

HORYZONTY AI, cyberbezpieczeństwo i technologie przyszłości

Redakcja:
Zdzisława Dacko-Pikiewicz,
Katarzyna Szczepańska-Woszczyna, Marcin Lis

HORYZONTY AI,
cyberbezpieczeństwo
i technologie przyszłości

HORYZONTY AI, cyberbezpieczeństwo i technologie przyszłości

REDAKCJA

Zdzisława Dacko-Pikiewicz
Katarzyna Szczepańska-Woszczyna
Marcin Lis

Akademia WSB
WSB University

Dąbrowa Górnicza 2026

Horyzonty AI, cyberbezpieczeństwo i technologie przyszłości

Redakcja

dr hab. Zdzisława Dacko-Pikiewicz, prof. AWSB

dr hab. Katarzyna Szczepańska-Woszczyna, prof. AWSB

dr hab. Marcin Lis, prof. AWSB

Recenzja

dr hab. Barbara Siuta-Tokarska, prof. UEK

Korekta

Anna Zdonek

Projekt okładki

Wojciech Ciągło Studio DTP

DTP publikacji

Wojciech Ciągło Studio DTP, www.dtp-studio.pl

ISBN 978-83-67673-78-5

e-ISBN 978-83-67673-85-3

Wydawca

Akademia WSB

ul. Cieplaka 1c, 41-300 Dąbrowa Górnicza, tel. (32) 295 93 59

e-mail: wydawnictwo@wsb.edu.pl, www.wsb.edu.pl

© Copyright by Akademia WSB

Kopiowanie w całości lub we fragmentach zabronione

Dąbrowa Górnicza 2026

Druk i oprawa

Totem.com.pl

Spis treści

EDYTA ABRAMEK

Rozwój kompetencji logistyków pod wpływem AI: ocena programów kształcenia metodą AHP 8

Development of Logistics Professionals' Competencies Under the Influence of AI: Evaluation of Educational Programs Using the AHP Method

GRZEGORZ GAWĘŁ

Sztuczna inteligencja jako narzędzie do zarządzania procesem utrzymania ruchu w zakresie wind zabudowanych na stacjach i przystankach w Polsce. Kontekst innowacyjności podejścia do infrastruktury transportowej 30

Artificial Intelligence as a Tool for Managing the Maintenance Process of Elevators Installed at Railway Stations and Stops in Poland: An Innovative Approach to Transport Infrastructure

JUDYTA KABUS

Dynamika innowacyjności przedsiębiorstw w Polsce w ujęciu regionalnym 38

The Dynamics of Innovation in Polish Enterprises in Regional Terms

MARCIN CYWIŃSKI

Poprawa efektywności procesu kompletacji towarów przy wykorzystaniu Internetu rzeczy oraz protokołu RFID, Z-Wave i 6LoWPAN 58

Improving the Efficiency of the Picking Process Using the Internet of Things, RFID, Z-Wave and 6LoWPAN Protocols

ADAM FILIPOWSKI, MARTA RAPACZ-MAHMOUD

Ograniczenia decyzyjne polskich funduszy venture capital w finansowaniu przełomowych technologii: studium przypadku ElevenLabs 70

Decision-Making Constraints of Polish Venture Capital Funds in Financing Breakthrough Technologies: ElevenLabs Case Study

MARTYNA OSTERN

Transformacja ról pracowników Centrów Usług Wspólnych w erze AI: trendy i nowe kompetencje 84

The Transformation of Employee Roles in Shared Services Centers in the Age of AI: Trends and Emerging Competencies

PAWEŁ JANAS

Mały uczeń w wielkim świecie AI: wyzwania i szanse cyfrowej rewolucji w edukacji wczesnodziecięcej 102

The Young Learner in the Vast World of AI: Challenges and Opportunities of the Digital Revolution in Early Childhood Education

MACIEJ RYFA

Lean Management, cyfryzacja i cyberbezpieczeństwo – nowoczesne podejście do doskonalenia produkcji 116

Lean Management, Digitization and Cybersecurity – A Modern Approach to Production Excellence

DOTA SZYMBORSKA, JULIA JASTRZĘBSKA

Między pamiętnikiem a chatbotem: o kobiecym głosie w epoce generatywnej AI 128

Between the Diary and the Chatbot: On the Female Voice in the Age of Generative AI

MAGDALENA SZCZYRBA

Europejska Przestrzeń Danych Zdrowotnych w Polsce – kontekst innowacyjności i cyfryzacji w świetle modelu Serendipity 144

Serendipity, Data and Innovation: EHDS in the Polish Context

KRZYSZTOF WRANA

Sztuczna inteligencja we wzmacnianiu partycypacji mieszkańców w zarządzaniu rozwojem miasta – perspektywy i ryzyka 162

Artificial Intelligence in Strengthening Citizen Participation in Urban Development Governance – Perspectives and Risks

MONIKA PŁOŃSKA, GRZEGORZ KĄDZIELAWSKI, MARCIN LIS

AI i blockchain w zarządzaniu systemami certyfikacji offsetów węglowych 180

AI and Blockchain in the Management of Carbon Offset Certification Systems

KAROL JĘDRASIAK

Od widma do znaczenia:

analiza cech semantycznych i spektralnych w detekcji deepfake 192

From Spectrum to Meaning:

Analysis of Semantic and Spectral Features in Deepfake Detection

JAROSŁAW KARSKI

**Wykorzystanie sztucznej inteligencji (AI) w zarządzaniu finansami
przedsiębiorstwa przy wsparciu systemu SAP – analiza wybranych
zastosowań i przypadków 226**

The Use of Artificial Intelligence (AI) in Enterprise Financial

Management Supported by the SAP System: Analysis of Selected

Applications and Case Studies

GRZEGORZ PŁACZEK

Zarządzanie wdrożeniami sztucznej inteligencji (ZT-AI)

w przedsiębiorstwach sektora energetycznego 244

Management of ZT-AI Implementations in Energy Companies

MARIUSZ RACZEK

**Industy 5.0 i sztuczna inteligencja jako źródło wyzwań dla polityki
rozwoju regionalnego województwa śląskiego 258**

Industry 5.0 and Artificial Intelligence as a Source of Challenges

for the Regional Development Policy of the Silesian Voivodeship

WALDEMAR BOJARUN

**Katowice w cyfrowej transformacji administracji samorządowej –
rola Centrów Usług Wspólnych w rozwoju miasta przyszłości 274**

Katowice in the Digital Transformation of Local Government

Administration – the Role of Shared Services Centres in the Development
of the City of the Future

Rozwój kompetencji logistyków pod wpływem AI: ocena programów kształcenia metodą AHP

Streszczenie

Tematem opracowania są kompetencje logistyków kształtowane przez rozwój sztucznej inteligencji (AI). Praca realizuje cele poznawcze, cel metodyczny oraz cele praktyczne. Cele poznawcze obejmują zidentyfikowanie kluczowych trendów i rynkowych wymagań wobec logistyki wspieranej AI oraz kompetencji AI niezbędnych w nowoczesnej logistyce, a także wskazanie obszarów logistyki, w których technologie AI zostały zaimplementowane lub powinny być rozwijane. Cel metodyczny polega na opracowaniu metodyki oceny programów studiów pod kątem wykorzystania AI w logistyce, opartej na metodzie AHP. Cele praktyczne obejmują przeprowadzenie oceny wybranych programów nauczania oraz opracowanie rekomendacji dla instytucji edukacyjnych, dotyczących uzupełnienia i aktualizacji treści kształcenia logistyków pragnących zdobyć wiedzę i umiejętności z zakresu AI oraz wskazanie im kolejnych ścieżek rozwoju. Wyniki badań potwierdzają, że AI rewolucjonizuje logistykę, automatyzując procesy i generując nowe potrzeby kompetencyjne, zarówno techniczne, jak i miękkie. Szkoły wyższe odpowiadają na te wyzwania, wprowadzając do oferty moduły AI. Do oceny dwóch wybranych programów studiów szkoły wyższej, tj. Sztuczna inteligencja i Data Science oraz Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie, zastosowano metodę AHP (Analytic Hierarchy Process). Według analizy AHP program Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie uzyskał wyższą ocenę niż Sztuczna inteligencja i Data Science, przy uwzględnieniu czterech kryteriów: rynku, innowacyjności, praktyczności i trendów. Wnioski wskazują na konieczność uzupełnienia programów nauczania w kontekście logistyków o Computer vision (stosowane w procesach magazynowych), Predictive analytics (stosowane w zaopatrzeniu, sprzedaży i transporcie, planowaniu zapasów, prognozowaniu popytu, optymalizacji floty, wykrywaniu zakłóceń, takich jak awarie, opóźnienia lub zdarzenia losowe), Edge Computing, wiedzę na temat Cyfrowych Bliźniaków, IIoT oraz z zakresu zarządzania ryzykiem z wykorzystaniem AI. Dzięki temu absolwenci logistyki będą mogli utrzymać konkurencyjność na rynku pracy.

Słowa kluczowe: logistyka, programy nauczania logistyki, sztuczna inteligencja, AHP

¹ Uniwersytet Ekonomiczny w Katowicach

Development of Logistics Professionals' Competencies Under the Influence of AI: Evaluation of Educational Programs Using the AHP Method

Abstract

The subject of this study is the competencies of logisticians shaped by the development of artificial intelligence (AI). The study pursues cognitive, methodological and practical objectives. The cognitive objectives include identifying key trends and market requirements for AI-supported logistics and AI competencies necessary in modern logistics, as well as indicating areas of logistics where AI technologies have been implemented or should be developed. The methodological objective is to develop a methodology for evaluating study programmes in terms of the use of AI in logistics, based on the AHP method. The practical objectives include conducting an assessment of selected curricula and developing recommendations for educational institutions on supplementing and updating the content of education for logisticians who wish to acquire knowledge and skills in the field of AI, as well as indicating further development paths for them. The results of the research confirm that AI is revolutionising logistics by automating processes and generating new competence needs, both technical and soft. Higher education institutions are responding to these challenges by introducing AI modules into their programmes. The Analytic Hierarchy Process (AHP) method was used to evaluate two selected university programmes, namely Artificial Intelligence and Data Science and Master of Business Administration (MBA) – Artificial Intelligence and Cybersecurity in Modern Leadership. According to the AHP analysis, the Master of Business Administration (MBA) – Artificial Intelligence and Cybersecurity in Modern Leadership programme received a higher rating than Artificial Intelligence and Data Science, taking into account four criteria: market, innovation, practicality and trends. The conclusions indicate the need to supplement the curricula in the context of logistics with Computer vision (used in warehouse processes), Predictive analytics (used in procurement, sales and transport, inventory planning, demand forecasting, fleet optimisation, detection of disruptions such as breakdowns, delays or random events), Edge Computing, knowledge of Digital Twins, IIoT and risk management using AI. This will enable logistics graduates to remain competitive in the job market.

Keywords: logistics, logistics education programs, artificial intelligence, AHP

Wstęp

Kontekst badawczy

Rozwój sztucznej inteligencji staje się kluczowym czynnikiem transformującym łańcuchy dostaw i procesy logistyczne. Według raportu DHL (Dohrmann et al., 2024) firmy logistyczne już wdrażają rozwiązania AI, co napędza potrzebę pogłębionych badań nad technologią AI w programach nauczania logistyki.

Luka badawcza

Według danych z 2024 roku (Główny Urząd Statystyczny, 2024, s. 98) jedynie 5,9% (najmniej w logistyce – 0,6%) przedsiębiorstw w Polsce deklarowało wykorzystanie technologii sztucznej inteligencji, co wskazuje na istotną lukę kompetencyjną, wymagającą jej wypełnienia poprzez uwzględnienie w programach nauczania przedmiotów związanych ze sztuczną inteligencją (Artificial Intelligence, AI).

Na podstawie literatury przedmiotu wyróżnić można sześć głównych obszarów, w których AI przynosi korzyści dla logistyki. Są to:

- Cyfrowe bliźniaki (Digital Twins) z IIoT (Intelligent IoT) – umożliwiają symulacje operacji oraz monitoring procesów w czasie rzeczywistym, co poprawia efektywność planowania i eksploatacji zasobów (Nicoletti, 2025; Dohrmann et al., 2024).
- Odporność łańcucha dostaw dzięki predykcji – wykorzystanie dużych zbiorów danych (big data) i opartych na nich modeli predykcyjnych wzmacnia odporność łańcucha dostaw, szczególnie w fazach gotowości, reakcji i adaptacji na zakłócenia (Zamani et al., 2022; Paksoy et al., 2020).
- Wizja komputerowa (Computer Vision) i związana z tym automatyzacja magazynu – gwałtowny wzrost zastosowań wizji komputerowej obecny jest w sortowaniu, kontroli jakości i integracji z robotyką magazynową (Nicoletti, 2025; BITO, 2025).
- Analityka predykcyjna (Predictive Analytics) – np. w transporcie, gdzie modele szeregów czasowych, takie jak ARIMA i LSTM, usprawniają prognozowanie popytu oraz optymalizację tras, przekładając to na obniżenie kosztów i czasu dostaw (szerzej na temat AI w SCM: Boute & Udenio, 2023).
- Edge Computing – jest to przetwarzanie danych blisko źródła, tj. bliżej sensorów i urządzeń brzegowych (Shi et al., 2016), np. w magazynie lub pojeździe (BITO, 2025). Jest to kluczowe dla utrzymania niskich opóźnień i niezawodności systemów monitoringu.
- Zarządzanie ryzykiem z wykorzystaniem AI – podnosi poziom bezpieczeństwa operacji logistycznych, wspierając planowanie reakcji w czasie rzeczywistym. Ryzyka AI w logistyce mogą być technologiczne, operacyjne, regulacyjne i etyczne (Nicoletti, 2025).

Pomimo intensywnego rozwoju technologii AI w logistyce (optymalizacja tras i czasu dostaw, prognozowanie popytu i zarządzanie zapasami, automatyzacja magazynów, planowanie pracy personelu, obsługa klientów za pomocą chatbotów, utrzymanie pojazdów, obliczanie emisji CO₂) nieuwzględnienie tych rozwiązań w programach kształcenia tworzy lukę kompetencyjną, która może osłabiać konkurencyjność absolwentów na rynku pracy. Szkoły wyższe w Polsce, dostosowując swoje sylabusy do nowych wymagań branżowych, podejmują wysiłki, aby wypełnić tę lukę poprzez włączanie rozwiązań opartych na AI do oferty studiów na kierunkach ekonomicznych i inżynierskich.

Założenia badawcze i hipoteza

W pracy przyjęto hipotezę, że istnieją zasadnicze różnice między programami nauczania Sztuczna inteligencja i Data Science (A1) oraz MBA – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (A2) wybranej szkoły wyższej o profilu ekonomicznym w Polsce. Różnice te dotyczą zarówno stopnia integracji aktualnych trendów rynkowych, jak i nabywania umiejętności praktycznych w zakresie zastosowania AI w logistyce. Hipoteza ta zostanie zweryfikowana przy użyciu opracowanej metodyki oceny opartej na metodzie wielokryterialnej analizy hierarchicznej (Analytic Hierarchy Proces, AHP) i na zastosowaniu technik obliczeniowych wykonanych w arkuszu kalkulacyjnym oraz języku Python.

Cele badawcze

Na potrzeby badania sformułowano cele poznawcze, cel metodyczny oraz cele praktyczne.

Cele poznawcze polegają na:

- identyfikacji kluczowych trendów oraz rynkowych wymagań wobec logistyki wspomaganej sztuczną inteligencją,
- określeniu podstawowych kompetencji AI niezbędnych w nowoczesnej logistyce,
- wskazaniu obszarów działalności logistycznej, w których technologie AI zostały już zaimplementowane oraz tych, w których powinny być rozwijane.

Cel metodyczny obejmuje opracowanie metodyki oceny programów studiów pod kątem umiejętności późniejszego wykorzystania przez absolwentów uczelni narzędzi sztucznej inteligencji w logistyce.

Cele praktyczne (użytkowe) polegają na:

- przeprowadzeniu oceny dwóch wybranych programów nauczania szkoły wyższej ekonomicznej według opracowanej metodyki,
- sformułowaniu rekomendacji dla instytucji edukacyjnych dotyczących uzupełnienia i aktualizacji treści kształcenia logistyków w erze AI,

- wskazaniu ścieżki rozwoju zawodowego dla specjalistów logistyki, uwzględniającej kluczowe obszary zastosowań sztucznej inteligencji.

Pytania badawcze

Osiągnięcie celów badawczych polega na ich weryfikacji i udzieleniu odpowiedzi na następujące pytania badawcze:

- P1. Jakie są kluczowe trendy oraz rynkowe wymagania wobec logistyki wspomaganej sztuczną inteligencją?
- P2. W których obszarach procesów logistycznych AI odgrywa największą rolę?
- P3. Jakie rozwiązania AI są najczęściej wymieniane w kontekście logistyki?
- P4. Jak szkoły wyższe włączają AI w programy studiów oraz jakie rekomendacje można na tej podstawie sformułować dla uczelni i ścieżek rozwoju zawodowego logistyków?
- P5. W jaki sposób opracować metodykę oceny programów studiów uwzględniającą metodę AHP pod kątem wykorzystania narzędzi AI w logistyce?
- P6. Który z wybranych programów nauczania jest lepiej dostosowany do potrzeb logistyków, od których będzie się wymagać znajomości AI?

Przegląd literatury

Współcześnie w logistyce obserwujemy dynamiczne zmiany wynikające z rozwoju sztucznej inteligencji (Paksoy et al., 2020). Zmiany te obejmują trzy główne obszary:

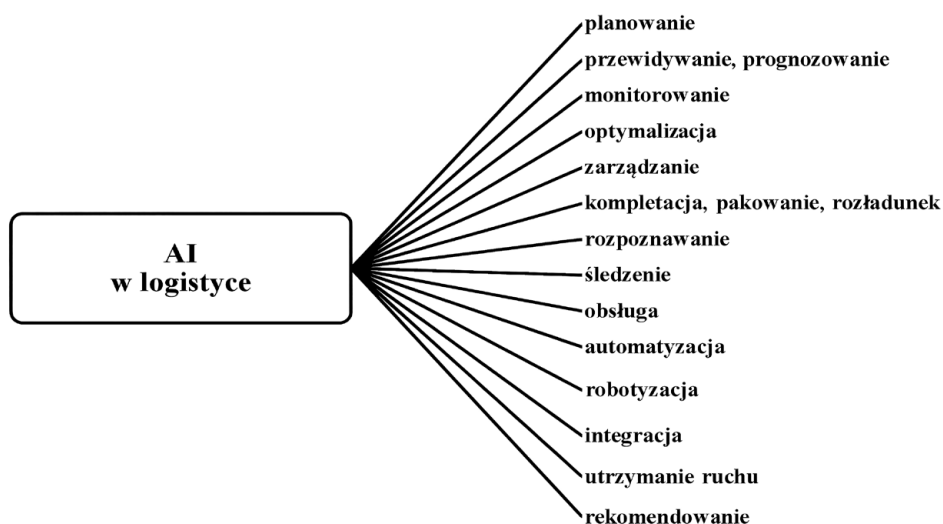
1. automatyzację² i optymalizację procesów (inteligentne magazyny, autonomiczne pojazdy i drony, robotyka wspierana AI, inteligentne okulary),
2. rozwiązania w obszarze łańcucha dostaw (AI zintegrowana z systemami informatycznymi, wirtualni asystenci, podejmowanie decyzji na podstawie danych z czujników IoT) (Sweeney, 2021),
3. predykcję (planowanie tras, generowanie zamówień, prognozowanie zapotrzebowania na zasoby) (Zamani et al., 2022).

Sztuczna inteligencja w logistyce wspomaga działania takie jak: pakowanie, kompletowanie, rozładunek czy wykrywanie błędów, co przekłada się na redukcję kosztów, oszczędność czasu oraz możliwość symulowania rozwiązań przed ich wdrożeniem (Nicoletti, 2025). AI wspiera również przewidywanie przyszłych działań, podejmowanie decyzji na podstawie analizy danych historycznych i big data, a także programowanie ścieżek jazdy autonomicznych pojazdów (BITO, 2025). Z kolei rozwój technologii kwantowych wskazuje na potencjał w zakresie rozwiązywania złożonych problemów optymalizacyjnych, identyfikowania

2 Należy odróżniać automatyzację od autonomii (Boute & Udenio, 2023).

wzorców, harmonogramowania, planowania tras oraz predykcji konserwacji pojazdów i maszyn (Boger, 2025; Gunashekar & Chis, 2025; Gunashekar et al., 2022). Ponadto AI przyczynia się do redukcji zapasów, optymalizowania tras przejazdu pojazdów, obsługi procesów zwrotów towarów oraz automatyzacji generowania etykiet wysyłkowych (Lindquist, 2024) itp. Rysunek 1 syntetyzuje i ilustruje czynności logistyczne, w których możliwe jest zastosowanie sztucznej inteligencji.

Rysunek 1. Działania w logistyce korzystające z AI



Źródło: opracowanie własne.

Jako kluczowe kompetencje AI wymagane w logistyce należy wymienić umiejętność korzystania (Lindquist, 2024; Tavana et al., 2023; Paksoy et al., 2020; Shad et al., 2014):

- z modeli generatywnych do automatyzacji zarządzania dokumentami i generowania dokumentów – w komunikacji e-mail, tworzeniu podsumowań zamówień, etykiet wysyłkowych itp.;
- z przetwarzania języka naturalnego (NLP) – w tworzeniu chatbotów i rozwiązań automatyzujących obsługę zapytań;
- z systemów wizyjnych (Computer Vision) – do sortowania, kontroli jakości i rozpoznawania obiektów;
- z rozwiązań optymalizacyjnych wspierających procesy logistyczne;
- z analityki predykcyjnej (Predictive Analytics) – do prognozowania (np. zapotrzebowania) i wykrywania (np. ryzyka, anomalii), w tym do budowy i trenowania modeli szeregów czasowych (ARIMA, LSTM) oraz algorytmów klasyfikacji i regresji (por. Tavana et al., 2023; Shad et al., 2014).

W związku z tym kluczowe trendy i wymagania rynkowe w logistyce wspomaganej przez AI obejmują (Nicoletti, 2025; BITO, 2025; Lindquist, 2024; Waters & Rinsler, 2014):

- automatyzację procesów (zamówienia, zwroty, kontrola jakości i inne),
- obsługę klientów i partnerów logistycznych (chatboty do zarządzania zapytaniami i reklamacjami),
- transparentność łańcucha dostaw, widoczność przepływu towarów (monitorowanie lokalizacji, monitorowanie stanu ładunku, sterowanie działaniami w razie zakłóceń, reagowanie na zakłócenia),
- optymalizację tras przejazdu, transportu i ruchu drogowego (z uwzględnieniem prognoz pogody, cen paliwa, czasu dostawy, śladu węglowego, emisji dwutlenku węgla i emisji gazów w wyniku spalania paliwa),
- prognozowanie (zapasów, zwrotów, popytu).

Zestawienie kluczowych kompetencji AI w logistyce wraz z odpowiadającymi im kluczowymi trendami i wymaganiami rynkowymi w logistyce wspomaganej AI przedstawiono w tabeli 1.

Tabela 1. Zestawienie trendów rynkowych w logistyce z kompetencjami AI

Lp.	Trend rynkowy	Wymagana kompetencja AI	Przykład zastosowania
1.	Automatyzacja	Computer Vision (CV) i Predictive Analytics (PA)	Ocena przyczyn zwrotów towarów
2.	Obsługa	Przetwarzanie języka naturalnego (NLP) i Generative AI	Ocena opinii klientów w czasie rzeczywistym z systemów ERP i mediów społecznościowych
3.	Transparentność	Integracja IoT i Edge Computing	Ocena strumieni danych z czujników w pojazdach lub w magazynach
4.	Optymalizacja	Metody optymalizacji i algorytmy decyzyjne	Planowanie tras w czasie rzeczywistym
5.	Prognozowanie	Predictive Analytics (PA)	Modele ARIMA i LSTM do przewidywania popytu

Źródło: opracowanie własne.

Wzrost znaczenia AI w logistyce powoduje, że od pracowników wymaga się nabywania nowych kompetencji (zarówno w obszarze wiedzy, umiejętności oraz postaw). Ewolucję kompetencji, a co za tym idzie kształcenia w obszarze logistyki, technologii, a obecnie także AI można zilustrować „ewolucją od STEM/SEAM, poprzez SMAC, aż do DARQ” następująco:

- STEM (Science, Technology, Engineering, Mathematics) – ukazuje, iż do tej pory dominowały klasyczne przedmioty ścisłe i techniczne, kładące nacisk na fundamenty wiedzy;
- STEAM (STEM + Arts) – STEAM zostało uzupełnione o aspekt kreatywności i projektowania, w celu rozwijania wśród ludzi kreatywnego myślenia oraz myślenia interdyscyplinarnego;
- SMAC (Social, Mobile, Analytics, Cloud) – wymienione kierunki w rozwoju kompetencji zdobyły popularność za sprawą mediów społecznościowych i technologii mobilnych, analizy i analityki danych, technologii chmurowych, co pozwoliło m.in. na stworzenie pierwszych rozwiązań dla logistyki cyfrowej;
- DARQ (Distributed Ledger Technology, AI, Extended Reality, Quantum Computing/Technology) – jest to najnowsza fala innowacji, obejmująca blockchain (transparentność łańcucha dostaw), sztuczną inteligencję (AI), Extended Reality (XR, popularne w szkoleniach pracowników, planowaniu przestrzeni) oraz Quantum Computing (znajdujący zastosowanie w rozwiązywaniu złożonych problemów optymalizacyjnych, np. w systemach SCM).

Podsumowując, wymienione transformacje w obszarze kompetencji wpływają na programy kształcenia, które muszą wyposażać logistyków nie tylko w podstawowe umiejętności STEM, ale też nauczyć ich kreatywności (STEAM), praktycznych rozwiązań analitycznych i chmurowych (SMAC) oraz zaawansowanych technologii (DARQ).

Szkoły wyższe starają się nadążać za rozwojem AI, co jest widoczne w programach studiów (tabela 2). Podążają za trendami, dostosowując lub tworząc programy kształcenia (licencjackie, inżynierskie, magisterskie, podyplomowe oraz MBA), które doskonalą w obszarze sztucznej inteligencji, dostarczając kompetencji niezbędnych do utrzymania konkurencyjności na rynku pracy.

Tabela 2. Wybrane programy nauczania z zakresu AI lub AI ukierunkowanego na potrzeby logistyki, realizowane w szkołach wyższych ekonomicznych w Polsce

Lp.	Nazwa szkoły wyższej	Wybrane programy nauczania z zakresu AI
1.	Uniwersytet Ekonomiczny w Katowicach	– Studia podyplomowe Sztuczna Inteligencja w Analizie Danych – kształtowanie kompetencji związanych ze sztuczną inteligencją, w tym związanych z uczeniem maszynowym i sieciami neuronowymi. – Studia podyplomowe Uczenie Maszynowe i Data Science – kształtowanie kompetencji z zakresu uczenia maszynowego i baz danych, m.in. poprzez przedmiot Podstawy sztucznej inteligencji. – Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie – trzyletni program studiów podyplomowych łączący kompetencje menedżerskie z AI i cyberbezpieczeństwem, prawem i etyką w AI, zarządzaniem innowacjami i transformacją cyfrową oraz z praktycznymi projektami poprzez laboratoria Łukasiewicz – AI, kontakt z realnymi wdrożeniami, a w ramach sprawdzenia zdobytych kompetencji uczestnictwo w hackathonie AI dla biznesu. – Kierunek Sztuczna Inteligencja i Data Science (I stopień) – interdyscyplinarny kierunek łączący elementy AI i analizy danych z elementami ekonomii, finansów, zarządzania i gospodarki przestrzennej, z praktycznymi modułami uczenia maszynowego, NLP czy AI w różnych sektorach gospodarki. W ramach studiów oferowany jest m.in. przedmiot Inteligentne systemy transportowe i logistyczne.
2.	Szkoła Główna Handlowa w Warszawie	– Studia podyplomowe Sztuczna inteligencja w biznesie i sektorze publicznym – wprowadzenie techniczne, wybrane zastosowania, wpływ AI, wyzwania prawne i etyczne, wykorzystanie AI, np. w prognozowaniu, w codziennej pracy w korporacji. – Studia podyplomowe Business Intelligence i sztuczna inteligencja w podejmowaniu decyzji – wiedza i praktyczne umiejętności w zakresie metod analizy danych oraz generatywnej sztucznej inteligencji w biznesie. – Studia magisterskie Analiza danych – big data, specjalność Sztuczna inteligencja w biznesie – kształcenie specjalistów w zakresie pozyskiwania danych z różnych źródeł oraz ich analizy.

Lp.	Nazwa szkoły wyższej	Wybrane programy nauczania z zakresu AI
3.	Uniwersytet Ekonomiczny w Poznaniu	– Studia podyplomowe Sztuczna inteligencja (AI) w biznesie i administracji – przeznaczone dla osób zainteresowanych wdrożeniem rozwiązań sztucznej inteligencji wspierających działalność biznesową oraz działalność jednostek administracji publicznej z certyfikatem audytora wewnętrznego systemu zarządzania sztuczną inteligencją (AI) według normy PN-EN ISO/IEC 42001. – Studia magisterskie Cyfrowe Łańcuchy Dostaw – oferują m.in. przedmiot Wykorzystanie sztucznej inteligencji w logistyce. – Studia pierwszego stopnia (licencjackie) – kierunek Informatyka i analityka danych, m.in. przedmiot Sztuczna inteligencja. – Studia drugiego stopnia – kierunek Nadzór i kontrola, m.in. przedmiot Sztuczna inteligencja w gospodarce i finansach firm. – Studia drugiego stopnia – kierunek Marketing, m.in. przedmiot Sztuczna inteligencja w marketingu.
4.	Uniwersytet Ekonomiczny we Wrocławiu	– Studia podyplomowe Transformacja technologiczna zakładów produkcyjnych w kierunku Industry 5.0 – oferują m.in. przedmioty Sztuczna inteligencja (AI) w transformacji technologicznej, Automatyzacja i robotyzacja zakładów produkcyjnych, Założenia koncepcji Industry 5.0. – Studia podyplomowe Analityk Biznesowy IT Business – oferują m.in. przedmiot Sztuczna inteligencja we wspomaganiu procesów biznesowych. – Studia I stopnia Informatyka w biznesie – oferują m.in. przedmiot Sztuczna inteligencja. – Szkolenia specjalistyczne Prompt Engineering.
5.	Uniwersytet Ekonomiczny w Krakowie	– Studia podyplomowe Logistyka i optymalizacja procesów w łańcuchu dostaw – kształcenie obejmuje trendy i innowacje w logistyce, m.in. z zakresu Green Logistics, systemów ERP i zastosowań sztucznej inteligencji w logistyce. – Studia podyplomowe Nowoczesne technologie i cyfrowa transformacja – w programie m.in. strategię wdrażania AI w organizacji.

Źródło: opracowanie własne na podstawie danych ze stron internetowych szkół wyższych ekonomicznych, według stanu na 2025 rok.

Najbardziej kompleksowe i pogłębione podejście do AI prezentuje Uniwersytet Ekonomiczny w Katowicach, ponieważ rozwija kompetencje zarówno w wymiarze technicznym, jak i menedżerskim. Uniwersytet Ekonomiczny we Wrocławiu

oraz Uniwersytet Ekonomiczny w Krakowie oferują studia podyplomowe z zakresu AI, które koncentrują się na zastosowaniach AI. Natomiast na Uniwersytecie Ekonomicznym w Poznaniu oraz w SGH w Warszawie sztuczna inteligencja postrzegana jest szerzej, jako rozwiązanie wspomagające biznes, zarządzanie publiczne, administrację.

Metodyka badań – opis procedury oceny programów nauczania z zastosowaniem metody AHP i języka Python

Źródła danych

Badaniem objęto dwa programy studiów szkoły wyższej, tj. Sztuczna inteligencja i Data Science (oznaczonego jako A1) oraz Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (oznaczonego jako A2). Jako narzędzia badawcze zastosowano MS Excel oraz język Python. Weryfikacja w języku Python po obliczeniach w Excelu była celowym zabiegiem walidacyjnym.

Procedura badawcza

W pracy przedstawiono metodykę porównania i oceny programów studiów z zastosowaniem metody AHP (Tavana et al., 2023; Leão et al., 2019; Ishizaka, 2019; Saaty, 2008). Metodyka obejmuje etapy od definicji problemu, przez analizę, aż do interpretacji i rekomendacji.

Na metodykę składają się następujące etapy:

1. Dobór i wstępna charakterystyka badanych programów studiów, dzięki którym logistycy mogą zdobyć kompetencje z zakresu AI;
2. Formułowanie hierarchii – od celów badania, poprzez alternatywy (wybór dwóch programów nauczania A1 i A2) oraz określenie kryteriów oceny programów studiów (w badaniu użyto takich kryteriów, jak: rynek, innowacje, praktyczność, trendy);
3. Ocena programów nauczania na podstawie informacji ze stron internetowych szkół wyższych, ze względu na podane kryteria oraz w celu stworzenia macierzy porównań;
4. Obliczenia AHP z wykorzystaniem arkusza kalkulacyjnego MS Excel;
5. Ocena spójności wyników – obliczenie wskaźnika CR i ewentualna korekta macierzy, jeżeli byłaby konieczna ($CR < 0,10$);
6. Analiza wrażliwości – test wpływu zmian wag na ostateczne rankingi;
7. Weryfikacja wyników w języku Python – automatyzacja obliczeń AHP,
8. Wizualizacja (prezentacja) wyników – tabele, wykresy i diagramy;
9. Identyfikacja rekomendacji – opracowanie na podstawie wyników propozycji uzupełnienia programów studiów;

10. Ukazanie ewentualnych ograniczeń badania;

11. Generalizacja wniosków i metodyki.

Etapy 2–6 są zgodne z kanonem metody Analytic Hierarchy Process (zweryfikowanie spójności $CR < 0.10$ to fundamentalny wymóg Saaty'ego). Na etapie drugim posłużono się kryteriami modelu 4P, tj.: rynek (potential), innowacje (prospects), praktyczność (practice), trendy (progress). Model 4P stanowi teoretyczną podstawę oceny programów studiów. Model ten skutecznie identyfikuje różnice jakościowe między programami studiów w kontekście wymagań Przemysłu 4.0. Weryfikacja wyników obliczeń w dwóch niezależnych narzędziach (MS Excel i za pomocą języka Python) zwiększa ich wiarygodność. Podsumowując całościowo tę metodykę, można ją zastosować w ocenie innych, wybranych programów studiów, np. ze względu na jej powtarzalność oraz możliwość zautomatyzowania obliczeń.

