Politechnika Częstochowska

Justyna Żywiołek

Organizacje w czasach cyberzagrożeń 6.0

Wpływ AI na bezpieczeństwo informacyjne i odporność operacyjną

Monografia

Wydawnictwo Politechniki Częstochowskiej

Częstochowa 2026

Recenzent
prof. dr hab. inż. Mieczysław Pawlisiak

Redakcja
Joanna Jasińska

Redakcja techniczna
Robert Świerczewski

Projekt okładki
Justyna Żywiołek z użyciem AI

ISBN 978-83-65976-56-7
ISBN 978-83-65976-57-4 (pdf)
DOI: 10.17512/CUT/9788365976574

Wydawnictwo Politechniki Częstochowskiej, 42-202 Częstochowa, Al. Armii Krajowej 36 B
redakcja – wydawnictwo.pcz.pl, tel. 34 325 04 80, e-mail: wydawnictwo@pcz.pl
dystrybucja – sklep.pcz.pl, tel. 34 325 03 93, e-mail: sklep@pcz.pl

Wstęp

Szybkie przemiany technologiczne stanowią dziś jedno z najważniejszych źródeł przeobrażeń systemów społecznych, gospodarczych i organizacyjnych, determinując kierunki rozwoju współczesnych państw i przedsiębiorstw. Postępująca cyfryzacja, automatyzacja oraz wdrażanie systemów wykorzystujących sztuczną inteligencję stają się fundamentem nowych modeli działania organizacji, szczególnie w sektorach przemysłowych, które jako pierwsze wprowadzają rozwiązania zdolne do przetwarzania danych w czasie rzeczywistym, autonomicznego podejmowania decyzji, przewidywania trendów lub generowania treści. Jednocześnie to właśnie w tych środowiskach pojawiają się najwcześniej symptomy ryzyk, których źródła, dynamika i konsekwencje odbiegają od klasycznych zagrożeń znanych naukom o bezpieczeństwie. Transformacja przemysłu dokonująca się na naszych oczach prowadzi do powstania nowej kategorii ryzyka – cyberzagrożeń 6.0, będących złożoną kombinacją zagrożeń informacyjnych, technicznych i społecznych, generowanych w coraz większym stopniu przez autonomiczne i adaptacyjne systemy sztucznej inteligencji.

Współczesne organizacje przemysłowe, funkcjonujące w ramach paradygmatu Industry 4.0, a coraz częściej Industry 5.0, wykorzystują technologie cyfrowe nie tylko jako narzędzia produkcyjne, lecz także jako element systemów decyzyjnych, komunikacyjnych i analitycznych. Wraz z ich rozwojem rośnie znaczenie informacji jako zasobu strategicznego, co prowadzi do redefinicji relacji pomiędzy człowiekiem, technologią a bezpieczeństwem. W przeciwieństwie do wcześniejszych rewolucji przemysłowych, w których kluczowe były zmiany techniczne, obecna faza transformacji charakteryzuje się wzrostem autonomii maszyn, zdolnością systemów do generowania nowych treści oraz możliwością wpływania na zachowania ludzkie. Wprowadza to do przestrzeni przemysłowej nowe obszary podatności oraz powoduje, że tradycyjne systemy bezpieczeństwa stają się niewystarczające.

Tematem niniejszej monografii jest analiza roli sztucznej inteligencji w kształtowaniu bezpieczeństwa informacyjnego oraz odporności operacyjnej organizacji przemysłowych w kontekście transformacji do epoki określanej mianem cyberzagrożeń 6.0. Koncepcja ta zakłada, że w kolejnej fazie rozwoju technologicznego zagrożenia nie będą wynikały jedynie z celowych działań cyberprzestępców czy błędów technicznych, lecz także z autonomicznego, nie zawsze przewidywalnego działania systemów generatywnych, które mogą tworzyć treści, decyzje i rekomendacje przekraczające zamierzone cele projektowe. Na tym tle kluczowego znaczenia nabiera świadomość społeczna – rozumiana nie jako zjawisko kulturowe, lecz jako zdolność pracowników i interesariuszy do rozumienia zagrożeń, interpretowania danych, krytycznej oceny

działań algorytmów oraz rozpoznawania manipulacji informacyjnych generowanych przez sztuczną inteligencję.

Wybór tematu monografii znajduje silne uzasadnienie w ramach nauk o bezpieczeństwie, które od lat rozwijają podejścia do zarządzania ryzykiem, ochrony informacji i wzmacniania odporności organizacyjnej. W obliczu ekspansji technologii generatywnych, które mogą tworzyć zniekształcone komunikaty, wzmacniać dezinformację lub inicjować niepożądane procesy operacyjne, obszar bezpieczeństwa wymaga rozszerzenia o analizę mechanizmów działania AI oraz o społeczne aspekty percepcji ryzyka. Nauki o bezpieczeństwie, badając złożone systemy zagrożeń, stanowią właściwy paradygmat do eksploracji zależności pomiędzy człowiekiem, technologią i organizacją w świecie rosnącej autonomii systemów cyfrowych. Z tego względu przedstawiona w monografii problematyka jest nie tylko aktualna, ale wręcz konieczna dla zrozumienia współczesnych procesów zachodzących w sektorach strategicznych, w tym w przemyśle, energetyce, logistyce czy infrastrukturze krytycznej.

W szczególności na uwagę zasługuje fakt, że sztuczna inteligencja znacząco zmienia charakter zagrożeń informacyjnych. Systemy generatywne mogą prowadzić do tworzenia treści o charakterze manipulacyjnym, dezinformacyjnym, dyskryminacyjnym lub destabilizującym, które wpływają zarówno na bezpieczeństwo operacyjne, jak i reputacyjne organizacji. Przedsiębiorstwa stają się odbiorcą oraz nadawcą komunikatów generowanych przez AI, co oznacza, że błędy algorytmów mogą przekształcić się w rzeczywiste incydenty, zagrażające funkcjonowaniu organizacji. Dlatego jednym z fundamentalnych zagadnień staje się kwestia nadzoru nad systemami sztucznej inteligencji, ich audytowalności, transparentności oraz ich zdolności do działania w sposób zgodny z oczekiwaniami użytkowników i regulatorów.

Drugim kluczowym elementem, wskazywanym w literaturze i rozwijanym w niniejszej monografii, jest rola świadomości społecznej. W warunkach rosnącej złożoności systemów przemysłowych pracownicy oraz interesariusze muszą posiadać nie tylko kompetencje techniczne, ale przede wszystkim zdolność interpretowania zagrożeń wynikających z działania autonomicznych systemów. Świadomość ta nie ma charakteru ogólnokulturowego – odnosi się do zdolności poznawczych związanych z oceną informacji, wiarygodności danych oraz rozumienia zasad działania sztucznej inteligencji. Pracownik nie musi być programistą, aby potrafił rozpoznać nieprawidłowości w działaniu algorytmu; musi jednak posiadać wiedzę na temat potencjalnych konsekwencji błędnych decyzji podejmowanych przez systemy.

Ważnym punktem wyjścia dla tej monografii jest także luka badawcza dotycząca relacji pomiędzy technologicznymi systemami bezpieczeństwa a percepcją społeczną ryzyka w organizacjach przemysłowych. Dotychczasowe badania koncentrowały się głównie na aspektach technicznych, takich jak cyberbezpieczeństwo systemów OT (operacyjnej technologii), podatności w IoT (Internetu Rzeczy), zabezpieczenia sieci przemysłowych czy analiza malware. Znacznie rzadziej podejmowano refleksję nad tym, jak zagrożenia generowane przez AI (sztuczną inteligencję) są rozumiane przez

pracowników i jakie znaczenie ma ich percepcja dla efektywności systemów ochrony. Tymczasem w środowisku przemysłowym to właśnie niewłaściwa reakcja użytkowników na ryzyko informacyjne może prowadzić do eskalacji incydentów, nawet wtedy, gdy system techniczny działa poprawnie.

Celem monografii jest ukazanie, w jaki sposób sztuczna inteligencja, funkcjonując w ramach zaawansowanych systemów przemysłowych, wpływa na bezpieczeństwo informacyjne i operacyjne oraz jak świadomość społeczna pracowników i interesariuszy może wzmacniać lub osłabiać odporność organizacji. W tym kontekście postawione zostają trzy kluczowe pytania badawcze:

PB1: *W jaki sposób systemy AI w sektorze przemysłowym generują nowe kategorie zagrożeń informacyjnych i operacyjnych?*

PB2: *Jakie kompetencje i formy świadomości społecznej są niezbędne, aby te zagrożenia rozpoznawać i nimi zarządzać?*

PB3: *Jakie modele integracji technologii, regulacji i edukacji mogą zwiększyć odporność organizacyjną w warunkach rosnącej autonomii systemów AI?*

Główna hipoteza badawcza zakłada, że w organizacjach przemysłowych funkcjonujących w warunkach rosnącej autonomii systemów sztucznej inteligencji pojawia się jakościowo nowa kategoria zagrożeń informacyjnych i operacyjnych, których materializacja jest w istotnym stopniu wzmacniana przez niewystarczającą świadomość społeczną pracowników i interesariuszy w zakresie ryzyka związanego z działaniem AI. Oznacza to, że sama obecność zaawansowanych technologicznie systemów bezpieczeństwa nie jest wystarczającym warunkiem zapewnienia odporności organizacyjnej, jeżeli towarzyszy jej niski poziom rozumienia mechanizmów działania algorytmów, ograniczona umiejętność interpretacji generowanych przez nie treści oraz brak kompetencji pozwalających na właściwe reagowanie na sygnały zagrożeń. Innymi słowy, hipoteza główna przyjmuje, że relacja między sztuczną inteligencją a bezpieczeństwem organizacyjnym ma charakter socjotechniczny: ryzyka generowane przez AI ulegają amplifikacji w środowiskach, w których świadomość informacyjna i technologiczna użytkowników jest niewystarczająco rozwinięta.

Uszczegółowieniem hipotezy głównej jest hipoteza, zgodnie z którą systemy sztucznej inteligencji wykorzystywane w organizacjach przemysłowych generują nowe kategorie zagrożeń informacyjnych i operacyjnych, które nie są w pełni identyfikowane i klasyfikowane przez tradycyjne modele bezpieczeństwa oparte na paradygmacie cyberbezpieczeństwa infrastruktury IT i OT. Zakłada się, że zagrożenia algorytmiczne – obejmujące między innymi błędne rekomendacje, zniekształcone komunikaty, emergentne zachowania modeli czy nieprzewidziane interakcje z innymi systemami – pozostają poza zasięgiem klasycznych procedur analizy ryzyka, co prowadzi do powstania obszarów ryzyka ukrytego, niewidocznego w standardowych procesach audytu.

Kolejna hipoteza szczegółowa odnosi się do percepcji ryzyka i świadomości społecznej pracowników. Zakłada ona, że niski poziom świadomości informacyjnej i technologicznej pracowników organizacji przemysłowych istotnie zwiększa podatność tych

organizacji na zagrożenia generowane przez AI, zarówno w wymiarze informacyjnym, jak i operacyjnym. Innymi słowy, przyjmuje się, że brak zrozumienia zasad działania systemów AI, ich ograniczeń, możliwych błędów oraz konsekwencji błędnych rekomendacji prowadzi do nieadekwatnej reakcji na sygnały ostrzegawcze, do nadmiernego zaufania do algorytmów lub przeciwnie – do nieuzasadnionego odrzucania ich wskazań. Oba te skrajne wzorce zachowań tworzą podatność na incydenty, które w sprzyjających warunkach mogą eskalować do poziomu wydarzeń krytycznych dla bezpieczeństwa organizacyjnego.

