online: Data you should never share,” (Dostęp 2025-03-26). [Online]. Available: https://telefonicatech.com/en/blog/data-you-should-never-share-how-to-protect-it.

[24] “Protect your personal information from hackers and scammers,” (Dostęp 2025-04-03). [Online]. Available: https://consumer.ftc.gov/articles/protect-your-personal-information-hackers-and-scammers.

[25] “Zastrzeż swój numer pesel lub cofnij zastrzeżenie,” (Dostęp 2025-03-30). [Online]. Available: https://www.gov.pl/web/gov/zastrzez-swoj-numer-pesel-lub-cofnij-zastrzezenie.

[26] “Co zrobić, gdy doszło do wycieku danych osobowych?” (Dostęp 2025-04-03). [Online]. Available: https://www.stockwatch.pl/wiadomosci/co-zrobic-gdy-doszlo-do-wycieku-danych-osobowych,notatki,326519.

WYKAZ RYSUNKÓW

5.1 Post na Facebooku 27
5.2 Ogłoszenie sprzedaży bazy danych klientów empiku 28
5.3 Próbka ogłoszenia 28
5.4 Protesty pod domem Angeli Dunn 29
5.5 Akt wandalizmu na drzwiach senatora 30
5.6 Fałszywe wpisy oskarżające o pedofilię 31
5.7 Fragment nagrania zamieszek na Kapitolu 32
5.8 Hibiki Sato 33