Wyniki badań

Pierwszym z ocenianych programów studiów jest Sztuczna inteligencja i Data Science (oznaczony jako A1). Najwyżej pod względem praktyczności, innowacyjności i zgodności z trendami należy ocenić przedmioty: Podstawy programowania w języku Python oraz Uczenie maszynowe/NLP/Smart Cities/Startupy. Mocnymi stronami programu są interdyscyplinarność, zapowiadana współpraca z firmami (IBM, ING, PwC) oraz elastyczność ścieżek kształcenia. Słabszymi punktami są przedmioty ogólne, które mają mniejszy wpływ na rozwój umiejętności praktycznych, ale mogą być ważne dla zrozumienia szerszego kontekstu. Język angielski i laboratoria AI zyskują w kierunku praktyczności i rynku, co odzwierciedla ich rolę wspierającą.

Drugim z ocenianych programów studiów jest program Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (oznaczony jako A2). Najwyżej pod względem praktyczności, innowacyjności i zgodności z trendami należy ocenić przedmioty: Przetwarzanie języka naturalnego, Generatywne modele AI, AI w cyberbezpieczeństwie (wykorzystanie ML do wykrywania zagrożeń), Zarządzanie innowacjami i wdrożeniem Edge AI w biznesie (Edge AI napędza IoT jako nowy trend w automatyce przemysłowej i produkcji), Automatykę przemysłową i systemy OT (istotne dla cyfryzacji fabryk i infrastruktury). Ponadto warsztaty praktyczne (np. hackathony, moduły projektowe i capstones) oraz przedmiot Decyzje menedżerskie wspierane AI, który stanowi nowoczesny paradygmat przywództwa, zgodny z trendami. Mocną stroną programu jest połączenie kompetencji technicznych z menedżerskimi i prawnymi. Słabszymi punktami są: ograniczona liczba przedmiotów logistycznych czy branżowych (np. Supply Chain AI), mniej intensywny blok programistyczny,

mniej nacisk na fundamenty matematyczne i statystyczne, brak modułów na temat Emerging Tech (np. Extended Reality, Quantum Computing).

Oba programy zostały ocenione według czterech kryteriów: rynek pracy (potencjał dla rozwoju kompetencji przyszłości, popyt na umiejętności), innowacyjność (nowatorstwo treści, kształtowanie kreatywnego myślenia, perspektywy), praktyczność (użyteczność w projektach, integracja teorii i praktyki w rzeczywistym środowisku), zgodność z trendami (postęp, zgodność z rozwojem branży, zdolność programu nauczania do adaptacji do zmian). W ocenie posłużono się skalą punktową od 1 do 10. Programy studiów A1 i A2 zaprezentowane w tabeli 3 różnią się profilem odbiorcy i stopniem zaawansowania technicznego. Program A1 to ścieżka techniczna z naciskiem na praktyczne umiejętności. Program A2 to ścieżka menedżersko-technologiczna.

Tabela 3. Macierz oceny alternatyw A1 i A2

Kierunek	Rynek	Innowacje	Praktyczność	Trendy
Sztuczna inteligencja i Data Science (A1)	10, specjaliści Data Science są poszukiwani na rynku	9, moduły z Deep Learningiem, big data, IoT, Smart Cities i Edge AI, modele generatywne	8, laboratoria i staże, capstones	10, NLP, generatywne modele AI, big data, AutoML, ML Ops
Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (A2)	9, łączenie umiejętności technicznych z zarządzaniem	8, kombinacja AI, generatywnych modeli, cyberbezpieczeństwa (security) i przywództwa (leadership)	9, warsztaty R&D, realne przypadki laboratoria, hackathony	9, generatywne AI, moduły no-code, AI w SecOps (SecOps w łańcuchu dostaw) i Edge AI, uzupełnione o nowe modele przywództwa – strategiczne zarządzanie AI
RAZEM	19	17	17	19

Źródło: opracowanie własne.

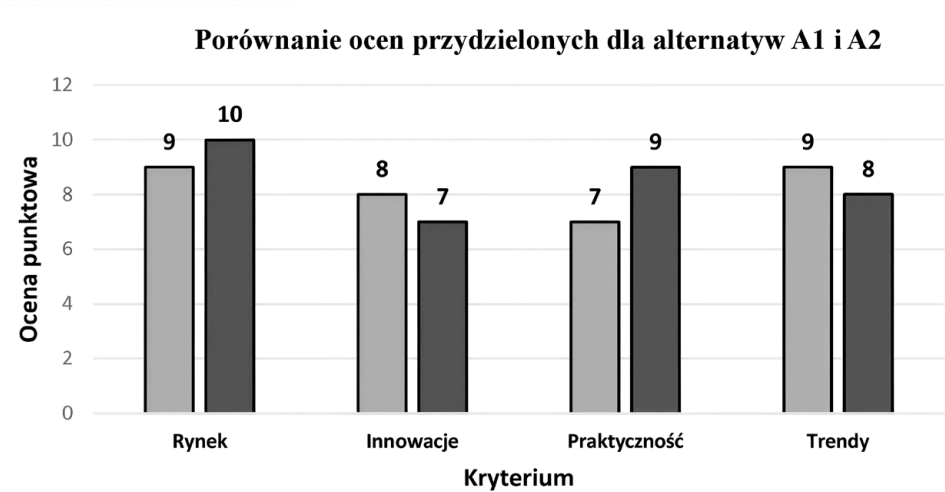
Tabela 4. Macierz ocen alternatyw A1 i A2 pod względem dopasowania do zawodu logistyka z kompetencjami w zakresie sztucznej inteligencji

Kierunek	Rynek	Innowacje	Praktyczność	Trendy
Sztuczna inteligencja i Data Science (A1)	9, logistyka i AI to rosnący trend	8, mało odniesień do branży logistycznej (digital twins, blockchain w łańcuchu dostaw)	7, logiści potrzebują gotowych narzędzi (do planowania, optymalizacji itp.)	9, mniejszy nacisk na rozwiązania supply-chain
Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (A2)	10, kadra zarządzająca logistyką poszukuje liderów potrafiących wdrażać AI w zarządzaniu zmianą, ryzykiem itp.	7, brak głębszych pakietów AI-DS; typowe przypadki z cyberbezpieczeństwa nie zawsze trafiają w potrzeby logistyczne	9, osadzone w biznesie logistycznym (zarządzanie ryzykiem, digitalizacja procesów)	8, brakuje modułów specyficznych dla logistyki (np. IoT w magazynie, blockchain w śledzeniu dostaw)
RAZEM	19 (bez zmian)	15 (-2 w stosunku do tabeli 3)	16 (-1 w stosunku do tabeli 3)	17 (-2 w stosunku do tabeli 3)

Źródło: opracowanie własne.

Dane w tabeli 4 zostały zweryfikowane w stosunku do ich pierwotnych ocen w tabeli 3, ze względu na ich przydatność dla zawodu logistyka wyposażonego w kompetencje AI. Program A2 zyskał przewagę w kryterium rynku i praktyczności w tym kontekście, podczas gdy program A1 zyskał wyższe noty w innowacjach i trendach. Jako dane wejściowe wzięto pod uwagę: dla alternatywy A1 – oceny [9, 8, 7, 9] odpowiednio dla kryteriów Rynek, Innowacje, Praktyczność, Trendy oraz dla alternatywy A2 – oceny [10, 7, 9, 8] odpowiednio dla kryteriów Rynek, Innowacje, Praktyczność, Trendy, co zilustrowano na rysunku 2.

Rysunek 2. Oceny alternatyw A1 i A2



Źródło: opracowanie własne.

Dla porównania programów studiów wykonano obliczenia metodą AHP, przy założeniu równych wag dla kryteriów (każde kryterium po 0,25). Zakładamy, że kryteria mają jednakową wagę, więc macierz porównań kryteriów to macierz jednostkowa (wszystkie elementy to 1). Wagi kryteriów są następujące: $w(\text{Rynek}) = 0,25$, $w(\text{Innowacje}) = 0,25$, $w(\text{Praktyczność}) = 0,25$, $w(\text{Trendy}) = 0,25$. Dla każdego kryterium obliczono średnią znormalizowanych wartości, co dało lokalne priorytety. W tabeli 5 przedstawiono macierz, gdzie wiersze to kryteria, a kolumny to alternatywy.

Tabela 5. Wektory lokalnych priorytetów (normalizacja)

Kryterium	Ocena A1	Ocena A2	Suma	Waga lokalna A1	Waga lokalna A2
Rynek	9	10	19	0,4737	0,5263
Innowacje	8	7	15	0,5333	0,4667
Praktyczność	7	9	16	0,4375	0,5625
Trendy	9	8	17	0,5294	0,4706
RAZEM				1,9739	2,0261

Źródło: opracowanie własne.

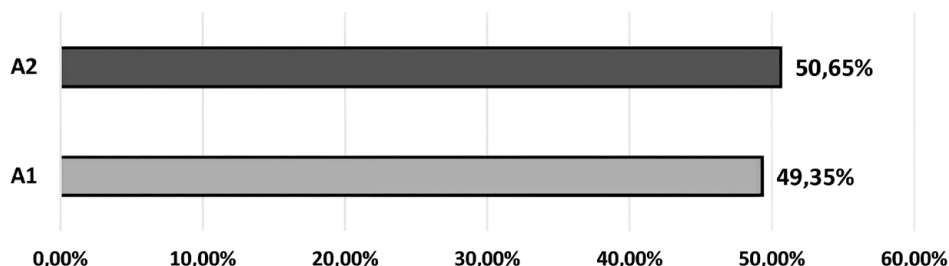
W tabeli 6 ukazano wynik obliczenia wag globalnych. Waga globalna jest iloczynem wagi kryterium (0,25) i wagi lokalnej z tabeli 5.

Tabela 6. Macierz ważona

Kryterium	A1	A2
Rynek (0,25)	0,1184	0,1316
Innowacje (0,25)	0,1333	0,1167
Praktyczność (0,25)	0,1094	0,1406
Trendy (0,25)	0,1324	0,1176
RAZEM	49,35%	50,65%

Źródło: opracowanie własne.

Otrzymano ostatecznie następujący ranking alternatyw: miejsce pierwsze dla A2 z wynikiem 50,65% oraz miejsce 2 dla A1 z wynikiem 49,35% (rysunek 3).

Rysunek 3. Ostateczny ranking alternatyw A1 i A2**Wykres porównania w poszczególnych kategoriach**

Źródło: opracowanie własne.

Wyjaśnienia wymagają jeszcze wskaźniki spójności w metodzie AHP. Maksymalna wartość własna to $\lambda_{\max}$. Dla macierzy kryteriów 4×4 , w której wszystkie przekątne mają wartość 1 i wszystkie pozadiagonalne relacje są równe 1 (czyli przyjmujemy jednakową wagę każdego z czterech kryteriów), mamy $n = 4$. Dlatego $\lambda_{\max} = n = 4$, ponieważ macierz jest doskonale spójna. Wskaźnik Spójności (Consistency Index, CI) jest równy zero, co wskazuje na pełną spójność macierzy. $CI = (\lambda_{\max} - n)/(n-1) = (4 - 4)/(4-1) = 0/3 = 0$. Gdyby CI było większe od 0 wówczas byłoby brak spójności. Losowy wskaźnik spójności (Random Index, RI) dla $n = 4$ wynosi 0,90 (z tablicy wartości Saaty'ego). Współczynnik spójności (Consistency Ratio, CR) wynosi zero, $CR = CI/RI = 0/0,90 = 0$. Interpretacja jest następująca: jeżeli CR jest mniejsze niż 0,10 ($CR < 0,10$), stąd macierz jest spójna i wyniki AHP wiarygodne. CR (Consistency Ratio) poniżej 0,10 (10%) świadczy o spójności ocen. Gdyby CR było większe lub równe 0,10 wówczas wymagana byłaby korekta.

W celu weryfikacji obliczeń wykonanych metodą AHP w arkuszu kalkulacyjnym Excel posłużono się dodatkowo kodem w języku Python. Stworzono plik `ahp_analysis.py` w Visual Studio Code. Konieczne było użycie bibliotek `pandas`, `numpy` i `matplotlib`. Wynik obliczeń został zapisany w pliku `wynik_AHP.csv` (rysunek 4).

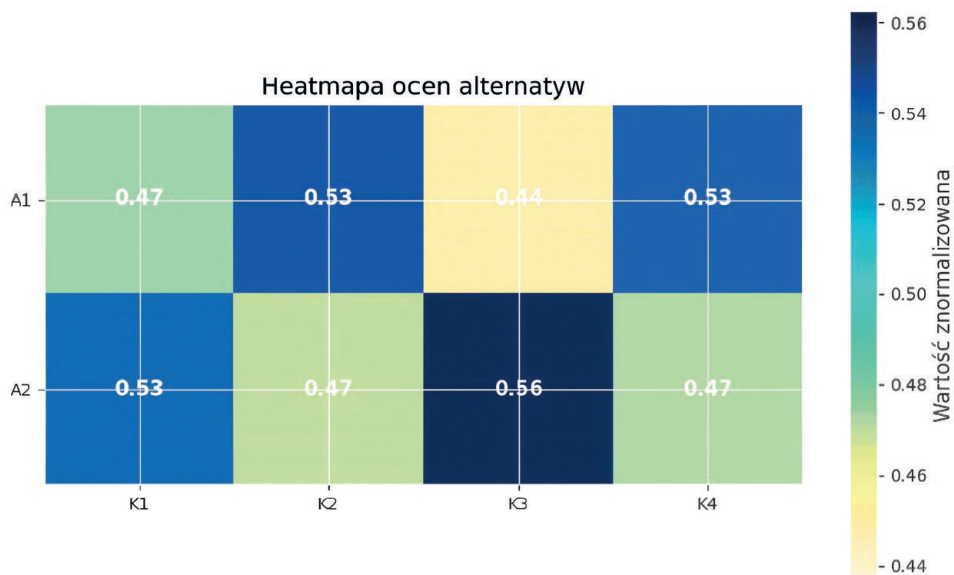
Rysunek 4. Wyniki alternatyw zapisane w arkuszu kalkulacyjnym na podstawie obliczeń wykonanych za pomocą algorytmu w języku Python

	A	B	C	D	E
1	Alternatywa,Wynik,Procent				
2	A1	0.49348232714138285	49.348232714138284		
3	A2	0.5065176728586172	50.651767285861716		

Źródło: zrzut ekranu z arkusza kalkulacyjnego.

Dodatkowo został wygenerowany raport (`raport_AHP.pdf`) z czterema wykresami, zawierający: Wagi kryteriów AHP, Wyniki alternatyw, Wykres radarowy i Heatmapę ocen alternatyw (ten ostatni ukazano na rysunku 5).

Rysunek 5. Heatmapa wygenerowana za pomocą języka Python



Źródło: zrzut ekranu z raportu `raport_AHP.pdf`.

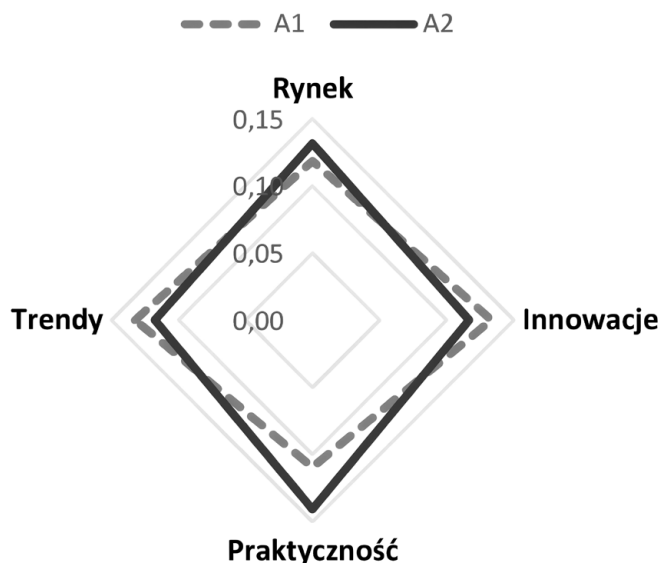
Liczby widoczne na rysunku 5 są zgodne z danymi znajdującymi się w tabeli 5.

Wnioski

Przeprowadzona analiza AHP, uwzględniająca kryteria rynku, innowacyjności, praktyczności (praktyczne wdrożenia) oraz trendów, wskazuje, że program Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (A2) w większym stopniu odpowiada na potrzeby specjalistów i menedżerów odpowiadających za wdrażanie i rozwój w organizacjach strategii AI w logistyce niż program Sztuczna inteligencja i Data Science (A1).

Rysunek 6. Profile alternatyw A1 i A2

Profile alternatyw (wykres radarowy)



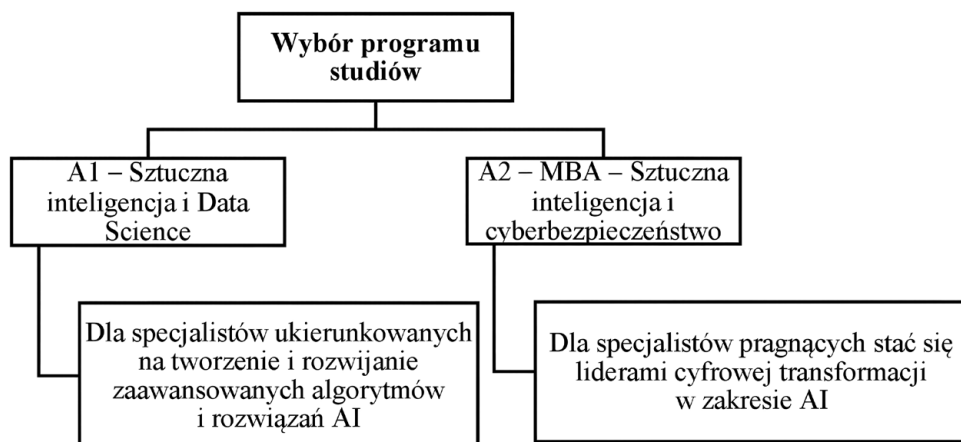
Źródło: opracowanie własne.

Przewaga programu A2 wynika przede wszystkim z wyższych ocen uzyskanych w obszarach praktycznych zastosowań AI i dostosowania firm do wymagań rynku w związku z rozwojem AI. Program A2 jest bliższy rynkowi. Natomiast A1 kładzie większy nacisk na aspekty techniczne, trendy, innowacje, w tym Data Science (rysunek 6).

Podsumowując, program A1 będzie lepszy od A2 dla logistyków ukierunkowanych na rozwijanie zaawansowanych algorytmów i rozwiązań AI – czyli odpowiedzialnych za działania operacyjne, w tym za analizowanie danych. Program A2 będzie lepszy od A1 dla logistyków pragnących stać się liderami cyfrowej

transformacji w zakresie AI – czyli odpowiedzialnych za działania taktyczne lub strategiczne, w tym w zakresie cyberbezpieczeństwa i etyki AI (rysunek 7).

Rysunek 7. Wizualne porównanie wybranych programów studiów



Źródło: opracowanie własne.

Dyskusja

Analiza obu programów pod kątem potrzeb logistyki z elementami AI ujawnia różnice w ich profilu merytorycznym i zastosowaniu.

- Specyfika programu Sztuczna inteligencja i Data Science (A1) zwraca uwagę na dobre przygotowanie techniczne w ML, NLP, języku Python i big data, stanowiące fundament dla zaawansowanych modeli predykcyjnych i analityki w logistyce. Brak jest jednak modułów dedykowanych procesom magazynowym czy transportowym, co może ograniczać bezpośrednie przełożenie wiedzy na zadania logistyczne (np. symulacje digital twins, wykonywanie operacji w systemach WMS lub TMS).
- Specyfika programu Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie (A2) kładzie nacisk na rzeczywiste wdrożenia – case study, warsztaty z integracji AI z systemami np. ERP/SCM oraz na bezpieczeństwo danych np. w łańcuchu dostaw. Program charakteryzuje mocny komponent przywódczy i strategiczny, ułatwiający logistykowi zarządzanie cyfrową transformacją, chociaż kosztem mniejszego zaawansowania algorytmicznego (mniej godzin kodowania i teorii ML). Podsumowując ocenę użyteczności obu programów w kontekście logistyki, należy zauważyć, że program A1 uzyskuje przewagę w obszarze

innowacji i trendów, co wynika z jego lepszego dopasowania do zagadnień automatyzacji widzenia maszynowego oraz zaawansowanego uczenia maszynowego. Z kolei program A2 dominuje w zakresie praktycznych wdrożeń i gotowości rynkowej, co jest konsekwencją większego nacisku na warsztaty oraz projekty z praktykami.

Podsumowanie i rekomendacje

Programy studiów dla przyszłych logistyków powinny uwzględniać elementy z zakresu AI i uczenia maszynowego, aby absolwenci zachowali konkurencyjność na rynku pracy. Z punktu widzenia logistyka chcącego posiadać umiejętności AI program Master of Business Administration (MBA) – Sztuczna inteligencja i cyberbezpieczeństwo w nowoczesnym przywództwie prezentuje się korzystniej niż Sztuczna inteligencja i Data Science, głównie dzięki większej użyteczności praktycznej związanej z rynkiem, dzięki lepszemu dopasowaniu do zarządzania procesami łańcucha dostaw oraz naciskowi na wiedzę z zakresu wdrożeń.

Programy studiów należałoby wzbogacić o moduły z zakresu Computer Vision (CV), Predictive Analytics (PA), a także o elementy IoT, Edge Computing oraz blockchain (Nicoletti, 2025; Dohrmann et al., 2024). Widzenie komputerowe znajduje bowiem zastosowanie w automatyzacji procesów magazynowych, w tym w kontroli jakości (np. do wykrywania uszkodzeń lub odchyłeń od wzorca), w rozpoznawaniu zawartości palet czy sortowaniu, a także w digital twins. Analityka predykcyjna (Predictive Analytics, PA) może wspierać prognozowanie popytu, optymalizację floty i wykrywanie zakłóceń w transporcie (Tavana et al., 2023; Shad et al., 2014). Włączenie do programów nauczania treści dotyczących CV i PA należy uznać za strategiczny element przygotowania przyszłych logistyków do efektywnego zarządzania procesami, przy jednoczesnej optymalizacji czasu i kosztów operacyjnych tych procesów. Z kolei dodanie elementów IoT i Edge Computingu jest istotne dla monitorowania floty (Shi et al., 2016), a blockchain do śledzenia przesyłek (Nicoletti, 2025).

Rozszerzenie programów nauczania o rozwiązania dedykowane logistyce, oparte na zaawansowanej sztucznej inteligencji, wzmocniłoby profil absolwenta, czyniąc go liderem cyfrowej logistyki. Szczególne znaczenie miałoby operowanie podczas studiów na rzeczywistych danych oraz rozwijanie nowoczesnych form dydaktyki, takich jak: living lab (czyli otwarte środowisko testowe – testowanie pomysłów w prawdziwym świecie), capstone (projekt integrujący wiedzę z całego programu), czy hackathony (intensywne zawody w celu szybkiego wygenerowania rozwiązań w odpowiedzi na realne wyzwania). Wspierają one kształcenie poprzez łączenie teorii z praktyką i umożliwiają dynamiczne reagowanie na wyzwania cyfrowej transformacji łańcucha dostaw (Rios et al., 2020).

W dalszej kolejności przedstawione badania mogą zostać rozwinięte o analizę ofert pracy z serwisów dla logistyków, m.in. Pracuj.pl, Pracujwlogistyce.pl, Aplikuj.pl, Asistwork.pl, aby szczegółowo poznać bieżące zapotrzebowanie rynku na specjalistów do spraw logistyki z umiejętnością wykorzystania AI. Ponadto szkoły wyższe powinny monitorować wskaźniki efektywności absolwentów (zatrudnienie, wdrożone projekty AI), aby szybko reagować na nowe wyzwania branży, a także regularnie aktualizować sylabusy na podstawie raportów (np. McKinsey & Company, 2025; World Economic Forum, 2025) i/lub ankiet wśród praktyków logistyki. Realizacja tych rekomendacji zapewni absolwentom logistyki nie tylko solidne fundamenty AI, lecz także umiejętność praktycznego zastosowania nowoczesnych narzędzi, co jest kluczem do utrzymania konkurencyjności na dynamicznie zmieniającym się rynku pracy.

Bibliografia

- Boute, R. N., & Udenio, M. (2023). AI in Logistics and Supply Chain Management. In R. Merkert & K. Hoberg (Eds.), *Global Logistics and Supply Chain Strategies for the 2020s*, (pp. 49–65). Springer, Cham.
- Ishizaka, A. (2019). Analytic Hierarchy Process and Its Extensions. In M. Doumpos, J. Figueira, S. Greco & C. Zopounidis (Eds.), *New Perspectives in Multiple Criteria Decision Making* (pp. 81–93). Springer, Cham.
- Leão, H. A. T., Canedo, E. D., Costa, P. H. T., Okimoto, M. V., & Santos, G. A. (2019). Use of AHP and Promethee for Research Project Portfolio Selection. In S. Misra, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, C. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, & E. Tarantino (Eds.), *Computational Science and Its Applications – ICCSA 2019 (Lecture Notes in Computer Science)*, 11619, 504–517.
- Nicoletti, B. (2025). Logistics 5.0. In *Artificial Intelligence for Logistics 5.0. From Foundation Models to Agent AI*. Palgrave Macmillan, Cham.
- Paksoy, T., Kochan, C. G., & Ali, S. S. (Eds.). (2020). *Logistics 4.0: Digital Transformation of Supply Chain Management*. CRC Press (Taylor & Francis Group).
- Rios, M. G., Herrera, D. J. C., Lucio, K. S. N., & Vásquez, M. Y. L. (2020). AHP for a Comparative Study of Tools Used for Programming Learning. In S. Nazir, T. Ahram, W. Karwowski (Eds.), *Advances in Human Factors in Training, Education, and Learning Sciences*. AHFE 2020. Advances in Intelligent Systems and Computing, 1211, 356–362.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), 83–98.
- Shad, Z., Roghanian, E., & Mojibian, F. (2014). Integration of QFD, AHP, and LPP methods in supplier development problems under uncertainty. *Journal of Industrial Engineering International*, 10(51), 23–34.

- Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and Challenges. *IEEE Internet of Things Journal*, 3(5), 637–646.
- Tavana, M., Soltanifar, M., & Santos-Arteaga, F. J. (2023). Analytical hierarchy process: revolution and evolution. *Annals of Operations Research*, 326, 879–907.
- Sweeney, E., Waters, D. (2021). *Global logistics. New directions in Supply Chain Management* (8th Edition). Kogan Page.
- Zamani, E. D., Smyth, C., Gupta, S., & Dennehy, D. (2022). Artificial intelligence and big data analytics for supply chain resilience: a systematic literature review. *Annals of Operations Research*, 327(2), 605–632. Springer.

Raporty i źródła branżowe

- Główny Urząd Statystyczny. (2024). Społeczeństwo informacyjne w Polsce w 2024 r. Urząd Statystyczny w Szczecinie. <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spoleczenstwo-informacyjne/spoleczenstwo-informacyjne/spoleczenstwo-informacyjne-w-polsce-w-2024-roku,1,18.html>
- Gunashekar, S., & Chis, T. (2025). Navigating skills and talent development for quantum technology. Current insights and future horizons. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA3889-1.html
- Gunashekar, S., Flanagan, I., d'Angelo, C., Motsi-Omoijiade, I. D., Virdee, M., Feijao, C., & Porter, S. (2022). Using Quantum Computers and Simulators in the Life Sciences. Current Trends and Future Prospects. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1899-1.html
- McKinsey & Company. (2025). *The Year of Quantum: From concept to reality in 2025*. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-year-of-quantum-from-concept-to-reality-in-2025>
- World Economic Forum. (2025). Future of Jobs Report 2025. https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf

Netografia

- BITO. (2025). *Logistics trends 2025*. <https://www.bito.com/en-gb/news-know-how/detail/logistics-trends-2025/>
- Boger, Y. (2025). *How Quantum Computing Will Transform Logistics*. Forbes. <https://www.forbes.com/councils/forbesbusinessdevelopmentcouncil/2025/01/23/how-quantum-computing-will-transform-logistics/>
- Dohrmann, K. et al. (2024). *Logistics Trend Radar 7.0*. DHL Group. <https://www.dhl.com/discover/en-global/news-and-insights/reports-and-press-releases/logistics-trend-radar-2024>
- Lindquist, M. (2024). AI in logistics: Potential Benefits and Applications. Oracle. <https://www.oracle.com/scm/ai-in-logistics/>

Sztuczna inteligencja jako narzędzie do zarządzania procesem utrzymania ruchu w zakresie wind zabudowanych na stacjach i przystankach w Polsce. Kontekst innowacyjności podejścia do infrastruktury transportowej

Streszczenie

Współczesny świat coraz bardziej polega na technologiach, które mają na celu poprawę jakości życia obywateli. Jednym z kluczowych elementów infrastruktury transportowej są windy, które zapewniają dostępność osobom z ograniczoną mobilnością. W Polsce na dworcach kolejowych windy odgrywają kluczową rolę w umożliwianiu podróży wszystkim pasażerom. Niestety, często spotykamy się z problemami związanymi z ich niezawodnością i dostępnością. Sztuczna inteligencja (AI) może być rozwiązaniem, które przyniesie rewolucję w tej dziedzinie.

Widoczny trend dostosowywania stacji i przystanków kolejowych przez Narodowego Zarządcę Infrastruktury Kolejowej PKP Polskie Linie Kolejowe S.A. do obsługi osób o ograniczonej możliwości poruszania się poprzez budowę wind osobowych wiąże się również z koniecznością ich właściwego utrzymania i zarządzania. Narzędziem wspomagającym proces utrzymania może być sztuczna inteligencja, ale także systemy utrzymania ruchu. Jednym z takich systemów jest CMMS QRMAINT, dzięki któremu poprzez wykorzystanie czujników IoT możemy mieć pełną transparentność danych na temat działania danego urządzenia. Bogata funkcjonalność systemu CMMS QRMAINT obejmuje możliwość zgłaszania usterek z wykorzystaniem kodu QR, agregowanie danych z możliwością przetworzenia do dashboard TV wraz z prezentacją wskaźników KPI, na geolokalizacji kończąc. Tak starannie dobrane i właściwie zastosowane oprogramowanie może być nieocenionym narzędziem w utrzymaniu urządzeń do transportu osób o ograniczonej możliwości poruszania się.

Słowa kluczowe: sztuczna inteligencja, utrzymanie ruchu, infrastruktura kolejowa, dostępność transportu

1 PKP PLK S.A.

Artificial Intelligence as a Tool for Managing the Maintenance Process of Elevators Installed at Railway Stations and Stops in Poland: An Innovative Approach to Transport Infrastructure

Abstract

The modern world increasingly relies on technologies that aim to improve the quality of life of citizens. One of the key elements of transport infrastructure are lifts, which ensure accessibility for people with reduced mobility. In Poland, lifts play a crucial role in enabling all passengers to travel at railway stations. Unfortunately, we often encounter problems related to their reliability and availability. Artificial intelligence (AI) may be the solution that will revolutionise this field. The visible trend of adapting railway stations and stops by the National Railway Infrastructure Administrator PKP Polskie Linie Kolejowe S.A. to serve people with reduced mobility through the construction of passenger lifts also involves the need for their proper maintenance and management. Artificial intelligence can be a tool to support the maintenance process, but so can traffic maintenance systems. One such system is CMMS QRMAINT, which allows us to have full transparency of data on the operation of a given device through the use of IoT sensors. One such system is CMMS QRMAINT, which allows us to achieve full transparency of data on the operation of a given device through the use of IoT sensors. The comprehensive functionality of the CMMS QRMAINT system includes the ability to report faults using a QR code, aggregate data with the option of processing it into a TV dashboard with KPI indicators, and geolocation. Such carefully selected and properly applied software can be an invaluable tool in the maintenance of equipment for transporting people with reduced mobility.

Keywords: artificial intelligence, maintenance management, railway infrastructure, transport accessibility

Rola sztucznej inteligencji w poprawie dostępności i niezawodności wind

Sztuczna inteligencja	Dziedzina informatyki zajmująca się tworzeniem systemów i programów komputerowych, które potrafią wykonywać zadania wymagające ludzkiej inteligencji. Obejmuje m.in. uczenie maszynowe, rozpoznawanie obrazów i mowy, analizę danych, planowanie oraz podejmowanie decyzji. Celem sztucznej inteligencji jest umożliwienie maszynom „uczenia się” i dostosowywania do zmieniających się warunków.
Innowacyjność	Zdolność do tworzenia i wdrażania nowych pomysłów, technologii, produktów, usług lub metod działania, które przynoszą wartość dodaną. Innowacyjność może dotyczyć zarówno przedsiębiorstw, jak i całych sektorów gospodarki. Oznacza ciągłe poszukiwanie ulepszeń i kreatywne podejście do rozwiązywania problemów.
Infrastruktura transportowa	Zespół obiektów, urządzeń i systemów umożliwiających funkcjonowanie transportu. Obejmuje drogi, linie kolejowe, porty morskie, lotniska, mosty, tunele, przystanki, terminale oraz sieci komunikacyjne. Dobrze rozwinięta infrastruktura transportowa jest kluczowa dla sprawnego przemieszczania osób i towarów oraz rozwoju gospodarczego kraju.
Zarządzanie procesem utrzymania ruchu	Działania mające na celu zapewnienie nieprzerwanego i efektywnego funkcjonowania maszyn, urządzeń oraz systemów technicznych w przedsiębiorstwie. Obejmuje planowanie, monitorowanie, diagnozowanie awarii, naprawy oraz profilaktykę techniczną. Efektywne zarządzanie utrzymaniem ruchu minimalizuje przestoje produkcyjne, zwiększa bezpieczeństwo i wydłuża żywotność sprzętu.