Następna hipoteza szczegółowa dotyczy roli świadomości społecznej jako czynnika modyfikującego skuteczność systemów bezpieczeństwa technologicznego. Zakłada się, że wysoki poziom świadomości informacyjnej, rozumianej jako zdolność do krytycznej analizy danych, rozpoznawania manipulacji, oceny wiarygodności treści generowanych przez AI oraz rozumienia podstaw działania algorytmów, wzmacnia efektywność technicznych systemów bezpieczeństwa w organizacjach przemysłowych. Przyjmuje się zatem, że organizacje charakteryzujące się rozwiniętą świadomością społeczną wykazują wyższy poziom odporności na zagrożenia algorytmiczne, ponieważ użytkownicy są w stanie wcześniej identyfikować anomalie, sygnały ostrzegawcze oraz wskazania systemów, które odbiegają od normy operacyjnej.

Ważnym uzupełnieniem powyższych założeń jest hipoteza dotycząca roli zintegrowanych modeli zarządzania bezpieczeństwem. Zakłada ona, że organizacje przemysłowe, które wdrażają systemowe podejście do zarządzania zagrożeniami AI, integrujące komponent technologiczny (rozwiązania techniczne i architekturę bezpieczeństwa), regulacyjny (normy, wytyczne, procedury) oraz społeczny (świadomość, kompetencje, edukację pracowników), osiągają wyższy poziom odporności operacyjnej niż organizacje, w których bezpieczeństwo postrzegane jest głównie jako domena infrastruktury technicznej. Hipoteza ta odzwierciedla założenie, że odporność organizacyjna w erze cyberzagrożeń 6.0 ma charakter wielowymiarowy i wymaga spójnej integracji praktyk technicznych, organizacyjnych i edukacyjnych.

Ostatnia z proponowanych hipotez szczegółowych odnosi się do związku pomiędzy obecnością systemów sztucznej inteligencji a strukturą ryzyka w organizacjach przemysłowych. Przyjmuje się, że im większy zakres autonomii przyznany systemom AI w obszarach kluczowych dla funkcjonowania organizacji (sterowanie procesami, analityka predykcyjna, rekomendacje decyzyjne, komunikacja operacyjna), tym większe jest prawdopodobieństwo, że błędy algorytmiczne lub nieprawidłowości w danych wejściowych doprowadzą do powstania incydentów bezpieczeństwa o charakterze wielowymiarowym, obejmujących zarówno szkody operacyjne, jak i informacyjne oraz reputacyjne. Hipoteza ta zakłada, że poziom ryzyka nie jest prostą funkcją samego wdrożenia technologii, lecz zależy od stopnia jej autonomii, jakości nadzoru oraz zdolności organizacji do monitorowania i korygowania działania systemów AI.

Przedstawiony zestaw hipotez porządkuje przyjętą w monografii perspektywę badawczą. Hipoteza główna wiąże w jednym założeniu dwa kluczowe wątki rozważań:

rolę sztucznej inteligencji jako generatora nowych kategorii zagrożeń oraz znaczenie świadomości społecznej jako czynnika warunkującego skuteczność systemów bezpieczeństwa. Hipotezy szczegółowe rozwijają to założenie, umożliwiając analizę: po pierwsze – specyfiki zagrożeń algorytmicznych, po drugie – zależności między świadomością informacyjną a podatnością organizacji na te zagrożenia, po trzecie – wpływu zintegrowanych modeli zarządzania bezpieczeństwem na odporność organizacyjną oraz po czwarte – relacji między zakresem autonomii systemów AI a strukturą ryzyka w organizacjach przemysłowych. Dzięki temu monografia uzyskuje wyraźną, spójną ramę hipotez, którą można następnie weryfikować zarówno na poziomie analiz teoretycznych, jak i badań empirycznych.

Struktura monografii obejmuje pięć rozdziałów, z których każdy koncentruje się na innym wymiarze analizowanego zjawiska. Pierwszy rozdział stanowi fundament teoretyczny, opisując ewolucję przemysłu i pojawienie się koncepcji cyberzagrożeń 6.0. W rozdziale drugim zaprezentowano analizę technologicznych źródeł ryzyka generowanych przez AI. W rozdziale trzecim omówiona została rola świadomości społecznej jako bariery bezpieczeństwa. W rozdziale czwartym skoncentrowano się na modelach przeciwdziałania zagrożeniom, natomiast w rozdziale piątym ukazano scenariusze przyszłości i kierunki rozwoju bezpieczeństwa w przemyśle w perspektywie Industry 6.0.

Wartość naukowa monografii polega na wprowadzeniu interdyscyplinarnego ujęcia, łączącego nauki o bezpieczeństwie z analizą technologii, zarządzania ryzykiem informacyjnym i edukacją cyfrową pracowników. Wartość praktyczna wynika z opracowania modeli, które mogą być wykorzystane przez organizacje przemysłowe do budowania odporności operacyjnej w środowisku rosnącej autonomii systemów sztucznej inteligencji.

W ostatnich latach obserwujemy intensywny rozwój badań poświęconych zależnościom pomiędzy sztuczną inteligencją, bezpieczeństwem informacyjnym oraz odpornością organizacji[1]. Dyskurs naukowy koncentruje się przede wszystkim na analizie rosnącej roli systemów algorytmicznych w przestrzeni operacyjnej, komunikacyjnej i decyzyjnej przedsiębiorstw[2], a także na zagrożeniach wynikających z ich autonomii oraz zdolności do generowania treści i rekomendacji. W literaturze światowej wyraźnie uwidacznia się tendencja do odejścia od klasycznego postrzegania bezpieczeństwa technologicznego jako zbioru środków ochrony infrastruktury technicznej[3]. W zamian akcentuje się konieczność systemowego ujęcia zagrożeń obejmujących nie tylko elementy techniczne, ale również informacyjne, społeczne i organizacyjne. Badacze

[1] B.R. Kuc, Z. Ścibiorek (2018): *Zarys metodologii nauk o bezpieczeństwie*, Toruń: Wydawnictwo Adam Marszałek.

[2] A. Lekka-Kowalik, *Metodologia nauk*, [w:] S. Janeczek, M. Walczak, A. Starościc (red.), *Metodologia nauk*, Katolicki Uniwersytet Lubelski, Lublin 2019, s. 309-332.

[3] J. Gierszewski, A. Pieczywok (2020): *Metodologiczne podstawy badania problemów bezpieczeństwa*, Warszawa: Difin.

podkreślają, że to właśnie wzajemne oddziaływanie tych elementów staje się najważniejszym czynnikiem determinującym bezpieczeństwo w erze sztucznej inteligencji[4].

Analizy poświęcone systemom opartym na AI wskazują, że ich rozwój w sektorach przemysłowych prowadzi do zmiany charakteru ryzyka. Modele uczenia maszynowego i systemy generatywne nie tylko wspierają procesy produkcyjne, logistyczne i decyzyjne, lecz także stają się nowym źródłem zagrożeń informacyjnych. Literatura zwraca uwagę, że systemy te mogą generować treści wykraczające poza pierwotne założenia projektowe oraz produkować rekomendacje[5], które w sposób niezamierzony prowadzą do destabilizacji procesów operacyjnych. W badaniach empirycznych wskazuje się, że nieprawidłowości takie mogą występować nawet w dobrze zabezpieczonych systemach, jeżeli dane wejściowe lub struktura modelu AI zawierają błędy lub stronniczość. Konsekwencją jest powstanie nowej kategorii zagrożeń, których nie można przypisać ani wyłącznie czynnikowi ludzkiemu, ani wyłącznie technicznemu. Są to zagrożenia hybrydowe, o charakterze informacyjno-technologicznym[6].

W przeglądach systematycznych podkreśla się również znaczne tempo wzrostu liczby publikacji dotyczących wpływu sztucznej inteligencji na procesy przemysłowe oraz bezpieczeństwo informacyjne. Analizy bibliometryczne wyraźnie wskazują, że po roku 2020 nastąpił gwałtowny wzrost liczby badań poświęconych etyce AI, dezinformacji generowanej przez algorytmy, autonomii modeli decyzyjnych oraz ich wpływowi na bezpieczeństwo organizacyjne. Tendencja ta widoczna jest we wszystkich głównych bazach naukowych, takich jak Scopus, Web of Science czy IEEE Xplore. Wynika ona przede wszystkim z upowszechnienia generatywnych modeli językowych oraz systemów zdolnych do samodzielnego tworzenia treści, które weszły do powszechnego użytku zarówno w biznesie, jak i w instytucjach administracyjnych.

W literaturze naukowej coraz częściej wskazuje się także na zagrożenia operacyjne wynikające z zastosowania AI w przemysłowych systemach sterowania, zarządzania procesami i monitorowania infrastruktury[7]. Modele predykcyjne stosowane w diagnostyce maszyn, analityce produkcji czy zarządzaniu zużyciem energii mogą prowadzić do błędnych rekomendacji, jeżeli oparto je na nieadekwatnych danych. W sektorach produkcyjnych, gdzie procesy są silnie zautomatyzowane, błędna decyzja algorytmu może skutkować przerwaniem produkcji, uszkodzeniem infrastruktury, stratami finansowymi lub ryzykiem dla zdrowia pracowników. Badania wskazują, że zagrożenia

---

[4] A. Dawidczyk, J. Jurczak (2022): *Metodologia bezpieczeństwa w przykładach i zastosowaniach. Podręcznik akademicki*, Warszawa: Difin.

[5] R. Kamprowski (2019): *Metodologia badań nad bezpieczeństwem miasta. W kierunku transdyscyplinarności*, „Public Administration and Regional Development", No. 4, s. 344-355. DOI: 10.34132/pard2019.04.06.

[6] P. Kobis, A. Kisiołek (2018): *Zarządzanie bezpieczeństwem danych w przedsiębiorstwach MSP z uwzględnieniem czynnika ludzkiego – wyniki badań*, „Przegląd Organizacji", nr 8(943), s. 44-52. DOI: 10.33141/po.2018.08.07.

[7] T. Gajewski (2020): *Towards resilience. European cybersecurity strategic framework*, „Ante Portas – Security Studies", No. 1(14), s. 103-122. DOI: 10.33674/3201911.

takie występują nie tylko w wyniku błędów technicznych, ale również w wyniku zjawisk związanych z tzw. błędami systemowymi, obejmującymi niewłaściwe projektowanie modeli, brak nadzoru nad procesem uczenia się oraz nieodpowiednie rozumienie wyników generowanych przez systemy AI.

Ważną kategorię badań stanowią analizy poświęcone zagrożeniom informacyjnym wynikającym z działania sztucznej inteligencji[8]. Systemy generatywne są zdolne do tworzenia treści o charakterze manipulacyjnym, naruszającym normy społeczne lub destabilizującym komunikację organizacyjną. W przemyśle zagrożenia te mają szczególne znaczenie, ponieważ nieprawidłowe komunikaty mogą prowadzić do błędnych decyzji kadry zarządzającej, zakłócenia komunikacji wewnętrznej, utraty zaufania publicznego, szkód reputacyjnych lub dezorganizacji pracy zespołów. Coraz więcej badań wskazuje, że dezinformacja generowana przez sztuczną inteligencję może mieć bezpośrednie konsekwencje operacyjne, zwłaszcza w sytuacjach kryzysowych, gdy czas reakcji oraz dostęp do rzetelnej informacji mają kluczowe znaczenie dla funkcjonowania przedsiębiorstwa.