Współczesne technologie sztucznej inteligencji znajdują coraz szersze zastosowanie w różnych aspektach naszego codziennego życia. Jednym z obszarów, gdzie AI zaczyna odgrywać kluczową rolę, są systemy transportu publicznego, w tym windy i schody ruchome na stacjach oraz przystankach kolejowych w Europie. Dzięki zastosowaniu AI możliwe jest podniesienie komfortu, bezpieczeństwa oraz efektywności operacyjnej tych obiektów. Sztuczna inteligencja może znacząco poprawić funkcjonowanie wind na dworcach kolejowych w Polsce. Dzięki zaawansowanym algorytmom i technologiom AI może monitorować i analizować działanie wind w czasie rzeczywistym, co pozwala na szybkie wykrywanie i diagnozowanie problemów. Dzięki zastosowaniu czujników i systemów monitoringu w czasie rzeczywistym AI może zbierać dane na temat działania wind, takie jak: częstotliwość używania, wibracje, temperatura i inne parametry. Analiza tych danych pozwala na wczesne wykrywanie nieprawidłowości i zapobieganie awariom. AI może również wspierać predykcyjną konserwację, która polega na przewidywaniu potencjalnych awarii na podstawie analizy zebranych danych. Dzięki temu

możliwe jest planowanie konserwacji w optymalnych momentach, co minimalizuje ryzyko przestojów i poprawia niezawodność wind. AI może również pomóc w poprawie dostępności wind poprzez analizę ruchu pasażerów i optymalizację działania wind w oparciu o rzeczywiste potrzeby. Na przykład w godzinach szczytu windy mogą być skonfigurowane tak, aby priorytetowo obsługiwać osoby z ograniczoną mobilnością, co zapewnia im lepszy dostęp do infrastruktury. Przykładem zastosowania sztucznej inteligencji w urządzeniach wspomagających transport osób w przestrzeni publicznej są schody ruchome prowadzące na perony w stacji Katowice. Katowice, dynamiczne serce Górnego Śląska, mogą pochwalić się jednym z najnowocześniejszych dworców kolejowych w Polsce. Inwestycje w infrastrukturę i technologie sprawiają, że podróżni mogą cieszyć się komfortem i bezpieczeństwem na najwyższym poziomie. Systemy AI na schodach ruchomych monitorują ich działanie w czasie rzeczywistym, wychwytyując wszelkie nieprawidłowości i potencjalne zagrożenia. Dzięki temu możliwe jest szybkie reagowanie na awarie i ich naprawa, zanim staną się poważnym problemem. W przypadku wykrycia jakiegokolwiek problemu system AI automatycznie wysyła powiadomienia do odpowiednich służb technicznych. To pozwala na natychmiastową interwencję i minimalizację ryzyka dla pasażerów. Innowacyjne schody ruchome na stacji kolejowej w Katowicach to przykład wykorzystania AI w celu poprawy komfortu i bezpieczeństwa podróżnych. Dzięki inteligentnemu zarządzaniu ruchem, energooszczędności oraz zaawansowanym systemom monitorowania i powiadamiania, stanowią one doskonałe rozwiązanie dla nowoczesnej infrastruktury transportowej. Katowice, jako jedno z najważniejszych miast w Polsce, pokazują, że inwestycja w nowe technologie przynosi korzyści wszystkim użytkownikom dworca kolejowego.

Predykcyjna konserwacja w windach zabudowanych. W drodze ku niezawodności i dostępności

Predykcyjna konserwacja to nowoczesna metoda zarządzania infrastrukturą, która znacząco poprawia niezawodność i bezpieczeństwo użytkowania urządzeń takich jak windy. W kontekście polskich stacji i przystanków kolejowych technologia ta zyskuje na znaczeniu, biorąc pod uwagę rosnące wymagania pasażerów oraz potrzebę utrzymania ciągłości operacyjnej w miejscach o wysokim natężeniu ruchu. Predykcyjna konserwacja, znana również jako konserwacja przewidująca, opiera się na zbieraniu i analizie danych w czasie rzeczywistym z różnych czujników zamontowanych na urządzeniach. W przypadku wind mogą to być czujniki monitorujące wibracje, temperaturę silnika, zużycie kabli czy czas pracy poszczególnych komponentów. Dzięki analizie tych danych za pomocą zaawansowanych algorytmów i technologii uczenia maszynowego możliwe jest przewidzenie awarii,

zanim do niej dojdzie. Na stacjach i przystankach kolejowych w Polsce zainstalowane są setki wind, które codziennie obsługują tysiące pasażerów. Niezawodne działanie tych urządzeń jest kluczowe dla zapewnienia komfortu i bezpieczeństwa podróżnych, zwłaszcza osób starszych, niepełnosprawnych oraz rodzin z małymi dziećmi. Wprowadzenie predykcyjnej konserwacji w oparciu o zaproponowany system CMMS QRMAINT na stacjach i przystankach kolejowych w Polsce ma na celu:

- minimalizację niespodziewanych awarii i związanych z nimi przestojów,
- obniżenie kosztów konserwacji poprzez bardziej efektywne zarządzanie zasobami,
- poprawę bezpieczeństwa użytkowników poprzez szybkie reagowanie na potencjalne problemy,
- wydłużenie żywotności urządzeń dzięki bardziej precyzyjnym interwencjom serwisowym.

Wdrażanie predykcyjnej konserwacji w windach wymaga zastosowania nowoczesnych technologii, takich jak:

- czujniki IoT (Internet of Things), które zbierają dane w czasie rzeczywistym,
- systemy SCADA (Supervisory Control and Data Acquisition) do monitorowania i zarządzania danymi,
- algorytmy uczenia maszynowego, które analizują zebrane dane i przewidują potencjalne awarie,
- aplikacje mobilne dla serwisantów, umożliwiające szybki dostęp do informacji o stanie technicznym wind.

Wprowadzenie predykcyjnej konserwacji wind na stacjach kolejowych przynosi liczne korzyści zarówno operatorom, jak i pasażerom. Do najważniejszych należą:

- redukcja kosztów operacyjnych dzięki bardziej efektywnemu planowaniu konserwacji,
- zmniejszenie liczby przestojów i opóźnień, co przekłada się na większą satysfakcję pasażerów,
- podniesienie poziomu bezpieczeństwa, co jest szczególnie istotne w miejscach o dużym natężeniu ruchu,
- poprawa dostępności dla osób o ograniczonej mobilności, co ma pozytywny wpływ na społeczną inkluzję.

Zastosowanie sztucznej inteligencji w windach i schodach ruchomych na stacjach i przystankach kolejowych w Europie staje się coraz powszechniejsze. Dzięki AI możliwe jest poprawienie zarządzania ruchem pasażerów, predykcji awarii, zwiększenie bezpieczeństwa oraz efektywności energetycznej. Technologie te nie tylko podnoszą komfort podróżnych, ale także przyczyniają się do zrównoważonego rozwoju systemów transportu publicznego.

Predykcyjna konserwacja wind na stacjach i przystankach kolejowych w Polsce to krok w przyszłość, który przynosi liczne korzyści, zarówno ekonomiczne, jak i społeczne. Dzięki nowoczesnym technologiom i systemom zarządzania możliwe jest zapewnienie wysokiego poziomu niezawodności i bezpieczeństwa urządzeń, co przekłada się na komfort i satysfakcję pasażerów. W miarę rozwoju tych technologii można spodziewać się dalszych usprawnień i innowacji, które uczynią polskie koleje jeszcze bardziej przyjaznymi dla użytkowników. W celu oceny potencjału rozwiązań technologicznych, pozwalającej na identyfikację zarówno ich atutów, jak i ograniczeń oraz zagrożeń płynących z otoczenia, przeprowadzono analizę SWOT dla wdrożenia systemu CMMS QRMAINT. W przypadku systemu CMMS QRMAINT szczególnie istotne jest uwzględnienie zarówno aspektów technicznych, jak i organizacyjnych, gdyż implementacja tego typu rozwiązań wykracza poza obszar *stricte* informatyczny, wpływając bezpośrednio na procesy utrzymania ruchu, efektywność pracy oraz kulturę organizacyjną przedsiębiorstwa.

Do mocnych stron (Strengths) analizowanego systemu należy zaliczyć przede wszystkim intuicyjny interfejs użytkownika oraz mobilność, przejawiającą się m.in. w możliwości wykorzystania kodów QR do szybkiej identyfikacji zasobów w terenie. Ponadto QRMAINT umożliwia automatyzację planowania i monitorowania prac serwisowych oraz oferuje zaawansowane funkcje raportowania i analizy danych, co przekłada się na redukcję przestojów maszyn oraz wzrost efektywności procesów produkcyjnych.

Należy jednak podkreślić obecność słabych stron (Weaknesses): relatywnie wysokie koszty wdrożenia – istotne szczególnie z perspektywy małych i średnich przedsiębiorstw – oraz konieczność odpowiedniego przeszkolenia personelu. System jest ponadto silnie uzależniony od stabilności infrastruktury teleinformatycznej i dostępu do Internetu, a w dużych organizacjach może wymagać złożonej i czasochłonnej konfiguracji.

W obszarze szans (Opportunities) należy wskazać potencjał dalszego rozwoju sztucznej inteligencji i koncepcji predictive maintenance, które pozwalają przewidywać awarie i optymalizować cykle konserwacji. Dodatkowym atutem jest możliwość integracji systemu QRMAINT z innymi rozwiązaniami klasy ERP, IoT czy BI, co sprzyja budowie spójnych ekosystemów cyfrowych, zgodnych z paradygmatem Przemysłu 4.0. Rosnąca świadomość znaczenia prewencyjnego utrzymania ruchu w przedsiębiorstwach dodatkowo wzmacnia pozycję rynkową tego typu rozwiązań.

Jednocześnie należy mieć świadomość zagrożeń (Threats). Wśród nich szczególne znaczenie ma ryzyko cyberataków i związane z tym problemy bezpieczeństwa danych, które stają się jednym z kluczowych wyzwań współczesnych systemów informatycznych. Należy także uwzględnić możliwy opór pracowników

wobec nowych technologii, presję konkurencyjnych rozwiązań dostępnych na rynku oraz ryzyko podejmowania błędnych decyzji w oparciu o niepełne lub niewłaściwie zinterpretowane dane.

Przeprowadzona analiza SWOT wskazuje, że system CMMS QRMAINT posiada istotny potencjał rozwojowy i może stanowić kluczowe narzędzie w procesie cyfryzacji utrzymania ruchu. Jednakże jego skuteczna implementacja wymaga uwzględnienia zarówno ograniczeń technologicznych, jak i czynników organizacyjnych oraz społecznych, które mogą determinować sukces lub porażkę wdrożenia.

Bibliografia

- Kowalski, J. (2021). Wdrażanie systemów CMMS w kontekście Przemysłu 4.0. *Zarządzanie i Inżynieria Produkcji*, 12(3), 45–56.
- Nowak, A. (2022). Analiza SWOT jako narzędzie oceny wdrożeń systemów informatycznych. *Ekonomika i Organizacja Przedsiębiorstwa*, 7(989), 22–31.
- Sitarz, M. (red.). (2023). *Transport w epoce zróżnicowanych potrzeb. Przesiadka bez barier*. Wydawnictwo Naukowe Akademii WSB w Dąbrowie Górniczej.
- Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2021/782 z dnia 29 kwietnia 2021 r. dotyczące praw i obowiązków pasażerów w ruchu kolejowym.
- Rozporządzenie Komisji (UE) nr 454/2011 z dnia 5 maja 2011 r. w sprawie technicznej specyfikacji interoperacyjności odnoszącej się do podsystemu „Aplikacje telematyczne dla przewozów pasażerskich” transeuropejskiego systemu kolei.
- Rozporządzenie Komisji (UE) nr 1299/2014 z dnia 18 listopada 2014 r. dotyczące technicznych specyfikacji interoperacyjności podsystemu „Infrastruktura” systemu kolei w Unii Europejskiej.
- <https://qrmaint.pl/>
- <https://www.kone.pl/uslugi/konserwacja-urzadzen/inteligentny-serwis-KONE-247-connect/>

Dynamika innowacyjności przedsiębiorstw w Polsce w ujęciu regionalnym

Streszczenie

Celem artykułu jest analiza dynamiki działalności innowacyjnej przedsiębiorstw przemysłowych w Polsce w latach 2020–2022, z uwzględnieniem zróżnicowania regionalnego oraz podziału na klasy wielkości firm. Badanie koncentruje się na wdrażaniu nowych lub ulepszonych produktów oraz procesów biznesowych. Zastosowano metodę desk research, bazującą na danych wtórnych pochodzących z Banku Danych Lokalnych GUS. Analiza obejmowała podział na przedsiębiorstwa małe, średnie i duże. Wyniki wskazują na silne zróżnicowanie regionalne oraz istotny wpływ wielkości przedsiębiorstwa na intensywność działań innowacyjnych. Duże przedsiębiorstwa wykazywały największą stabilność innowacyjną, natomiast firmy małe i średnie były bardziej podatne na czynniki zewnętrzne, w tym pandemię COVID-19. W 2022 roku nastąpiło odbicie po spadku innowacyjności w 2021 roku. Analiza wskazuje na potrzebę dalszego wspierania rozwoju innowacji w regionach słabiej rozwiniętych oraz wzmacniania potencjału MŚP.

Słowa kluczowe: innowacyjność, przedsiębiorstwa przemysłowe, regionalne zróżnicowanie, innowacje produktowe, innowacje procesowe

1 Uniwersytet Jana Długosza w Częstochowie

The Dynamics of Innovation in Polish Enterprises in Regional Terms

Abstract

The aim of this article is to analyse the dynamics of innovative activity of industrial enterprises in Poland in 2020–2022, taking into account regional differences and the division into company size classes. The study focuses on the implementation of new or improved products and business processes. A desk research method was used, based on secondary data from the Local Data Bank of the Central Statistical Office. The analysis covered a division into small, medium and large enterprises. The results indicate strong regional differences and a significant impact of company size on the intensity of innovation activities. Large companies showed the greatest innovation stability, while small and medium-sized companies were more vulnerable to external factors, including the COVID-19 pandemic. In 2022, there was a rebound after a decline in innovation in 2021. The analysis points to the need to continue supporting innovation in less developed regions and strengthening the potential of SMEs.

Keywords: innovation, industrial enterprises, regional diversity, product innovation, process innovation

Wstęp

Współczesna gospodarka opiera się na innowacyjności, która stała się jednym z kluczowych czynników rozwoju (Rykała, 2024). Zagadnienie to zyskuje coraz większe znaczenie, szczególnie w kontekście nauk o zarządzaniu, gdzie obserwuje się złożone i wielowymiarowe podejście do innowacji (Sopińska, 2025). Wraz z ewolucją rozumienia samego pojęcia dominującym kierunkiem stało się podejście interdyscyplinarne (Olejniczuk-Merta, 2013; Schumpeter, 2006). W literaturze z zakresu zarządzania coraz rzadziej traktuje się innowację wyłącznie jako nową technologię. Obecnie postrzegana jest ona jako szerokie typologiczne pojęcie, obejmujące różne obszary działalności człowieka – od zarządzania organizacją, przez strategie marketingowe, po relacje z otoczeniem i sposoby funkcjonowania przedsiębiorstw (Barcik, 2020).

Na podstawie analizy literatury przedmiotu w artykule przyjęto, że innowacja jest procesem złożonym z zespołu działań zmierzających do wprowadzenia nowych bądź udoskonalonych produktów, technologii lub rozwiązań organizacyjnych.

Zgodnie z Europejskim Rankiem Innowacyjności 2023 Polska to najbardziej dynamicznie rozwijające się państwo europejskie w obszarze innowacyjnej gospodarki, wykazujące znaczący postęp w takich dziedzinach jak: cyfryzacja, badania i rozwój, transfer technologii oraz wdrażanie nowoczesnych rozwiązań w sektorze przedsiębiorstw i administracji publicznej (Kleszcz, 2021; Wroński et al., 2025; GUS, 2023). Dane Eurostatu wskazują, że do końca 2021 roku kraj przeznaczył ponad 8,3 miliarda euro na działalność badawczo-rozwojową (Eurostat, 2025). Dla porównania, w chwili akcesji do Unii Europejskiej wartość ta była aż o 624% niższa, co czyni Polskę jednym z liderów pod względem tempa wzrostu nakładów na B+R wśród państw członkowskich (PARP, 2024).

Celem artykułu jest analiza przestrzenno-czasowej dynamiki innowacyjności przedsiębiorstw w Polsce w latach 2020–2022, z uwzględnieniem zróżnicowania regionalnego oraz klasy wielkości przedsiębiorstw. Badanie koncentruje się na dwóch podstawowych formach działalności innowacyjnej: wdrażaniu nowych lub ulepszonych produktów oraz nowych lub ulepszonych procesów biznesowych w przedsiębiorstwach przemysłowych. Analiza została przeprowadzona w podziale na trzy klasy wielkości firm: małe (10–49 pracowników), średnie (50–249) oraz duże (250 i więcej). Dane źródłowe zaczerpnięto z Banku Danych Lokalnych GUS, a metodologia badawcza opiera się na jakościowej analizie porównawczej wskaźników innowacyjności w układzie wojewódzkim. Opracowanie uwzględnia także zmiany w czasie, co pozwala uchwycić wpływ pandemii COVID-19 i jej następstw na zdolność przedsiębiorstw do podejmowania działań innowacyjnych.

Wyniki przeprowadzonej analizy ujawniają silne zróżnicowanie terytorialne w zakresie aktywności innowacyjnej przedsiębiorstw, przy jednoczesnym wpływie klasy wielkości firmy na skalę wdrażanych innowacji. Duże przedsiębiorstwa cechowały się największą stabilnością i intensywnością działań innowacyjnych, niezależnie od lokalizacji, podczas gdy firmy małe i średnie wykazywały większą wrażliwość na czynniki zewnętrzne i różnice regionalne. Szczególnie interesującym zjawiskiem jest odbudowa aktywności innowacyjnej w 2022 roku, po spadkach zanotowanych w 2021 roku, co odzwierciedla adaptacyjność sektora przedsiębiorstw do zmiennych warunków otoczenia społeczno-gospodarczego.

Nowatorstwo opracowania polega na ujęciu danych z ostatniego dostępnego trzyletniego okresu, z rozróżnieniem zarówno na rodzaj innowacji, jak i strukturę przedsiębiorstw według wielkości. Przedstawione wyniki ukazują nie tylko statystyczny obraz innowacyjności, ale również jej dynamikę i odporność na zakłócenia o charakterze systemowym, jakim była pandemia. Analiza ta wnosi potencjalny wkład do badań nad regionalnym rozwojem gospodarczym oraz polityką innowacyjną i może znaleźć zastosowanie zarówno w pracach akademickich, jak i w praktyce administracyjno-gospodarczej.

Metodyka badania

Analiza przedstawiona w niniejszej pracy została podporządkowana głównemu problemowi badawczemu, który sformułowano w postaci pytania: Czy w latach 2020–2022 wystąpiły zróżnicowania regionalne oraz zmiany w poziomie i dynamice innowacyjności przedsiębiorstw przemysłowych w Polsce, z uwzględnieniem wielkości firm i wpływu pandemii COVID-19?

Zastosowano metodę desk research, która polega na analizie danych wtórnych/zastanych, już istniejących i wcześniej zebranych przez instytucje zewnętrzne, celem sformułowania nowych wniosków badawczych (Bednarowska, 2015). Metoda ta została wykorzystana do interpretacji informacji statystycznych dotyczących innowacyjności przedsiębiorstw w Polsce.

Badanie obejmuje działalność innowacyjną przedsiębiorstw przemysłowych w Polsce, ze szczególnym uwzględnieniem wprowadzania innowacji produktowych oraz nowych i ulepszonych procesów biznesowych. Zakresem podmiotowym objęto przedsiębiorstwa zatrudniające co najmniej 10 pracowników, przy czym analizę przeprowadzono w trzech klasach wielkości: 10–49, 50–249 oraz 250 i więcej zatrudnionych.

Zakres czasowy obejmuje lata 2020, 2021 i 2022, co pozwala na ocenę dynamiki zmian oraz uchwycenie wpływu pandemii COVID-19 na aktywność innowacyjną przedsiębiorstw. Analiza dotyczy całego terytorium Polski i została

przeprowadzona z uwzględnieniem podziału na 16 województw, co umożliwia wskazanie regionalnych różnicowań w poziomie innowacyjności.

Głównym źródłem danych była baza Banku Danych Lokalnych (BDL) Głównego Urzędu Statystycznego, z której pozyskano wskaźniki opisujące udział przedsiębiorstw przemysłowych wprowadzających nowe lub ulepszone produkty oraz nowe lub ulepszone procesy biznesowe w ogólnej liczbie firm.

Dane statystyczne zostały poddane weryfikacji krzyżowej oraz zorganizowane w tabele porównawcze i wykresy.

Należy zaznaczyć, iż badanie działalności innowacyjnej przedsiębiorstw w Polsce prowadzone przez Główny Urząd Statystyczny opiera się na zharmonizowanej metodologii międzynarodowej (GUS, 2023). Od 2018 roku wdrożono zrewidowaną, czwartą edycję *Podręcznika Oslo*, opracowanego przez Organizację Współpracy Gospodarczej i Rozwoju (OECD) i Eurostat. Zmiana ta przyniosła modyfikację klasyfikacji innowacji, polegającą na wprowadzeniu nowej kategorii – innowacji procesów biznesowych – oraz rezygnacji z wcześniejszego podziału na innowacje procesowe, organizacyjne i marketingowe (OECD/Unia Europejska, 2018). W efekcie obecnie za działalność innowacyjną uznaje się wszelkie działania naukowe, techniczne, organizacyjne, finansowe i komercyjne, które rzeczywiście prowadzą lub mają na celu prowadzenie do wdrożenia innowacji (Suchocka, Kaim-Mierzwicka, 2022). Działania te mogą mieć bezpośrednio charakter nowości lub stanowić niezbędne elementy procesu wdrażania innowacji. Do działalności innowacyjnej zalicza się również działalność badawczo-rozwojową (B+R), niezależnie od tego, czy jest ona bezpośrednio związana z opracowaniem konkretnej innowacji (OECD/Unia Europejska, 2018).

GUS od 2023 roku badanie działalności innowacyjnej prowadzi w cyklu dwuletnim, co oznacza, że dane będą publikowane wyłącznie za lata parzyste. Wcześniejsze badania prowadzono corocznie. W latach 2020–2022, objętych analizą w niniejszym artykule, objęto pełnym badaniem wszystkie przedsiębiorstwa przemysłowe zatrudniające co najmniej 50 pracowników, natomiast przedsiębiorstwa przemysłowe zatrudniające od 10 do 49 osób objęto badaniem na próbie reprezentacyjnej.

Warto również podkreślić, że jednym z istotnych ograniczeń badania jest zakres czasowy dostępnych danych statystycznych. Dane udostępniane przez Główny Urząd Statystyczny dotyczą aktywności innowacyjnej przedsiębiorstw jedynie do roku 2022, co – biorąc pod uwagę dynamiczne zmiany w gospodarce krajowej i globalnej – nie pozwala na pełną ocenę długofalowych trendów oraz skutków polityk publicznych wdrażanych po pandemii COVID-19 i w kontekście nowych wyzwań, takich jak wojna w Ukrainie, inflacja czy transformacja energetyczna. Brak aktualizacji danych za rok 2023 i kolejne lata ogranicza możliwość przeprowadzenia analizy opartej na pełnym i aktualnym obrazie rzeczywistości gospodarczej, a tym

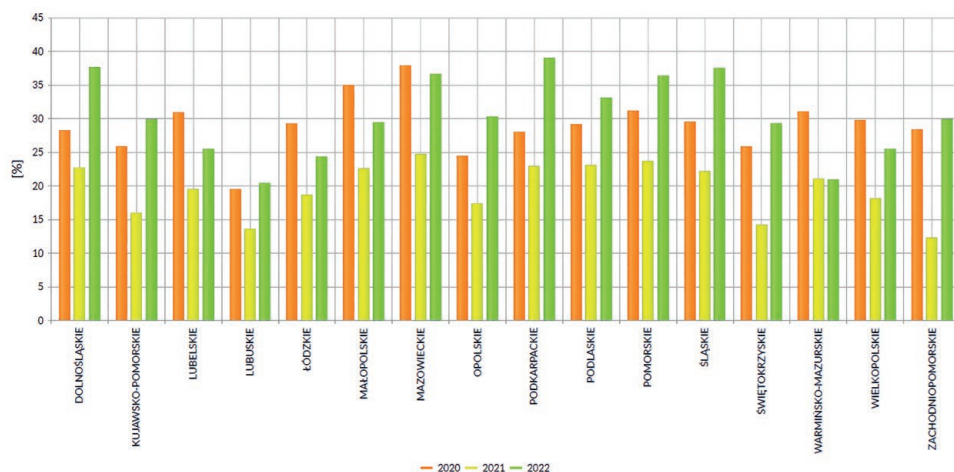
samym wpływa na potencjalną trafność prognoz oraz wniosków odnoszących się do długoterminowych zmian strukturalnych w systemie innowacji w Polsce.

Wyniki i dyskusja

Aktywność innowacyjna przedsiębiorstw stanowi istotny wskaźnik rozwoju regionalnego i potencjału gospodarczego danego obszaru (Reformat, 2018). W Polsce obserwuje się znaczne zróżnicowanie w poziomie innowacyjności firm, zarówno w ujęciu przestrzennym, jak i czasowym. Analiza danych statystycznych dotyczących udziału przedsiębiorstw innowacyjnych w ogólnej liczbie firm pozwala ocenić, które województwa wyróżniają się pod względem dynamiki rozwoju innowacji, a które wymagają dodatkowego wsparcia.

Wykres 1 przedstawia średni udział przedsiębiorstw innowacyjnych w ogólnej liczbie firm w poszczególnych województwach Polski w latach 2020, 2021 i 2022. Dane te umożliwiają analizę przestrzenną i czasową aktywności innowacyjnej sektora przedsiębiorstw w Polsce oraz identyfikację regionów o najwyższym i najniższym poziomie zaangażowania w działania innowacyjne.

Wykres 1. Średni udział przedsiębiorstw innowacyjnych w ogólnej liczbie przedsiębiorstw w Polsce w latach 2020–2022



Źródło: GUS, *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*. GUS, <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spoleczenstwo-informacyjne/nauka-i-technika/dzialalnosc-innowacyjna-przedsiębiorstw-w-latach-2020-2022,2,22.html> [dostęp: 04.07.2025]

W 2022 roku najwyższy poziom innowacyjności osiągnęły województwa: śląskie, mazowieckie oraz pomorskie – w tych regionach udział przedsiębiorstw innowacyjnych wyniósł od około 37% do 39%. Ich wysoka pozycja nie jest

przypadkowa – są to obszary o silnie rozwiniętych ośrodkach akademickich, instytucjach badawczo-rozwojowych oraz dużym udziale przemysłu zaawansowanego technologicznie. Dodatkowo sprzyjają im dobre warunki infrastrukturalne, dostęp do wykwalifikowanej kadry oraz większe możliwości pozyskiwania na innowacje środków zewnętrznych, zarówno krajowych, jak i unijnych.

Z kolei najniższy udział przedsiębiorstw innowacyjnych w 2022 roku odnotowano w województwach: zachodniopomorskim, kujawsko-pomorskim i lubuskim – wartości te nie przekroczyły 20%. Tak niski poziom może wskazywać na słabiej rozwinięte zaplecze naukowo-badawcze, mniejsze zaangażowanie firm w badania i rozwój oraz ograniczony dostęp do źródeł finansowania innowacji. W regionach tych szczególnie istotne może być wsparcie polityki regionalnej w zakresie rozwoju ekosystemów innowacyjnych i współpracy nauki z biznesem.

Analizując dane w ujęciu czasowym, zauważyć można ciekawą dynamikę zmian. W wielu województwach w 2021 roku nastąpił wyraźny spadek udziału przedsiębiorstw innowacyjnych względem roku 2020. Było to zjawisko ogólnokrajowe, obserwowane niezależnie od poziomu wyjściowego danego regionu. Spadek ten można powiązać przede wszystkim z wpływem pandemii COVID-19 na działalność gospodarczą w Polsce.

W 2021 roku wiele firm koncentrowało się na działaniach związanych z przetrwaniem – redukcji kosztów, utrzymaniu zatrudnienia i płynności finansowej – co naturalnie ograniczyło ich zdolność do inwestowania w innowacje. Projekty badawczo-rozwojowe często wymagały współpracy z partnerami zewnętrznymi, co było utrudnione przez lockdowny, ograniczenia kontaktów i pracę zdalną. Wiele działań innowacyjnych zostało zawieszonych lub odłożonych w czasie.

W roku 2021 przypadł również okres przejściowy między unijnymi perspektywami finansowymi 2014–2020 i 2021–2027. To przejście wiązało się z mniejszą dostępnością środków europejskich na badania, rozwój i wdrażanie innowacji. Przedsiębiorstwa miały zatem ograniczone możliwości pozyskania zewnętrznego wsparcia na realizację projektów innowacyjnych. Dodatkowo popyt na nowe rozwiązania technologiczne i produkty innowacyjne uległ zmniejszeniu – konsumenci i inne firmy w warunkach kryzysu preferowali stabilne, znane i często tańsze rozwiązania.

Do przyczyn spadku innowacyjności należy również zaliczyć problemy organizacyjne – niedobory kadrowe, konieczność reorganizacji pracy, a także ograniczony dostęp do specjalistów, którzy są kluczowi w procesie tworzenia i wdrażania innowacji.

W 2022 roku sytuacja zaczęła się poprawiać. W większości województw odnotowano wzrost udziału firm innowacyjnych, co może świadczyć o odbudowie aktywności po kryzysie pandemicznym, większej stabilizacji gospodarczej oraz rozpoczęciu wdrażania nowych środków z unijnej perspektywy 2021–2027.

Przedsiębiorstwa zaczęły ponownie podejmować działania rozwojowe, zwiększał się popyt na innowacyjne rozwiązania, a otoczenie instytucjonalne i finansowe stopniowo wracało do normalnego funkcjonowania.

Innowacje produktowe

W analizie regionalnej innowacyjności przedsiębiorstw przemysłowych w Polsce warto zwrócić uwagę na udział firm wdrażających nowe lub ulepszone produkty, ze szczególnym uwzględnieniem małych przedsiębiorstw, zatrudniających od 10 do 49 osób (tabela 1).

Tabela 1. Udział innowacji produktowych w małych przedsiębiorstwach w Polsce w latach 2020–2022 według województw (w %)

Województwo	2020 rok	2021 rok	2022 rok
dolnośląskie	13,7	9,3	11,2
kujawsko-pomorskie	7,1	4,6	3,9
lubelskie	15,3	7,3	8,3
lubuskie	11,5	5,5	7,1
łódzkie	14,1	8,9	12,0
małopolskie	13,3	7,8	7,7
mazowieckie	12,8	7,4	8,9
opolskie	9,1	5,1	6,4
podkarpackie	15,5	15,2	13,3
podlaskie	13,1	11,7	10,1
pomorskie	13,9	12,0	11,7
śląskie	13,9	10,1	10,0
świętokrzyskie	9,2	6,7	4,1
warmińsko-mazurskie	8,9	8,4	9,1
wielkopolskie	15,8	7,3	8,9
zachodniopomorskie	12,3	4,2	3,6

Źródło: opracowanie własne na podstawie: *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*. GUS, <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spolescenstwo-informacyjne/nauka-i-technika/dzialalnosc-innowacyjna-przedsiębiorstw-w-latach-2020-2022,2,22.html> [dostęp: 04.07.2025]

Dane przedstawione w tabeli obrazują odsetek przedsiębiorstw przemysłowych, które wdrożyły innowacje produktowe w latach 2020–2022, z podziałem na województwa. Innowacje produktowe oznaczają wprowadzenie na rynek nowych lub istotnie ulepszonych wyrobów lub usług – zarówno pod względem ich cech technicznych, funkcjonalności, przeznaczenia, jak i technologii wykonania.

W 2020 roku w większości regionów udział przedsiębiorstw wdrażających innowacje był stosunkowo wysoki, a w niektórych przypadkach osiągał bardzo dobre wyniki – np. w województwie wielkopolskim (15,8%), podkarpackim (15,5%), lubelskim (15,3%) czy łódzkim (14,1%). Świadczy to, że mimo pandemii wiele firm podjęło działania w zakresie modernizacji oferty produktowej.

Jednak już w 2021 roku odnotowano wyraźny spadek udziału przedsiębiorstw innowacyjnych w niemal wszystkich województwach. Największe spadki miały miejsce w zachodniopomorskim (z 12,3% do 4,2%), kujawsko-pomorskim (z 7,1% do 4,6%), lubuskim (z 11,5% do 5,5%) i wielkopolskim (z 15,8% do 7,3%). Można to interpretować jako efekt niepewności gospodarczej, spadku inwestycji oraz ograniczonego popytu, co zniechęcało firmy do ponoszenia kosztów związanych z wdrażaniem nowych produktów.

W 2022 roku sytuacja częściowo się poprawiła – wiele województw zanotowało wzrost udziału firm wprowadzających innowacje produktowe. Przykładowo, w dolnośląskim wzrost z 9,3% do 11,2%, w łódzkim z 8,9% do 12,0%, a w warmińsko-mazurskim z 8,4% do 9,1%. Mimo to tylko nieliczne regiony powróciły do poziomu z 2020 roku, a w wielu udział firm innowacyjnych nadal był niższy niż przed pandemią.

Najbardziej wyróżniającym się województwem było podkarpackie, które jako jedyne utrzymało wysoki i względnie stabilny poziom innowacyjności w całym analizowanym okresie (15,5% → 15,2% → 13,3%). Wysoki udział utrzymywały także województwa pomorskie i podlaskie. Z kolei najniższy odsetek firm wprowadzających innowacje produktowe w 2022 roku odnotowano w województwie zachodniopomorskim (3,6%) oraz kujawsko-pomorskim (3,9%).

Średnie przedsiębiorstwa (zatrudniające od 50 do 249 osób) odgrywają również istotną rolę w polskiej gospodarce, zarówno jako pracodawcy, jak i podmioty zdolne do prowadzenia działalności innowacyjnej na większą skalę niż mikro- i małe firmy. Analiza aktywności innowacyjnej tej grupy przedsiębiorstw, mierzonej udziałem firm wdrażających nowe lub ulepszone produkty, wskazuje na stosunkowo wysoki i stabilny poziom innowacyjności w latach 2020–2022, pomimo występowania globalnych zawirowań gospodarczych (tabela 2).

Tabela 2. Udział innowacji produktowych w średnich przedsiębiorstwach w Polsce w latach 2020-2022 według województw (w %)

Województwo	2020 rok	2021 rok	2022 rok
dolnośląskie	23,5	23,3	24,5
kujawsko-pomorskie	28,4	28,7	29,9
lubelskie	25,7	23,1	24,0
lubuskie	22,5	20,1	20,0
łódzkie	31,3	28,2	27,3
małopolskie	28,7	25,1	24,9
mazowieckie	28,3	28,5	30,8
opolskie	29,0	27,6	28,3
podkarpackie	36,5	32,4	33,3
podlaskie	34,7	33,5	35,7
pomorskie	28,6	25,6	26,4
śląskie	27,9	25,2	25,3
świętokrzyskie	24,2	24,7	24,7
warmińsko-mazurskie	31,6	30,9	29,6
wielkopolskie	25,4	24,0	23,0
zachodniopomorskie	22,0	20,1	19,3

Źródło: opracowanie własne na podstawie: *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*. GUS, <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spoleczenstwo-informacyjne/nauka-i-technika/dzialalnosc-innowacyjna-przedsiębiorstw-w-latach-2020-2022,2,22.html> [dostęp: 04.07.2025].

W 2020 roku, mimo wybuchu pandemii COVID-19, średnie przedsiębiorstwa w wielu regionach Polski wykazywały silną aktywność innowacyjną. Najwyższe wskaźniki odnotowano w województwach: podkarpackim (ok. 36–37%), łódzkim, podlaskim i warmińsko-mazurskim (ok. 30–33%). Dane te sugerują, że firmy tej wielkości były w stanie relatywnie szybko reagować na zmieniające się warunki

rynkowe, m.in. dostosowując ofertę produktową, wdrażając rozwiązania cyfrowe lub innowacje procesowe przekładające się na nowe produkty.

Rok 2021 przyniósł tylko nieznaczny spadek aktywności innowacyjnej średnich przedsiębiorstw. W większości województw utrzymano poziomy zbliżone do roku poprzedniego, a różnice były niewielkie. W porównaniu do sektora małych przedsiębiorstw, które w tym samym okresie odnotowały wyraźne załamanie aktywności innowacyjnej, firmy średnie wykazały się większą odpornością na negatywne skutki pandemii. Może to wynikać z ich stabilniejszej pozycji finansowej, bardziej rozwiniętych struktur zarządzania oraz lepszego dostępu do zewnętrznych źródeł finansowania.

W 2022 roku nastąpiła lekka poprawa wskaźników w części województw, takich jak: podlaskie, podkarpackie i mazowieckie, gdzie odsetek firm wdrażających nowe produkty ponownie wzrósł. W wielu innych regionach utrzymano poziomy zbliżone do tych z lat poprzednich. Najniższe wartości w analizowanym okresie występowały konsekwentnie w województwach: zachodniopomorskim, wielkopolskim i lubuskim, gdzie wskaźnik innowacyjności produktowej wśród średnich przedsiębiorstw nie przekraczał 24–25%.

Przyglądając się wynikom regionalnym, można zauważyć wyraźne zróżnicowanie w potencjale innowacyjnym średnich firm. Regiony wschodniej Polski (np. podkarpackie, podlaskie) oraz częściowo centralne (łódzkie, mazowieckie) prezentowały znacznie wyższy poziom wdrażania innowacji niż regiony zachodnie i północno-zachodnie. Wskazuje to na skuteczność lokalnych polityk wspierających innowacje, a także potencjalne różnice w strukturze branżowej firm i stopniu ich powiązań z sektorem badawczo-rozwojowym.

Średnie przedsiębiorstwa w Polsce charakteryzują się relatywnie wysoką i odporną na kryzysy aktywnością innowacyjną, szczególnie w zakresie wdrażania nowych lub ulepszonych produktów. Pomimo globalnych wyzwań w latach 2020–2022 w wielu regionach utrzymały lub nawet zwiększyły swoją innowacyjność. Wyniki te sugerują, że firmy średniej wielkości pełnią kluczową rolę w procesie dyfuzji innowacji w gospodarce regionalnej i powinny być istotnym adresatem polityki wspierającej innowacyjność w Polsce.

Duże przedsiębiorstwa pełnią kluczową rolę w systemie innowacyjnym każdego kraju, ze względu na posiadane zasoby, zdolność do prowadzenia działalności badawczo-rozwojowej oraz oddziaływanie na otoczenie biznesowe i społeczne. Analiza danych Głównego Urzędu Statystycznego dotyczących udziału dużych firm wdrażających nowe lub ulepszone produkty w latach 2020–2022 wskazuje na stosunkowo wysoki poziom aktywności innowacyjnej, choć z zauważalnymi różnicami regionalnymi oraz lekką tendencją spadkową w skali kraju (tabela 3).

Tabela 3. Udział innowacji produktowych w dużych przedsiębiorstwach w Polsce w latach 2020–2022 według województw (w %)

Województwo	2020 rok	2021 rok	2022 rok
dolnośląskie	45,3	40,5	38,2
kujawsko-pomorskie	45,9	47,3	48,5
lubelskie	41,8	45,0	48,2
lubuskie	38,4	35,1	34,0
łódzkie	42,7	41,0	45,3
małopolskie	60,2	59,5	58,7
mazowieckie	52,0	50,3	49,5
opolskie	55,7	54,1	53,0
podkarpackie	62,0	61,0	59,6
podlaskie	61,5	60,5	59,8
pomorskie	49,1	48,9	48,7
śląskie	44,7	42,5	45,1
świętokrzyskie	41,5	41,0	42,0
warmińsko-mazurskie	48,0	50,9	52,5
wielkopolskie	50,8	51,3	50,2
zachodniopomorskie	41,0	39,5	33,1

Źródło: opracowanie własne na podstawie: *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*. GUS, <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spoleczenstwo-informacyjne/nauka-i-technika/dzialalnosc-innowacyjna-przedsiębiorstw-w-latach-2020-2022,2,22.html> [dostęp: 04.07.2025].

W 2020 roku w niemal wszystkich województwach ponad 40% dużych przedsiębiorstw deklarowało wdrożenie innowacji produktowych. Szczególnie wysoki poziom odnotowano w województwach: podkarpackim, podlaskim, małopolskim i warmińsko-mazurskim, gdzie wskaźnik ten przekraczał 60%. Wysoka aktywność w tych regionach może być wynikiem koncentracji dużych firm przemysłowych oraz dobrze rozwiniętej współpracy z sektorem nauki i jednostkami badawczo-rozwojowymi.

W roku 2021 poziom innowacyjności utrzymał się na zbliżonym poziomie w większości województw, przy czym w niektórych regionach (np. mazowieckim, lubelskim, podkarpackim) zaobserwowano nieznaczne spadki. Mimo to ogólna aktywność innowacyjna dużych firm pozostała wysoka i przewyższała wyniki dla małych i średnich przedsiębiorstw. Oznacza to, że duże podmioty lepiej poradziły sobie z wyzwaniami pandemicznymi i kontynuowały działania rozwojowe mimo niepewności gospodarczej.

W 2022 roku obserwowano zróżnicowane trendy. W części województw – takich jak kujawsko-pomorskie, lubelskie, łódzkie czy warmińsko-mazurskie – nastąpił nawet wzrost udziału firm wdrażających nowe produkty. W innych – zwłaszcza zachodniopomorskim, dolnośląskim, lubuskim – aktywność ta spadła do poziomu około 35–38%, co oznacza zauważalny regres względem roku 2020. Przyczyn takiego zróżnicowania można upatrywać m.in. w strukturze gospodarki regionalnej, sile lokalnych ekosystemów innowacji oraz dostępności wykwalifikowanych kadr i funduszy.

Najwyższy poziom innowacyjności w całym analizowanym okresie utrzymywał się w regionach z rozwiniętym przemysłem wysokich technologii i aktywnym udziałem firm w międzynarodowych łańcuchach wartości – np. podkarpackie, małopolskie, podlaskie. Z kolei województwa o relatywnie niskim udziale innowacyjnych dużych firm to m.in. zachodniopomorskie, świętokrzyskie i lubuskie, gdzie wskaźniki innowacyjności produktowej nie przekraczały 40% w 2022 roku.

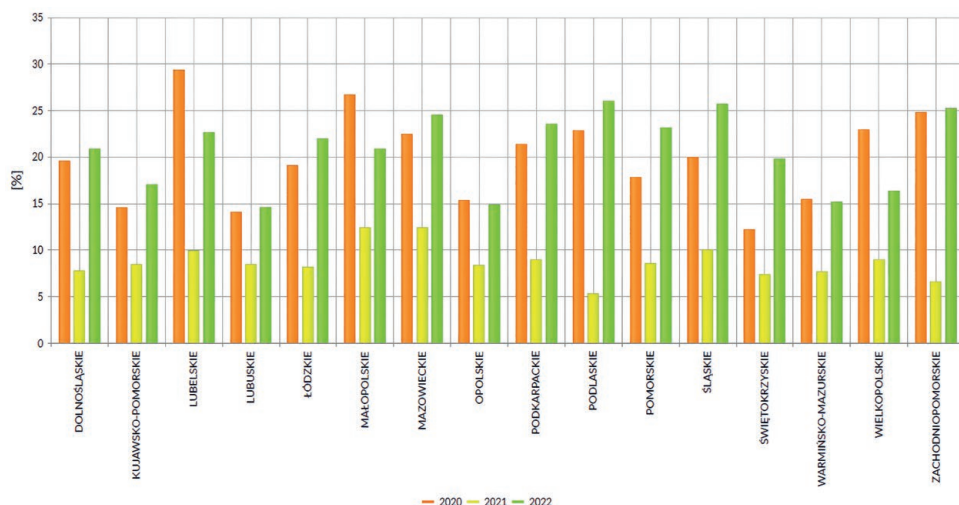
W latach 2020–2022 duże przedsiębiorstwa w Polsce charakteryzowały się wysoką aktywnością innowacyjną, zwłaszcza w zakresie wdrażania nowych lub ulepszonych produktów. Mimo globalnych i lokalnych wyzwań gospodarczych w wielu województwach poziom tej aktywności utrzymywał się na stabilnym poziomie, a w niektórych regionach nawet wzrastał. Równocześnie dane ujawniają potrzebę dalszego wspierania rozwoju innowacji w regionach słabszych gospodarczo oraz tworzenia warunków sprzyjających trwałej współpracy dużych firm z sektorem nauki i otoczeniem innowacyjnym.

Innowacje procesowe

Innowacje procesowe, rozumiane jako wdrażanie nowych lub istotnie ulepszonych metod organizacyjnych, operacyjnych lub zarządczych, stanowią ważny element podnoszenia efektywności przedsiębiorstw.

W przypadku małych firm pełnią one szczególną rolę jako relatywnie mniej kosztowne alternatywy dla innowacji produktowych. Dane Głównego Urzędu Statystycznego przedstawione na wykresie 2 ukazują poziom aktywności innowacyjnej w tym zakresie wśród małych firm w Polsce w latach 2020–2022, z rozbiorem na województwa.

Wykres 2. Nowe lub ulepszone procesy biznesowe w małych przedsiębiorstwach w latach 2020–2022 (w %)



Źródło: GUS, *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*, <https://bdl.stat.gov.pl/bdl/dane/podgrup/wykres> [dostęp: 30.06.2025]

W 2020 roku w wielu województwach udział małych firm wdrażających nowe procesy był stosunkowo wysoki – szczególnie wyróżniały się województwa: lubelskie (29%), małopolskie, łódzkie i zachodniopomorskie, gdzie wskaźnik ten wyniósł od 25 do 29%. Dane te mogą sugerować, że w początkowej fazie pandemii COVID-19 firmy podejmowały szybkie działania reorganizacyjne, m.in. przechodząc na pracę zdalną, wdrażając nowe systemy komunikacji czy dostosowując procedury do warunków sanitarnych.

W 2021 roku odnotowano jednak wyraźny spadek aktywności innowacyjnej w zakresie procesów biznesowych w niemal wszystkich województwach. Przykładowo, w województwach: podlaskim, pomorskim, świętokrzyskim i warmińsko-mazurskim udział firm innowacyjnych procesowo spadł nawet poniżej 10%. Spadek ten można tłumaczyć zmęczeniem organizacyjnym, wyczerpaniem zasobów firm oraz skupieniem się na utrzymaniu bieżącej działalności w warunkach długotrwałej niepewności.

Rok 2022 przyniósł jednak znaczną poprawę w wielu regionach. W województwach: śląskim, zachodniopomorskim, podkarpackim, dolnośląskim i mazowieckim wskaźnik wdrażania innowacji procesowych wzrósł do poziomu powyżej 20%, osiągając wartości zbliżone lub nawet wyższe niż w 2020 roku. Może to świadczyć o adaptacji firm do nowych warunków oraz ponownym uruchomieniu działań rozwojowych po okresie stagnacji. Z drugiej strony, województwa takie jak: podlaskie, warmińsko-mazurskie i świętokrzyskie nadal pozostają w tyle,

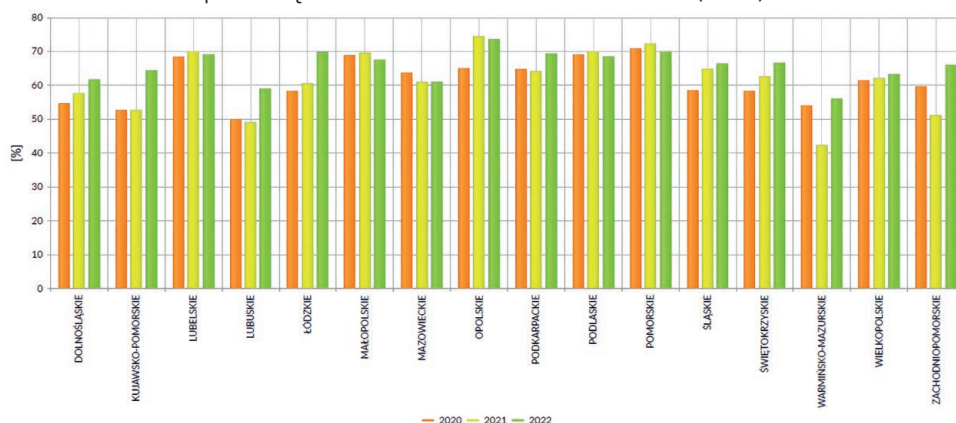
z wynikami poniżej 10%, co sugeruje utrzymujące się trudności we wdrażaniu zmian organizacyjnych.

W ujęciu przestrzennym widoczna jest wyraźna polaryzacja regionalna. Wysoką aktywność innowacyjną procesową wykazują regiony o większej liczbie ośrodków miejskich i lepiej rozwiniętej infrastrukturze technologicznej (np. śląskie, mazowieckie, dolnośląskie), natomiast regiony o charakterze peryferyjnym, takie jak warmińsko-mazurskie czy podlaskie, pozostają w tyle, co może wskazywać na potrzebę silniejszego wsparcia dla procesów modernizacyjnych w tych obszarach.

Powyższe dane ukazują niestabilny, ale w ostatnim okresie pozytywny trend w zakresie wdrażania innowacji procesowych przez małe przedsiębiorstwa w Polsce. Po spadku aktywności w 2021 roku, wywołanym prawdopodobnie przez przedłużający się kryzys pandemiczny, w 2022 roku zaobserwowano ponowne zwiększenie dynamiki zmian organizacyjnych w firmach. Wyniki te potwierdzają znaczenie elastyczności i zdolności adaptacyjnych małych firm oraz rolę, jaką pełnią innowacje procesowe w podnoszeniu ich konkurencyjności, szczególnie w warunkach kryzysowych i postkryzysowych.

Z kolei w latach 2020–2022 średnie przedsiębiorstwa w Polsce cechował stosunkowo wysoki poziom wdrażania innowacji procesowych. W większości województw odsetek firm podejmujących takie działania utrzymywał się na poziomie powyżej 35%, a w niektórych przypadkach – jak w województwach: lubelskim, podlaskim, podkarpackim czy warmińsko-mazurskim – przekraczał nawet 40%, co świadczy o dużej skłonności firm do doskonalenia wewnętrznych procesów zarządczych, operacyjnych i organizacyjnych (wykres 3).

Wykres 3. Nowe lub ulepszone procesy biznesowe w średnich przedsiębiorstwach w latach 2020–2022 (w %)



Źródło: GUS, *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*, <https://bdl.stat.gov.pl/bdl/dane/podgrup/wykres> [dostęp: 30.06.2025]

W 2020 roku, pomimo rozpoczęcia pandemii COVID-19, wiele firm średniej wielkości podjęło działania reorganizacyjne, co znalazło odzwierciedlenie w stosunkowo wysokim wskaźniku wdrożeń procesowych. Największą aktywność odnotowano w województwach: opolskim, podkarpackim i warmińsko-mazurskim (od 43% do 46%). Sytuacja ta może być wynikiem potrzeby szybkiego dostosowania się do nowych realiów gospodarczych, takich jak: praca zdalna, cyfryzacja procesów czy zmiany logistyczne.

Rok 2021 przyniósł niewielkie wahania w poziomie aktywności innowacyjnej, jednak nie zaobserwowano tak drastycznych spadków, jak w sektorze małych firm. W niektórych województwach – np. świętokrzyskim, śląskim, wielkopolskim – odsetek firm wdrażających procesy był nieco niższy niż w roku poprzednim, co mogło być skutkiem trwającej niepewności i ograniczonego dostępu do zasobów. Z kolei w takich regionach jak: lubelskie, podlaskie czy kujawsko-pomorskie aktywność ta wzrosła, osiągając najwyższe poziomy w całym badanym okresie.

W 2022 roku innowacyjność procesowa średnich firm ponownie wzrosła w wielu województwach. Szczególnie wyróżniały się: lubelskie (ok. 44%), podlaskie (ok. 43%) oraz podkarpackie i warmińsko-mazurskie, gdzie wskaźniki utrzymywały się na bardzo wysokim poziomie. Dane te wskazują na rosnącą świadomość znaczenia innowacji organizacyjnych jako kluczowego elementu poprawy efektywności działania oraz przygotowania firm na przyszłe wyzwania rynkowe.

Warto również zauważyć, że w żadnym województwie wskaźnik wdrażania innowacji procesowych nie spadł poniżej 30% w żadnym roku analizowanego okresu, co świadczy o stosunkowo wysokim i stabilnym potencjale innowacyjnym w tym segmencie przedsiębiorstw. Jednocześnie różnice regionalne utrzymują się – np. województwa: świętokrzyskie, lubuskie i zachodniopomorskie pozostają na nieco niższym poziomie aktywności, co może wynikać z mniejszego dostępu do wyspecjalizowanego personelu, ograniczonych inwestycji technologicznych lub słabszego powiązania z ośrodkami naukowo-badawczymi.

Średnie przedsiębiorstwa w Polsce wykazują wysoką i stosunkowo równomierną w skali kraju aktywność innowacyjną w zakresie wdrażania nowych lub ulepszonych procesów biznesowych. Utrzymujący się na przestrzeni trzech lat stabilny poziom tych działań, mimo zmieniających się warunków makroekonomicznych, wskazuje na ich istotną rolę jako nośników zmian organizacyjnych i technologicznych w skali regionalnej. Tego rodzaju innowacje mogą stanowić kluczowy obszar rozwoju gospodarczego i powinny być wspierane zarówno na poziomie krajowym, jak i lokalnym.

W przypadku dużych firm wdrażanie nowych lub istotnie ulepszonych procesów biznesowych wiąże się często z ważnymi inwestycjami, ale również przynosi mierzalne korzyści w postaci lepszej organizacji pracy, obniżenia kosztów

czy zwiększenia jakości usług. Analiza danych statystycznych dla lat 2020–2022 pokazuje, że duże przedsiębiorstwa w Polsce wykazują bardzo wysoki poziom aktywności innowacyjnej procesowej, znacznie przewyższający wyniki firm małych i średnich.

W 2020 roku, mimo pandemii COVID-19, odsetek dużych przedsiębiorstw wdrażających innowacje procesowe był bardzo wysoki. W większości województw przekraczał 55–60%, a w niektórych, jak: małopolskie, opolskie, podkarpackie i podlaskie, osiągał nawet 70%. Dane te wskazują, że duże firmy jako pierwsze zareagowały na kryzys, dostosowując swoje struktury, modele zarządzania i systemy operacyjne do nowych warunków funkcjonowania. Skala i zasoby tych firm umożliwiły im szybką reorganizację, m.in. poprzez cyfryzację, automatyzację, wdrażanie nowych procedur zarządczych i reorganizację łańcuchów dostaw.

Rok 2021 przyniósł utrzymanie wysokiego poziomu innowacyjności, choć w niektórych województwach – warmińsko-mazurskim, zachodniopomorskim czy kujawsko-pomorskim – zaobserwowano nieznaczne spadki aktywności. Nadal jednak w skali kraju dominował poziom powyżej 60%, co oznacza, że duże przedsiębiorstwa konsekwentnie kontynuowały działania modernizacyjne, niezależnie od wyzwań gospodarczych i organizacyjnych.

W 2022 roku nastąpił ponowny wzrost udziału dużych firm wprowadzających nowe procesy, zwłaszcza w województwach, które wcześniej wykazywały nieco niższą aktywność. Najwyższe wartości osiągnęły ponownie: małopolskie, opolskie, podkarpackie i warmińsko-mazurskie, gdzie odsetek firm sięgał 70–75%. Nawet w województwach relatywnie słabszych pod względem innowacyjności – takich jak: lubuskie, kujawsko-pomorskie czy zachodniopomorskie – wskaźniki te zbliżały się do 60% lub przekraczały tę granicę, co świadczy o ogólnokrajowym trendzie intensyfikacji innowacji procesowych w dużym biznesie.

Charakterystyczną cechą innowacyjności procesowej dużych przedsiębiorstw w Polsce jest stosunkowo małe zróżnicowanie regionalne. O ile w przypadku firm małych i średnich obserwuje się silne rozwarstwienie między regionami wysoko rozwiniętymi a peryferyjnymi, o tyle duże przedsiębiorstwa wykazują wysoką aktywność innowacyjną niemal niezależnie od lokalizacji. Wynika to m.in. z faktu, że duże firmy często działają w skali ponadregionalnej lub międzynarodowej, dysponują lepszym dostępem do technologii, kapitału i kadr zarządzających oraz są bardziej zintegrowane z globalnymi łańcuchami wartości.

Dane z lat 2020–2022 potwierdzają, że duże przedsiębiorstwa w Polsce odgrywają kluczową rolę w rozwoju innowacji procesowych, niezależnie od sytuacji gospodarczej. Ich wysoka aktywność innowacyjna, stabilność oraz zdolność do adaptacji w warunkach kryzysu stanowią ważny filar nowoczesnej gospodarki, opartej na wiedzy i efektywnym zarządzaniu. Równocześnie ich rola jako

liderów zmian organizacyjnych może stanowić punkt odniesienia i źródło dobrych praktyk dla mniejszych podmiotów, które często potrzebują impulsu do wdrażania podobnych rozwiązań.

Podsumowanie

Przedstawione badanie dostarczyło diagnozę regionalnego zróżnicowania i dynamiki innowacyjności przedsiębiorstw w Polsce w latach 2020–2022, z uwzględnieniem ich wielkości oraz rodzaju wdrażanych innowacji. Uzyskane wyniki wskazują jednoznacznie, że największą i najbardziej stabilną aktywność innowacyjną wykazują przedsiębiorstwa duże, podczas gdy firmy małe są najbardziej podatne na zmiany otoczenia, co ujawniło się szczególnie w 2021 roku, w okresie bezpośrednich skutków pandemii COVID-19. Wyniki potwierdzają także, że województwa o rozwiniętej infrastrukturze instytucjonalnej i badawczej (np. mazowieckie, małopolskie) konsekwentnie przodują w zakresie wdrażania innowacji, podczas gdy regiony peryferyjne (np. świętokrzyskie, lubuskie) utrzymują się na niższym poziomie aktywności innowacyjnej.

Badanie może służyć jako baza porównawcza dla kolejnych analiz, obejmujących nowe okresy programowania unijnego, a także jako materiał źródłowy do projektowania narzędzi wsparcia w ramach regionalnych strategii inteligentnych specjalizacji (RIS3). Wnioski płynące z niniejszego opracowania wskazują na potrzebę wzmacniania zdolności innowacyjnej, szczególnie w segmencie małych firm, oraz wspierania rozwoju instytucji otoczenia biznesu w regionach o słabszym potencjale endogenicznym.

W przyszłych badaniach niezbędne będzie uwzględnienie najnowszych danych oraz kontynuowanie monitoringu zmian w zachowaniach innowacyjnych przedsiębiorstw w warunkach postpandemicznych i gospodarki opartej na transformacji cyfrowej i zielonej.

Bibliografia

- Barcik, A. (2020). Compliance systems as a determinant for the functioning of social responsibility instruments in an organization. *Zeszyty Naukowe. Organizacja i Zarządzanie*, 147, 9–10. Politechnika Śląska.
- Bednarowska, Z. (2015). Desk research – wykorzystanie potencjału danych zastanych w prowadzeniu badań marketingowych i społecznych. *Marketing i Rynek*, (7), 18–26.
- Cymerys, M. (2025). Cyfrowy model biznesu jako narzędzie innowacji w przedsiębiorstwie. W A. Weinert, I. Dudziuk (red.), *New Level. Zarządzanie organizacją i aspekty społeczne z zakresu prawa, bezpieczeństwa i socjologii*, 41. Wydawnictwo Naukowe FNCE.

- Główny Urząd Statystyczny (2023). *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*, <https://bdl.stat.gov.pl/bdl/dane/podgrup/wykres> [dostęp: 30.06.2025]
- Główny Urząd Statystyczny. (2023). *Działalność innowacyjna przedsiębiorstw w Polsce w latach 2020–2022*. GUS, <https://stat.gov.pl/obszary-tematyczne/nauka-i-technika-spoleczenstwo-informacyjne/nauka-i-technika/dzialalnosc-innowacyjna-przedsiębiorstw-w-latach-2020-2022,2,22.html> [dostęp: 04.07.2025]
- https://mfiles.pl/pl/index.php/Desk_research [dostęp: 03.07.2025]
- Kleszcz, A. (2021). Główne składowe innowacyjności w krajach Unii Europejskiej. *Wiomości Statystyczne. The Polish Statistician*, 66(8), 24–45.
- OECD/Unia Europejska. (2018). *Podręcznik Oslo 2018: Zalecenia dotyczące pozyskiwania, prezentowania i wykorzystywania danych dotyczących innowacji*. https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5496/18/1/1/podrecznik_oslo_2018_internet.pdf
- Okoń-Horodyńska, E. (2024). Kryzys i innowacje. Co jest grane? *Optimum. Economic Studies*, 115(1), 45–67.
- Olejniczuk-Merta, A. (2013). Innowacje społeczne. *Konsumpcja i Rozwój*, 1(4), 21.
- Reformat, B. (2018). Modele procesów innowacyjnych a stadia rozwoju współczesnej gospodarki. *Zeszyty Naukowe Politechniki Śląskiej. Seria: Organizacja i Zarządzanie*, (30), 473.
- Rykała, P. (2024). Wpływ instytucji otoczenia biznesu na rozwój gospodarczy miast województwa śląskiego. *Acta Universitatis Lodziensis. Folia Oeconomica*, 1(366), 62–83.
- Schumpeter, J. A. (2006). *Theorie der wirtschaftlichen Entwicklung* (ss. 104–105). Duncker & Humblot GmbH.
- Sopińska, A. (2025). Ekoinnowacje w MŚP jako element zrównoważonego rozwoju. Determinanty działalności ekoinnowacyjnej. *Studia i Prace Kolegium Zarządzania i Finansów*, (200), 127–145.
- Suchocka, E., & Kaim-Mierzwicka, K. (2022). *Znaczenie definicji innowacyjności*, <https://crido.pl/blog-business/znaczenie-definicji-innowacyjnosci/> [dostęp: 06.07.2025]

Poprawa efektywności procesu kompletacji towarów przy wykorzystaniu Internetu rzeczy oraz protokołu RFID, Z-Wave i 6LoWPAN

Streszczenie

Większość podmiotów na rynku poszukuje sposobów na optymalizację procesów transportowych i gospodarowania zasobami magazynowymi. Ich celem jest ograniczenie kosztów i poprawa efektywności działania. Organizacje na rynku zmagają się z rosnącą presją czasu i kosztów. Poszukują narzędzi, które umożliwią działalność efektywną ekonomicznie. Celem artykułu jest analiza i ocena zastosowania Internetu rzeczy w gospodarce magazynowej. Analizy dokonano w oparciu o prognozy eksperckie oraz badania nad zastosowaniem i funkcjonalnością systemów IoT w laboratorium logistycznym. W wyniku przeprowadzonych badań udowodniono, jak systemy oparte na protokole RFID, Z-Wave oraz 6LoWPAN poprawiają efektywność procesu logistycznego, przyczyniając się tym samym do skrócenia czasu kompletacji. W obliczu zmian technologicznych i rozwoju urządzeń ICT system IoT nabiera istotnego znaczenia w procesach magazynowych i kompletacyjnych. W związku z powyższym implikacja wskazanych rozwiązań istotnie przyczynia się do skrócenia czasu procesów logistycznych w magazynie.

Słowa kluczowe: logistyka, Internet and communication technology (ICT), Internet rzeczy (IoT)

1 Akademia im. Jakuba z Paradyża w Gorzowie Wielkopolskim

Improving the Efficiency of the Picking Process Using the Internet of Things, RFID, Z-Wave and 6LoWPAN Protocols

Abstract

Most entities on the market are looking for ways to optimize transport processes and warehouse resource management. Their goal is to reduce costs and improve operational efficiency. Organizations on the market are struggling with increasing time and cost pressure. They are looking for tools that will enable economically efficient operations. The aim of the article is to analyze and evaluate the use of the Internet of Things in warehouse management. The analysis was made based on expert forecasts and research on the use and functionality of IoT systems in the logistics laboratory. As a result of the conducted research, it was proven how systems based on the RFID, Z-Wave and 6LoWPAN protocols improve the efficiency of the logistics process, thus contributing to shortening the picking time. In the face of technological changes and the development of ICT devices, the IoT system is gaining significant importance in warehouse and picking processes. In connection with the above, the implication of the indicated solutions significantly contributes to shortening the time of logistics processes in the warehouse.

Keyword: logistics, Internet and communication technology (ICT), Internet of Things (IoT)

Wstęp

W obliczu zmian technologicznych i rozwoju technologii teleinformatycznych wynikających z rozwoju sieci Internet, sieci komórkowej oraz miniaturyzacji urządzeń do jej obsługi, obserwuje się zmiany w procesach logistycznych, transportowych oraz magazynowych. Zmiany w gospodarce zachodzą szybko, a postęp technologiczny powoduje, że przedsiębiorstwa funkcjonują w trudno przewidywalnych warunkach. Zmienne otoczenie gospodarcze, splecione siecią powiązań, cechuje złożona dynamika, powodująca konieczność ewaluowania systemów zarządzania procesami.

Przedsiębiorstwa próbują wdrażać rozwiązania, które przyczynią się do ich rozwoju oraz jak najlepszej pozycji rynkowej, koncentrując się przy tym na działaniach umożliwiających dostosowanie rozwiązań i oferty do oczekiwań współczesnego klienta.

Zarządzanie procesami magazynowymi angażuje integrację każdego z uczestników w łańcuchu dostaw – od producenta, poprzez centra magazynowe, aż po końcowego odbiorcę. Wzajemne integracje mogą być skomplikowane, szczególnie w sytuacji, gdy oferta składa się z wielu komponentów. W związku z tym logistyka procesu powinna być budowana w sposób efektywny i racjonalny, a dokonywana przy pomocy nowoczesnych narzędzi, umożliwiających optymalny kosztowo układ wzajemnych współzależności.

Kooperacja i przepływ informacji w łańcuchu dostaw, a tym samym dążenie do efektywności procesu przepływu towarów, są możliwe dzięki rozwojowi technologicznemu i ciągłemu rozwojowi technologii tworzących sieć ICT. Szukanie sposobu na ograniczanie kosztów funkcjonowania organizacji z jednej strony, a dążenie do zaspokajania rosnących oczekiwań konsumentów z drugiej – wymusza poszukiwanie najbardziej efektywnego, zbilansowanego systemu wykorzystania dostępnych zasobów. Wydaje się to pożądanym kierunkiem, który może sprostać zmianom w większości dziedzin życia.

Działalność transportowa i gospodarowanie zasobami magazynowymi to jedno z najważniejszych składników kosztowych dzisiejszej logistyki, która podlega wpływowi czynników rynkowych i pozarynkowych. Rozwój technologii wpływa i nadal będzie wpływać na organizowanie procesów, choćby z uwagi na automatyzację i robotyzację magazynów, rozwój systemów informatycznych, konieczność analizy dużych pakietów danych lub informacji, ciągłej integracji procesów magazynowych z logistyką ostatniej mili i rynkiem e-commerce oraz z brakami personelu na rynku pracy. Te ostatnie są szczególnie istotne z punktu widzenia procesu. Człowiek jest ułomny, ma swoje słabości i ograniczenia, a presja rynku związana ze spełnianiem oczekiwań konsumenta wzrasta. Wymusza zwiększenie wydajności operacyjnej, redukcję błędów ludzkich, obniżenie kosztów, niekiedy

zastępowanie człowieka mobilnymi robotami, wdrażanie automatyki magazynowej, wdrażanie systemów optymalizacji rozmieszczenia i śledzenia zapasów etc.

Celem artykułu jest analiza i ocena zastosowania technologii Internetu rzeczy w gospodarce magazynowej i diagnoza, jak komunikacja ICT usprawnia procesy przepływu towarów i informacji w łańcuchu dostaw. Analizy przeprowadzono w oparciu o źródła wtórne i autorskie analizy wykorzystania technologii ICT oraz IoT w laboratorium logistycznym.

Rozwój sieci ICT

Początki globalnej sieci internetowej sięgają lat 60. XX wieku. Od prostych komputerów i krótkich wiadomości dotarliśmy jako użytkownicy do sieci 5G, algorytmów sztucznej inteligencji i humanoidów, wykorzystujących dane krążące w globalnej sieci Internet. Współcześnie, przy wykorzystaniu sieci, każdy użytkownik jest w stanie odszukać, ocenić i dokonać zakupu niemalże dowolnego produktu w prawie każdej lokalizacji. Największą presją dla producentów i operatorów jest czas i cena. Relacje i współpraca może odbywać się w kilku konfiguracjach, m.in.: relacje pomiędzy przedsiębiorstwami (B2B), pomiędzy konsumentami (C2C) oraz relacje pomiędzy producentem a konsumentem (B2C). Coraz więcej procesów odbywa się i jest sterowanych cyfrowo.