W przeglądach badań dotyczących bezpieczeństwa organizacyjnego wskazuje się także na coraz silniejszy wpływ regulacji prawnych, które wymuszają na przedsiębiorstwach wdrażanie odpowiednich systemów nadzoru nad działaniem AI. Europejski Akt o Sztucznej Inteligencji (AI Act), rozporządzenie NIS2 oraz inne akty regulujące bezpieczeństwo informacji i technologii nakładają na przedsiębiorstwa obowiązek dokumentowania, monitorowania i kontrolowania systemów opartych na algorytmach, szczególnie w sytuacjach, gdy systemy takie mogą wywierać wpływ na decyzje o charakterze społecznym lub operacyjnym. Badania wskazują, że organizacje przemysłowe wciąż mają trudności z właściwą implementacją tych wymogów, głównie ze względu na brak kompetencji technicznych oraz brak precyzyjnych procedur nadzoru nad działaniem inteligentnych algorytmów.

W literaturze podejmuje się również problem odporności organizacyjnej w kontekście rosnącej autonomii systemów technologicznych[9]. Tradycyjne modele odporności koncentrowały się na zdolności organizacji do reakcji na zakłócenia techniczne, takie jak awarie maszyn czy przerwy w dostawach energii. Obecnie jednak podkreśla się, że odporność musi być rozumiana jako systemowa zdolność do adaptacji wobec złożonych zagrożeń generowanych przez technologie cyfrowe. Oznacza to konieczność uwzględnienia nie tylko infrastruktury technicznej, ale również kompetencji użytkowników, jakości informacji, przejrzystości modeli decyzyjnych, wiarygodności danych oraz zdolności organizacji do monitorowania i interpretowania sygnałów generowanych przez algorytmy. W tym kontekście pojawia się rosnąca liczba badań wskazujących,

[8] A. Alotaibi, M.A. Rassam (2023): *Adversarial machine learning attacks against intrusion detection systems: A survey on strategies and defense*, „Future Internet", No. 15(2), s. 62. DOI: 10.3390/fi15020062.

[9] W. Danielak, P. Niewiadomski, D. Sobotkiewicz (2023): *Od teorii do praktyki zarządzania. Wybrane zagadnienia nauki o zarządzaniu i jakości oraz nauki o bezpieczeństwie*, Zielona Góra: Oficyna Wydawnicza Uniwersytetu Zielonogórskiego.

że czynniki społeczne odgrywają fundamentalną rolę w funkcjonowaniu systemów zabezpieczających.

Analizy badań wskazują wyraźnie, że świadomość społeczna w kontekście systemów sztucznej inteligencji jest obszarem niedostatecznie rozwiniętym, choć kluczowym dla bezpieczeństwa organizacyjnego. W literaturze zauważa się, że pracownicy posiadają często wysokie kompetencje operacyjne, ale brakuje im wiedzy dotyczącej tego, jak funkcjonują modele AI, jakie są ich ograniczenia oraz w jaki sposób interpretować wyniki generowane przez algorytmy. Badania wykazują również istnienie zjawisk takich jak nadmierne lub niewystarczające zaufanie do systemów AI, co może prowadzić do błędnych decyzji operacyjnych. Zjawiska te mają bezpośrednie przełożenie na bezpieczeństwo informacyjne oraz odporność organizacyjną.

Warto podkreślić, że w literaturze nie istnieje dotąd model teoretyczny, który w sposób kompleksowy ujmowałby zależności między technologicznymi zagrożeniami generowanymi przez AI, świadomością społeczną pracowników oraz odpornością organizacyjną w sektorach przemysłowych. Jest to wyraźna luka badawcza, którą niniejsza monografia zamierza wypełnić. Analizy teoretyczne wskazują, że systemy bezpieczeństwa oparte wyłącznie na technologii są niewystarczające wobec zagrożeń o charakterze informacyjno-algorytmicznym. Badania prowadzone w ostatnich latach pokazują, że konieczne jest zintegrowanie technologii, edukacji oraz regulacji, aby stworzyć wielowymiarowe modele bezpieczeństwa odpowiadające realiom przemysłu.

Na tle powyższych analiz można stwierdzić, że kierunek badań dotyczących AI, bezpieczeństwa i przemysłu zmierza w stronę identyfikacji zależności między działaniem systemów generatywnych a czynnikami społecznymi. Wskazuje się, że efektywność systemów bezpieczeństwa zależy nie tylko od jakości technologii, ale przede wszystkim od zdolności ludzi do ich rozumienia, nadzorowania oraz właściwej interpretacji wyników. Tendencja ta jest zgodna z założeniami Industry 5.0 i 6.0, które podkreślają współpracę człowieka z technologią oraz etyczne i społeczne aspekty funkcjonowania sztucznej inteligencji. W tym kontekście bezpieczeństwo staje się zagadnieniem interdyscyplinarnym, obejmującym zarówno inżynierię systemów, zarządzanie technologią, jak i nauki o bezpieczeństwie.

Rozważania dotyczące miejsca sztucznej inteligencji w systemach bezpieczeństwa organizacji przemysłowych wymagają osadzenia w szerszych ramach teoretycznych. Nauki o bezpieczeństwie, które od dekad analizują źródła zagrożeń, mechanizmy ich materializacji oraz sposoby budowania odporności, stoją obecnie przed koniecznością redefinicji swoich podstawowych założeń. Wynika to z faktu, że nowoczesne technologie nie tylko umożliwiają nowe formy komunikacji i zarządzania procesami operacyjnymi, ale stają się autonomicznymi podmiotami wpływającymi na strukturę ryzyka. W tradycyjnych teoriach bezpieczeństwa zakładano, że zagrożenia mają charakter zewnętrzny lub wynikają z błędów ludzkich i technicznych. W przypadku systemów opartych na sztucznej inteligencji podział ten ulega zatarciu, a część zagrożeń powstaje na skutek interakcji pomiędzy danymi, algorytmem, środowiskiem technicznym oraz

użytkownikiem. Tego rodzaju zagrożenia nie mogą być analizowane wyłącznie poprzez pryzmat klasycznych koncepcji cyberbezpieczeństwa, gdyż obejmują elementy, które wymykają się jednoznacznej klasyfikacji.

Jednym z punktów wyjścia dla zrozumienia tej problematyki jest teoria systemów złożonych, wskazująca, że systemy techniczne o wysokim stopniu autonomii stają się źródłem zachowań emergentnych. W kontekście przemysłowym oznacza to, że działania algorytmu mogą tworzyć nowe zagrożenia, których nie da się przewidzieć poprzez analizę pojedynczych składowych. Teoria ta podkreśla, że zachowanie systemów z udziałem AI jest wypadkową dynamicznych interakcji oraz nieustannego uczenia się modeli. W konsekwencji bezpieczeństwo takich systemów musi być analizowane jako proces ciągłej adaptacji, a nie jako stan osiągany poprzez zastosowanie określonych środków technicznych. Z tego względu nauki o bezpieczeństwie muszą rozszerzyć swoje zainteresowania na mechanizmy uczenia maszynowego, przetwarzania danych oraz generowania treści, które stają się integralną częścią systemów organizacyjnych.

Istotnym uzupełnieniem powyższego podejścia jest teoria ryzyka technologicznego, która zakłada, że rozwój technologii zawsze generuje nowe formy podatności. Teoria ta wprowadza pojęcie ryzyka rezydualnego, odnoszącego się do zagrożeń, których nie można wyeliminować nawet poprzez zastosowanie najbardziej zaawansowanych systemów ochrony. W przypadku sztucznej inteligencji ryzyko takie obejmuje m.in. błędy modelu, stronniczość danych, nieprzewidywalność treści generowanych przez algorytmy czy możliwość niezamierzonych konsekwencji decyzji podejmowanych przez systemy autonomiczne. Badania teoretyczne wskazują, że im większa autonomia technologii, tym bardziej rośnie obszar ryzyka rezydualnego, co stanowi fundamentalne wyzwanie dla bezpieczeństwa organizacyjnego.

W literaturze dotyczącej informacji oraz bezpieczeństwa informacyjnego kluczowe znaczenie ma teoria zarządzania informacją, według której jakość, wiarygodność i interpretacja informacji determinują prawidłowość procesów decyzyjnych. Systemy sztucznej inteligencji, które przetwarzają dane w sposób autonomiczny lub częściowo autonomiczny, wpływają bezpośrednio na treść, strukturę i przepływ informacji w organizacji. Oznacza to, że na poziomie teoretycznym bezpieczeństwo informacyjne nie może być już analizowane jako ochrona danych przed nieautoryzowanym dostępem. Musi zostać rozszerzone o kwestię tego, jak dane są generowane, przetwarzane, analizowane i interpretowane przez systemy algorytmiczne oraz jak te procesy wpływają na percepcję użytkowników. Teoria zarządzania informacją wskazuje, że ryzyko informacyjne może wynikać zarówno z błędnej interpretacji informacji przez człowieka, jak i z błędnej interpretacji danych przez systemy AI.

Kolejnym istotnym nurtem teoretycznym jest teoria odpowiedzialności algorytmicznej, która analizuje relacje między twórcami technologii, jej użytkownikami oraz

skutkami działania systemów autonomicznych[10]. Wprowadza ona pojęcie odpowiedzialności wielopodmiotowej, obejmującej programistów, projektantów, operatorów, decydentów oraz użytkowników końcowych. Teoria ta podkreśla, że systemy sztucznej inteligencji tworzą nowy rodzaj odpowiedzialności zbiorowej, w której granice między kontrolą człowieka a autonomią maszyny nie są jasno określone. Z perspektywy nauk o bezpieczeństwie teoria ta jest niezwykle istotna, ponieważ wskazuje, że bezpieczeństwo w erze AI nie może być przypisane jednemu sektorowi organizacji – musi być przedmiotem współpracy technologów, specjalistów od ryzyka, kadr zarządzających oraz pracowników odpowiedzialnych za operacyjne wykorzystanie systemów.

W kontekście niniejszej monografii szczególną rolę odgrywa również teoria odporności organizacyjnej, która pozwala zrozumieć mechanizmy adaptacji przedsiębiorstw do zakłóceń generowanych przez systemy technologiczne. Teoria ta zakłada, że organizacja odporna to taka, która potrafi przewidywać, rozpoznawać, absorbować i adaptować się do zagrożeń. W przypadku zagrożeń generowanych przez AI szczególne znaczenie ma zdolność wczesnego wykrywania sygnałów ostrzegawczych oraz rozumienie, w jaki sposób decyzje i rekomendacje algorytmów mogą wpływać na funkcjonowanie organizacji. Badania wskazują, że odporność organizacyjna w środowisku silnie zinformatyzowanym zależy zarówno od jakości technologii, jak i od kompetencji ludzi odpowiedzialnych za nadzór nad systemami. Oznacza to, że odporność technologiczna musi być ściśle powiązana z odpornością społeczną.

Jednym z kluczowych wniosków wynikających z powyższych teorii jest to, że bezpieczeństwo w erze sztucznej inteligencji nie może być definiowane wyłącznie w kategoriach ochrony infrastruktury technicznej. Musi być rozumiane jako proces dynamiczny, obejmujący interakcję pomiędzy technologią, informacją i użytkownikiem. Z tego powodu fundamentalne znaczenie ma teoria świadomości informacyjnej, która analizuje zdolność jednostek i grup do interpretowania zagrożeń oraz podejmowania decyzji na podstawie danych. Teoria ta wskazuje, że interpretacja informacji zależy od wiedzy, doświadczenia, umiejętności analitycznych oraz zdolności do identyfikacji manipulacji. W kontekście systemów AI świadomość informacyjna staje się jednym z najważniejszych elementów bezpieczeństwa, ponieważ użytkownicy muszą być zdolni do oceny, czy generowane przez algorytmy informacje są wiarygodne, adekwatne i bezpieczne w użyciu.

Połączenie powyższych teorii prowadzi do wyodrębnienia zestawu zależności kluczowych dla niniejszej monografii. Po pierwsze, systemy AI generują zagrożenia informacyjne i operacyjne w sposób odmienny niż tradycyjne technologie, co wymaga zmiany podejścia badawczego. Po drugie, zdolność organizacji do reagowania na te zagrożenia zależy od poziomu świadomości społecznej, rozumianej jako kompetencja poznawcza umożliwiająca właściwą interpretację danych generowanych przez

[10] K.M. Eisenhardt (1989): *Building theories from case study research*, „The Academy of Management Review", Vol. 14(4), s. 532-550. DOI: 10.2307/258557.

algorytmy. Po trzecie, odporność organizacyjna w środowisku AI jest zależna od współdziałania technologii, wiedzy i struktur decyzyjnych, co stanowi istotny wkład w rozwój nauk o bezpieczeństwie. Po czwarte, regulacje prawne dotyczące sztucznej inteligencji muszą być analizowane w szerszym kontekście teoretycznym, uwzględniającym zarówno kwestie techniczne, jak i społeczne.

Z powyższych analiz płynie wniosek, że istnieje pilna potrzeba integracji dorobku nauk o bezpieczeństwie, teorii informacji, teorii systemów oraz badań nad sztuczną inteligencją. Współczesne zagrożenia mają charakter wielowymiarowy i obejmują zarówno błędy techniczne, jak i błędne interpretacje informacji przez użytkowników. Tego typu zagrożenia wymagają nowego podejścia teoretycznego, które umożliwi analizę systemów autonomicznych w sposób uwzględniający ich wpływ na bezpieczeństwo organizacyjne.

Rozwój sztucznej inteligencji, autonomicznych systemów decyzyjnych i narzędzi generatywnych spowodował powstanie jakościowo nowych wyzwań dla współczesnych organizacji funkcjonujących w środowisku przemysłowym. W efekcie pojawiła się konieczność zrewidowania podstawowych pytań dotyczących bezpieczeństwa informacyjnego, bezpieczeństwa technologicznego oraz odporności organizacyjnej. Dotychczasowe modele analizy zagrożeń koncentrowały się na czynnikach technicznych, cyberatakach czy błędach ludzkich. Tymczasem współczesna rzeczywistość przemysłowa wskazuje, że coraz większa liczba incydentów wynika z interakcji pomiędzy algorytmem, użytkownikiem a środowiskiem technicznym, w którym system działa. Oznacza to, że bezpieczeństwo nie jest już tożsame jedynie z cyberbezpieczeństwem. Z perspektywy nauk o bezpieczeństwie obserwujemy wyraźne przesunięcie ciężaru analizy z ochrony infrastruktury na ochronę procesów informacyjnych, decyzyjnych i organizacyjnych, które są kształtowane przez technologie o wysokim stopniu autonomii.

Monografia wychodzi naprzeciw tej zmianie paradygmatu. Jej celem jest przedstawienie wielowymiarowej analizy relacji zachodzących pomiędzy sztuczną inteligencją, bezpieczeństwem informacyjnym, odpornością operacyjną oraz świadomością społeczną pracowników i interesariuszy organizacji przemysłowych. Analiza ta opiera się na założeniu, że bezpieczeństwo organizacji w erze autonomicznych systemów nie zależy jedynie od technologii czy procedur, ale od synergii pomiędzy technologią, regulacjami, kompetencjami użytkowników oraz właściwie zintegrowanymi procesami nadzoru. Tego typu spojrzenie jest zgodne z najnowszymi trendami teoretycznymi w naukach o bezpieczeństwie, które podkreślają znaczenie podejścia systemowego, holistycznego oraz interdyscyplinarnego.

Głównym celem monografii jest zbudowanie spójnego modelu teoretycznego umożliwiającego analizę i ocenę bezpieczeństwa organizacji przemysłowych w kontekście rosnącej autonomii systemów sztucznej inteligencji. Cel ten obejmuje kilka szczegółowych aspektów.

Pierwszym jest określenie, w jaki sposób systemy AI wpływają na procesy informacyjne, komunikacyjne i operacyjne w organizacjach przemysłowych. Sztuczna inteligencja staje się elementem codziennego funkcjonowania przedsiębiorstwa, jednak jej działanie – zarówno poprawne, jak i błędne – może generować nowe ryzyka. Zrozumienie mechanizmów ich powstawania wymaga interdyscyplinarnego spojrzenia łączącego teorię informacji, teorię systemów, teorię ryzyka technologicznego oraz teorie odpowiedzialności algorytmicznej.

Drugim celem jest analiza świadomości społecznej jako czynnika wpływającego na bezpieczeństwo technologiczne. W monografii założono, że kompetencje użytkowników, ich zdolność do rozumienia działania algorytmów, interpretowania danych oraz rozpoznawania manipulacji generowanych przez systemy cyfrowe stanowią kluczowy element odporności organizacyjnej.

Trzecim celem jest ocena aktualnych modeli regulacyjnych, normatywnych i organizacyjnych dotyczących sztucznej inteligencji w kontekście bezpieczeństwa. W szczególności chodzi o analizę AI Act, NIS2, regulacji dotyczących ochrony informacji oraz norm branżowych, które mają wpływ na systemy zarządzania ryzykiem.

Celem czwartym jest wskazanie kierunków rozwoju bezpieczeństwa organizacyjnego w erze Industry 6.0, w której systemy sztucznej inteligencji stanowią aktywne elementy infrastruktury decyzyjnej, operacyjnej i komunikacyjnej.

W ramach monografii sformułowano zestaw pytań badawczych, które wyznaczają kierunek analizy teoretycznej i pozwalają na uporządkowanie zagadnień kluczowych dla bezpieczeństwa organizacyjnego.

Pierwsze pytanie dotyczy charakteru zagrożeń powstających w wyniku działania autonomicznych systemów sztucznej inteligencji: Jakie typy zagrożeń informacyjnych, decyzyjnych i operacyjnych są generowane przez współczesne systemy AI w środowisku przemysłowym? Pytanie to odnosi się do potrzeby klasyfikacji zagrożeń, które nie wynikają ani z klasycznych cyberataków, ani z błędów ludzkich, lecz z dynamiki działania algorytmów, jakości danych wejściowych lub niezamierzonych procesów uczenia maszynowego.

Drugie pytanie dotyczy wpływu użytkowników na bezpieczeństwo technologiczne: W jaki sposób świadomość społeczna pracowników i interesariuszy wpływa na zdolność organizacji do rozpoznawania i interpretowania zagrożeń generowanych przez AI? Pytanie to wymaga analizy kompetencji informacyjnych pracowników, poziomu zaufania do systemów AI, percepcji ryzyka oraz umiejętności krytycznej interpretacji treści.

Trzecie pytanie dotyczy regulacji i nadzoru: Na ile obecne regulacje, normy i wytyczne są adekwatne do zarządzania zagrożeniami wynikającymi z autonomii systemów AI? Dotyczy to oceny zgodności pomiędzy rzeczywistym działaniem systemów technologicznych a wymaganiami formalnymi narzucanymi przez państwo i instytucje ponadnarodowe.

Czwarte pytanie dotyczy odporności organizacyjnej: Jak może wyglądać model odporności organizacyjnej 6.0, obejmujący integrację technologii, procedur, kompetencji oraz nadzoru nad autonomicznymi systemami AI?

Analiza literatury pokazuje, że badania dotyczące sztucznej inteligencji w przemyśle koncentrują się najczęściej na zagadnieniach technicznych – na przykład bezpieczeństwie systemów OT–IT, detekcji zagrożeń, analizie podatności lub automatyzacji procesów. Innym nurtem są badania dotyczące etyki AI, stronniczości danych lub algorytmicznych form dyskryminacji. Wciąż jednak brakuje prac, które w kompleksowy sposób analizują zależności między:
1) technologicznymi zagrożeniami generowanymi przez AI,
2) bezpieczeństwem informacyjnym i procesowym organizacji przemysłowych,
3) świadomością społeczną pracowników,
4) odpornością organizacyjną i systemową.

Brakuje również badań wskazujących, jak systemy AI oddziałują na procesy organizacyjne, jak modyfikują przepływ informacji oraz jakie nowe role pełnią użytkownicy w ekosystemach, w których systemy cyfrowe generują autonomiczne treści. Luka ta wynika częściowo z szybkości zmian technologicznych oraz z trudności w interdyscyplinarnym ujęciu problemu.

Niniejsza monografia wnosi do nauk o bezpieczeństwie kilka kluczowych wartości.

Po pierwsze, proponuje nowe ujęcie bezpieczeństwa organizacyjnego, obejmujące zagrożenia wynikające z działania systemów autonomicznych. Ujęcie to może stanowić podstawę do dalszych badań teoretycznych nad bezpieczeństwem informacyjnym oraz bezpieczeństwem technologicznym.

Po drugie, monografia wprowadza model analityczny bezpieczeństwa łączący technologię, regulacje i świadomość społeczną. W literaturze brak jest spójnego ujęcia integrującego te trzy obszary, mimo że są one fundamentalne dla funkcjonowania współczesnych organizacji.

Po trzecie, praca ta rozwija obszar resilience studies w kierunku uwzględniającym autonomiczne systemy technologiczne. Odporność organizacyjna nie może być dłużej analizowana wyłącznie w kontekście zakłóceń zewnętrznych – musi obejmować również zdolność organizacji do adaptacji wobec błędów generowanych przez systemy cyfrowe.

Po czwarte, monografia wnosi wkład do teorii informacji w naukach o bezpieczeństwie, wskazując, że bezpieczeństwo nie zależy wyłącznie od kontroli nad dostępem do danych, ale od ich interpretacji, zrozumienia i jakości. W erze sztucznej inteligencji to właśnie interpretacja informacji – zarówno przez ludzi, jak i przez systemy algorytmiczne – staje się kluczową determinantą bezpieczeństwa.

Po piąte, praca ta wnosi istotny wkład metodologiczny, pokazując, że analiza zagrożeń generowanych przez AI wymaga łączenia metod ilościowych, jakościowych, analizy

systemowej i studiów przypadków. W interdyscyplinarnym polu nauk o bezpieczeństwie takie podejście jest szczególnie potrzebne.

Tworzenie monografii o tak złożonym charakterze wymaga zastosowania podejścia wielowątkowego. Zastosowano metodę analizy systemowej[11], która umożliwia holistyczne spojrzenie na relacje między technologią, użytkownikami i procesami organizacyjnymi. Wykorzystano przegląd literatury o charakterze systematycznym i krytycznym, aby zidentyfikować kluczowe teorie, modele i luki badawcze. Zastosowano także elementy analizy foresightowej, koniecznej do określenia przyszłych scenariuszy rozwoju zagrożeń i odporności organizacyjnej. Złożony charakter problematyki poruszanej w niniejszej monografii wymagał przyjęcia podejścia metodologicznego właściwego dla nauk o bezpieczeństwie, które z definicji operują na styku wielu dyscyplin i poziomów analizy[12]. Badanie zagrożeń generowanych przez systemy sztucznej inteligencji w środowisku organizacji przemysłowych nie może być ograniczone do jednego paradygmatu badawczego ani do pojedynczej metody analizy. Bezpieczeństwo w warunkach rosnącej autonomii systemów technologicznych ma bowiem charakter systemowy, dynamiczny i wielowymiarowy, co implikuje konieczność zastosowania triangulacji metod badawczych oraz integracji metod teoretycznych i empirycznych[13].