Za sprawą rozwoju technologii teleinformatycznych kluczową rolę na rynku odgrywa hyperconnectivity (powszechna łączność). Relacje sieciowe coraz częściej kierują się w stronę budowania cyfrowo-neuronowych sieci łączących urządzenia i maszyny w jeden wielki cyfrowy organizm, bazujący na automatycznej wymianie danych. Literatura przedmiotu wymianę danych przy użyciu dużej liczby sensorów definiuje jako Internet rzeczy (ang. Internet of Things). Urządzenia funkcjonujące w świecie rzeczywistym wymieniają informacje ze światem cyfrowym. Systemy zarządzania ruchem miejskim pojazdów, systemy świetlne, zarówno drogowe, jak i oświetleniowe, urządzenia grzewcze, automaty kontroli przepływu wody, urządzenia GPS będące wyposażeniem pojazdów, urządzenia AGD, czy powszechnie obecne smartfony wymieniają zbierane dane z siecią cyfrową, umożliwiając tym samym rozwój cyfrowego ekosystemu, w którym interakcje użytkowników wykraczają poza powszechnie przyjęte postrzeganie. Cyfrowy ekosystem łączy narzędzia komunikacji elektronicznej, głównie urządzenia i aplikacje mobilne. Do 2016 roku prawie połowa populacji globu (3,5 mld osób) miała dostęp do Internetu (Raport Meltwater, 2024). Z końcem 2023 roku dostęp do sieci miało już 5,35 mld osób, co stanowi 66,2% całej populacji globu, a dostęp do sieci w Unii Europejskiej posiada już 93% jej populacji (Eurostat). Głównym czynnikiem rozwoju jest liczba urządzeń mobilnych, która szacunkowo przekracza już liczbę podłączonych do sieci urządzeń stacjonarnych (Pieregud, Gajewski,

2016, s. 17). Ogromna liczba urządzeń cyfrowych, wymieniających między sobą terabajty danych, tworzy cyfrowy ekosystem w chmurze (ang. Cloud), przyczynia się dodatkowo do rozwoju modeli rozproszonego przetwarzania danych, opartej na użytkowaniu usługi dostarczanej przez podmioty zewnętrzne (Allegro, 2022), który umożliwia zupełnie nowy poziom outsourcingu, niewymagający budowania własnego centrum obliczeniowego z kosztowną i rozbudowaną infrastrukturą IT. Coraz częściej obserwuje się konkurowanie między sobą organizacji funkcjonujących klasycznie i podmiotów cyfrowych, oferujących usługi multi-services, które w swojej działalności przenikają różne branże, docierając do klientów innymi kanałami w układzie ekosystemu piramidalnego, jak giganci cyfrowi (Amazon, Google), tym samym przyczyniając się do rozbudowy systemów logistycznych (5PL) (Cywiński, Skwarek, 2020, s. 94).

Użytkownicy są świadkami procesu cyfryzacji życia społeczno-ekonomicznego, w którym świat cyfrowy miesza się ze światem rzeczywistym. Postępujący proces cyfryzacji powoduje powstawanie nieznanych dotąd rozwiązań, systemów i narzędzi, do których zaliczyć można między innymi:

- Internet of Things (IoT) – sieć połączonych ze sobą urządzeń, które prowadzą między sobą wymianę danych, bez udziału operatora,
- Hyperconnectivity – powszechna łączność w sieci każdego użytkownika z każdym systemem,
- cloud computing – chmura obliczeniowa, w której wykorzystywane są aplikacje, urządzenia mobilne i systemy IT,
- automatyzacja i robotyzacja procesów (roboty i humanoidy),
- big data analytics – analityka dużych zbiorów matematycznych i danych przesyłanych z sensorów w czasie rzeczywistym,
- wielokanałowe modele dystrybucji towarów i usług (Cywiński, 2019, s. 66).

Dla starszego pokolenia radykalny, dla młodego pokolenia przełomowy charakter zmian wnosi nowe wartości dla przedsiębiorstw funkcjonujących na rynku, w tym również dla klientów. Zmiany zachodzące na rynku w pewnym stopniu wymuszają transformację cyfrową, która staje się istotnym czynnikiem funkcjonowania organizacji. Umiejętność wykorzystania technologii cyfrowych w celu poprawy efektywności procesów w organizacji przynosi wymierne korzyści. Łatwiejsze staje się zrozumienie potrzeb klientów, definiowanie optymalnych kanałów komunikacji i dystrybucji, poprawa środowiska pracy i wydajności procesów wewnętrznych (Walicka, Czemieli-Grzybowska, 2023, s. 111).

Transformację gospodarki i zmiany działalności organizacji na rynku można podzielić na kilka obszarów, w ramach których stosowane są określone technologie (Berger):

- a. komunikacja (Cloud Computing, IT, Digital Products, Intelligent Flow management),

- b. dane cyfrowe (big data, IoT, Security Bufors, Prognozy popytu),
- c. automatyka (drony, robotyka, pojazdy autonomiczne),
- d. cyfrowy konsument w cyfrowej gospodarce (5PL, e-commerce, Mobile App, Social Network).

Rynek usług logistycznych w obliczu rozwoju ICT

Wszechobecne zmiany w otoczeniu wpływają na rozwój usług logistycznych, które na przestrzeni lat ewoluowały, zmieniając model biznesowy tych usług od 1PL do 5PL.

Model 1PL, czyli First Party Logistics, nazywany również In-House Logistics, charakteryzował się uczestnictwem w realizowaniu usług logistycznych tylko jednego podmiotu, organizacji macierzystej, która od początku do końca zajmowała się całym procesem logistycznym i była za niego odpowiedzialna (Burinskiene, Lingaitiene, 2023, s. 45). Model ten, traktowany jako bazowy do dalszego rozwoju, niestety miał wady. Jedną z nich była konieczność posiadania niezbędnych zasobów, potrzebnych do zrealizowania wszystkich procesów w łańcuchu dostaw, za które odpowiedzialny był tylko producent. Podstawowym celem korzystania z tego modelu było zaspokojenie własnych potrzeb.

W drugim modelu – 2PL (Second Party Logistics) wykorzystano outsourcing. Część elementów procesu logistycznego, takich jak transport, przeładunek, magazynowanie, realizowana była przez podmiot zewnętrzny (Leończuk, Golubniewska, 2021, s. 229). Był to zazwyczaj wybrany przez producenta jeden obszar, najbardziej kosztowny dla podmiotu macierzystego. Do zadań producenta należała natomiast obsługa transakcji. Brak było pełnej kontroli nad elementami procesu logistycznego, ale rekompensowały to niższe koszty ponoszone przez producentów, wynikające z braku konieczności zakupu na przykład środków transportu.

Logistyka 3PL (Third party Logistics) oferowała usługi związane ze zintegrowanym magazynowaniem i transportem (Hosie et al., 2012, s. 303). Operatorzy logistyczni byli pośrednikami pomiędzy organizacją macierzystą a klientem końcowym. Za obsługę transakcji odpowiadał producent lub sprzedawca.

Model 4PL (Fourth Party Logistics) oprócz rozwiązań poprzedniego modelu oferował dodatkowo logistykę operacyjną transportu, zarządzanie zapasami, czyli zarządzał całym łańcuchem dostaw. Zasadniczo 4PL to „integrator, który gromadzi własne zasoby, możliwości i technologię swojej i innych organizacji w celu zaprojektowania, zbudowania i uruchomienia kompleksowej obsługi łańcucha dostaw” (Li, Liu et al., s. 838). Duży nacisk kładzie się na to, aby dostawca 4PL był jedynym punktem kontaktowym dla nadawcy, integralną częścią rozwiązania biznesowego dla klienta.

Najbardziej zaawansowany pod względem innowacyjności jest model 5PL (Fifth Party Logistics). Oprócz rozwiązań 4PL wykorzystuje nowoczesne technologie, takie jak algorytmy sztucznej inteligencji, by prognozować popyt, optymalizować logistykę dostaw poprzez śledzenie i monitorowanie przesyłek. Jest to rozwiązanie zamieniające w perspektywie klienta łańcuch dostaw w funkcję, która jest całkowicie napędzana technologią, posiada aktywa fizyczne, rozległą wiedzę oraz wykorzystuje najnowocześniejsze technologie (Rocha, Lehenbauer, 2020).

Dystrybucja oparta na piątym poziomie dostawcy usług logistycznych z punktu widzenia ekonomii to proces doskonalenia i rozwoju przemysłu. Nowe podejście do organizacji zaopatrzenia, produkcji i dystrybucji kładzie nacisk na wzmacnianie łańcuchów dostaw w celu osiągnięcia zwiększonych efektów ekonomicznych i technicznych (np. wzrost wydajności). Należy pamiętać, że przechodząc z poziomu komunikacji maszyna – system IT w procesie logistycznym, powinno się szukać koegzystencji operatora i maszyny. Takie połączenie wykorzystuje technologie cyfrowe i pozwala niwelować zakłócenia w łańcuchach dostaw. Wykorzystując rozwój technologiczny, przedsiębiorstwa coraz częściej sięgają po algorytmy sztucznej inteligencji.

Powszechnie wiadomo, że łańcuchy dostaw są efektem stale rozwijających się rozwiązań technologicznych na globalnym rynku, które mają za zadanie wspomagać procesy efektywnego zarządzania nimi. Potrzeba uzyskania przewagi konkurencyjnej na rynku determinuje szukanie okazji tworzenia wartości dodanej dla nabywcy (Kryś, 2020, s. 40). Współczesne łańcuchy dostaw mają strukturę podobną do połączonych ze sobą zależnościami dostaw grup przedsiębiorstw oraz nabywców, których celem jest generowanie wartości w postaci produktów (Skowron-Grabowski, 2016, s. 4). Działania podejmowane w logistycznych łańcuchach dostaw przy uwzględnieniu 5PL, oprócz planowania, kontrolowania i realizacji przepływów towarów oraz informacji, uwzględniają zaawansowane systemy zbierania informacji i prognozowanie komputerowe. Na początku ery kapitalizmu w Polsce realizacja zamówień i dostaw przy uwzględnieniu własnych zasobów przedsiębiorstwa wymagała jedynie subiektywnych prognoz eksperckich. Włączenie do systemu zaopatrzenia i dystrybucji partnerów wyspecjalizowanych w szeroko rozumianym outsourcingu wymuszało zastosowanie analiz statystycznych i matematycznych przy użyciu prostych systemów komputerowych. Obecnie, w obliczu idei Industry 5.0, konceptu Logistics 4.0 i 5.0, przy ogromnej ilości zbieranych informacji, system wymaga ogromnych mocy obliczeniowych. Z pomocą logistynom przychodzi sztuczna inteligencja i analizy w chmurze. W obecnych warunkach społeczno-gospodarczych tylko niewielu organizacjom udaje się wykorzystać potencjał technologiczny i innowacyjny cyfrowej transformacji. Liderami na rynku, z zyskami, które potwierdzają osiągnięty rozwój technologiczny, są między innymi Amazon i Google z ich wiodącymi na świecie laboratoriami

sztucznej inteligencji, jak Google Brain i Deep Mind. To jeden skoncentrowany ekosystem, który może poszczycić się jednymi z największych przełomów w dziedzinie AI, neurobiologii, inżynierii, matematyki, infrastruktury obliczeniowej i symulacji. Efekty implikacji AI widać w wynikach finansowych wspomnianych spółek. Amazon w 2024 roku zarobił 638 mld USD (Amazon, 2024), a Google 350,02 mld USD (Alphabet Revenue 2015–2024).

Internet rzeczy w procesach magazynowych

Pojęcie Internetu rzeczy (ang. Internet of Things, IoT) zostało przedstawione ponad dwie dekady temu w USA. Od tamtego czasu jest skutecznie wdrażane w sieciach ICT w kolejnych gospodarkach poszczególnych krajów. Jest ściśle skorelowane z poziomem rozwoju technologii teleinformatycznych w danym kraju. IoT to globalna sieć, łącząca ze sobą zasoby wielu urządzeń elektronicznych, takich jak komputery, rzeczy lub przedmioty wyposażone w czujniki lub sensory, wykorzystujące systemy radiowej komunikacji do wymiany informacji. Te elementy systemu wyposażone w czujniki, moduły komunikacyjne i oprogramowanie tworzą rozproszone systemy cybernetyczno-fizyczne, które funkcjonują w czasie rzeczywistym. Od kilku lat zyskują na znaczeniu dzięki miniaturyzacji elektroniki, rozwojowi sieci 5G, popularyzacji chmury obliczeniowej. Znaczący rozwój IoT rozpoczął się w pierwszej dekadzie XXI wieku, dzięki dużemu zainteresowaniu smartfonami w codziennym użytkowaniu. Każdy użytkownik telefonu może posiadać dziesiątki aplikacji, za pomocą których może się komunikować lub sterować urządzeniami. Standardowa architektura systemu Internetu rzeczy składa się z trzech głównych warstw: warstwy percepcyjnej – czujniki i urządzenia zbierające dane, drugiej warstwy sieciowej, gdzie odbywa się transmisja danych, oraz warstwy aplikacyjnej, gdzie realizowana jest analityka danych wejściowych poprzez interfejsy użytkownika, wspomagana przez automatyzację procesu. Integracja tych warstw umożliwia skuteczne gromadzenie i analizę danych oraz podejmowanie działań w odpowiedzi na określone stany środowiskowe lub operacyjne.

W sektorze przemysłowym IoT znajduje zastosowanie przede wszystkim w monitorowaniu stanu maszyn, optymalizacji łańcucha dostaw, a także w tzw. predykcyjnym utrzymaniu ruchu (predictive maintenance) (Lee, Bagheri, Kao, 2015, s. 19). Dzięki wdrożeniu czujników i analizy danych możliwe jest minimalizowanie przestojów produkcyjnych oraz zwiększenie efektywności energetycznej zakładów.

Rozwiązania IoT oferują ogromne możliwości. W medycynie umożliwiają zdalne monitorowanie parametrów życiowych pacjentów (np. ciśnienie krwi, poziom glukozy, EKG) w czasie rzeczywistym, przy wykorzystaniu urządzeń typu wearable, czyli smartwatche, opaski fitness, monitory EKG, tętna, glukozy, google

glass, odzież z sensorami etc. Wspomagają diagnostykę, rehabilitację i opiekę nad osobami starszymi, poprawiając jakość życia i umożliwiając szybką interwencję w przypadku wykrycia nieprawidłowości. Co istotne, urządzenia, oprócz zbierania danych w czasie rzeczywistym, przesyłają dane do kolejnego urządzenia zorganizowanego w sieci, umożliwiając ich operatorowi (np. lekarz) podejmowanie decyzji istotnych dla zdrowia i życia pacjenta.

W miastach IoT wspomaga zarządzanie transportem, oświetleniem ulicznym, gospodarką odpadami czy jakością powietrza. Inteligentne systemy monitorujące ruch drogowy i wykorzystanie przestrzeni publicznej przyczyniają się do poprawy płynności komunikacyjnej, ograniczenia zjawiska kongestii, redukcji emisji CO₂ oraz podnoszenia poziomu bezpieczeństwa mieszkańców miast.

Istotnym wsparciem technologicznym są technologie komunikacji bezprzewodowej (Smejda, 2016, s. 45):

- komunikacja Wi-Fi, która umożliwia łatwą komunikację z narzędziami do mobilnej obsługi i sterowania (niestety o małym zasięgu),
- sieć komórkowa (5G), dzięki której możliwa jest komunikacja z każdym urządzeniem w technologii GSM, UMTS lub LTE,
- komunikacja Bluetooth, czyli protokół komunikacji bezprzewodowej przeznaczony do wymiany informacji pomiędzy urządzeniami na małych odległościach, ale za to o szybkiej prędkości wymiany danych,
- technologia RFID, korzystająca z fal radiowych. przesyłająca niewielkie pakiety danych,
- Z-Wave, bezprzewodowy protokół przeznaczony do systemów automatyki,
- 6LoWPAN², czyli bezprzewodowy system sieci dla systemów wbudowanych.

Jednymi z istotnych urządzeń w systemach IoT są małe urządzenia działające na bateriach, wysyłające sygnał radiowy i komunikujące się z odbiornikami (głównie smartfonami) za pomocą protokołu Bluetooth, zwane potocznie beacons. To urządzenia, które wyświetlają potrzebne informacje w momencie, kiedy użytkownik systemu jest w ich pobliżu. Wyświetlają informacje o produktach, danych uzyskanych z sensorów lub czujników (EKG, poziom glukozy), stanie urządzenia (pojazdy, automaty, roboty), dzieła sztuki, budynki architektoniczne etc. Beacon emituje sygnał w technologii Bluetooth Low Energy. Dzięki przesyłowi danych możliwe jest tworzenie elektronicznej dokumentacji.

Jedną z największych barier w systemie IoT jest ograniczona samowystarczalność energetyczna sensorów. Do tego dochodzą ograniczone możliwości zarządzania urządzeniami według protokołu IP. W Polsce większość urządzeń funkcjonuje w oparciu o protokół IPv4, a tylko część z nich w oparciu o protokół IPv6. To powoduje, że nie wszystkie sensory i urządzenia są ze sobą kompatybilne.

2 IPv6 over Low Power WireLess Personal Area Networks.

Kolejnym problemem jest ograniczona dostępność gospodarstw domowych do szerokopasmowego Internetu, który jak wiadomo jest dostępny głównie w miastach dużych i średniej wielkości. Ważnym czynnikiem powodującym trudności we wdrażaniu i rozwoju sieci IoT jest opór w stosunku do technologii. Ponadto przeszkodą mogą być niejasne korzyści z ich wykorzystania, a także potencjalny stopień ich skomplikowania. Warto także zaznaczyć, że wielu obywateli można zaliczyć do wykluczonych cyfrowo. Pojawiają się bariery mentalnościowe i obawy związane z zagrożeniem prywatności.

Jednak zastosowanie koncepcji IoT powinno być rozpatrywane z perspektywy pojedynczej jednostki, społeczeństwa oraz podmiotów medycznych czy gospodarczych. Każda z tych grup oczekuje realizacji innych celów, wobec czego spodziewa się innych korzyści płynących z implementacji sieci IoT. Do korzyści pojedynczych jednostek w systemie IoT można zaliczyć między innymi (Kwiatkowska, 2014, s. 67):

- poprawę bezpieczeństwa,
- automatyzację procesów,
- optymalizację kosztów związanych z utrzymaniem jednostki.

Podsumowanie

Współczesne społeczeństwo staje na progu dużych zmian dotyczących jego funkcjonowania. Wynikają one z rosnącego rozwoju technologii teleinformatycznych, rozwoju sieci telefonii komórkowej i miniaturyzacji urządzeń do obsługi sieci. Dynamiczny postęp technologiczny powoduje rozwój infrastruktury sieci ICT. Tym samym rozwiązania Internetu rzeczy zaczynają odgrywać coraz większą rolę w różnych dziedzinach gospodarki, w tym również w efektywności procesu kompletowania towarów w gospodarce magazynowej. Zastosowanie systemów teleinformatycznych w centrach magazynowych i dystrybucyjnych poprzez użycie skanerów optycznych, sensorów (Z-Wave), czujników (6LoWPan), technologii radiowej (RFID), kodów wielowymiarowych (QR), automatyki i robotyki (manipulatory), a także algorytmów sztucznej inteligencji, która umożliwia lepsze przewidywanie i prognozowanie zamówień, pozwala na znaczne przyspieszenie procesów magazynowych i ograniczanie zbędnych procedur, tym samym przekłada się na zmniejszanie kosztów gospodarki magazynowej (w tym kompletacji towarów), a dodatkowo ukazuje szanse poprawy pozycji konkurencyjnej na rynku.

Pomimo licznych zalet rozwój IoT wiąże się z istotnymi wyzwaniem. Najważniejsze z nich to bezpieczeństwo danych, interoperacyjność urządzeń różnych producentów oraz zapewnienie niezawodności komunikacji w warunkach dużego obciążenia sieciowego. Urządzenia są często słabo zabezpieczone, co niesie ryzyko ingerencji w nie przez osoby trzecie. Bazy danych sieci IoT gromadzą

ogromne ilości danych, co może powodować ryzyko śledzenia użytkowników bez ich wiedzy, profilowanie na potrzeby marketingowe, kontrolę społeczną etc. W przyszłości kluczowe będzie wdrożenie jednolitych standardów oraz rozwój edge computing, czyli przetwarzania danych na obrzeżach sieci, co pozwoli zmniejszyć opóźnienia i obciążenie centrów danych. Różnorodność protokołów i systemów operacyjnych prowadzi do problemów ze współpracą urządzeń różnych producentów, trudności w integracji systemów w dużej skali i fragmentacji świadczonych usług.

Bibliografia

- Burinskiene, A., Lingaitiene, O. (2023). Actualities of Supply Chain Concept Evolution. *Journal of Service, Innovation and Sustainable Development*, 4(2).
- Cywiński, M. (2019). Koncepcja Internetu rzeczy (IoT) i idei Industry 4.0 w rozwoju inteligentnego miasta. W: W. Januszkiewicz, M. Cywiński, M. Chojnacka (red.), *Idea smart city w miastach średniej wielkości*. Akademia im. Jakuba z Paradyża, Gorzów Wielkopolski.
- Cywiński, M., Skwarek, A. (2020). *Wybrane strategie przedsiębiorstw sektora TSL*. Akademia im. Jakuba z Paradyża, Gorzów Wielkopolski.
- Hosie, P., Sundarakani, B., Wee Kwan Tan, A., & Koźlak, A. (2012). Determinants of Fifth Party Logistics (5PL): service providers for Supply Chain Management. *Logistics Systems and Management*, 13(3).
- Kryś, P. (2020). *Metoda oceny funkcjonowania centrów logistycznych*. Oficyna Wydawnicza Politechniki Warszawskiej.
- Kwiatkowska, E. M. (2014). Rozwój Internetu rzeczy – szanse i zagrożenia. *Kwartalnik Antymonopolowy i Regulacyjny*, 8(3).
- Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3.
- Leończuk, D., Golubniewska, W. (2021). Kierunki rozwoju usługodawców logistycznych i ich wpływ na funkcjonowanie współczesnych łańcuchów dostaw. *Akademia Zarządzania*, 5.
- Pieriegud, J., Gajewski, J. (red.). (2016). *Cyfryzacja gospodarki i społeczeństwa*. Europejski Kongres Finansowy, Gdańsk.
- Skowron-Grabowski, B. (2016). Management of Supply Chain and networks of enterprises. *Gospodarka Magazynowa i Logistyka*, 12.
- Smejda, P. (2016). Internet rzeczy we współczesnej gospodarce. Rola, zadania i bariery rozwoju. *Zeszyty Naukowe Politechniki Łódzkiej*, 1208.
- Walicka, M., Czemieli-Grzybowska, W. (2023). Sztuczna inteligencja w zarządzaniu kapitałem przedsiębiorstwa w dobie Przemysłu 5.0. *Akademia Zarządzania*, 7(4).

Źródła internetowe

- Amazon przebija prognozy analityków we wszystkich segmentach.* <http://strefainwestorow.pl/strefa-global-market/spolki/amazon-aws-1q2024/> [dostęp: 10.04.2025]
- Alphabet Revenue 2015–2024.* <http://bullfincher.io/summary-table-annual-revenue/> [dostęp: 09.04.2025]
- Berger, R., *The digital transformation of industry*, BDI. https://roland-berger.com/media/pdf/roland_berger_digital_transformation_of_industry_20150315.pdf [dostęp: 10.05.2025]
- https://ec.europa.eu/eurostat/statistics_explained/index.php?title=Archive:Dane_statystyczne_dotyczące_gospodarki_cyfrowej_i_społeczeństwa_cyfrowego_-_gospodarstwa_domowe_i_osoby_fizyczne, [dostęp: 09.05.2025]
- Li, X., Liu, W., Lei, L., Zhao, W., & Ren, S. (2003). *The design and realization of four party logistics, systems, man and cybernetics*. IEEE Conference, USA, s. 838. <https://ieeexplore.ieee.org/document/1243919> [dostęp: 09.04.2025]
- Rocha, J., Lehenbauer, K. (2020). *Third Party Logistics and Beyond*, RAIS. <https://rais.education/wp-content/uploads/2019/11/0017JRK.pdf> [dostęp: 09.04.2025]
- Raport Allegro 2022, s. 4. https://about.allegro.eu/system/files-encrypted/nasdaq_kms/assets/2023/04/04/9-10-25/PL_Raport%20ESG%202022.pdf [dostęp: 11.05.2025]
- Raport Meltwater, *Global Digital Report 2024*. <https://meltwater.com/en/global-digital-trends/> [dostęp: 09.05.2025]

Adam Filipowski¹

Marta Rapacz-Mahmoud²

Ograniczenia decyzyjne polskich funduszy venture capital w finansowaniu przełomowych technologii: studium przypadku ElevenLabs

Streszczenie

Brak zaangażowania polskich funduszy venture capital (VC) w przełomowe projekty z obszaru technologii przyszłości umożliwia wgląd w granice rozwoju rodzimego rynku inwestycyjnego. Badanie przyczyn pasywności inwestorów instytucjonalnych staje się więc konieczne do stworzenia ustrukturyzowanych rekomendacji dla polskiego ekosystemu VC. Niniejszy artykuł stanowi próbę zdefiniowania ograniczeń decyzyjnych występujących w funduszach specjalizujących się w finansowaniu inwestycji załączkowych i oparte jest na studium przypadku ElevenLabs, start-upu o polskim kapitale założycielskim, rozwijającym generatywną syntezę mowy. Badanie ma charakter wyjaśniający, a przeprowadzone zostało przy wykorzystaniu triangulacji danych: analizy literatury z zakresu teorii decyzji i zdolności dynamicznych, raportów branżowych oraz oficjalnych komunikatów spółki ElevenLabs i współpracujących z nią partnerów biznesowych. Niniejszy artykuł uwzględnia dane z lat 2022–2025, gdy start-up pozyskał 281 milionów dolarów wyłącznie od zagranicznych inwestorów, w tym renomowanych funduszy Andreessen Horowitz i Sequoia Capital. Analiza ujawnia trzy komplementarne bariery blokujące decyzję inwestycyjną: ograniczenia poznawcze, organizacyjne oraz kulturowe. Zderzenie tych barier z praktykami funduszy globalnych pokazuje, że kluczowym czynnikiem sukcesu jest zdolność do szybkiego łączenia wiedzy technicznej, kapitału sieciowego i procesów decyzyjnych opartych na tolerancji ryzyka. Wkład teoretyczny artykułu polega na rozszerzeniu zastosowania koncepcji racjonalności ograniczonej oraz zdolności dynamicznych w kontekście zarządzania inwestycjami w sektorze deep tech. Z praktycznego punktu widzenia badanie wskazuje na potrzebę rozwoju kompetencji technologicznych w funduszach VC, reorganizacji procesów decyzyjnych oraz transformacji kultury organizacyjnej, sprzyjającej inwestowaniu w projekty o wysokim ryzyku i potencjale technologicznym.

Słowa kluczowe: venture capital, ograniczenia decyzyjne, racjonalność ograniczona, zdolności dynamiczne, deep tech, ElevenLabs

1 Akademia WSB w Dąbrowie Górniczej

2 Akademia WSB w Dąbrowie Górniczej

Decision-Making Constraints of Polish Venture Capital Funds in Financing Breakthrough Technologies: ElevenLabs Case Study

Abstract

The limited involvement of Polish venture capital funds (VC) in pioneering future technology projects offers insight into the constraints on the development of the domestic investment market. Exploring the reasons behind institutional investors' passivity is therefore essential to develop structured recommendations for the Polish VC ecosystem. This article aims to define the decision-making constraints faced by funds specialising in pre-seed investment. The analysis is based on a case study of ElevenLabs, a start-up with Polish founding capital that develops generative voice synthesis technology. This explanatory case study employs data triangulation, incorporating literature on decision theory and dynamic capabilities, industry reports, and official communications from ElevenLabs and its partners. It considers data from 2022 to 2025, during which the start-up raised a total of \$281 million exclusively from foreign investors, including prominent funds such as Andreessen Horowitz and Sequoia Capital. The analysis identifies three interconnected barriers hindering investment decisions: cognitive, organisational, and cultural constraints. The interaction between these barriers and the practices of global funds demonstrates that the key success factor is the ability to skillfully combine technical knowledge, network capital, and risk-tolerant decision-making processes. The theoretical contribution of the article is to extend the application of the conceptual framework of bounded rationality and dynamic capabilities in the context of investment management in the deep tech sector. From a practical point of view, the research indicates the need to develop technological competencies in VC funds, reorganise decision-making processes and transform the organisational culture to encourage investment in high-risk projects with high technological potential.

Keywords: venture capital, decision constraints, bounded rationality, dynamic capabilities, deep tech, ElevenLabs

Wprowadzenie

Dynamika zmian następujących w świecie technologii znajduje odzwierciedlenie na rynku finansowym. Zaobserwować bowiem można bezprecedensowy zwrot venture capital w kierunku inwestycji w projekty oparte na sztucznej inteligencji. Udział przedsięwzięć związanych z AI wzrósł już do ponad 25% w globalnym wolumenie transakcji VC (Dealroom, 2023). Równolegle, mimo że polski ekosystem venture capital osiągnął pułap dofinansowań w łącznej kwocie przekraczającej 2,1 miliarda złotych, udział projektów z obszaru sztucznej inteligencji pozostaje marginalny (PFR Ventures, 2024). Ta strukturalna dysproporcja stała się szczególnie widoczna w 2022 roku, kiedy ElevenLabs, start-up rozwijający generatywne technologie syntezy mowy, założony przez Mateusza Staniszewskiego i Piotra Dąbkowskiego, pozyskał łącznie 281 milionów dolarów od wiodących funduszy VC, takich jak Andreessen Horowitz oraz Sequoia Capital, przy całkowitym braku zaangażowania polskich inwestorów instytucjonalnych. Sytuacja ta wskazuje na istotny problem dotyczący ograniczeń decyzyjnych rodzimych funduszy venture capital wobec projektów o wysokim poziomie niepewności technologicznej. Dotychczasowe badania nad VC skupiały się przede wszystkim na czynnikach sukcesu finansowego oraz strukturach sieciowych syndykatów inwestycyjnych. Stosunkowo niewiele uwagi poświęcono jednak analizie negatywnych decyzji inwestycyjnych, zwłaszcza w kontekście ich poznawczo-organizacyjnych determinantów (Cliff, Park, Bock, 2022). Dlatego tak ważne jest zbadanie, jakie uwarunkowania poznawcze, organizacyjne oraz kulturowe determinują decyzje polskich inwestorów. Niniejszy artykuł jest próbą wypełnienia tej luki badawczej poprzez pogłębioną analizę przypadku start-upu ElevenLabs. Podejście to jest osadzone w koncepcji racjonalności ograniczonej (Simon, 1979) i jej zastosowania w obszarze decyzji inwestycyjnych w sektorze deep tech poprzez identyfikację konkretnych heurystyk i barier poznawczych. Co więcej, niniejsze badanie stanowić ma uzupełnienie teorii zdolności dynamicznych o empiryczny wgląd w ograniczenia procesów absorpcji okazji inwestycyjnych w kontekście funduszy działających w rynkach wschodzących (Heugens, Kapoor, 2024) oraz ukazanie roli kultury organizacyjnej jako filtra decyzyjnego wpływającego na percepcję ryzyka oraz preferencje geograficzne inwestorów (Bottazzi, Giorcelli, 2021).

Ramy teoretyczne

Proces inwestycyjny funduszy venture capital w projekty z obszaru technologii przyszłości wymaga podejścia nieco wykraczającego poza tradycyjne modele maksymalizacji stopy zwrotu. Charakterystyczna dla sektora deep tech niepewność technologiczna oraz rynkowa czyni decyzje inwestycyjne podatnymi na ograniczenia poznawcze, deficyty organizacyjne oraz czynniki kulturowe. Podstawą

więc dla zrozumienia zachowań decyzyjnych w takich warunkach jest koncepcja racjonalności ograniczonej, zgodnie z którą decydenci w obliczu złożoności sytuacji, niedostatecznej informacji oraz ograniczonych zasobów poznawczych nie dokonują wyborów optymalnych, lecz zadowolających (Simon, 1979). W przypadku funduszy venture capital, które funkcjonują w dynamicznym ekosystemie innowacji, decyzje inwestycyjne często są podejmowane w oparciu o heurystyki poznawcze. Jedną z najczęściej występujących jest dopasowanie wzorców (ang. pattern-matching), polegające na zestawieniu nowych propozycji z wewnętrznym zbiorem atrybutów wcześniejszych sukcesów inwestycyjnych (Cliff, Park, Bock, 2022). Tego typu uproszczenia poznawcze mogą prowadzić do systematycznego odrzucania projektów, jeśli nie odpowiadają zakodowanym wcześniej wzorcom, mimo że posiadają potencjał, by stać się przełomowymi. Szczególnie istotne jest to w sektorze deep tech, gdzie wartość projektu materializuje się często ze znacznym opóźnieniem, a jego zrozumienie wymaga specjalistycznej wiedzy technologicznej. Brak kompetencji w zakresie oceny złożonych algorytmów czy też architektury sztucznej inteligencji może więc skutkować percepcją ryzyka jako niekwantyfikowalnego i w konsekwencji nieakceptowalnego.

Jednocześnie przy inwestowaniu w projekty z obszaru technologii przyszłości tradycyjny model dywersyfikacji ryzyka w portfelu venture capital okazuje się niewystarczający. Mając na uwadze, że projekty deep tech charakteryzuje długi horyzont komercjalizacji oraz znaczące koszty rekurencyjne, wymagają one bardziej zaawansowanych procedur due diligence, opartych na wiedzy specjalistów z dziedziny AI. Obecnie skuteczne fundusze wprowadzają w takich warunkach strategię zarządzania opcjonalnością: akceptują wysoką zmienność wyników częściowych, lecz równolegle budują mechanizmy aktywnego wpływu na rozwój start-upu, m.in. poprzez wsparcie ścieżek rekrutacji talentu czy też udostępnianie własnej infrastruktury obliczeniowej (McKinsey, 2023). Zarządzanie ryzykiem w takich warunkach wymaga odejścia od wspomnianych już klasycznych modeli dywersyfikacji portfelowej na rzecz strategii uwzględniających naturę niepewności technologicznej i regulacyjnej (Miller, Lessard, 2000). W kontekście sztucznej inteligencji dodatkowe trudności wynikają również z braku regulacji branżowych, a także ze słabej porównywalności między poszczególnymi start-upami.

Na poziomie organizacyjnym kluczowe znaczenie zyskuje teoria zdolności dynamicznych, według której skuteczna alokacja zasobów wymaga umiejętności identyfikowania okazji, a także ich uchwycenia oraz rekonfiguracji posiadanych zasobów (Teece, 2020). Fundusze, które budują więc zespoły interdyscyplinarne, skracają cykle decyzyjne i tworzą elastyczne struktury współpracy, zwiększając tym samym prawdopodobieństwo udanych inwestycji w obszarach wysokiego ryzyka. Z kolei ograniczenia organizacyjne, takie jak mała skala operacyjna, brak ekspertów z obszaru deep tech czy zbiurokratyzowane i przeciągłe procedury,

redukują zdolność do szybkiego reagowania na okazje inwestycyjne i przyczyniają się do ich utraty.