Podstawą metodologiczną monografii jest podejście systemowe, które umożliwia analizę organizacji przemysłowych jako złożonych układów społeczno-technicznych, w których technologia, informacja, człowiek i struktura organizacyjna pozostają w ciągłej interakcji. Przyjęcie tej perspektywy pozwala odejść od redukcjonistycznego ujmowania bezpieczeństwa wyłącznie w kategoriach ochrony infrastruktury technicznej i skupić się na procesach generowania, interpretowania oraz wykorzystywania informacji w środowisku wspieranym przez sztuczną inteligencję. Podejście systemowe jest szczególnie adekwatne do analizy cyberzagrożeń 6.0, które nie są jednorodne ani statyczne, lecz wynikają z emergentnych relacji pomiędzy algorytmami, danymi i użytkownikami[14].

W warstwie teoretycznej zastosowano zestaw klasycznych metod badawczych powszechnie wykorzystywanych w naukach o bezpieczeństwie. Analiza została użyta w celu dekompozycji złożonego problemu bezpieczeństwa AI na elementy składowe, takie jak zagrożenia informacyjne, ryzyka algorytmiczne, czynniki społeczne, regulacje prawne oraz mechanizmy organizacyjne. Metoda ta umożliwiła identyfikację

[11] M. Cieślarczyk (2003): *Metody, techniki i narzędzia badawcze oraz elementy statystyki stosowane w pracach magisterskich i doktorskich*, Warszawa: Akademia Obrony Narodowej.
[12] A. Dawidczyk, J. Jurczak, P. Łuka (2019): *Metody, techniki, narzędzia nauk o bezpieczeństwie*, Warszawa: Difin.
[13] C. Frankfort-Nachmias, D. Nachmias, E. Hornowska, M. Zakrzewska (2001): *Metody badawcze w naukach społecznych*, Poznań: Zysk i S-ka.
[14] S. Janeczek, M. Walczak, A. Starościc (red.) (2019): *Metodologia nauk*, Lublin: Katolicki Uniwersytet Lubelski.

kluczowych obszarów podatności oraz określenie relacji zachodzących pomiędzy poszczególnymi komponentami systemu bezpieczeństwa. Analiza była stosowana zarówno w odniesieniu do literatury naukowej, jak i do dokumentów normatywnych, regulacyjnych oraz raportów branżowych.

Uzupełnieniem analizy była synteza, która pozwoliła na integrację rozproszonych wyników badań, koncepcji teoretycznych oraz obserwacji empirycznych w spójny model interpretacyjny. Zastosowanie syntezy było niezbędne ze względu na interdyscyplinarny charakter problematyki, obejmującej elementy nauk o bezpieczeństwie, teorii informacji, zarządzania ryzykiem, studiów nad sztuczną inteligencją oraz badań nad odpornością organizacyjną. Dzięki syntezie możliwe było wypracowanie autorskiej koncepcji cyberzagrożeń 6.0 oraz wskazanie mechanizmów, poprzez które AI wpływa na bezpieczeństwo informacyjne i operacyjne organizacji przemysłowych.

Metoda porównawcza została wykorzystana w celu zestawienia tradycyjnych modeli bezpieczeństwa organizacyjnego z modelami uwzględniającymi obecność systemów autonomicznych. Porównanie objęło zarówno ujęcia teoretyczne, jak i praktyczne rozwiązania stosowane w organizacjach przemysłowych. Analiza porównawcza pozwoliła wykazać, że klasyczne podejścia, skoncentrowane na cyberbezpieczeństwie infrastruktury IT i OT, nie obejmują w wystarczającym stopniu zagrożeń wynikających z autonomii algorytmów, generowania treści oraz wpływu AI na procesy decyzyjne. Metoda ta umożliwiła również identyfikację luk pomiędzy wymaganiami regulacyjnymi a praktyką funkcjonowania systemów bezpieczeństwa w organizacjach.

Istotnym elementem warsztatu badawczego było także abstrahowanie, rozumiane jako wyodrębnienie kluczowych cech i mechanizmów bezpieczeństwa AI niezależnie od specyfiki branżowej czy technologicznej[15]. Dzięki tej metodzie możliwe było sformułowanie uogólnionych wniosków odnoszących się do organizacji przemysłowych jako kategorii analitycznej, a nie jedynie do pojedynczych przypadków. Abstrahowanie pozwoliło również na oderwanie analizy od konkretnych rozwiązań technologicznych i skupienie się na uniwersalnych mechanizmach ryzyka, percepcji zagrożeń oraz odporności organizacyjnej.

Wnioskowanie logiczne, zarówno dedukcyjne, jak i indukcyjne, stanowiło podstawę budowy hipotez badawczych oraz interpretacji wyników analiz. Dedukcja została wykorzystana do wyprowadzenia konsekwencji teoretycznych z istniejących koncepcji bezpieczeństwa, ryzyka technologicznego i teorii systemów. Indukcja natomiast umożliwiła formułowanie uogólnień na podstawie obserwacji empirycznych, studiów przypadków oraz analizy praktyk organizacyjnych. Zastosowanie obu typów wnioskowania było konieczne, aby połączyć rozważania teoretyczne z realiami funkcjonowania organizacji przemysłowych[16].

[15] R. Kamprowski (2029): *Metodologia badań…*, op. cit.
[16] M. Cieślarczyk (2003): *Metody, techniki…*, op. cit.

Metoda analogii została użyta w celu porównania rozwoju zagrożeń generowanych przez sztuczną inteligencję z wcześniejszymi fazami rozwoju zagrożeń technologicznych, takimi jak automatyzacja, informatyzacja czy cyfryzacja procesów przemysłowych. Analogie te pozwoliły lepiej zrozumieć dynamikę zmian oraz wskazać, które elementy wcześniejszych modeli bezpieczeństwa mogą zostać zaadaptowane do nowych warunków, a które wymagają zasadniczej redefinicji. Wykorzystanie analogii miało również charakter prognostyczny, umożliwiając formułowanie scenariuszy rozwoju cyberzagrożeń w perspektywie Industry 6.0.

Warstwa empiryczna monografii opiera się na jakościowym podejściu badawczym, właściwym dla eksploracji zjawisk nowych, złożonych i niedostatecznie opisanych w literaturze. Zastosowano studia przypadków organizacji przemysłowych wykorzystujących systemy sztucznej inteligencji w obszarach krytycznych dla bezpieczeństwa informacyjnego i operacyjnego[17]. Studium przypadku umożliwiło analizę rzeczywistych sytuacji, w których działanie algorytmów prowadziło do zakłóceń, incydentów lub zwiększenia podatności organizacji. Metoda ta pozwoliła uchwycić kontekst organizacyjny, decyzyjny i społeczny, którego nie da się w pełni odzwierciedlić za pomocą badań ilościowych[18].

Równolegle przeprowadzono analizę dokumentów, obejmującą regulacje prawne, normy branżowe, wytyczne instytucjonalne oraz strategie bezpieczeństwa stosowane w organizacjach przemysłowych[19]. Analiza ta miała na celu ocenę stopnia przygotowania organizacji do zarządzania ryzykiem algorytmicznym oraz identyfikację rozbieżności pomiędzy wymaganiami formalnymi a praktyką operacyjną. Szczególną uwagę poświęcono aktom regulacyjnym dotyczącym sztucznej inteligencji, cyberbezpieczeństwa oraz ochrony informacji, które kształtują ramy funkcjonowania systemów AI w środowisku przemysłowym[20].

Istotnym elementem badań empirycznych była także jakościowa ocena ryzyka, rozumiana nie jako formalna analiza ilościowa, lecz jako identyfikacja i klasyfikacja potencjalnych scenariuszy zagrożeń wynikających z działania systemów AI[21]. Ocena ta uwzględniała zarówno czynniki technologiczne, jak i społeczne, w tym poziom świadomości użytkowników, sposób interpretacji informacji oraz mechanizmy nadzoru nad algorytmami. Dzięki temu możliwe było ukazanie, w jaki sposób brak kompetencji

---

[17] A. Czop (2015): *Methodological basis of research in security studies*, „Security Dimensions. International & National Studies", Vol. 3(15), s. 49-67.

[18] A. Dawidczyk, J. Jurczak (2022): *Metodologia bezpieczeństwa…*, op. cit.

[19] T. Pawłuszko (2021): *Nauki o bezpieczeństwie. Budowanie szkoły naukowej*, Kraków: Księgarnia Akademicka. DOI: 10.12797/9788381384612.

[20] A. Dębska, T. Soluch, S. Szostakowska (2016): *Metodologia badań nad przyjmowaniem perspektywy poznawczej podczas dialogu*, „Lingwistyka Stosowana", nr 5(20), s. 31-49. DOI: 10.32612/uw.20804814.2016.5.pp.31-49.

[21] Z. Sabak (2021): *Wstęp do metodologii badań bezpieczeństwa*, Biała Podlaska: Państwowa Szkoła Wyższa im. Papieża Jana Pawła II w Białej Podlaskiej.

informacyjnych i niewystarczająca świadomość społeczna mogą prowadzić do eskalacji incydentów, nawet w warunkach technicznie poprawnego działania systemów[22].

Przyjęta metodologia została dobrana w sposób świadomy i celowy, z uwzględnieniem specyfiki badań nad bezpieczeństwem w erze sztucznej inteligencji. Zastosowanie wyłącznie metod ilościowych byłoby niewystarczające ze względu na nowość zjawisk, brak ustabilizowanych definicji oraz dynamiczny charakter zagrożeń. Z kolei podejście wyłącznie teoretyczne nie pozwoliłoby na odniesienie rozważań do realiów funkcjonowania organizacji przemysłowych. Połączenie metod teoretycznych i empirycznych umożliwiło kompleksowe ujęcie problemu oraz zwiększyło wartość poznawczą i aplikacyjną monografii.

Kontynuacją przyjętego podejścia metodologicznego jest wyraźne uporządkowanie procesu badawczego w formie sekwencyjnych etapów, które umożliwiają systematyczne przejście od identyfikacji problemu do sformułowania wniosków teoretycznych i praktycznych. Zastosowanie etapowego modelu badań wynika z potrzeby zachowania logicznej spójności pomiędzy analizą literatury, diagnozą luk badawczych i praktycznych, eksploracją zjawisk empirycznych oraz budową autorskich ujęć interpretacyjnych. W naukach o bezpieczeństwie, w których badane zjawiska mają charakter dynamiczny i wielopoziomowy, takie podejście pozwala na stopniowe pogłębianie analizy oraz kontrolę poprawności wnioskowania.

Pierwszym etapem badań była systematyczna analiza literatury przedmiotu, obejmująca publikacje z zakresu nauk o bezpieczeństwie, cyberbezpieczeństwa, zarządzania ryzykiem, teorii informacji, kultury organizacyjnej oraz badań nad sztuczną inteligencją. Analiza ta miała na celu identyfikację dominujących paradygmatów badawczych, sposobów definiowania bezpieczeństwa informacyjnego oraz zakresu dotychczasowych badań nad zagrożeniami generowanymi przez systemy AI. Szczególną uwagę poświęcono publikacjom podejmującym problem autonomii algorytmów, generatywności systemów oraz ich wpływu na procesy decyzyjne i komunikacyjne w organizacjach. Efektem tego etapu było zidentyfikowanie luki badawczej i teoretycznej, polegającej na braku ujęć integrujących perspektywę technologiczną z analizą percepcji ryzyka i świadomości społecznej użytkowników systemów AI.

Drugim etapem badań była analiza regulacji prawnych, norm oraz dokumentów strategicznych odnoszących się do bezpieczeństwa informacji i wykorzystania sztucznej inteligencji. Etap ten obejmował analizę aktów prawnych, wytycznych instytucjonalnych oraz standardów odnoszących się do cyberbezpieczeństwa, ochrony danych, zarządzania ryzykiem oraz odpowiedzialności algorytmicznej. Celem tego etapu było określenie, w jakim stopniu istniejące ramy regulacyjne uwzględniają specyfikę zagrożeń algorytmicznych oraz czy dostarczają organizacjom narzędzi umożliwiających skuteczny nadzór nad systemami AI. Analiza regulacyjna pozwoliła również

[22] J. Kozłowski (2024): *Metody, techniki i narzędzia analityczne*, Warszawa: Wydawnictwo Akademii Sztuki Wojennej.

na identyfikację niespójności pomiędzy wymaganiami formalnymi a realnymi praktykami organizacyjnymi, co stanowi istotny element luki praktycznej w obszarze bezpieczeństwa.

Trzecim etapem badań było przeprowadzenie studiów przypadków organizacji przemysłowych wykorzystujących systemy sztucznej inteligencji w obszarach kluczowych dla bezpieczeństwa informacyjnego i operacyjnego. Studium przypadku umożliwiło analizę konkretnych sytuacji, w których działanie algorytmów wpływało na procesy decyzyjne, komunikacyjne lub operacyjne, ujawniając nowe kategorie zagrożeń. Analiza przypadków pozwoliła na uchwycenie złożonych relacji pomiędzy technologią, strukturą organizacyjną oraz zachowaniami pracowników. W szczególności umożliwiła identyfikację mechanizmów eskalacji ryzyka wynikających z niewystarczającej świadomości informacyjnej, nadmiernego zaufania do systemów AI lub braku procedur nadzoru nad ich działaniem.

Czwartym etapem była jakościowa ocena ryzyka, prowadzona na podstawie zidentyfikowanych scenariuszy zagrożeń. Ocena ta miała charakter eksploracyjny i koncentrowała się na identyfikacji potencjalnych konsekwencji błędów algorytmicznych, nieprawidłowej interpretacji danych oraz zakłóceń w przepływie informacji. Uwzględniono zarówno ryzyka o charakterze technicznym, jak i społecznym, w tym ryzyka wynikające z zachowań użytkowników, braku kompetencji informacyjnych oraz niedojrzałej kultury bezpieczeństwa. Etap ten pozwolił na powiązanie analiz teoretycznych z praktyką funkcjonowania organizacji oraz na wskazanie obszarów szczególnie podatnych na zagrożenia w warunkach wykorzystania sztucznej inteligencji.

Ostatnim etapem procesu badawczego była synteza uzyskanych wyników, prowadząca do sformułowania wniosków teoretycznych oraz rekomendacji praktycznych. Synteza umożliwiła integrację wyników analizy literatury, regulacji, studiów przypadków i oceny ryzyka w spójny model interpretacyjny opisujący bezpieczeństwo organizacji w erze cyberzagrożeń generowanych przez AI. Na tym etapie sformułowano również propozycje modeli przeciwdziałania zagrożeniom, uwzględniających integrację technologii, procedur oraz świadomości społecznej jako filarów odporności organizacyjnej. Sekwencyjny charakter badań pozwolił na zachowanie przejrzystości metodologicznej oraz zwiększył wiarygodność uzyskanych wniosków.

Przyjęta metodologia umożliwia nie tylko opis i wyjaśnienie analizowanych zjawisk, lecz także ich interpretację w perspektywie prognostycznej. W naukach o bezpieczeństwie szczególnie istotne jest bowiem nie tylko rozpoznanie aktualnych zagrożeń, lecz również zdolność do identyfikowania trendów i potencjalnych scenariuszy rozwoju ryzyka. Zastosowanie metod jakościowych, podejścia systemowego oraz analizy etapowej pozwala na uchwycenie procesualnego charakteru bezpieczeństwa organizacyjnego oraz na wskazanie kierunków dalszych badań w obszarze zagrożeń generowanych przez sztuczną inteligencję. W celu uporządkowania zastosowanego podejścia metodologicznego oraz zapewnienia przejrzystości procesu badawczego zestawiono metody teoretyczne i empiryczne wykorzystane w monografii. Ich zakres zastosowania,

cele poznawcze oraz powiązanie z poszczególnymi etapami badań przedstawiono w ***tabeli W.1***.

Tabela W.1. Zastosowane metody badawcze i ich rola w realizacji celów monografii

Metoda badawcza	Zakres zastosowania	Cel poznawczy	Efekt badawczy	Powiązanie z rozdziałami
Analiza	Literatura naukowa, raporty, dokumenty strategiczne	Identyfikacja zagrożeń, pojęć i luk badawczych	Wyodrębnienie kluczowych kategorii ryzyka AI	Wstęp, Rozdział 1
Synteza	Wyniki analiz teoretycznych i empirycznych	Integracja wiedzy interdyscyplinarnej	Autorski model bezpieczeństwa organizacyjnego	Rozdziały 1-5
Porównanie	Klasyczne i nowe modele bezpieczeństwa	Ocena adekwatności istniejących podejść	Wskazanie ograniczeń tradycyjnych systemów	Rozdział 2
Abstrahowanie	Zjawiska bezpieczeństwa AI niezależne od branży	Uogólnienie mechanizmów zagrożeń	Model uniwersalny dla organizacji przemysłowych	Rozdział 3
Wnioskowanie dedukcyjne i indukcyjne	Teoria i dane empiryczne	Budowa i weryfikacja hipotez	Uzasadnienie zależności między AI i ryzykiem	Wstęp, Rozdział 4
Analogia	Porównanie etapów rozwoju zagrożeń technologicznych	Identyfikacja trendów i scenariuszy	Prognoza cyberzagrożeń 6.0	Rozdział 5
Studium przypadku	Organizacje przemysłowe wykorzystujące AI	Analiza rzeczywistych incydentów	Identyfikacja luk praktycznych	Rozdział 3
Analiza regulacji	Prawo, normy, wytyczne	Ocena ram formalnych bezpieczeństwa	Wskazanie niespójności regulacyjnych	Rozdział 4
Jakościowa ocena ryzyka	Scenariusze zagrożeń algorytmicznych	Identyfikacja podatności systemowych	Mapowanie ryzyk AI	Rozdział 4

W ***tabeli W.1*** zaprezentowano zestaw metod badawczych wykorzystanych w monografii wraz z ich zakresem zastosowania, celem poznawczym oraz efektami badawczymi. Zestawienie ukazuje komplementarność metod teoretycznych i empirycznych oraz ich powiązanie z poszczególnymi rozdziałami pracy, co potwierdza systemowy charakter przyjętej metodologii.

Uzasadnienie doboru materiału badawczego i źródeł wykorzystanych w niniejszej monografii wynika bezpośrednio z charakteru analizowanego problemu badawczego oraz z przyjętej perspektywy nauk o bezpieczeństwie. Badanie zagrożeń generowanych przez systemy sztucznej inteligencji w organizacjach przemysłowych wymaga

bowiem sięgnięcia do zróżnicowanych kategorii źródeł, które odzwierciedlają zarówno normatywny, technologiczny, jak i społeczno-organizacyjny wymiar bezpieczeństwa. Ograniczenie materiału badawczego wyłącznie do jednej grupy źródeł, na przykład literatury naukowej lub dokumentów technicznych, prowadziłoby do redukcjonistycznego ujęcia zjawiska i uniemożliwiłoby uchwycenie jego systemowego charakteru.

Pierwszą zasadniczą kategorię materiału badawczego stanowią dokumenty regulacyjne i normatywne, w szczególności akty prawne oraz dokumenty strategiczne odnoszące się do bezpieczeństwa informacji, cyberbezpieczeństwa oraz wykorzystania sztucznej inteligencji. Włączenie tej grupy źródeł do analizy jest uzasadnione faktem, że regulacje prawne wyznaczają formalne ramy funkcjonowania organizacji, określają obowiązki w zakresie zarządzania ryzykiem oraz wpływają na sposób projektowania i eksploatacji systemów AI. Analiza dokumentów takich jak unijny akt o sztucznej inteligencji, dyrektywa NIS2 czy powiązane regulacje dotyczące ochrony danych i infrastruktury krytycznej pozwala na ocenę, w jakim stopniu obowiązujące i projektowane normy odpowiadają rzeczywistym zagrożeniom algorytmicznym oraz czy dostarczają organizacjom narzędzi umożliwiających skuteczny nadzór nad systemami autonomicznymi. Dokumenty regulacyjne są w tym kontekście nie tylko źródłem informacji normatywnej, lecz także istotnym elementem systemu bezpieczeństwa, który wpływa na praktyki organizacyjne i strukturę odpowiedzialności.

Drugą istotną grupę materiału badawczego stanowią raporty dotyczące incydentów związanych z wykorzystaniem systemów sztucznej inteligencji oraz szerzej rozumianych zdarzeń cyberbezpieczeństwa. Wybór tej kategorii źródeł wynika z potrzeby osadzenia rozważań teoretycznych w realnych przykładach funkcjonowania organizacji w warunkach zagrożeń. Raporty z incydentów, publikowane przez instytucje publiczne, zespoły reagowania na incydenty bezpieczeństwa, organizacje branżowe oraz podmioty zajmujące się analizą ryzyka technologicznego, umożliwiają identyfikację powtarzalnych schematów zagrożeń, mechanizmów eskalacji ryzyka oraz konsekwencji błędów algorytmicznych i organizacyjnych. Materiały te pozwalają również na analizę rozbieżności pomiędzy deklarowanym poziomem bezpieczeństwa a rzeczywistą odpornością organizacji na zagrożenia generowane przez systemy AI.

Uwzględnienie raportów incydentalnych jest szczególnie istotne w kontekście nauk o bezpieczeństwie, które kładą nacisk na analizę zdarzeń krytycznych jako źródła wiedzy o słabościach systemów ochrony. Incydenty związane z błędnymi rekomendacjami algorytmicznymi, niewłaściwą interpretacją danych generowanych przez AI, automatyzacją błędów decyzyjnych czy zakłóceniami w procesach przemysłowych ujawniają ograniczenia zarówno technologiczne, jak i organizacyjne. Analiza tych materiałów umożliwia przejście od abstrakcyjnych rozważań o ryzyku do konkretnych scenariuszy zagrożeń, które mogą materializować się w środowisku przemysłowym.

Trzecią kluczową kategorię źródeł stanowi literatura naukowa, obejmująca publikacje z lat 2020-2024. Ograniczenie zakresu czasowego do ostatnich lat jest uzasadnione dynamicznym rozwojem technologii sztucznej inteligencji oraz zmianami w podejściu

do bezpieczeństwa informacyjnego i cyberbezpieczeństwa. Wcześniejsze opracowania, choć istotne z punktu widzenia rozwoju teorii, nie uwzględniają w pełni zjawisk związanych z generatywnymi modelami AI, rosnącą autonomią algorytmów oraz ich integracją z systemami przemysłowymi. Dobór literatury z najnowszego okresu pozwala na uchwycenie aktualnych trendów badawczych, identyfikację nowych kategorii zagrożeń oraz ocenę, w jakim stopniu współczesne teorie bezpieczeństwa odpowiadają wyzwaniom wynikającym z transformacji technologicznej.