Nie mniej istotnym czynnikiem determinującym decyzje inwestycyjne są uwarunkowania kulturowe. Jak wskazują Bottazzi i Giorcelli (2021), fundusze operujące w Europie Środkowo-Wschodniej wykazują tendencję do faworyzowania projektów o niskim profilu ryzyka i krótkim horyzoncie inwestycyjnym, co znajduje odzwierciedlenie w ich strategiach selekcji projektów. Warto jednak nadmienić, że wartość ekosystemu inwestycyjnego EEC wyceniana jest aktualnie na 243 miliardy euro (Dealroom.co, 2025). Nie zmienia to jednak faktu, że dostrzegalny jest charakterystyczny dla tej części Europy wysoki poziom unikania niepewności, a także preferencja projektów lokalnych. To jednak sprawia, że fundusze venture capital mają mniejszą skłonność do angażowania się w przedsięwzięcia o globalnym zasięgu, zwłaszcza te, które nie generują natychmiastowych przychodów.

Zestawienie trzech powyższych perspektyw: racjonalności ograniczonej, zarządzania ryzykiem innowacyjnym oraz dynamicznych i kulturowych zdolności adaptacyjnych funduszy VC umożliwia zbudowanie spójnych ram koncepcyjnych do analizy decyzji inwestycyjnych w sektorze technologii przyszłości. Proces decyzyjny w funduszach venture capital, zwłaszcza przy inwestowaniu w technologie przyszłości, rzadko bowiem przebiega w warunkach pełnej świadomości informacyjnej, przewidywalności i racjonalności ekonomicznej. Decyzje te podejmowane są w rzeczywistości nie tylko złożonej i niepewnej, ale często także nieustrukturyzowanej. To zaś wymaga szerszego podejścia badawczego i szczegółowej analizy inwestycji dokonanych przez rodzime venture capital, jak również projektów z polskim kapitałem, lecz dofinansowanych z zagranicznych źródeł.

Metodyka badawcza

Niniejszy artykuł jest pojedynczym studium przypadku o charakterze wyjaśniającym, uznawanym za najbardziej adekwatny, gdy zamierzeniem jest pogłębione zrozumienie złożonych procesów organizacyjnych osadzonych w specyficznym kontekście (Yin, 2018). Przypadek ElevenLabs, spółki, która zrewolucjonizowała świat syntezatorów mowy, spełnia kryterium *krytycznego studium przypadku* (Flyvbjerg, 2001). Należy zwrócić uwagę na jego skrajny charakter: pomimo polskiego rodowodu założycieli oraz światowego sukcesu technologii, żaden krajowy fundusz VC nie uczestniczył w którejkolwiek fazie finansowania, nawet w rundzie pre-seed, w której rolę inwestora wiodącego objął czeski Credo Ventures. Ta wyraźna luka kapitałowa pozwala prześledzić czynniki wpływające na brak decyzji o inwestycji w start-up ElevenLabs po stronie polskich podmiotów i tym samym stanowi wyjątkowo klarowny materiał do weryfikacji wyżej zaproponowanych ram teoretycznych.

Badanie ma charakter jakościowy i eksploracyjny, gdyż głównym celem jest zrozumienie mechanizmów decyzyjnych, a nie testowanie ilościowych zależności. Dobór danych został oparty na triangulacji: dostępną literaturę naukową wykorzystano do zbudowania ram teoretycznych, raporty branżowe PFR Ventures, Inovo VC, PitchBook i Dealroom posłużyły do odtworzenia kontekstu rynku, natomiast oficjalne komunikaty zaangażowanych funduszy oraz start-upu ElevenLabs stanowiły materiał opisujący przebieg rund finansowania. Analiza przebiegała z kolei zgodnie z logiką studium przypadku wyjaśniającego, w którym priorytetem jest syntetyczne odtworzenie ciągu zdarzeń i decyzji (Yin, 2018). Głównym ograniczeniem niniejszego badania pozostaje brak dostępu do niepublikowanych, wewnętrznych informacji funduszy, co uniemożliwia pełne odtworzenie motywów poszczególnych decydentów. Triangulacja źródeł wtórnych zminimalizowała jednak ryzyko stronniczości, a uzyskane wyniki dostarczają wiarygodnych podstaw do dalszych badań porównawczych, które mogłyby objąć wywiady pogłębione z partnerami polskich funduszy VC oraz wielokrotne studia przypadków odnoszące się zarówno do sukcesów, jak i porażek inwestycyjnych w sektorze deep tech.

Dane interpretowano w trybie iteracyjnym: najpierw rekonstrukcja osi czasu zdarzeń inwestycyjnych, a następnie przyporządkowanie kluczowych obserwacji do trzech wcześniej zdefiniowanych kategorii ograniczeń decyzyjnych: poznawczych, organizacyjnych i kulturowych. Tak skonstruowana w formie rozbudowanej tabeli matryca porównawcza umożliwiła powiązanie zjawisk empirycznych z ramami teoretycznymi i bezpośrednio odniesienie do pytania badawczego artykułu. Niniejsza praca naukowa jest bowiem poszukiwaniem odpowiedzi na pytanie o ograniczenia decyzyjne polskich funduszy venture capital w finansowaniu przełomowych technologii.

Studium przypadku: ElevenLabs

ElevenLabs powstało na początku 2022 roku jako spin-out Mateusza Staniszewskiego i Piotra Dąbkowskiego, dwóch polskich inżynierów, którzy doświadczenie w branży technologicznej zdobywali w Google Research i Palantir. Od początku przyświecał im cel stworzenia firmy o globalnym zasięgu. Założyciele trafnie zidentyfikowali niszę w obszarze generatywnej syntezy mowy – segmencie rynku, który przeżywał wtedy, po debiucie dużych modeli językowych, gwałtowny wzrost zainteresowania inwestorów. Mimo że globalne inwestycje w generatywną sztuczną inteligencję wzrosły z 29,1 miliarda dolarów w 2023 roku do rekordowych 56 miliardów dolarów w 2024, to przedsiębiorstwa spoza rynku amerykańskiego pozyskały jedynie 6,2 miliarda dolarów. To z kolei świadczy o relatywnym niedofinansowaniu tego sektora w Europie (PitchBook, 2024). Technologicznie firma

opiera się na autorskiej architekturze modułowych sieci neuronowych uczonych na wielojęzycznych korpusach, co daje przewagę kosztową i jakościową nad konkurentami. W terminologii zarządzania innowacjami jest to klasyczny przykład skoku technologicznego, gdzie wysoka niepewność technologiczna rekompensowana jest potencjałem tworzenia zupełnie nowych rynków (McKinsey, 2023).

Spółka zamknęła w styczniu 2023 rundę pre-seed, czyli wstępne finansowanie, mające na celu wsparcie budowy prototypu i weryfikację koncepcji rynkowej, kwotą 2 milionów dolarów. Inwestorem w tej fazie został czeski fundusz Credo Ventures. W odróżnieniu od większości polskich podmiotów, które specjalizują się w dokapitalizowywaniu wczesnych faz, Credo Ventures od lat rozwija kompetencje technologiczne i procedury szybkiej oceny projektów deep tech, co w przypadku ElevenLabs umożliwiło finalizację transakcji w ciągu zaledwie sześciu tygodni (Sifted, 2023). Następnie, w czerwcu 2023 roku, spółka przeszła do rundy A, czyli pierwszej większej rundy wzrostowej ukierunkowanej na skalowanie modelu biznesowego, pozyskując łącznie 19 milionów dolarów. Liderem tej fazy został amerykański gigant Andreessen Horowitz, a syndykat uzupełniły Sequoia Capital i Concept Ventures. Z perspektywy teorii strategii inwestycyjnej uruchomiony został mechanizm certyfikacji lidera. Polega on na udziale renomowanego funduszu, co znacząco obniża postrzegane ryzyko technologiczne i przyciąga kolejnych inwestorów (Gornall, Strebulaev, 2021). Od rozpoczęcia działalności aż do stycznia 2025 spółka ElevenLabs pozyskała łącznie 281 milionów dolarów, a jej aktualna wycena to 3,3 miliarda dolarów, dając tym samym ElevenLabs tytuł potrójnego jednorożca (TechCrunch, 2025).

W latach 2023–2025 udział projektów z obszaru sztucznej inteligencji w globalnej wartości transakcji venture capital wzrósł z 14% do 28% (Preqin, 2025), podczas gdy w Europie Środkowo-Wschodniej inwestycje skupiały się głównie na fintechach i platformach rynkowych, a segment deep tech nie przekraczał 7% rocznego wolumenu (PFR Ventures, 2024). W efekcie, pomimo rejestracji w Stanach Zjednoczonych, często motywowanej kwestiami regulacyjnymi, projekt ElevenLabs nie odpowiadał ani geograficznemu, ani branżowemu kryterium inwestycyjnym większości polskich funduszy. Co istotne, zarządy polskich funduszy VC dysponujące kapitałem mieszanym, w którym dominującą rolę odgrywały środki publiczne, były zobligowane do spełnienia wymogów limited partners, czyli podmiotów dostarczających kapitał do funduszu, dotyczących minimalnego poziomu generowanych przychodów oraz osiągnięcia kolejnych kamieni milowych Technology Readiness Level, czyli oceny dojrzałości technologii. Model biznesowy ElevenLabs przewidywał wielomiesięczny okres zużycia kapitału (ang. burn rate) bez istotnych przychodów aż do rundy B, co znacząco podnosiło próg wejścia dla inwestorów o krótszym horyzoncie zwrotu. Ponadto tempo procesu due diligence w Stanach Zjednoczonych, wynoszące średnio 8 tygodni,

wyraźnie kontrastowało z ponad 14-tygodniową weryfikacją w polskich funduszach (EIF, 2024). To ostatecznie zamknęło lokalnym inwestorom możliwość dołączenia do inwestycji po rundzie A i skazało ich na rolę późnych obserwatorów, niespełniających wymaganego profilu wartości dodanej (Gornall, Strebulaev, 2021). Z perspektywy nauk o zarządzaniu przypadek ElevenLabs ilustruje, jak w praktyce działa sprzężenie racjonalności ograniczonej decydentów, niewykształconych dynamicznych zdolności organizacyjnych w postaci braku szybkich procesów i funkcjonowania syndykatów oraz stronniczości kulturowej i preferencji aktywów krajowych. To właśnie ta triada, a nie jedynie niedobór kapitału, zdeterminowała absencję polskich funduszy venture capital we wszystkich czterech rundach finansowania spółki ElevenLabs.

Analizując jednak przyczyny nieobecności polskich VC w kolejnych rundach finansowania ElevenLabs, należy odwołać się do trzech spletających się poziomów ograniczeń: poznawczego, organizacyjnego i kulturowego, którym towarzyszy słabość otoczenia instytucjonalnego. Pierwszym wymiarem są deficyty kompetencyjne, prowadzące do zjawiska ograniczonej racjonalności, ponieważ fundusze operują w sytuacji niedoboru wiedzy o sztucznej inteligencji, przez co przy ocenie start-upów deep tech uciekają się do prostych reguł podobieństwa, znanych jako wspomniane wcześniej pattern-matching. Cliff, Park i Bock (2022) wykazali, że w takich warunkach dominują heurystyki eliminujące projekty trudne do zrozumienia technologicznie lub nieposiadające mierzalnych przychodów. W praktyce oznaczało to, iż partnerzy inwestycyjni, dysponujący niewielkim kapitałem ludzkim z wykształceniem STEM, uznali przewagi konkurencyjne ElevenLabs za niemożliwe do rzetelnej wyceny i na wstępnym etapie wykluczyli spółkę z leja inwestycyjnego.

Drugi poziom ograniczeń ma charakter organizacyjny i wiąże się ze stopniem rozwoju zdolności adaptacyjnych funduszy. Heugens i Kapoor (2024) argumentują, że inwestorzy finansowi, aby skutecznie dokapitalizować przełomowe technologie, muszą łączyć kompetencje identyfikacji okazji (ang. sensing), zagospodarowania możliwości (ang. seizing) i przekształcania zasobów (ang. reconfiguring). Polskie VC dysponują relatywnie niewielkimi aktywami – średnia wartość inwestycji załączkowych (ang. pre-seed) w 2024 roku wyniosła 1,8 miliona dolarów, a wartość kolejnej rundy oscylowała zwykle na poziomie 4,1 miliona dolarów. Jednocześnie polskie venture capital zrealizowały w 2024 roku 39 transakcji zagranicznych o łącznej wartości 66 milionów dolarów (PFR Ventures, 2025). Taka wysokość aktywów ogranicza wartość pojedynczej transakcji i zmusza fundusze do współfinansowania dużych rund. Tymczasem słabo rozwinięta sieć syndykatów w Polsce oraz udział środków publicznych w kapitale funduszy ograniczają elastyczność i wydłużają proces due diligence nawet do kilkunastu tygodni. Gdy ElevenLabs w styczniu 2024 roku zamykało rundę B na poziomie 80 milionów dolarów w ciągu niespełna sześciu tygodni, rodzime fundusze nie

posiadały ani wystarczającej skali finansowej, ani procesów umożliwiających tak szybkie podjęcie decyzji.

Trzecim, najtrudniej modyfikowalnym wymiarem jest kultura organizacyjna, w której dominuje awersja do niepewności i silne preferencje inwestycji w projekty o skali regionalnej, a nie globalnej. Bottazzi i Giorcelli (2021) wykazali, że w funduszach Europy Środkowo-Wschodniej przeważa krótkoterminowa logika transakcyjna, premiująca projekty z powtarzalnymi przychodami i wyraźnym rynkiem w regionie CEE. Z kolei dane ankietowe Inovo VC (2024), wskazujące na to, że ponad 70% polskich funduszy VC wymaga co najmniej 85 tysięcy złotych miesięcznego przychodu przed wejściem kapitałowym, potwierdzają, że nastawiony globalnie, pozbawiony wczesnych przychodów model ElevenLabs nie mieścił się w lokalnych kryteriach selekcji. Jednocześnie należy zwrócić uwagę na tendencję do nadmiernego inwestowania w aktywa krajowe. To, w połączeniu z lękiem przed porażką związaną z inwestycją w segment o niejasnych regulacjach, zbudowało dodatkową barierę.

Ostatnim, lecz równie istotnym czynnikiem jest otoczenie instytucjonalne. W Polsce brakuje rozwiniętego rynku wtórnego własności intelektualnej oraz płynnych ścieżek wyjścia z inwestycji poprzez giełdę, czy przejęcia w sektorze głębokich technologii. Niedostatek krajowych nabywców własności intelektualnej powoduje, że potencjalne wyjście ElevenLabs z inwestycji musiało być postrzegane jako obce i niezrozumiałe wydarzenie zagraniczne. Wynika to również z przewagi informacyjnej funduszy z Doliny Krzemowej, na którą składa się łatwiejszy dostęp do kluczowych danych rynkowych, sieci kontaktów, wyników pilotaży czy wewnętrznych analiz technologicznych, co w efekcie pozwala na dokładniejszą i szybszą ocenę wartości inwestycji z perspektywy inwestora instytucjonalnego oraz prywatnego.

Z perspektywy nauk o zarządzaniu przypadek ElevenLabs ujawnia, że na rynkach wschodzących sama dostępność kapitału nie przesądza o zdolności do finansowania przełomowych innowacji. Kluczowe okazują się bowiem kompetencje poznawcze, zdolności adaptacyjne i kultura organizacyjna pozwalająca na akceptację warunków o wysokiej niepewności inwestycyjnej. Te wyniki potwierdzają ogólną tezę Simona o granicach racjonalności, a zarazem otwierają drogę do dalszych badań porównawczych nad rozwojem zdolności adaptacyjnych w funduszach VC różnych jurysdykcji.

Warto również zwrócić uwagę na rozbieżność między decyzjami polskich i zagranicznych funduszy venture capital wobec ElevenLabs, która przejawia się odmiennymi zdolnościami adaptacyjnymi, różną praktyką zarządzania wiedzą czy standardem jakości procesów inwestycyjnych. Najbardziej elitarne fundusze z Doliny Krzemowej, w tym wspomniane wyżej Andreessen Horowitz i Sequoia Capital, od ponad dekady budują model określany w literaturze jako *platforma*

venture capital, w którym finansowanie stanowi jedynie część pakietu wartości (Colombo, Grilli, Piva, 2022). Ich struktury organizacyjne obejmują rozbudowane interdyscyplinarne zespoły partnerów współpracujących, a wśród nich znajdują się m.in. inżynierowie sztucznej inteligencji, eksperci produktu oraz specjaliści rekrutacji czy doradcy prawni. Takie podejście pozwala funduszowi pełnić funkcję quasi-korporacyjnej działalności badawczo-rozwojowej i ograniczać ryzyko technologiczne na poziomie portfela funduszu. Tego typu kompetencje pozwalają m.in. na przeprowadzenie procesu due diligence, który nie tylko trwa krócej, ale również pozwala na dogłębną walidację danych, tylko w niewielkim stopniu opierającą się na metrykach przychodowych. Gornall i Strebulaev (2021) dowiedli, że fundusze funkcjonujące zgodnie z powyższymi zaleceniami potrafią prawidłowo konwertować technologiczną niepewność na elastyczność strategiczną. Co więcej, dysponują portfelem wystarczających rozmiarów, aby ograniczyć ekspozycję na ryzyko związane z wysokim odsetkiem nieudanych projektów.

Z punktu widzenia wspomnianej wcześniej teorii zdolności dynamicznych fundusze zagraniczne wykazują więc większą sprawność w fazach identyfikacji okazji i gospodarowania możliwości, ze względu na to, że potrafią zidentyfikować wczesne sygnały, a następnie sprawnie przekształcić je w potencjalnie korzystną decyzję inwestycyjną (Heugens, Kapoor, 2024). W praktyce oznacza to, że ElevenLabs otrzymało od funduszu Andreessen Horowitz warunki inwestycji w mniej niż sześć tygodni. Takie tempo jest niemożliwe do osiągnięcia w strukturach polskich funduszy, obciążonych procedurami oraz słabszą siecią syndykatów. Kluczową rolę odgrywa też kapitał relacyjny: Sequoia Capital zapewnia dostęp do globalnych kanałów dystrybucji i dalszego kapitału wzrostowego, co istotnie poprawia prognozy potencjalnego wyjścia z inwestycji i pozytywnie wpływa na percepcję ryzyka. W przypadku polskiego rynku venture capital tego rodzaju zasoby sieciowe są w fazie kształtowania, a brak ich rozwoju może pogłębić niechęć rodzimych inwestorów do podejmowania ryzyka i globalnej ekspansji. Co istotne, w amerykańskim ekosystemie wysoka tolerancja na niepewność technologiczną idzie w parze z systemowym uczeniem organizacji, gdyż fundusze prowadzą retrospektywne analizy projektów, gromadząc tym samym wiedzę o przyczynach niepowodzeń i aktualizując bazy doświadczeń. To z kolei podnosi jakość kolejnych decyzji inwestycyjnych (McKinsey, 2023). W Polsce natomiast kultura inwestycyjna premiuje ostrożność, albowiem mechanizmy zbierania informacji zwrotnej są niewystarczające, a koszty wizerunkowe błędnych decyzji przypisywane są indywidualnym partnerom funduszy. W rezultacie dominuje podejście minimalizujące ryzyko oraz awersja do strat, podczas gdy renomowane podmioty zagraniczne rekomendują maksymalizację wartości opcji strategicznych (Botsari i Lang, 2024). Co więcej, na poziomie instytucjonalnym przewagę zagranicznych funduszy wzmacnia dojrzałość rynków kapitałowych oraz infrastruktury ochrony

własności intelektualnej. Możliwość szybkiej wtórnej sprzedaży udziałów lub debiutu giełdowego skraca oczekiwany czas wyjścia z inwestycji, obniżając jednocześnie koszt finansowania i pozwalając akceptować wyższe ryzyko międzyokresowe. Tymczasem polskie fundusze, funkcjonujące w warunkach ograniczonych opcji wyjścia, uwzględniają w kalkulacjach wyższą premię za płynność i preferują projekty o krótszym horyzoncie zwrotu. W rezultacie to zagraniczne podmioty venture capital dysponują kompetencjami umożliwiającymi finansowanie przedsięwzięć z obszaru deep tech takich jak ElevenLabs, podczas gdy fundusze krajowe pozostają poza tym kręgiem.

Z perspektywy nauk o zarządzaniu ten kontrast potwierdza, że sukces inwestycyjny w sektorze technologicznym jest w równym stopniu funkcją jakości infrastruktury decyzyjnej, rozumianej jako zestaw praktyk, standardów oraz zdolności organizacyjnych, co samej wielkości zasobów finansowych.

Podsumowanie

Studium przypadku ElevenLabs ukazuje wielopłaszczyznowy charakter ograniczeń, które uniemożliwiły polskim funduszom venture capital partycypację w finansowaniu przełomowego projektu z obszaru sztucznej inteligencji. Kluczowe jednak jest wskazanie, że zasadniczą barierą, która wpłynęła na brak zaangażowania we wspomniany projekt, nie był niedobór kapitału. O wiele większe znaczenie miały mechanizmy zakorzenione w strukturze poznawczej, organizacyjnej i kulturowej, ponieważ właśnie te składowe wpływają na sposób postrzegania i zarządzania ryzykiem inwestycyjnym w sektorze deep tech.

Po pierwsze, silne heurystyki aktywowane przez niedobory kompetencji technicznych w zespołach inwestycyjnych, w warunkach wysokiej złożoności technologicznej prowadzą do błędu zaniechania (Cliff, Park, Bock, 2022). W praktyce przekłada się to na brak zdolności funduszy do rzetelnej oceny projektów z obszaru deep tech oraz nadmierną zależność od uprzednich schematów inwestycyjnych. Po drugie, brak rozwiniętych zdolności adaptacyjnych objawia się m.in. w opóźnionych procesach decyzyjnych, które ograniczają możliwość szybkiego reagowania na okazje inwestycyjne i mogą zniechęcać innych uczestników syndykatów, wpływając pośrednio na dostępność kapitału. To zaś uniemożliwia skuteczne rozpoznawanie i wykorzystywanie okazji inwestycyjnych (Heugens, Kapoor, 2024). Po trzecie, kultura organizacyjna dominująca w polskim ekosystemie venture capital charakteryzuje się awersją do niepewności oraz lokalnym kierunkiem inwestycji, co wzmacnia preferencję projektów o przewidywalnych, krótkoterminowych zyskach. To zaś skutecznie eliminuje z portfela inwestycji przedsięwzięcia technologiczne o wysokim ryzyku (Bottazzi, Giorcelli, 2021). Dodatkowo, obserwowany w 2023 roku kryzys instytucjonalny pogłębił zapaść

rynku venture capital w Polsce. Brak kontynuacji programów publicznych, takich jak Bridge Alfa, doprowadził do likwidacji wielu funduszy i powstania luki kapitałowej, szczególnie dotkliwej dla projektów we wczesnych fazach rozwoju. Pomimo dynamicznego rozwoju gospodarczego Polska pozostaje na końcu zestawienia rozwiniętych gospodarek pod względem relacji wartości inwestycji venture capital do PKB, która wynosi 0,01%; w porównaniu do 0,14% w Estonii czy ponad 0,20% w Stanach Zjednoczonych i Kanadzie (PFR Ventures, 2024). Dalsze badania porównawcze pozwolą zidentyfikować uniwersalne i lokalnie uwarunkowane czynniki wpływające na skuteczność decyzji inwestycyjnych w projektach z obszaru technologii przyszłości.

W ujęciu teoretycznym niniejsze studium przypadku rozwija koncepcję ograniczonej racjonalności Herberta Simona, wskazując, że decyzje inwestycyjne funduszy venture capital opierają się na heurystykach, których charakter wynika z istniejących ograniczeń organizacyjnych oraz zinternalizowanych norm kulturowych. Jednocześnie znajduje tu potwierdzenie teza Teece'a (2020) o kluczowej roli zdolności dynamicznych w zarządzaniu zmianą technologiczną, wskazując, że brak zdolności reorganizacyjnych i adaptacyjnych uniemożliwia budowanie efektywnych syndykatów inwestycyjnych dla kolejnych realizowanych rund finansowania projektów z obszaru deep tech.

Z perspektywy praktyki zarządzania sformułować można trzy główne rekomendacje. Po pierwsze, wskazane jest systemowe wzmacnianie kompetencji technologicznych w funduszach venture capital poprzez zatrudnianie specjalistów z obszaru technologii oraz rozwijanie współpracy z uczelniami technicznymi. Działania te sprzyjają trafniejszej ocenie ryzyka inwestycyjnego w projektach o wysokim stopniu zaawansowania technologicznego. Po drugie, zalecana jest reorganizacja procesów inwestycyjnych w kierunku modeli hybrydowych, w których formalne procedury i analiza wskaźnikowa uzupełniane są doświadczeniem i ekspercką intuicją. Tego rodzaju podejście może prowadzić do skrócenia czasu podejmowania decyzji oraz zwiększenia konkurencyjności względem międzynarodowych podmiotów inwestycyjnych. Ponadto kluczowa jest długofalowa transformacja kulturowa, która zakłada instytucjonalizację narzędzi umożliwiających świadome funkcjonowanie w warunkach niepewności. Przykładem takich rozwiązań mogą być obowiązkowe analizy retrospektywne dotyczące globalnych inwestycji oraz transparentne rejestrowanie oraz analizowanie nie tylko nieudanych projektów, ale również tych, które się powiodły. Kompleksowa analiza pozwoli określić czynniki, które miały wpływ na rezultat realizowanego projektu i może stanowić punkt wyjściowy do dalszych rozważań. Co więcej, taki wspólny kapitał wiedzy organizacyjnej jest nieoceniony z punktu widzenia nie tylko funduszy venture capital i innych inwestorów instytucjonalnych, ale również samych potencjalnych założycieli start-upów z obszaru technologii przyszłości.

Niniejszy artykuł podkreśla potrzebę dalszych badań porównawczych i szerszej próby badawczej, aby poprawnie zweryfikować przedstawiony model w różnych kontekstach kulturowych i regulacyjnych. Tylko takie podejście umożliwi projektowanie skutecznych działań instytucjonalnych zwiększających gotowość funduszy venture capital do inwestowania w technologie przyszłości.

Bibliografia

- Antkiewicz, S. (2020) *Fundusze wysokiego ryzyka w Polsce na przykładzie Narodowych Funduszy Inwestycyjnych*. GWSH.
- Beauchamp, M. (2018). *Złota księga venture capital w Polsce 2018*. Fundacja Startup Poland.
- Bottazzi, L., Giorcelli, M. (2021). Local cultures and venture capital decision-making. *Entrepreneurship Theory and Practice*, 45(8), 1645–1672.
- Cenamor, J., Wiklund, J. (2023). Absorptive capacity in entrepreneurial finance. *Journal of Business Venturing*, 38(4), 106–129.
- Cliff, C., Park, J., & Bock, A. (2022). Heuristics in venture capital. *Journal of Business Research*, 147, 408–421.
- Colombo, M., Grilli, L., & Piva, E. (2022) Venture capital as ecosystem architect: organisational routines for value creation. *Research Policy*, 51(1), 104–178.
- Dealroom (2023). *Baltic Deep Tech Report 2023*. Amsterdam.
- Dealroom (2024). *Central and Eastern European startups 2024*. Amsterdam.
- Dealroom (2025). *Central and Eastern European startups 2025, Spotlight on scaleups*. Amsterdam.
- Drucker, P. F. (1962). *Innowacje i przedsiębiorczość. Praktyka i zasady*. Polskie Wydawnictwo Ekonomiczne.
- European Investment Fund (EIF) (2024). *Venture Capital Due Diligence Timelines Survey*, Luksemburg.
- European Investment Fund (2024), Botsari, A., Lang F., *EIF VC Survey 2024: Market sentiment. EIF Working Paper 2024/00, Market Assessment and Research*, Luksemburg.
- Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: notes on the Gioia methodology. *Organizational Research Methods*, 16(1), 15–31.
- Gornall, W., Strebulaev, I. (2021). The economic impact of venture capital: evidence from IPOs. *Oxford Review of Economic Policy*, 37(2), 185–203.
- Gornall, W., Strebulaev, I. A. (2021). Lead-certification and the dynamics of venture capital investment. *Journal of Financial Economics*, 141(1), 47–63.

- Heugens, P. P. M., & Kapoor, R. (2024). Dynamic capabilities in financial intermediaries: a multilevel synthesis. *Strategic Management Journal*, 45(3), 411–439.
- Inovo VC (2024). *Polish VC Preferences Report 2024*. Warszawa.
- Kokot-Stępień, P. (2016). Finansowanie działalności innowacyjnej przedsiębiorstw w Polsce. *Zeszyty Naukowe Politechniki Częstochowskiej. Zarządzanie*, nr 24.
- McKinsey & Company (2023). *European deep tech: What investors need to know*, Bruksela.
- Miller, R., Lessard, D. (2001). Understanding and managing risks in large engineering projects. *International Journal of Project Management*, 19(8), 437–443.
- PitchBook (2024). *AI & Machine Learning Annual Report 2024*. Seattle.
- PFR Ventures (2024). Raport: Inwestycje VC w Europie Środkowo-Wschodniej 2024. Warszawa.
- PFR Ventures (2024). *Polish VC Market Outlook 2024*. Warszawa.
- Preqin (2025). *Global Venture Capital Report 2025*. London.
- Rosa, A. (2008). Venture Capital w Polsce. *Zeszyty Naukowe Wydziału Nauk Ekonomicznych Politechniki Koszalińskiej*, 1(12), 133–143.
- Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review*, 69(4), 493–513.
- Stake, R. E. (2020). *The Art of Case Study Research*. SAGE Publications.
- Teece, D. J. (2020). *Dynamic Capabilities and Strategic Management: Organizing for Innovation and Growth*. Oxford University Press.
- Węclawski, J. (1997). *Venture capital. Nowy instrument finansowania przedsiębiorstw*. PWN.
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods. Sixth edition*. SAGE Publications.
- Zóltowski, M., Majchrak, O., & Zreda, M. (2017). *Fundusze Venture Capital – BRIDGE Alfa – zagadnienia prawne i merytoryczne*. NCBR.

Netografia

- Cao, L. et al. (2022). *Using Deep Learning to Find the Next Unicorn: A Practical Synthesis*, *arXiv preprint*. <https://arxiv.org/abs/2210.14195>
- Carniel, T. et al. (2023). *Investor-patent networks as mutualistic networks*, *arXiv preprint*. <https://arxiv.org/abs/2311.18625>
- Ozince, E., & Ihlamur, Y. (2024). *Automating Venture Capital: Founder assessment using LLM-powered segmentation*, *arXiv preprint*. <https://arxiv.org/abs/2407.04885>
- TechCrunch (2025). *ElevenLabs is launching its own speech-to-text model*. <https://techcrunch.com/2025/02/26/elevenlabs-is-launching-its-own-speech-to-text-model>
- Vestbee (2025). *Polish-founded voice AI platform ElevenLabs secures \$180 M....* <https://www.vestbee.com/blog/articles/eleven-labs-secures-180-m>

Transformacja ról pracowników Centrów Usług Wspólnych w erze AI: trendy i nowe kompetencje

Streszczenie

W dobie dynamicznego rozwoju technologii opartych na sztucznej inteligencji Centra Usług Wspólnych (Shared Services Centers, SSC) przechodzą głęboką transformację. Tradycyjnie kojarzone z wykonywaniem powtarzalnych zadań operacyjnych, dziś stają się strategicznymi jednostkami wspierającymi innowacje, analizę danych oraz cyfrową transformację całych organizacji. Zmiany te pociągają za sobą konieczność redefinicji ról pracowników SSC oraz rozwoju nowych kompetencji: cyfrowych, analitycznych i strategicznych.

Celem artykułu jest analiza aktualnej sytuacji oraz wpływu wdrażania technologii AI w branży usług biznesowych. Przedstawiono aktualne trendy w sektorze, przykłady zastosowań AI oraz wyzwania, jakie napotykają organizacje, szczególnie w kontekście wolniejszego tempa rozwoju kompetencji w stosunku do inwestycji technologicznych. W artykule zaproponowano czterofazowy model ewolucji roli pracownika SSC: od wykonawcy zadań, przez nadzorcę technologii, analityka danych i partnera biznesowego, aż po lidera innowacji i zmian.

Wnioski płynące z analizy podkreślają znaczenie integracji strategii technologicznej ze strategią zarządzania zasobami ludzkimi oraz istotę budowania kultury organizacyjnej opartej na ciągłym uczeniu się, by sprostać rosnącym oczekiwaniom wobec SSC w erze sztucznej inteligencji i stać się liderem innowacji i transformacji w swoich organizacjach.

Słowa kluczowe: Centra Usług Wspólnych, transformacja, sztuczna inteligencja, umiejętności przyszłości

1 Akademia WSB w Dąbrowie Górniczej

The Transformation of Employee Roles in Shared Services Centers in the Age of AI: Trends and Emerging Competencies

Abstract

In the era of rapid development of Artificial Intelligence (AI), Shared Services Centers (SSC) are undergoing a profound transformation. Traditionally associated with the execution of repetitive operational tasks, SSCs are now evolving into strategic units supporting innovation, analytics, and the digital transformation of entire organizations. These changes require a re-definition of employee roles and the development of new competencies, including digital, analytical, and strategic skills.

This article analyzes the actual situation and the impact of AI implementation on the Shared Services industry. It presents current trends in the sector, examples of AI adoption, and the key challenges organizations face, especially in the context of a growing gap between technological investment and the development of relevant employees' competencies. The article proposes a four-phase model of employee role evolution in SSCs: from task executor, through technology coordinator, data analyst and business partner to innovation leader and change agent.

The findings highlight the importance of integrating technology strategy with Human Resources management strategy, and emphasize the need to build an organizational culture based on continuous learning in order to meet the growing expectations placed towards SSCs in the era of AI and to become innovation and transformation leaders within their organizations.

Keywords: Shared Services, transformation, AI, skills of the future

Wykaz skrótów

AI	– Artificial Intelligence (Sztuczna inteligencja)
APEX GBS	– Alliance for Professional Excellence in Global Business Services
BPM	– Business Process Management (Zarządzanie procesami biznesowymi)
CoE	– Center of Excellence (Centrum doskonałości)
GBS	– Global Business Services (Globalne usługi biznesowe)
GenAI	– Generative Artificial Intelligence (Generatywna sztuczna inteligencja)
HR	– Human Resources (Zasoby ludzkie)
IDP	– Intelligent Document Processing (Inteligentne przetwarzanie dokumentów)
IT	– Information Technology (Technologie informacyjne)
ML	– Machine Learning (Uczenie maszynowe)
NLP	– Natural Language Processing (Przetwarzanie języka naturalnego)
RPA	– Robotic Process Automation (Zrobotyzowana automatyzacja procesów)
SSC	– Shared Services Center (Centrum Usług Wspólnych)
SLA	– Service Level Agreement (Umowa dotycząca poziomu świadczenia usług)
SSON	– Shared Services Outsourcing Network

Wstęp

Współczesne Centra Usług Wspólnych (Shared Services Centers, SSC) znajdują się w kluczowym momencie transformacji. Tradycyjnie pełniące funkcję jednostek operacyjnych realizujących powtarzalne, rutynowe zadania, dziś coraz częściej stają się dynamicznymi uczestnikami, a często wręcz liderami cyfrowej transformacji organizacji. Wraz z rosnącą dostępnością i dojrzałością technologii opartych na sztucznej inteligencji (Artificial Intelligence, AI), zrobotyzowanej automatyzacji procesów (Robotic Process Automation, RPA) oraz analityce predykcyjnej, zmienia się nie tylko sposób wykonywania pracy, ale także charakter i wartość dodana roli pracownika w strukturze SSC.