Literatura naukowa została wykorzystana nie tylko jako źródło definicji i koncepcji teoretycznych, lecz także jako punkt odniesienia do krytycznej analizy istniejących modeli bezpieczeństwa. Szczególną uwagę poświęcono publikacjom z zakresu nauk o bezpieczeństwie, zarządzania ryzykiem, teorii systemów, badań nad kulturą organizacyjną oraz studiów nad interakcją człowiek–AI. Tak dobrany korpus literatury umożliwia integrację różnych perspektyw badawczych i sprzyja budowie interdyscyplinarnego ujęcia bezpieczeństwa organizacyjnego.

Czwartą kategorię materiału badawczego stanowią wyniki analiz technicznych systemów sztucznej inteligencji, w szczególności opracowania dotyczące architektury modeli, mechanizmów uczenia, podatności algorytmicznych oraz ograniczeń systemów autonomicznych. Włączenie tej grupy źródeł jest niezbędne, aby uniknąć nadmiernie socjologicznego lub normatywnego ujęcia problemu bezpieczeństwa. Zrozumienie mechanizmów działania systemów AI, ich zależności od danych uczących, podatności na manipulacje oraz potencjalnych błędów jest warunkiem rzetelnej analizy ryzyka algorytmicznego. Analizy techniczne stanowią zatem uzupełnienie badań organizacyjnych i pozwalają na lepsze wyjaśnienie źródeł zagrożeń, które następnie manifestują się na poziomie operacyjnym i społecznym.

Dobór wskazanych kategorii źródeł został dokonany w sposób celowy i uzasadniony potrzebą triangulacji danych. Zastosowanie różnych typów materiałów badawczych pozwala na wzajemną weryfikację wniosków oraz ogranicza ryzyko jednostronnej interpretacji zjawisk. Dokumenty regulacyjne pokazują formalne ramy bezpieczeństwa, raporty incydentów ujawniają praktyczne problemy i słabości systemów, literatura naukowa dostarcza aparatu pojęciowego i koncepcji teoretycznych, natomiast analizy techniczne umożliwiają zrozumienie źródeł zagrożeń na poziomie algorytmicznym. Taka konfiguracja źródeł jest szczególnie adekwatna dla badań prowadzonych w naukach o bezpieczeństwie, gdzie kluczowe znaczenie ma łączenie perspektywy normatywnej, empirycznej i systemowej.

Przyjęty dobór materiału badawczego umożliwia realizację celów monografii zarówno w wymiarze poznawczym, jak i aplikacyjnym. Pozwala on na identyfikację luk w istniejących systemach bezpieczeństwa, ocenę skuteczności obowiązujących regulacji oraz sformułowanie rekomendacji odnoszących się do praktyki funkcjonowania organizacji przemysłowych. Jednocześnie zapewnia solidne podstawy do wnioskowania teoretycznego i budowy modeli interpretacyjnych, które mogą stanowić punkt

wyjścia do dalszych badań nad bezpieczeństwem organizacji w warunkach intensywnego wykorzystania sztucznej inteligencji.

Założenia koncepcyjne monografii wpisują się w ten globalny trend poszukiwania podejścia wielowymiarowego. Założono tutaj, że systemy sztucznej inteligencji działają jako aktorzy w organizacyjnych ekosystemach informacyjnych, a nie jako narzędzia, które można kontrolować w sposób całkowicie przewidywalny. Oznacza to, że technologia nie jest już wyłącznie elementem infrastruktury, ale staje się współtwórcą informacji, rekomendacji i decyzji. Takie ujęcie jest zgodne z koncepcjami teoriopoznawczymi, które wskazują na konieczność analizy technologii jako elementu wpływającego na procesy poznawcze i interpretacyjne pracowników. W praktyce oznacza to, że bezpieczeństwo organizacji zależy nie tylko od stabilności systemów technicznych, ale przede wszystkim od tego, jak użytkownicy interpretują generowane przez AI informacje oraz jakie decyzje podejmują na ich podstawie.

Z tego względu monografia przyjmuje, że kluczową rolę w bezpieczeństwie przemysłowym odgrywa świadomość społeczna. Świadomość ta – będąc jednocześnie poznawczą, informacyjną i operacyjną – umożliwia użytkownikom identyfikację potencjalnych zagrożeń, interpretację sygnałów generowanych przez algorytmy, a także ocenę jakości informacji. Jest to świadomość, która wykracza poza tradycyjną kulturę bezpieczeństwa, ponieważ koncentruje się nie na normach społecznych czy zachowaniach zbiorowych, ale na zdolności jednostek i zespołów do krytycznej analizy danych oraz zrozumienia zasad działania technologii. W sektorach przemysłowych, gdzie decyzje podejmowane są często pod presją czasu, a systemy technologiczne wprowadzają dodatkową złożoność, świadomość ta staje się nieodzownym elementem odporności organizacyjnej.

Istotna jest również analiza roli regulacji w kształtowaniu bezpieczeństwa organizacyjnego. Powstanie i rozwój aktu o sztucznej inteligencji (AI Act), NIS2, RODO oraz szeregu norm branżowych świadczą o tym, że państwa i organizacje ponadnarodowe dostrzegają konieczność wprowadzenia środków formalnych ograniczających ryzyka generowane przez autonomiczne systemy. Regulacje te pełnią funkcję stabilizującą, definiując standardy bezpieczeństwa, odpowiedzialności i nadzoru. Jednak ich skuteczność zależy od kilku czynników: od jakości implementacji, od zgodności między wymogami regulacyjnymi a rzeczywistym działaniem systemów technologicznych oraz od poziomu świadomości użytkowników. Regulacje nie zastąpią wiedzy, kompetencji i zdolności pracowników do właściwego reagowania na sygnały płynące z systemów AI. Z tego powodu monografia wskazuje na ścisły związek między regulacjami a świadomością społeczną, traktując je jako komplementarne elementy systemu bezpieczeństwa.

W niniejszej monografii przyjęto także, że przyszłość bezpieczeństwa organizacyjnego należy analizować w odniesieniu do koncepcji Industry 6.0, która zakłada pełną integrację człowieka, technologii i zrównoważonego rozwoju w celu tworzenia systemów odpornych, przejrzystych i odpowiedzialnych społecznie. Industry 6.0 nie jest

jedynie kolejną fazą automatyzacji, lecz paradygmatem, który włącza elementy etyczne, społeczne, środowiskowe i technologiczne do zarządzania systemami przemysłowymi. W tym kontekście sztuczna inteligencja nie jest celem sama w sobie, lecz narzędziem służącym do budowania systemów odpornych i zorientowanych na człowieka. Monografia rozwija te założenia, analizując, jakie kompetencje społeczne, regulacyjne i organizacyjne są niezbędne do funkcjonowania w takim modelu przemysłu oraz jakie zagrożenia mogą pojawić się na drodze do jego realizacji.

Ważną częścią niniejszej pracy jest zatem wskazanie, że współczesne zagrożenia informacyjne i technologiczne nie są statyczne. Mają charakter dynamiczny, adaptacyjny i często niejednoznaczny. Zagrożenia generowane przez AI mogą rozwijać się szybciej niż zdolności organizacji do ich identyfikacji, co tworzy przestrzeń dla incydentów trudnych do przewidzenia lub klasyfikacji. Oznacza to, że konieczne jest budowanie systemowej odporności, obejmującej zarówno odporność technologiczną (związaną z jakością i stabilnością systemów), jak i odporność społeczną (związaną ze świadomością, kompetencjami i dojrzałością informacyjną pracowników).

Na tej podstawie można stwierdzić, że niniejsza monografia rozwija nowatorskie spojrzenie na bezpieczeństwo organizacyjne, proponując podejście integrujące technologię, prawo, zarządzanie i wiedzę użytkowników. Jest to podejście, które wpisuje się w interdyscyplinarny charakter nauk o bezpieczeństwie i jednocześnie odpowiada na bieżące potrzeby sektora przemysłowego. W praktyce oznacza to, że w monografii nie tylko dokonano analizy zagrożenia i regulacji, ale przede wszystkim wskazano, jak organizacje mogą budować modele odporności adekwatne do wyzwań XXI wieku.

Ostatnią, kluczową funkcją tej części wstępu jest stworzenie pomostu między teorią a strukturą dalszych rozdziałów. Kolejne części monografii rozwijają szczegółowo przedstawione tu zagadnienia, pokazując, jak ewolucja przemysłu wpływa na charakter zagrożeń, w jaki sposób systemy AI generują nowe podatności, jaką rolę odgrywa świadomość społeczna oraz jak budować odporność organizacyjną w erze systemów autonomicznych. W rozdziale pierwszym czytelnik zostanie wprowadzony w zjawisko przemysłowej transformacji oraz genezę cyberzagrożeń 6.0. W rozdziale drugim przedstawiono technologiczne aspekty zagrożeń, w rozdziale trzecim skupiono się na świadomości społecznej, w rozdziale czwartym na regulacjach, w rozdziale piątym na modelach przeciwdziałania, a w rozdziale szóstym na przyszłych scenariuszach Industry 6.0. Taka struktura monografii umożliwia całościowe, wielowymiarowe spojrzenie na zagadnienie bezpieczeństwa w erze sztucznej inteligencji.

W miarę pogłębiania się procesów cyfrowej transformacji organizacje przemysłowe stoją przed szczególną koniecznością nie tylko wdrażania nowych technologii, lecz także rozwijania zdolności ciągłego monitorowania ich wpływu na bezpieczeństwo operacyjne. Sztuczna inteligencja staje się integralnym elementem infrastruktury organizacyjnej, a jej działanie wpływa na środowisko pracy, strukturę zarządzania informacją oraz sposób podejmowania decyzji. AI zmienia również to, jak organizacje rozumieją zagrożenia i jak je klasyfikują, ponieważ tradycyjne podejścia nie obejmują

takich zjawisk jak autonomiczne rekomendacje, samodzielne generowanie treści, interaktywne systemy komunikacyjne czy algorytmy predykcyjne, które mogą zmieniać decyzje operacyjne w czasie rzeczywistym.

W tej rzeczywistości powstaje potrzeba opracowania nowego podejścia teoretycznego, które pozwoli analizować systemy sztucznej inteligencji jako czynnik dynamicznie kształtujący bezpieczeństwo. Zagrożenia generowane przez AI nie są stałe, przewidywalne ani powtarzalne. Systemy te ulegają ciągłej modyfikacji poprzez procesy uczenia maszynowego, adaptacji do nowych danych, a także konfigurowania przez użytkowników, którzy nadają im kolejne cele operacyjne. Oznacza to, że współczesne zagrożenia technologiczne mają wymiar ewolucyjny i ich analiza wymaga elastycznych i wielowymiarowych modeli badawczych. Nauki o bezpieczeństwie muszą uwzględnić fakt, że bezpieczeństwo w erze AI jest procesem ciągłym, podatnym na dynamiczne zmiany i skłonności, które mogą ujawnić się dopiero w określonym kontekście operacyjnym.

Kolejnym istotnym aspektem jest rola, jaką odgrywa systemowa integracja technologii i człowieka. Sztuczna inteligencja przejmuje część funkcji poznawczych, analitycznych i decyzyjnych, ale nie eliminuje człowieka z systemu bezpieczeństwa. Przeciwnie – zwiększa jego rolę, ponieważ to użytkownicy muszą nadzorować działanie systemów, oceniać jakość danych, interpretować wyniki generowane przez algorytmy i identyfikować sygnały niewłaściwego lub ryzykownego działania AI. W literaturze podkreśla się, że w sytuacjach, gdy systemy AI generują treści nieadekwatne, nadmiernie uproszczone lub błędne, użytkownik staje się ostatnim elementem weryfikacji poprawności informacji. Jeśli jednak brakuje mu odpowiednich kompetencji, podejmuje decyzje na podstawie komunikatów algorytmicznych obarczonych błędem, a w konsekwencji ryzyko operacyjne radykalnie wzrasta.

W tym kontekście należy również zwrócić uwagę na zjawisko tzw. podwójnego obciążenia poznawczego. Współczesny pracownik przemysłowy musi równocześnie wykonywać określone zadania operacyjne oraz interpretować coraz bardziej złożone treści generowane przez systemy AI. Z jednej strony systemy te mają wspierać pracownika, odciążając go od analizy danych. Z drugiej strony wprowadzają dodatkowe bodźce informacyjne, które mogą skutkować przeciążeniem informacyjnym, błędnymi decyzjami lub niewłaściwymi interpretacjami. W literaturze wskazuje się, że organizacje, które nie kontrolują tego procesu, narażają się na wzrost liczby incydentów wynikających nie z błędów systemów, lecz z błędów wynikających z braku zdolności użytkowników do prawidłowej interpretacji wyników działania AI.

Nie można pominąć również istotnej roli jakości danych. Sztuczna inteligencja działa na podstawie informacji, które otrzymuje, dlatego jakość danych determinująca jakość działania modeli musi stać się przedmiotem analiz bezpieczeństwa. Dane błędne, niekompletne, nieaktualne, stronnicze lub zmanipulowane mogą prowadzić do powstania nieprawidłowych decyzji, które z kolei mogą destabilizować procesy organizacyjne. Konieczne jest więc opracowanie metod kontroli danych oraz ich walidacji, ponieważ

stanowią one fundament dla wszystkich modeli predykcyjnych i rekomendacyjnych. W monografii podkreślono, że bezpieczeństwo nie może być definiowane bez odniesienia do jakości informacji, które napędzają systemy AI, oraz bez analizy tego, w jaki sposób dane są gromadzone, przetwarzane i interpretowane.

W procesach przemysłowych szczególnie istotnym elementem jest wpływ AI na systemy monitorowania i kontroli produkcji[23]. Systemy sterowania mogą wykorzystywać algorytmy predykcyjne do przewidywania zużycia maszyn, diagnostyki awarii, optymalizacji pracy linii produkcyjnych lub analizowania danych środowiskowych. Jeśli algorytm wygeneruje błędne przewidywania, konsekwencje mogą być poważne – od przestojów produkcyjnych po zagrożenia w zakresie bezpieczeństwa pracy. W niniejszej monografii analizowane jest, w jaki sposób organizacje mogą wdrożyć mechanizmy nadzoru (human oversight) i kontroli, które będą w stanie wychwycić odchylenia w działaniu systemów predykcyjnych i zapobiec ich eskalacji do incydentów operacyjnych.

Kolejnym kluczowym obszarem analizy jest wpływ sztucznej inteligencji na komunikację organizacyjną. W coraz większej liczbie przedsiębiorstw systemy AI generują komunikaty operacyjne, raporty, alerty, rekomendacje lub instrukcje dla użytkowników. Systemy te mogą także pośredniczyć w komunikacji między pracownikami lub między zespołami. Jeśli komunikaty takie są błędne, nieprecyzyjne lub nieadekwatne, mogą wprowadzać chaos informacyjny, zniekształcać interpretację sytuacji, prowadzić do błędów operacyjnych lub utraty zaufania do systemów technologicznych. W tym kontekście w monografii dokonano analizy zagrożeń komunikacyjnych generowanych przez AI oraz roli świadomości użytkowników w ich identyfikacji.

Ważnym elementem jest również relacja między sztuczną inteligencją a regulacjami bezpieczeństwa. Wprowadzenie AI Act oznacza konieczność identyfikowania systemów wysokiego ryzyka, monitorowania ich działania, dokumentowania procesu trenowania algorytmów oraz wprowadzenia nadzoru człowieka. W monografii przeanalizowano, w jaki sposób przedsiębiorstwa przemysłowe mogą efektywnie integrować regulacje prawne z praktykami bezpieczeństwa informacyjnego, biorąc pod uwagę, że regulacje te wprowadzają nowe obowiązki, takie jak konieczność zapewnienia przejrzystości algorytmów, możliwości ich audytu oraz ograniczenia ryzyk wynikających z autonomicznego działania.

Rozwijając te zagadnienia, w monografii wskazano również, że systemy przemysłowe są szczególnie wrażliwe na zakłócenia wynikające z nieautoryzowanego wpływu na modele AI. Przykłady obejmują manipulację danymi wejściowymi, zatruwanie modeli (data poisoning), generowanie fałszywych sygnałów operacyjnych, zakłócanie procesów decyzyjnych oraz inicjowanie działań, które prowadzą do zaburzeń logistycznych. Dla nauk o bezpieczeństwie oznacza to, że zagrożenia muszą być

[23] B. Wiśniewski (2023): *Transfer wiedzy w bezpieczeństwie informacyjnym*, „Zeszyty Naukowe Pro Publico Bono", t. 1(1), s. 413-418. DOI: 10.5604/01.3001.0054.1732.

rozpatrywane nie tylko w kontekście cyberataków, ale także w kontekście błędów wynikających z interakcji pomiędzy algorytmem a użytkownikiem.

W perspektywie Industry 6.0, którego założeniem jest integracja człowieka i technologii w sposób zrównoważony, bezpieczny i przejrzysty, szczególne znaczenie ma budowanie kompetencji informacyjnych i technologicznych pracowników[24]. Oznacza to, że organizacje muszą inwestować nie tylko w infrastrukturę technologiczną, ale także w rozwój świadomości, wiedzy i zdolności interpretacyjnych użytkowników. Tylko wówczas można stworzyć środowisko zdolne do skutecznego funkcjonowania w warunkach wysokiej autonomii technologicznej.

Niniejsza monografia stanowi oryginalny i samodzielny wkład w rozwój dyscypliny nauk o bezpieczeństwie, zarówno w wymiarze teoretycznym, jak i metodologicznym oraz aplikacyjnym. Jej nowatorstwo polega na całościowym i systemowym ujęciu bezpieczeństwa informacyjnego w organizacjach przemysłowych funkcjonujących w warunkach transformacji określanej mianem Przemysłu 6.0, z uwzględnieniem roli sztucznej inteligencji jako czynnika generującego jakościowo nowe kategorie zagrożeń oraz znaczenia świadomości społecznej jako kluczowej determinanty skuteczności systemów bezpieczeństwa. W dostępnym dorobku naukowym nie występuje dotychczas monografia, która w sposób spójny i pogłębiony integrowałaby te trzy obszary badawcze w ramach jednej koncepcji analitycznej, osadzonej explicite w realiach organizacji przemysłowych.

Dotychczasowe opracowania koncentrują się fragmentarycznie na wybranych aspektach bezpieczeństwa, takich jak cyberbezpieczeństwo infrastruktury technicznej, ochrona systemów OT i IoT, ryzyka technologiczne związane z algorytmami lub regulacyjne uwarunkowania rozwoju sztucznej inteligencji. Brakuje jednak ujęć, które traktowałyby bezpieczeństwo organizacyjne jako zjawisko socjotechniczne, wynikające z interakcji pomiędzy technologią, strukturą organizacyjną i człowiekiem jako podmiotem interpretującym informacje i podejmującym decyzje. Niniejsza monografia wypełnia tę lukę, przesuwając punkt ciężkości analizy z samej infrastruktury technicznej na mechanizmy poznawcze, kulturowe i organizacyjne, które w praktyce decydują o odporności organizacji na zagrożenia generowane przez systemy sztucznej inteligencji.

Szczególny nowatorski charakter badań polega na wykazaniu, że zagrożenia algorytmiczne nie stanowią jedynie kolejnej kategorii zagrożeń technologicznych, lecz tworzą nową klasę ryzyka o charakterze systemowym, którego źródłem jest autonomia procesów decyzyjnych wspieranych przez AI oraz ograniczona zdolność użytkowników do krytycznej interpretacji wyników generowanych przez algorytmy. W monografii dowiedziono, że w warunkach Przemysłu 6.0 skuteczność systemów bezpieczeństwa nie jest funkcją wyłącznie poziomu zaawansowania technologicznego, lecz w decydującym stopniu zależy od poziomu świadomości społecznej, kompetencji

[24] T. Gajewski (2020): *Towards resilience…*, op. cit.

informacyjnych oraz kultury bezpieczeństwa funkcjonującej w organizacji. Takie ujęcie stanowi istotne rozszerzenie dotychczasowych paradygmatów nauk o bezpieczeństwie, w których czynnik ludzki był najczęściej analizowany w kategoriach błędu lub najsłabszego ogniwa systemu.

Wkład autorski monografii przejawia się również w opracowaniu autorskiej koncepcji relacji pomiędzy sztuczną inteligencją, percepcją ryzyka a odpornością organizacyjną. Koncepcja ta pozwala wyjaśnić mechanizmy, w których formalna zgodność z normami, regulacjami i standardami bezpieczeństwa nie przekłada się na realne bezpieczeństwo organizacji, a incydenty eskalują mimo poprawnego funkcjonowania systemów technicznych. Wykazano, że brak integracji technologii z procesami edukacyjnymi, komunikacyjnymi i kulturowymi prowadzi do powstawania ukrytych podatności, które nie są identyfikowane w tradycyjnych analizach ryzyka. Tym samym monografia wnosi do nauk o bezpieczeństwie nową perspektywę interpretacyjną, umożliwiającą bardziej realistyczną ocenę zagrożeń w środowiskach wysokiej automatyzacji.

Nowatorstwo pracy przejawia się także w wykorzystaniu koncepcji Przemysłu 6.0 jako ramy analitycznej dla badań nad bezpieczeństwem. W odróżnieniu od dominujących w literaturze ujęć technologicznych i futurologicznych Przemysł 6.0 został w niniejszej monografii potraktowany jako kontekst systemowy, w którym bezpieczeństwo informacyjne staje się kluczowym warunkiem ciągłości działania organizacji. Takie osadzenie badań pozwala na identyfikację zagrożeń, które nie są widoczne w ramach wcześniejszych modeli przemysłowych, a które wynikają z głębokiej integracji sztucznej inteligencji z procesami decyzyjnymi, komunikacyjnymi i zarządczymi.

Podsumowując: niniejsza monografia stanowi samodzielne, oryginalne opracowanie, które poszerza granice poznawcze nauk o bezpieczeństwie poprzez wprowadzenie zintegrowanego ujęcia bezpieczeństwa informacyjnego, zagrożeń algorytmicznych oraz świadomości społecznej w realiach Przemysłu 6.0. Zaproponowane podejście nie tylko uzupełnia istniejące luki teoretyczne i praktyczne, lecz także tworzy podstawy do dalszych badań nad bezpieczeństwem organizacji w warunkach rosnącej autonomii systemów sztucznej inteligencji. Tym samym monografia wnosi trwały wkład w rozwój dyscypliny, odpowiadając na jedno z kluczowych wyzwań współczesnego bezpieczeństwa organizacyjnego.

Monografia zawiera rysunki i tabele opracowane zarówno na podstawie źródeł literaturowych, jak i w wyniku własnych analiz autorki. Materiały wykorzystane z literatury zostały opatrzone odpowiednimi odwołaniami do źródeł. Rysunki i tabele nieposiadające oznaczenia źródła należy traktować jako opracowanie własne.