Adaptacja nowoczesnych technologii może przynieść znaczące korzyści – zwiększoną efektywność, lepszą jakość danych i szybsze podejmowanie decyzji. Jednak wdrażanie AI wiąże się także z szeregiem wyzwań, w tym z rosnącą luką kompetencyjną wśród pracowników. Podczas gdy organizacje inwestują w technologie, rozwój umiejętności cyfrowych pracowników często nie nadąża za tempem tych zmian. Powoduje to potrzebę redefinicji ról zawodowych, modeli kompetencyjnych oraz sposobu postrzegania pracy w SSC.

Celem niniejszego artykułu jest analiza aktualnej sytuacji w branży SSC z uwzględnieniem zastosowania AI, identyfikacja kluczowych trendów i barier transformacyjnych, a także przedstawienie czterech faz ewolucji roli pracownika w tym sektorze: od wykonawcy prostych zadań do innowatora i partnera strategicznego. Sztuczna inteligencja zmienia role pracowników, a nowe kompetencje stają się kluczowe, aby skutecznie wykorzystać jej potencjał oraz zapewnić ciągły rozwój organizacji i spełnić oczekiwania stawiane SSC.

Artykuł łączy zagadnienia z obszaru technologii, strategii organizacyjnej i zarządzania zasobami ludzkimi. Opiera się na przeglądzie literatury, analizie raportów branżowych i obserwacjach z praktyki biznesowej, oferując kompleksowe spojrzenie na kierunek, w jakim zmierzają Centra Usług Wspólnych w erze sztucznej inteligencji.

Sektor Centrów Usług Wspólnych: aktualna sytuacja i plany rozwoju

W dzisiejszych czasach wiele firm działa na skalę globalną, oferując produkty i usługi w licznych krajach, często na kilku kontynentach. Oprócz swojej głównej działalności firmy te mają w swojej strukturze szereg procesów niezbędnych do prowadzenia działalności, takich jak: zarządzanie zasobami ludzkimi (HR), finanse, IT, zarządzanie łańcuchem dostaw czy obsługa klientów. Procesy te są widziane jako procesy wspierające (tzw. back office). Firmy działające na dużą skalę

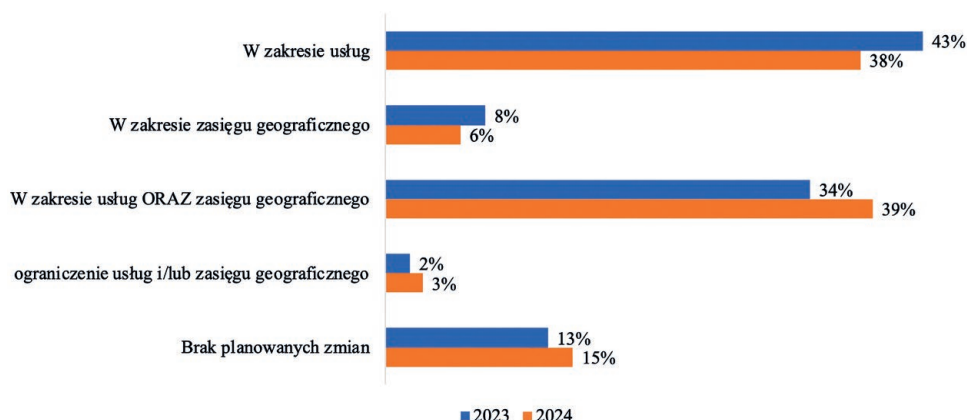
często decydują, aby procesy wspierające wykonywane były przez jednostki konsolidujące je w jednej lokalizacji, zamiast przez każdy oddział osobno. Konsolidacja może się odbywać np. poprzez centralizację usług i obsługę pozostałych jednostek z poziomu centrali, outsourcing procesów na zewnątrz organizacji lub poprzez utworzenie niezależnej jednostki wewnątrz organizacji, tzw. Centrum Usług Wspólnych (Shared Services Centers, SSC), obsługującej pozostałe departamenty jako klientów wewnętrznych. Centra łączące funkcje takie jak finanse, HR, IT i inne oraz oferujące zintegrowane podejście do świadczenia procesów wspierających nazywane są Globalnymi Centrami Usług (Global Business Services, GBS).

Centra Usług Wspólnych są niezależnymi jednostkami, zazwyczaj rozliczającymi zarówno świadczone usługi, jak i rezultaty ich działania na rzecz klientów wewnętrznych. Przed rozpoczęciem obsługi podpisywane są umowy dotyczące poziomu świadczenia usług (Service Level Agreements, SLA), w których uzgadniane są koszty, czas przewidywany na dostarczenie poszczególnych usług oraz mierniki jakościowe. Centra Usług Wspólnych są tworzone z zamiarem uzyskania szeregu korzyści. Są to między innymi: niższe koszty pracy, ekonomia skali, większa kontrola nad wykonywanymi procesami, standaryzacja procesów oraz potencjał wykorzystania technologii celem usprawniania pracy i podnoszenia efektywności. SSC należą do sektora nowoczesnych usług biznesowych². Według raportu Związku Liderów Sektora Usług Biznesowych (ABSL), opisującego sektor nowoczesnych usług biznesowych w Polsce w 2024 roku, „zatrudnienie w sektorze rosło szybciej niż ogólne zatrudnienie w gospodarce. Udział sektora w całkowitym zatrudnieniu w sektorze przedsiębiorstw w Polsce wzrósł do poziomu 7,0%, a udział w PKB – do 5,3%. Dwie lokalizacje – Kraków i Warszawa przekroczyły próg 100 000 miejsc pracy, a każda z nich zatrudnia w swych centrach 101 000 pracowników” (ABSL, 2024, s. 6). Zatrudnienie w sektorze wynosiło w 2024 roku w Polsce blisko pół miliona pracowników w prawie dwóch tysiącach centrów nowoczesnych usług, a nowe centra nadal są tworzone (ABSL, 2024, s. 34). Zgodnie z raportem opublikowanym przez Deloitte, Polska wraz z USA oraz Indiami znajdują się od wielu lat niezmiennie w światowej czołówce lokalizacji, w których tworzone są centra usług biznesowych (Deloitte, 2025, s. 11). Rozwój tego sektora gospodarki jest widoczny w skali globalnej. Shared Services Outsourcing Network (SSON) Research & Analytics opublikowało wyniki badania, w którym aż 38% przedstawicieli badanych organizacji z sektora Shared Services zadeklarowało w 2024 roku plany rozszerzenia świadczonych usług w ciągu najbliższego roku. Mimo że jest to wynik o 5% niższy niż w 2023 roku, świadczy nadal o rozwoju

² Sektor nowoczesnych usług biznesowych według Związku Liderów Sektora Usług Biznesowych (ABSL) obejmuje działalność centrów z obszaru outsourcingu procesów biznesowych, usług wspólnych, globalnych usług biznesowych, informatyki oraz badań i rozwoju zlokalizowanych w Polsce i zatrudniających co najmniej 25 osób.

tych organizacji i potencjale do oferowania coraz większej ilości usług dla swoich klientów (SSON Research & Analytics, 2025a, s. 6).

Wykres 1. Planowany rozwój usług w sektorze Centrów Usług Wspólnych w ciągu najbliższego roku (wg stanu na 2025 r.)



Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 6.

Dynamika sektora nowoczesnych usług biznesowych jest widoczna również w rotacji pracowników, co z pewnością stanowi wyzwanie dla wielu organizacji. Według ostatniego raportu opublikowanego przez firmę Antal dotyczącego zjawiska attrition, wskaźnik ten wynosił 20% w 2019 roku, w porównaniu do 10% w 2016 roku (Antal, 2019, s. 13). „Wyniki badania wskazują, że częściej obciążone najwyższym wskaźnikiem rotacji są grupy obejmujące osoby o krótszym stażu pracy w danej organizacji. 21% badanych pracodawców wskazało, że najwyższy wskaźnik attrition³ obserwują wśród pracowników zatrudnionych poniżej 3 miesięcy. 18% firm wskazuje jako grupę z najwyższym wskaźnikiem osoby zatrudnione powyżej 3 miesięcy, lecz poniżej 1 roku. Najrzadziej badani wskazują grupę pracowników zatrudnionych powyżej 3 lat” (Antal, 2019, s. 20). Tak częsta rotacja i zmienność w zespołach mają wpływ na planowanie pracy i jej rezultaty. Wpływa to również na rozwój kompetencji pracowników i utrzymanie tych kompetencji w zespołach. „Biorąc pod uwagę globalną niepewność oraz szybko zmieniające się uwarunkowania biznesowe, dla centrów ważniejsze jest skupienie się na zatrzymaniu pracowników, zapewnienie im narzędzi do ciągłego uczenia się i rozwoju zawodowego oraz zapewnienie elastyczności, której coraz częściej wymagają pracownicy wykwalifikowani” (ABSL, 2024, s. 8). Stabilna sytuacja kadrowa

3 Wskaźnik procentowy obliczany jako liczba pracowników, którzy odeszli w ciągu roku / ogólną liczbę pracowników.

i ugruntowana wiedza w zespołach mogą pozytywnie wpłynąć na realizację strategii rozszerzania usług świadczonych przez centra i ich plany rozwoju.

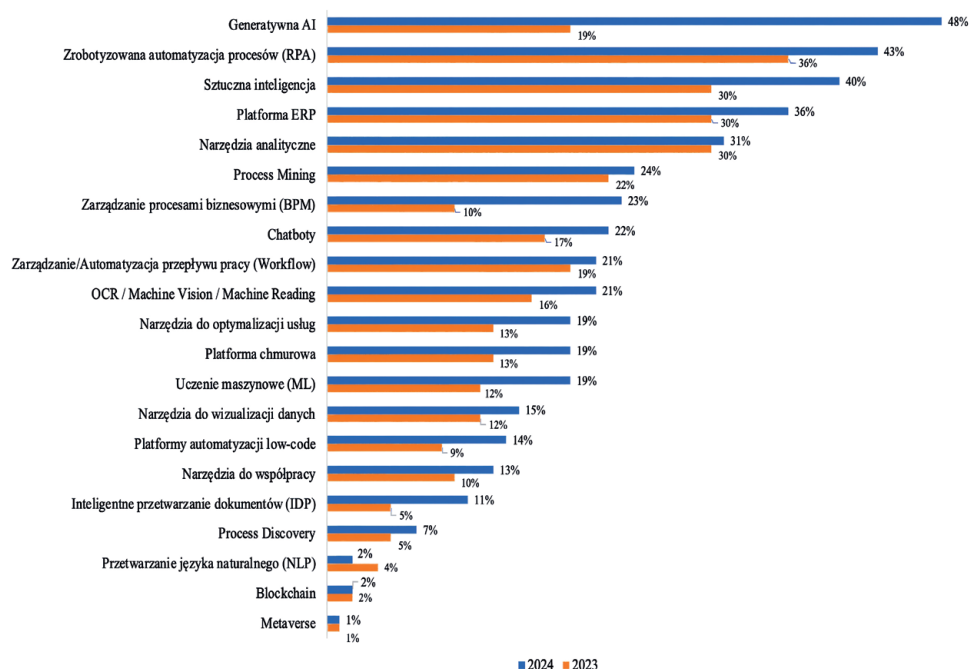
Rola Centrów Usług Biznesowych w transformacji cyfrowej – wyzwania i oczekiwania

Plany rozwojowe w sektorze Shared Services niosą też wyzwania, którym centra muszą sprostać. Wzrastają oczekiwania wobec ich roli – od funkcji jednostek koncentrujących się na redukcji kosztów i realizacji procesów transakcyjnych, do pełnienia funkcji partnera biznesowego. Oznacza to konieczność podnoszenia poziomu dojrzałości organizacyjnej oraz spełniania rosnących wymagań ze strony biznesu, przy jednoczesnym zachowaniu korzyści wynikających z optymalizacji kosztów obsługi procesów. W tym kontekście kluczowe znaczenie ma wdrażanie nowych technologii, w tym rozwiązań opartych na sztucznej inteligencji, a także postępująca cyfryzacja i automatyzacja procesów. Według danych pochodzących z raportu ABSL „jedna czwarta centrów twierdzi, że poziom automatyzacji przekracza 30%. 20% twierdzi, że już wykorzystuje sztuczną inteligencję w działach finansów, kadr i informatycznych” (ABSL, 2024, s. 8). Technologie te pomagają w wykonywaniu powtarzalnych zadań, tworzeniu treści, automatyzują obsługę klientów i wspomagają analizę danych. Pozwala to pracownikom skupić się na zadaniach wnoszących wartość dodaną oraz uwolnić ich czas na zadania o charakterze koncepcyjnym i strategicznym, jednocześnie przeznaczając mniej czasu na powtarzalne i rutynowe aktywności. Według raportu Deloitte 58% organizacji już rozpoczęło bądź planuje wykorzystanie generatywnej sztucznej inteligencji (Generative Artificial Intelligence, GenAI) w swoich procesach, najczęściej w ramach wdrożenia chatbotów, narzędzi AI, zarządzania fakturami oraz kontraktami, obsłudze klientów, analizie danych oraz zarządzaniem danymi (Deloitte, 2025, s. 28). Z kolei ABSL podaje, że oprócz generatywnej sztucznej inteligencji w sektorze wykorzystywana jest również adaptacyjna sztuczna inteligencja. Adaptacyjna sztuczna inteligencja koncentruje się na ciągłym uczeniu się z nowych danych i po zakończeniu etapu treningu dostosowuje swoje działania i prognozy. W sposób dynamiczny aktualizuje swoje algorytmy w odpowiedzi na zmieniające się warunki czy napływające informacje, co stopniowo podnosi jej zdolności decyzyjne i predykcyjne. Dzięki temu potrafi szybciej i skuteczniej się doskonalić, dostosowując się do materiału treningowego. Jej zastosowania widać m.in. w spersonalizowanych rekomendacjach treści czy adaptacyjnych systemach edukacyjnych. Rozwiązania tego typu stanowią ważny element koncepcji „sześciennu transformacji biznesowej” opisanego w raporcie ABSL Strategic Industry Foresight 2023 (raport uwzględnia strategiczną perspektywę na 10 lat+), ponieważ mogą wpływać na

wybór przyszłych modeli biznesowych centrów usług wspólnych. Adaptacyjna AI odgrywa tym samym istotną rolę w rozwoju hiperpersonalizacji usług i poprawie doświadczeń użytkowników (ABSL, 2024, s. 17).

Nowe technologie wpływają bezpośrednio na trendy inwestycyjne w organizacjach. Według danych pochodzących z raportu SSON Research & Analytics w 2024 roku najczęściej wymienianymi priorytetami w zakresie inwestycji technologicznych w branży były GenAI, RPA oraz AI. W przypadku generatywnej sztucznej inteligencji jest to wzrost o 29% w porównaniu do 2023 roku (SSON Research & Analytics 2025a, s. 20–21).

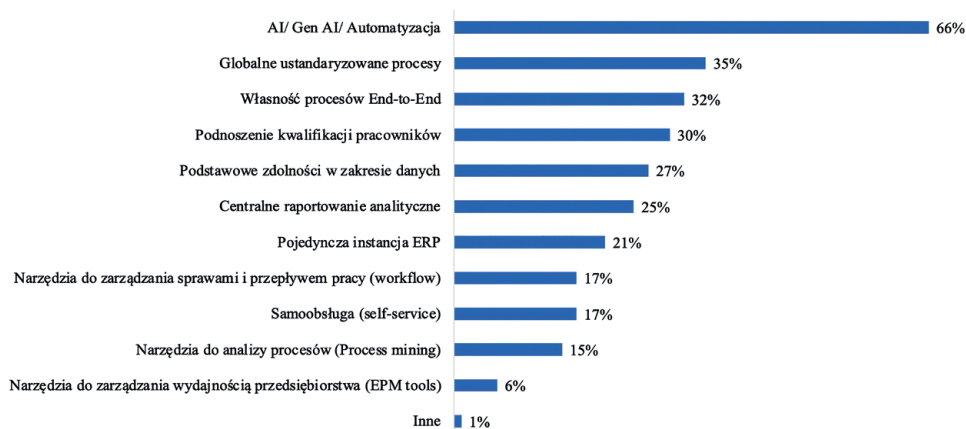
Wykres 2. Najważniejsze inwestycje technologiczne planowane w bieżącym roku



Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 20–21.

Deloitte przewiduje, że trend ten utrzyma się w najbliższym czasie, a inwestycje w AI, GenAI oraz automatyzację procesów będą odgrywały kluczową rolę w zakresie rozwoju organizacji z sektora nie tylko w najbliższym czasie, ale w perspektywie co najmniej trzech najbliższych lat (Deloitte 2025, s. 27).

Wykres 3. Znaczenie inwestycji w GenAI i automatyzację dla rozwoju GBS w perspektywie 1–3 lat (wg stanu na 2025 r.)

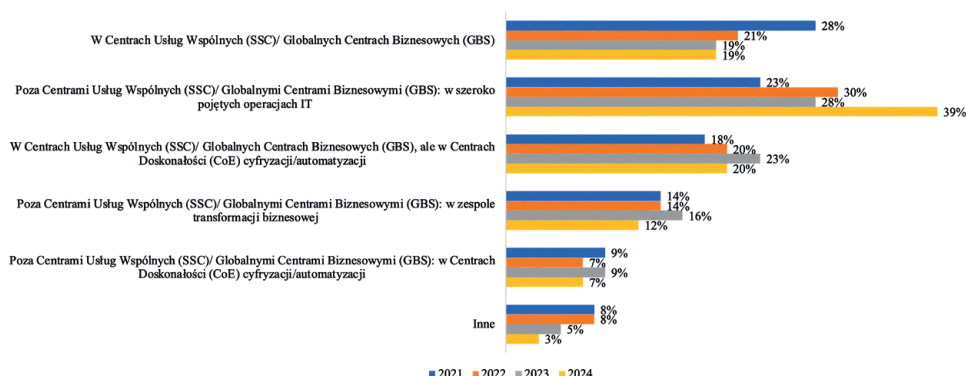


Suma przekracza 100%, ponieważ respondenci mogli zaznaczyć więcej niż jedną odpowiedź.

Źródło: opracowanie własne na podstawie Deloitte, 2025, s. 27.

Korzystanie z nowoczesnych technologii nie jest już więc wyborem a koniecznością. Centra Usług Wspólnych ze względu na swoją charakterystykę mogą i powinny stać się liderami transformacji technologicznej w swoich organizacjach. Niestety, udział centrów prowadzących projekty związane z cyfryzacją uległ zmniejszeniu w 2021 roku po pandemii COVID-19. Obecnie projekty te są w znacznej większości realizowane poza jednostkami usług biznesowych, najczęściej w działach związanych z szeroko pojętym IT (SSON Research & Analytics 2025a, s. 17).

Wykres 4. Lokalizacja projektów cyfryzacyjnych w strukturze organizacyjnej przedsiębiorstw, które uruchomiły Centra Usług Wspólnych



Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 17.

Wnioski pochodzące z raportu przygotowanego przez SSON Research & Analytics oraz firmę Capgemini wskazują potencjał dla organizacji oferujących usługi biznesowe do bycia katalizatorem zmian i rozwoju. Autorzy raportu przewidują jednak, że luki w zakresie kompetencji oraz opór po stronie kultury organizacyjnej mogą stanowić wyzwanie na tej drodze (SSON Research & Analytics 2025b, s. 14). Budowanie nowych kompetencji w organizacjach oraz przezwycięzenie oporu po stronie kultury organizacyjnej jest możliwe poprzez jednocześnie inwestycje w technologie i umiejętności pracowników posługujących się nimi w codziennej pracy. Zgodnie z wynikami ankiety przeprowadzonej przez Deloitte inwestycje w podnoszenie umiejętności pracowników (ang. upskilling) znajdują się na czwartym miejscu listy priorytetów inwestycyjnych organizacji z sektora w perspektywie 1–3 lat (wykres 3). Autorzy raportu *GBS as a Catalyst for Growth* przekonują, że kluczowa w odniesieniu sukcesu będzie koncentracja na budowaniu nowych kompetencji w centrach, wzmacnianie relacji z interesariuszami oraz wnoszenie wartości wykraczającej poza redukcję kosztów, a organizacje świadczące usługi biznesowe powinny stać się centrum innowacji. Będzie to możliwe dzięki wdrożeniu nowych technologii, takich jak wykorzystanie w procesach sztucznej inteligencji (SSON Research & Analytics 2025b, s. 14). Programy i projekty transformacyjne powinny być wdrażane w oparciu o strategię organizacji, w sposób zorganizowany, a nie przypadkowy. Raport *GBS as the Engine of Digital Transformation* potwierdza, że organizacje, które wdrażają AI w sposób strategiczny, są lepiej przygotowane do zwiększania wydajności, optymalizacji kosztów i poprawy doświadczeń użytkowników procesów. Przezwycięzenie luki we wdrażaniu AI nie jest już opcją – jest to niezbędny krok do utrzymania konkurencyjności i osiągnięcia długoterminowego sukcesu w świecie coraz bardziej nastawionym na nowoczesne technologie (SSON Research & Analytics 2025c, s. 4).

Ewolucja ról i kompetencji pracowników Centrów Usług Wspólnych w dobie AI

Mimo że rozwój kompetencji pracowników jest niezbędnym elementem transformacji cyfrowej i wdrożenia planowanych rozwiązań AI, według raportu SSON Research & Analytics w 2025 roku umiejętności związane z AI znalazły się dopiero na jedenastym miejscu umiejętności poszukiwanych w zakresie rekrutacji i szkoleń. Tuż za nimi znalazły się umiejętności związane z zarządzaniem zmianą, niezbędne do skutecznego wdrażania nowych technologii. Obecnie najbardziej poszukiwaną umiejętnością jest rozwiązywanie problemów, a zaraz za nią ogólnie pojęta biegłość technologiczna oraz zarządzanie danymi i ich analiza (SSON Research & Analytics 2025a, s. 14).

Wykres 5. Priorytetowe obszary kompetencyjne w zakresie rekrutacji i szkoleń w Centrach Usług Wspólnych

Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 14.

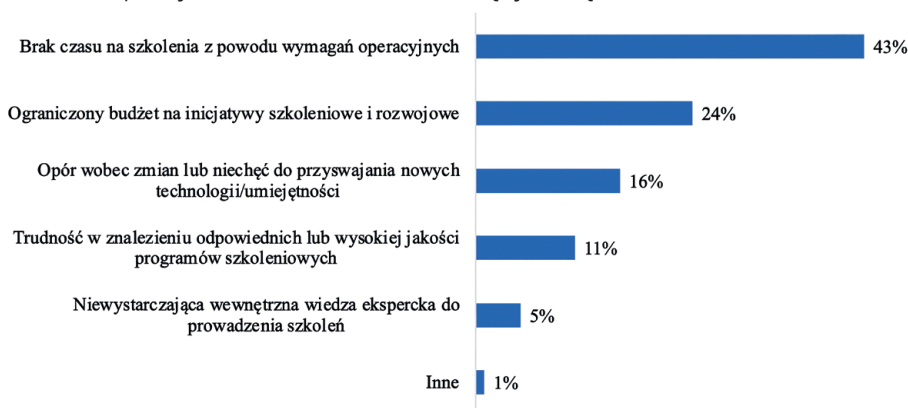
Wynik ten uwidacznia istotną lukę pomiędzy priorytetami inwestycyjnymi w technologie i kompetencje. Z raportu wynika, że w zakresie rozwoju umiejętności organizacje Shared Services bardziej skupiają się na bieżących potrzebach. Uczestnicy badania przewidują zdecydowanie większy nacisk na umiejętności związane z AI i automatyzacją w perspektywie 3–5 lat (SSON Research & Analytics 2025a, s. 16).

Wykres 6. Ewolucja potrzeb kompetencyjnych zespołów i planowane działania strategiczne w perspektywie 3–5 lat w Centrach Usług Wspólnych (wg stanu na 2025)

Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 16.

Firmy z sektora wskazują, że największą przeszkodą, z jaką się mierzą, żeby podnieść umiejętności pracowników i wyjść naprzeciw przyszłym potrzebom biznesowym, jest brak czasu na szkolenia, który z kolei spowodowany jest działalnością operacyjną. Kolejną przeszkodą są ograniczenia budżetowe na wydatki szkoleniowe, a następnie opór przed zmianą oraz opór związany z używaniem nowych technologii (SSON Research & Analytics 2025a, s. 15).

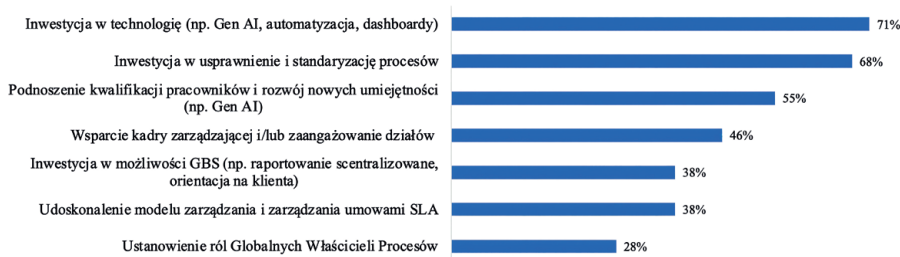
Wykres 7. Bariery w podnoszeniu kwalifikacji pracowników Centrów Usług Wspólnych w kontekście zmieniających się oczekiwań biznesu



Źródło: opracowanie własne na podstawie SSON Research & Analytics, 2025a, s. 15.

Mimo potencjalnych wyzwań warto jednak inwestować w podnoszenie umiejętności związanych ze sztuczną inteligencją. Raport Deloitte przewiduje, że podnoszenie i budowanie kompetencji w zakresie AI jest postrzegane jako element niezbędny do odniesienia kluczowych sukcesów, zaraz po inwestycjach w technologię oraz standaryzację i usprawnianie procesów (Deloitte, 2025, s. 22).

Wykres 8. Czynniki wspierające realizację celów strategicznych w organizacjach GBS

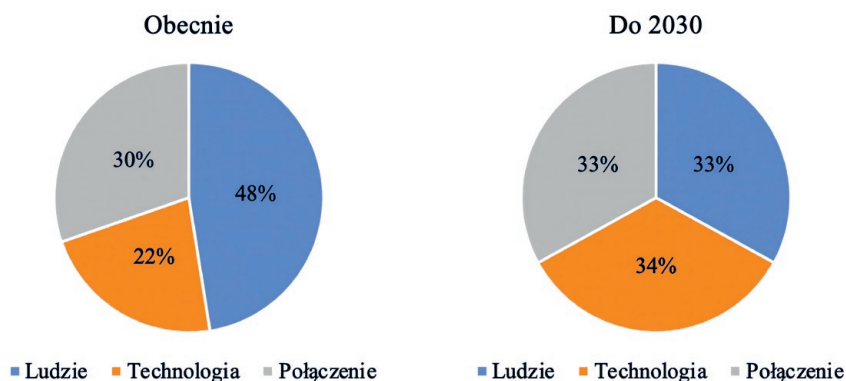


Suma przekracza 100%, ponieważ respondenci mogli zaznaczyć więcej niż jedną odpowiedź.

Źródło: opracowanie własne na podstawie Deloitte, 2025, s. 22.

Nasuwa się zatem pytanie, w jaki sposób firmy z sektora mogą przygotować się pod kątem rozwoju nowych kompetencji pracowników, aby zniwelować lukę pomiędzy już istniejącymi i planowanymi inwestycjami w technologie a kompetencjami potrzebnymi do ich obsługi? Alliance for Professional Excellence in Global Business Services (APEX GBS) podaje, że biorąc pod uwagę relatywnie krótki czas, od jakiego funkcjonują organizacje typu Shared Services, liderzy i menedżerowie muszą liczyć się z koniecznością rozwoju kompetencji pracowników (APEX GBS, 2024, s. 79–80). Rozwój ten może przebiegać zarówno poprzez formalne szkolenia, naukę od koleżanek i kolegów z większym doświadczeniem, coaching lub mentoring. Jednak w czasach tak szybkiego rozwoju technologicznego rozwój kompetencji w obszarze funkcyjnym (np. finanse, HR) nie jest już wystarczający do osiągnięcia strategicznych celów, wspierania innowacji czy maksymalizowania efektywności operacyjnej w centrach. Istotne stają się kompetencje przyszłości. World Economic Forum w swoim raporcie przewiduje w najbliższej przyszłości zanikanie ról związanych z wprowadzaniem danych, księgowością czy księgowaniem płac (World Economic Forum, 2025, s. 23). Autorzy raportu przewidują również, że w 2030 roku proporcje podziału zadań pomiędzy pracowników, technologie lub ich połączenie wyrównają się. Przewidywany spadek zadań wykonywanych przez pracowników to aż 14% w ciągu pięciu lat (World Economic Forum, 2025, s. 26). Ten sam trend może dotyczyć również Centrów Usług Wspólnych.

Wykres 9. Struktura podziału zadań wykonywanych przez człowieka, technologie lub ich połączenie w roku 2025 i prognoza na rok 2030



Źródło: opracowanie własne na podstawie World Economic Forum, 2025, s. 26.

W przypadku tak dynamicznej transformacji technologicznej istotne jest zaplanowanie zmiany ról pracowników SSC i rozwój nowych kompetencji w sposób ramowy, w co najmniej czterech fazach.

Faza 1. Wykonywanie rutynowych zadań, procesy wykonywane w sposób manualny

- Standardowe zadania w Centrach Usług Wspólnych, takie jak wprowadzanie danych, procesowanie dokumentów finansowych, obsługa zapytań.
- Pracownicy wykonują pracę manualnie, a poziom automatyzacji jest bardzo mały.
- Od pracowników oczekuje się postępowania zgodnie z procedurami i instrukcjami, dokładności i szybkości.

Faza 2. Współpraca z narzędziami automatyzującymi

- Wprowadzenie do procesów narzędzi (np. RPA, AI) automatyzujących powtarzalne zadania i operacje.
- Rola pracowników ewoluuje w kierunku nadzoru i kontroli wyników oraz obsługi wyjątków.
- Pojawia się konieczność rozwoju podstawowej wiedzy w zakresie automatyzacji, rozwoju kompetencji technicznych i współpracy z AI.

Faza 3. Partnerstwo biznesowe i analiza danych

- Pracownicy korzystają z analiz wygenerowanych przez AI, interpretują dane, a następnie rekomendują działania, stając się ekspertami wspierającymi decyzje biznesowe.
- Konieczny staje się rozwój kompetencji w zakresie analizy danych, komunikacji i rozwiązywania problemów.

Faza 4. Innowacyjność i przewodzenie zmianom

- Pracownicy projektują nowe rozwiązania wdrażające AI i wspierające rozwój cyfrowej transformacji, a ich kompetencje ewoluują w kierunku zarządzania projektami, zarządzania zmianą oraz innowacyjności i kreatywnego myślenia.
- Pracownicy stają się liderami zmian i ambasadorami kultury innowacji w całej organizacji, a kluczową rolę odgrywają kompetencje przywódcze.

Zaproponowany powyżej ramowy model można podsumować pod kątem zmiany ról pracowników, wpływu AI na ich pracę oraz kluczowych kompetencji w następujący sposób:

Tabela 1. Fazy transformacji roli pracowników Centrów Usług Wspólnych w świetle ich kluczowych kompetencji i wpływu sztucznej inteligencji (AI)

Faza	Rola pracownika SSC	Kluczowe kompetencje	Wpływ AI
1	Wykonawca rutynowych zadań	Dokładność, szybkość	Minimalny
2	Nadzorca automatyzacji	Podstawy AI, nadzór, współpraca z AI	Automatyzacja rutyny, kontrola jakości
3	Analitik i partner biznesowy	Analiza danych, komunikacja, rozwiązywanie problemów	Wsparcie decyzji, interpretacja danych
4	Lider innowacji	Przywództwo, innowacyjność i kreatywne myślenie, zarządzanie projektami i zmianą	Kreowanie nowych rozwiązań AI przez pracowników

Źródło: opracowanie własne.

Dla pracowników Centrów Usług Wspólnych oznacza to ewolucję roli pełnionej w organizacji: od transakcyjnej, poprzez hybrydową, aż do osiągnięcia roli o charakterze strategicznym. Początkowa koncentracja na zadaniach powtarzalnych przenosi się w stronę zadań analitycznych i decyzyjnych. Pewne role zanikają (np. związane z przetwarzaniem dokumentów czy wprowadzaniem danych), a w ich miejsce pojawiają się nowe (np. specjalista ds. automatyzacji, koordynator ds. danych, trener AI). Zmiany te wymagają nabycia nowych kompetencji, takich jak: wiedza o AI, współpraca interdyscyplinarna, krytyczne myślenie, innowacyjność, analiza danych czy przywództwo. Transformacja ról pokazuje, jak ważne jest uczenie się przez całe życie (ang. lifelong learning) oraz otwartość na zmiany i stały rozwój kompetencji. APEX GBS przekonuje, że transformacja umiejętności pracowników umożliwia liderom GBS przejście od bycia organizacjami zorientowanymi wyłącznie na usługi do stania się liderami w zakresie innowacji i pojawiających się trendów. Stowarzyszenie branżowe podkreśla, jak istotne jest stworzenie środowiska, które zachęca do eksperymentowania i uwzględnia pojawiające się nowe technologie. Zapewni to nie tylko rentowność organizacji GBS w przyszłości, ale także wspiera rozwój zawodo-
wy, umożliwiając pracownikom rozwój w stale ewoluującym biznesowo świecie (APEX GBS, 2024, s. 91).

Zakończenie

Transformacja Centrów Usług Wspólnych w kierunku jednostek opartych na zaawansowanych technologiach, takich jak sztuczna inteligencja, zmienia nie tylko charakter pracy, ale także redefiniuje rolę pracowników w strukturze organizacyjnej. Tradycyjne zadania operacyjne ustępują miejsca działaniom analitycznym, doradczym i innowacyjnym. Pewne stanowiska zanikają, a w ich miejsce pojawiają się nowe. Wraz z tą zmianą rośnie znaczenie kompetencji cyfrowych, komunikacyjnych oraz strategicznych.

Zaproponowany w artykule czterofazowy model transformacji ról pracowników Centrów Usług Wspólnych uwzględnia rosnącą złożoność ich pracy i wychodzi naprzeciw oczekiwaniom stawianym SSC. Skuteczna adaptacja do nowej rzeczywistości wymaga czegoś więcej, niż tylko wdrożenia technologii. Konieczne jest równoległe inwestowanie w rozwój pracowników. Luka kompetencyjna, która w wielu organizacjach narasta, może stać się poważnym ograniczeniem dla dalszego rozwoju SSC i pełnego wykorzystania potencjału AI. Z raportów branżowych wynika, że inwestycje technologiczne są obecnie prowadzone na szeroką skalę, zatem nowe kompetencje powinny być nabywane równoległe. Jednak ograniczenia czasowe i budżetowe, intensywność zmian, lęk pracowników przed utratą pracy oraz opór przed zmianami mogą znacznie opóźnić ten proces.

Aby SSC mogły pełnić rolę partnera strategicznego w swoich organizacjach, potrzebne jest zintegrowane podejście, łączące strategię technologiczną ze strategią zarządzania zasobami ludzkimi. Konieczna jest identyfikacja i wspieranie rozwoju kompetencji przyszłości oraz redefinicja ścieżek kariery i systemów ocen pracowniczych. Elastyczność, gotowość do eksperymentowania oraz otwartość na zmiany powinny być wspierane przez liderów i struktury HR, co umożliwi budowanie kultury organizacyjnej opartej na ciągłym uczeniu się. Organizacje powinny okresowo oceniać swój poziom zaawansowania technologicznego i kompetencyjnego, by skutecznie planować kolejne etapy rozwoju. Tylko w ten sposób możliwe będzie stworzenie elastycznych, odpornych i innowacyjnych zespołów, gotowych do działania w dynamicznie zmieniającym się otoczeniu biznesowym i technologicznym, jednocześnie pozwalając SSC sprostać rosnącym wymaganiom biznesu oraz pełnić rolę lidera transformacji w swoich organizacjach.

Bibliografia

- ABSL Association of Business Service Leaders. (2023). *Strategic Foresight in the Business Services sector 2023*, Warszawa. Association of Business Service Leaders.
- ABSL Association of Business Service Leaders. (2024). *Sektor Nowoczesnych Usług Biznesowych w Polsce 2024*, Warszawa. Association of Business Service Leaders.

- Antal. (2019). *Rotacja dobrowolna pracowników w Polsce. Edycja 2. 2019*. Warszawa, Antal.
- APEX GBS Alliance for Professional Excellence in Global Business Services. (2024). *APEX GBS Standards*, Stany Zjednoczone. Alliance for Professional Excellence in Global Business Services Standards Committee.
- Brodzicki, T., Kubacki, D. (2024). *European Business Services Sector 2024*. Warszawa. Association of Business Service Leaders.
- Deloitte (2011). *Shared Services Handbook. Hit the road. A practical guide to implementing shared services*. Deloitte MCS Limited. <https://www.deloitte.co.uk/makeconnections/assets/pdf/shared-services-handbook-hit-the-road.pdf>, [dostęp: 01.06.2025]
- Deloitte (2025). *2025 Deloitte's Global Business Services (GBS) Survey*. Deloitte Development LLC.
- Ferreira, C., Janssen, M. (2022). Shaping the Future of Shared Services Centers: Insights from a Delphi Study About SSC Transformation Towards 2030. *Journal of the Knowledge Economy* (2023), 14, 4828–4847. <https://link.springer.com/article/10.1007/s13132-022-01072-0>, [dostęp: 01.06.2025]
- SSON Research & Analytics. (2025a). *Global State of the Shared Services Industry 2025: Key Highlights*. Singapur, SSON Research & Analytics.
- SSON Research & Analytics. (2025b). *Research Insight Report. GBS as a Catalyst for Growth. From Cost Center to Strategic Value Driver. Survey Analysis*, Singapur, SSON Research & Analytics.
- SSON Research & Analytics. (2025c). *Research Insight Report. GBS as the engine of Digital Transformation*, Singapur, SSON Research & Analytics.
- SSON Research & Analytics (2025d). *Unlocking the Next Frontier of GBS Performance: 10 Key Metrics That Define Success*, Singapur. SSON Research & Analytics.
- Swientozielskyj, S. (2015). *Certificate in Global Business Services. The fundamentals of the main operating models*. Chartered Institute of Management Accountants.
- World Economic Forum. (2024). *Leveraging Generative AI for Job Augmentation and Workforce Productivity: Scenarios, Case Studies, and a Framework for Action. Insight Report November 2024*, Genewa. https://reports.weforum.org/docs/WEF_Leveraging_Generative_AI_for_Job_Augmentation_and_Workforce_Productivity_2024.pdf, [dostęp: 01.06.2025]
- World Economic Forum. (2025). *Future of Jobs Report 2025. Insight Report January 2025*, Genewa. https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf. [dostęp: 01.06.2025]

Mały uczeń w wielkim świecie AI: wyzwania i szanse cyfrowej rewolucji w edukacji wczesnodziecięcej

Streszczenie

Artykuł analizuje wpływ sztucznej inteligencji (AI) na edukację dzieci w wieku przedszkolnym i wczesnoszkolnym z perspektywy praktyka i badacza systemów opieki nad dziećmi. Omówiono zarówno szanse, jak i zagrożenia wynikające z wczesnego kontaktu dzieci z technologią AI, ze szczególnym uwzględnieniem aspektów pedagogicznych, psychologicznych, społecznych oraz cyberbezpieczeństwa. Przedstawiono dobre praktyki z wybranych krajów europejskich i arabskich, pokazując różnorodność podejść do wczesnej edukacji cyfrowej. Artykuł zawiera rekomendacje dotyczące odpowiedzialnego wdrażania AI w edukacji, podkreślając konieczność rozwoju kompetencji cyfrowych, etycznych i społecznych zarówno u dzieci, jak i kadry pedagogicznej. Wskazuje także na zmieniające się paradygmaty w wykorzystaniu technologii informatycznych w pedagogice na tle światowych trendów.

Słowa kluczowe: sztuczna inteligencja, edukacja wczesnodziecięca, cyberbezpieczeństwo, etyka AI, personalizacja nauczania, systemy edukacyjne

1 Akademia WSB w Dąbrowie Górniczej

The Young Learner in the Vast World of AI: Challenges and Opportunities of the Digital Revolution in Early Childhood Education

Abstract

This article analyzes the impact of artificial intelligence (AI) on early childhood and primary education from the perspective of a practitioner and researcher of global childcare systems. Both the opportunities and risks of early exposure to AI technologies are discussed, with particular attention to pedagogical, psychological, social, and cybersecurity aspects. Good practices from selected European and Arab countries are presented, illustrating the diversity of approaches to early digital education. The article offers recommendations for the responsible implementation of AI in education, emphasizing the need to develop digital, ethical, and social competencies among both children and teaching staff. The changing paradigms in the use of information technologies in pedagogy are also highlighted against the backdrop of global trends.

Keywords: artificial intelligence, early childhood education, cybersecurity, AI ethics, personalized learning, education systems

Wstęp

Współczesna edukacja dzieci w wieku przedszkolnym i wczesnoszkolnym staje w obliczu bezprecedensowej transformacji technologicznej. Sztuczna inteligencja, dotychczas domena nauk ścisłych i przemysłu, coraz śміiej wkracza do sal żłobkowych i przedszkolnych, stając się narzędziem codziennej pracy nauczycieli, wychowawców i dyrektorów placówek. Jako dyrektor niepublicznych żłobków i przedszkoli oraz badacz systemów opieki nad dziećmi na świecie, obserwuję zarówno entuzjazm, jak i obawy związane z tym procesem. Celem niniejszego artykułu jest krytyczna analiza szans i zagrożeń, jakie niesie cyfrowa rewolucja w edukacji najmłodszych, oraz sformułowanie praktycznych rekomendacji dla odpowiedzialnego wdrażania AI w pedagogice wczesnodziecięcej. W Polsce obserwuje się rosnące zainteresowanie integracją AI w edukacji, czemu towarzyszą zarówno entuzjazm, jak i obawy nauczycieli oraz rodziców dotyczące bezpieczeństwa, etyki i jakości kształcenia (Polska edukacja w cieniu AI, 2024).

Sztuczna inteligencja w edukacji wczesnodziecięcej – szanse i nowe możliwości

Personalizacja i indywidualizacja procesu nauczania

Jedną z najważniejszych zalet AI w edukacji najmłodszych jest możliwość personalizacji nauczania. Systemy oparte na AI potrafią analizować postępy dzieci w czasie rzeczywistym, dostosowując poziom trudności i formę przekazu do indywidualnych potrzeb każdego ucznia (How AI is transforming early childhood learning, 2025). Przykładem są aplikacje edukacyjne, które dynamicznie modyfikują treści w zależności od zaangażowania i zrozumienia dziecka, co sprzyja efektywnemu przyswajaniu wiedzy (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025). AI umożliwia szybsze wykrywanie trudności w nauce oraz identyfikację dzieci wymagających dodatkowego wsparcia. Systemy analityczne monitorują postępy i sugerują interwencje, zanim problemy staną się poważne, co pozwala na skuteczniejszą pomoc i zwiększa motywację do nauki (Zastosowanie sztucznej inteligencji w edukacji wczesnoszkolnej, 2024).

Rozwój kompetencji przyszłości

Zastosowanie AI sprzyja rozwojowi umiejętności krytycznego myślenia, rozwiązywania problemów oraz kreatywności. Interaktywne programy edukacyjne pozwalają dzieciom na eksplorację i podejmowanie decyzji w realistycznych, symulowanych sytuacjach, co przygotowuje je do wyzwań przyszłego rynku pracy (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025). W polskich przedszkolach i szkołach coraz częściej testuje się narzędzia AI

wspierające indywidualizację nauczania oraz automatyzację oceniania, choć skala wdrożeń jest wciąż ograniczona przez dostępność zasobów i kompetencji cyfrowych kadry (Sztuczna inteligencja w edukacji – raport z InnoLab, 2024).

Zagrożenia i wyzwania cyfrowej rewolucji

Wprowadzenie AI do edukacji wiąże się z koniecznością przetwarzania dużych ilości danych osobowych dzieci. Istnieje ryzyko naruszenia prywatności oraz nieuprawnionego wykorzystania informacji, co wymaga wdrożenia rygorystycznych procedur ochrony danych i edukowania kadry oraz rodziców w zakresie cyberbezpieczeństwa (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025; Zagrożenia przy wdrażaniu rozwiązań AI w edukacji, 2025).

AI, podobnie jak inne technologie cyfrowe, może prowadzić do uzależnienia od bodźców cyfrowych, zaburzeń koncentracji, a nawet wzrostu ryzyka depresji i izolacji społecznej u dzieci (AI dla dzieci pomoże w nauce. Ale zwiększa ryzyko depresji, 2025). Algorytmy projektowane w celu utrzymania uwagi mogą prowadzić do nadmiernego korzystania z urządzeń, dlatego kluczowe jest wprowadzanie ograniczeń czasowych i edukacja w zakresie higieny cyfrowej.

AI w edukacji stawia przed nauczycielami i dyrektorami nowe wyzwania etyczne. Algorytmy mogą nieświadomie wzmacniać stereotypy, prowadzić do dyskryminacji lub nieprzejrzystych decyzji dotyczących oceny czy ścieżki rozwoju dziecka (Etyka AI w edukacji – co warto wiedzieć?, 2025). Niezbędne jest budowanie świadomości etycznej zarówno wśród kadry, jak i dzieci, oraz promowanie transparentności stosowanych rozwiązań.

Przykłady dobrych praktyk – wybrane systemy edukacji na świecie

W krajach skandynawskich (Finlandia, Szwecja, Norwegia) oraz w części krajów Europy Zachodniej zapewniony jest powszechny dostęp do wysokiej jakości wczesnej edukacji już od 6–18 miesiąca życia (Kluczowe dane dotyczące wczesnej edukacji i opieki w Europie, 2022). Systemy te kładą nacisk na rozwój kompetencji cyfrowych i społecznych, przy jednoczesnym zachowaniu równowagi między technologią a bezpośrednią interakcją społeczną. Zjednoczone Emiraty Arabskie jako pierwszy kraj na świecie wprowadziły obowiązkowe zajęcia z AI już od przedszkola. Program nauczania jest zróżnicowany wiekowo i obejmuje kluczowe obszary związane z AI, etyką technologiczną i rozwojem kompetencji przyszłości (UAE makes AI classes mandatory from kindergarten, 2025). Inicjatywa ta ma na celu przygotowanie dzieci do funkcjonowania w społeczeństwie opartym na wiedzy i innowacji.

W krajach arabskich, takich jak Irak, Maroko i Oman, edukacja jest postrzegana jako narzędzie rozwoju społecznego i gospodarczego, jednak wyzwaniem

pozostaje dostępność i jakość edukacji cyfrowej (Edukacja w świecie arabskim na przykładzie Iraku, Maroka i Omanu, 2020). Wdrażanie AI wymaga tam nie tylko inwestycji technologicznych, ale także reform systemowych i podnoszenia kompetencji nauczycieli.

Rekomendacje dla bezpiecznego i odpowiedzialnego wdrażania AI w edukacji wczesnodziecięcej

Rozwój kompetencji cyfrowych i etycznych

Kadra pedagogiczna powinna być regularnie szkolona w zakresie nowych technologii, cyberbezpieczeństwa oraz etyki AI (Odpowiedzialne wykorzystanie sztucznej inteligencji w edukacji, 2025).

Personalizacja z zachowaniem równowagi

AI powinna wspierać, a nie zastępować nauczyciela. Kluczowe jest zachowanie równowagi między technologią a relacjami społecznymi i rozwojem emocjonalnym dzieci (How AI is transforming early childhood learning, 2025).

Transparentność i ochrona danych

Wdrażanie AI musi być zgodne z przepisami o ochronie danych osobowych, a rodzice i dzieci powinni być informowani o sposobie działania systemów AI i przetwarzania danych (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025).

Monitorowanie wpływu AI na zdrowie psychiczne

Należy wprowadzać limity czasu korzystania z AI oraz regularnie monitorować wpływ technologii na dobrostan dzieci (AI dla dzieci pomoże w nauce. Ale zwiększa ryzyko depresji, 2025).

Współpraca międzynarodowa

Warto korzystać z doświadczeń krajów, które już wdrożyły AI w edukacji, adaptując sprawdzone rozwiązania do lokalnych warunków (UAE makes AI classes mandatory from kindergarten, 2025).

Sztuczna inteligencja w edukacji wczesnodziecięcej to zarówno ogromna szansa na rozwój indywidualnych talentów i kompetencji przyszłości, jak i wyzwanie związane z bezpieczeństwem, etyką i zdrowiem psychicznym najmłodszych. Jako dyrektor niepublicznych placówek oraz badacz systemów edukacyjnych rekomenduję wdrażanie AI w sposób przemyślany, z poszanowaniem wartości pedagogicznych i praw dziecka. Kluczowe jest budowanie kultury odpowiedzialności cyfrowej oraz ciągłe doskonalenie kompetencji kadry

pedagogicznej, aby technologia służyła rozwojowi, a nie ograniczała potencjał najmłodszych. Praktyczne przewodniki i raporty, takie jak publikacje PARP czy KIRE, oferują nauczycielom konkretne wskazówki dotyczące wdrażania AI w procesie dydaktycznym, podkreślając znaczenie etyki i bezpieczeństwa (PARP, 2024; KIRE, 2024).

Wyzwania związane z wczesną edukacją dzieci w erze AI

Główne wyzwania to: ochrona prywatności i danych dzieci, nierówności w dostępie do AI, uzależnienie od technologii, brak kompetencji cyfrowych nauczycieli, ryzyko stronniczości algorytmów oraz ograniczenie rozwoju społecznego i krytycznego myślenia. Najnowsze polskie badania podkreślają, że nauczyciele często czują się nieprzygotowani do pracy z AI i oczekują wsparcia w zakresie doskonalenia zawodowego oraz praktycznych szkoleń (Potrzeba doskonalenia nauczycieli w obliczu edukacji z AI, 2025).

Systemy AI w edukacji wymagają gromadzenia i analizowania dużej ilości danych osobowych uczniów, co rodzi poważne obawy dotyczące ochrony prywatności dzieci oraz ryzyka niewłaściwego wykorzystania informacji (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025; Sztuczna inteligencja we współczesnej edukacji, 2025). Niezbędne są rygorystyczne zabezpieczenia oraz edukacja nauczycieli i rodziców w zakresie cyberbezpieczeństwa.

Nie wszystkie dzieci mają równy dostęp do nowoczesnych narzędzi AI, co może pogłębiać istniejące różnice edukacyjne i społeczne (Zastosowanie sztucznej inteligencji w edukacji wczesnoszkolnej, 2024; Jak sztuczna inteligencja wpłynie na edukację dzieci w przyszłości?, 2024). Koszty wdrożenia AI oraz brak infrastruktury technologicznej w niektórych regionach stanowią poważną barierę.

Nadmierne poleganie na AI może prowadzić do ograniczenia bezpośrednich kontaktów międzyludzkich, które są kluczowe dla rozwoju społeczno-emocjonalnego dzieci (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025). Dzieci mogą preferować interakcje z technologią zamiast z rówieśnikami, co sprzyja izolacji społecznej.

Personalizowane, angażujące aplikacje AI mogą prowadzić do nadmiernego korzystania z ekranów i uzależnienia od technologii, a także ograniczać rozwój krytycznego myślenia i samodzielności (Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie?, 2025).

Dzieci, które nie rozumieją, jak działa AI, mogą uwierzyć w nieprawdziwe lub zmyślane odpowiedzi generowane przez algorytmy. Istnieje ryzyko promowania nieobiektywnych treści i manipulacji informacją. Wykorzystanie AI w edukacji wymaga jasnych zasad etycznych dotyczących zbierania i wykorzystywania danych, a także przejrzystości działania algorytmów. Konieczne jest projektowanie

narzędzi AI z myślą o potrzebach i rozwoju dzieci oraz uwzględnianie głosu najmłodszych w procesie wdrażania technologii.

Brak kompetencji cyfrowych nauczycieli i rodziców

Wdrożenie AI wymaga odpowiedniego przeszkolenia kadry pedagogicznej, która często nie posiada wystarczających umiejętności cyfrowych, by efektywnie korzystać z nowych narzędzi (Zastosowanie sztucznej inteligencji w edukacji wczesnoszkolnej, 2024; Sztuczna inteligencja we współczesnej edukacji, 2025). Opór nauczycieli i rodziców wobec nowych technologii może utrudniać skuteczne wdrożenie AI w edukacji. Zaawansowane rozwiązania AI są kosztowne, co stanowi barierę zwłaszcza dla placówek niepublicznych i szkół w mniej rozwiniętych regionach. Wczesna edukacja w erze AI to szansa na personalizację nauczania i rozwój kompetencji przyszłości, ale wymaga rozwiązania poważnych wyzwań: ochrony prywatności, równości dostępu, zachowania równowagi między technologią a relacjami społecznymi, przeciwdziałania uzależnieniom, zapewnienia etyki i transparentności oraz inwestowania w kompetencje cyfrowe nauczycieli i rodziców.

W jaki sposób edukacja nauczycieli może przygotować dzieci do odpowiedzialnego korzystania z AI

Edukacja nauczycieli odgrywa kluczową rolę w przygotowaniu dzieci do odpowiedzialnego korzystania z AI. Odpowiednio przeszkoleni pedagodzy mogą świadomie wdrażać technologie, kształtować postawy etyczne i rozwijać kompetencje cyfrowe uczniów na kilku poziomach.

Nauczyciele, którzy rozumieją mechanizmy działania AI, mogą uczyć dzieci krytycznego myślenia, rozpoznawania dezinformacji oraz odpowiedzialnego korzystania z narzędzi cyfrowych. Stają się mentorami etyki, pomagając uczniom zrozumieć, jak bezpiecznie i świadomie korzystać z nowych technologii (Szkola Popów, 2025). Pedagodzy powinni wprowadzać i omawiać z uczniami zasady korzystania z AI, np. w formie klasowych porozumień (Classroom Agreement), które obejmują kwestie etyki, ochrony prywatności, krytycznego podejścia do treści generowanych przez AI oraz zdrowych nawyków cyfrowych (Omatmo, 2025). Podczas szkoleń nauczyciele uczą się wykorzystywać AI do personalizacji materiałów edukacyjnych, monitorowania postępów i dostosowywania ścieżek nauki do potrzeb każdego dziecka. Takie podejście zwiększa zaangażowanie uczniów i pozwala lepiej rozwijać ich kompetencje cyfrowe (Aktywny Nauczyciel, 2025). Edukacja nauczycieli obejmuje nie tylko aspekty techniczne, ale także umiejętność rozwijania u dzieci empatii, pracy zespołowej i komunikacji – kompetencji, których AI nie zastąpi. Nauczyciel staje się przewodnikiem po świecie nowych

technologii, dbając równocześnie o równowagę między rozwojem cyfrowym a społecznym (Szkoła Popów, 2025). Uczestnictwo w szkoleniach, kursach i społecznościach nauczycielskich pozwala pedagogom na bieżąco poznawać nowe narzędzia AI, dzielić się dobrymi praktykami i wspólnie wypracowywać strategie odpowiedzialnego wdrażania technologii w edukacji (AIProgres, 2025). Nawet najbardziej zaawansowane narzędzia AI nie zastąpią nauczyciela jako osoby wspierającej rozwój emocjonalny i społeczny ucznia. Szkolenia pomagają pedagogom zachować tę równowagę i świadomie wykorzystywać AI jako wsparcie, a nie zamiennik relacji międzyludzkich (KIRE, 2024). Edukacja nauczycieli w zakresie AI umożliwia im tworzenie bezpiecznego, etycznego i angażującego środowiska, w którym dzieci uczą się odpowiedzialnego korzystania z nowych technologii. To nauczyciele – dzięki wiedzy, refleksji i świadomemu podejściu – kształtują przyszłe pokolenia użytkowników AI, którzy będą korzystać z niej mądrze, krytycznie i z poszanowaniem wartości społecznych.

Jak przygotować nauczycieli do uczenia dzieci odpowiedzialnego korzystania z nowych technologii

Aby skutecznie przygotować nauczycieli do uczenia dzieci odpowiedzialnego korzystania z nowych technologii, należy wdrożyć kompleksowe działania, obejmujące rozwój kompetencji cyfrowych, metodycznych i etycznych. Kluczowe elementy tego procesu to szereg poniższych działań:

1. Regularne szkolenia z zakresu obsługi nowoczesnych narzędzi cyfrowych, w tym sztucznej inteligencji, programowania, korzystania z platform e-learningowych oraz aplikacji edukacyjnych.
2. Warsztaty dotyczące cyberbezpieczeństwa, profilaktyki e-uzależnień, rozpoznawania i przeciwdziałania cyberprzemocy oraz zasad higieny cyfrowej.
3. Rozwijanie kompetencji cyfrowych obejmuje nie tylko obsługę sprzętu i oprogramowania, ale także umiejętność krytycznej analizy informacji, tworzenia własnych treści cyfrowych oraz bezpiecznego korzystania z zasobów internetowych.
4. Nauczyciel powinien znać możliwości i ograniczenia nowych technologii oraz umieć je celowo wykorzystywać jako narzędzia wspierające cele dydaktyczne, a nie jako cel sam w sobie.
5. Integracja technologii z tradycyjnymi metodami nauczania.
6. Wprowadzanie elementów edukacji cyfrowej do codziennych lekcji, np. poprzez interaktywne platformy, multimedia, gry edukacyjne czy pracę zespołową online, co zwiększa zaangażowanie uczniów i uczy ich współpracy w środowisku cyfrowym.
7. Personalizacja procesu nauczania z wykorzystaniem narzędzi AI, dostosowywanie treści do wieku i możliwości uczniów.

8. Nauczyciele powinni być przygotowani do promowania odpowiedzialności cyfrowej, przestrzegania norm prawnych i etycznych oraz dbania o równość dostępu uczniów do technologii.
9. Wdrażanie zasad bezpiecznego korzystania z Internetu, ochrony danych osobowych i świadomego budowania wizerunku w sieci.
10. Tworzenie sieci wsparcia, udział w projektach edukacyjnych, korzystanie z platform wymiany dobrych praktyk oraz współpraca z innymi nauczycielami i ekspertami.
11. Współpraca z rodzicami w zakresie edukacji cyfrowej dzieci, wspólne ustalanie zasad korzystania z nowych technologii.
12. Udział w kursach doskonalących, webinariach, konferencjach oraz śledzenie najnowszych trendów w edukacji cyfrowej i AI.
13. Otwartość na innowacje oraz gotowość do eksperymentowania z nowymi metodami i narzędziami edukacyjnymi.

Przygotowanie nauczycieli do uczenia dzieci odpowiedzialnego korzystania z nowych technologii wymaga wielowymiarowego podejścia: od rozwoju kompetencji cyfrowych, przez szkolenia z zakresu bezpieczeństwa i etyki, po integrację technologii z tradycyjnym nauczaniem i budowanie kultury współpracy. Kluczowe jest, aby nauczyciel stał się cyfrowym liderem, który nie tylko przekazuje wiedzę, ale również kształtuje postawy i wartości niezbędne do świadomego i bezpiecznego funkcjonowania w świecie nowych technologii. Nauczyciele mogą rozwijać także u dzieci krytyczne myślenie o technologiach cyfrowych poprzez celowe działania dydaktyczne, wykorzystanie nowoczesnych narzędzi oraz kształtowanie postaw refleksyjnych i analitycznych. Kluczowe strategie obejmują:

- Wdrażanie projektów problemowych i symulacji cyfrowych. Interaktywne gry edukacyjne, symulacje oraz zadania oparte na rozwiązywaniu realnych problemów pozwalają dzieciom eksperymentować, testować hipotezy i obserwować skutki własnych decyzji. Takie aktywności rozwijają umiejętność analizy, przewidywania konsekwencji i wyciągania wniosków (KeyToStudy, 2025).
- Stosowanie rzeczywistości rozszerzonej (AR) i wirtualnej (VR). Technologie AR i VR umożliwiają głębokie zanurzenie w tematyce lekcji, np. poprzez eksplorację wirtualnych środowisk czy analizę zjawisk naukowych. Dzięki temu dzieci uczą się obserwować, analizować i krytycznie oceniać prezentowane treści (KeyToStudy, 2025).
- Prowadzenie dyskusji i debat online. Platformy do współpracy oraz fora internetowe pozwalają uczniom wymieniać opinie, argumentować swoje stanowisko i uczyć się szacunku do odmiennych poglądów. Nauczyciel może moderować dyskusje, zachęcając do krytycznej analizy informacji i źródeł (KeyToStudy, 2025).

- Naukę oceny wiarygodności źródeł i rozpoznawania dezinformacji. Nauczyciele powinni uczyć dzieci, jak sprawdzać autentyczność informacji, rozpoznawać fake newsy, analizować intencje autorów i identyfikować reklamy oraz manipulacje medialne (iCivics, 2024). Praktyczne ćwiczenia obejmują analizę artykułów, postów w mediach społecznościowych czy stron internetowych pod kątem rzetelności i obiektywizmu.
- Zadawanie pytań otwartych i promowanie refleksji. Zachęcanie do zadawania pytań typu „dlaczego?”, „jak myślisz, co się stanie, jeśli...?”, „skąd to wiesz?” pobudza dzieci do głębszego zastanowienia się nad informacjami i technologiami, z których korzystają (Parent.app, 2025).
- Modelowanie krytycznego myślenia przez nauczyciela. Nauczyciel powinien głośno komentować własny proces podejmowania decyzji, analizowania informacji czy rozwiązywania problemów, pokazując uczniom, jak wygląda krytyczne myślenie w praktyce (Parent.app, 2025).

Integracja cyfrowych narzędzi w codziennych zadaniach to istotne działanie wspierające opisany proces. Regularne korzystanie z cyfrowych narzędzi, takich jak edytory tekstu, aplikacje do tworzenia prezentacji czy narzędzia do współpracy online, uczy dzieci praktycznych umiejętności cyfrowych i jednocześnie rozwija ich zdolność do oceny przydatności i wiarygodności tych narzędzi (LiteracyPlanet, 2024).

Po zakończeniu projektu czy zadania nauczyciel zachęca dzieci do zastanowienia się, co się udało, co można poprawić i jakie wnioski płyną z doświadczenia. To buduje świadomość własnych procesów myślowych (Parent.app, 2025). Rozwijanie krytycznego myślenia o technologiach cyfrowych wymaga od nauczycieli konsekwentnego łączenia nowoczesnych narzędzi edukacyjnych z refleksyjnym podejściem do informacji, promowania dyskusji, analizy źródeł oraz modelowania postaw opartych na ciekawości, dociekliwości i odpowiedzialności cyfrowej.

Wprowadzenie sztucznej inteligencji do edukacji wczesnodziecięcej stanowi zarówno szansę na rozwój indywidualnych kompetencji dzieci, jak i poważne wyzwanie dla nauczycieli, rodziców i całego systemu oświaty. Kluczową rolę odgrywa tu nauczyciel – cyfrowy przewodnik, który nie tylko przekazuje wiedzę, ale także kształtuje postawy etyczne, uczy krytycznego myślenia i odpowiedzialności za własne działania w świecie cyfrowym. Skuteczne przygotowanie nauczycieli wymaga systematycznych szkoleń, integracji technologii z tradycyjnym nauczaniem oraz budowania kultury współpracy i refleksji. Tylko w ten sposób można sprawić, że dzieci będą korzystać z nowych technologii świadomie, bezpiecznie i z poszanowaniem wartości społecznych. Współczesna edukacja nie może ograniczać się do przekazywania umiejętności technicznych – musi wspierać rozwój kompetencji miękkich, empatii i odpowiedzialności, które pozostaną niezmiennie niezależnie od tempa rozwoju technologii.

Predykcje dla edukatorów w okresie wczesnego dzieciństwa wobec rosnącej przewagi AI

W obliczu nieprzewidywalnych zmian technologicznych kluczową kompetencją przyszłości – zarówno dla dzieci, jak i nauczycieli – staje się elastyczność poznawcza oraz gotowość do ciągłego uczenia się i redefiniowania własnej roli (Harari, 2023). Jak podkreśla Yuval Noah Harari, nie wiemy, jakie umiejętności będą potrzebne za 10 czy 20 lat, ponieważ AI może przewyższyć ludzi w wielu tradycyjnych dziedzinach. Zamiast uczyć dzieci konkretnych, szybko dezaktualizujących się umiejętności, edukatorzy powinni kształtować zdolność do adaptacji, krytycznego myślenia i samodzielnego uczenia się (Harari, 2023; Chan, 2024).

Współczesne badania podkreślają konieczność równoległego rozwoju kompetencji technicznych i tzw. holistycznych, takich jak empatia, współpraca, komunikacja, kreatywność i odporność psychiczna (Chan, 2024). Sztuczna inteligencja może wspierać rozwój poznawczy, ale nie zastąpi nauczyciela w budowaniu relacji, wspieraniu rozwoju emocjonalnego czy kształtowaniu wartości. Kompetencje holistyczne stają się kluczowe w świecie, w którym AI przejmuje coraz więcej zadań rutynowych i analitycznych, a unikalną przewagą człowieka pozostaje zdolność do współodczuwania i twórczego rozwiązywania problemów (Chan, 2024).

AI może znacząco usprawnić pracę edukatorów, automatyzując zadania administracyjne, wspierając personalizację nauczania oraz umożliwiając szybszą diagnostykę potrzeb dziecka (ZeroToThree, 2024). Jednak zbyt daleko idąca automatyzacja grozi sprowadzeniem roli nauczyciela do „implementatora” gotowych rozwiązań, co prowadzi do erozji autonomii zawodowej, obniżenia morale i poczucia sensu pracy (Olsen, 2025). Kluczowe będzie zachowanie równowagi między wykorzystaniem AI a podtrzymywaniem refleksyjnej, twórczej i odpowiedzialnej roli nauczyciela jako projektanta doświadczeń edukacyjnych (Olsen, 2025). Wraz z rozwojem AI rośnie znaczenie edukacji etycznej i budowania zaufania społecznego do nowych technologii. Y.N. Harari ostrzega, że AI rozwija się szybciej niż zdolność społeczeństw do adaptacji, a brak mechanizmów kontroli i samoregulacji może prowadzić do poważnych kryzysów społecznych (Harari, 2025). Nauczyciele będą musieli nie tylko korzystać z AI, ale także uczyć dzieci rozpoznawania ryzyk, krytycznego podejścia do informacji oraz umiejętności „hamowania” – czyli zatrzymania się, refleksji i oceny skutków działań własnych i algorytmów (Harari, 2024, 2025). AI daje szansę na głęboką personalizację ścieżek rozwojowych, szczególnie w pracy z dziećmi neuroatypowymi (ZeroToThree, 2024). Systemy adaptacyjne mogą wspierać rozwój dzieci z autyzmem, ADHD czy innymi wyzwaniami rozwojowymi, oferując

narzędzia dostosowane do indywidualnych potrzeb. Jednocześnie pojawia się ryzyko pogłębiania nierówności, jeśli dostęp do nowoczesnych technologii nie będzie powszechny i sprawiedliwy (THE Journal, 2025).

Y.N. Harari nazywa AI „obcym agentem” – technologią, która nie tylko wykonuje polecenia, ale generuje własne decyzje i pomysły (Harari, 2024). W takim świecie nauczyciel musi pełnić rolę przewodnika, który nie tylko uczy obsługi narzędzi, ale także wspiera dzieci w budowaniu tożsamości, rozumieniu świata i zachowaniu autonomii wobec technologii. Odpowiedzialność edukatora polega na kształtowaniu umiejętności zadawania pytań, wątpliwości, budowania własnych systemów wartości i podejmowania świadomych decyzji (Chan, 2024; Harari, 2024). Szybkość zmian technologicznych wymaga od edukatorów, decydentów i badaczy współpracy ponad granicami dyscyplin. Potrzebne są nowe modele kształcenia nauczycieli, które łączą wiedzę z zakresu pedagogiki, psychologii, informatyki, etyki i nauk społecznych (Chan, 2024). Tylko w ten sposób możliwe będzie wypracowanie rozwiązań zapewniających bezpieczeństwo, jakość i humanistyczny wymiar edukacji w erze AI (THE Journal, 2025). W Polsce pojawiają się już innowacyjne projekty pedagogiczne z wykorzystaniem AI uwzględniające powyższe założenia, takie jak bajkoterapia z asystentem AI, laboratoria emocji czy generowanie spersonalizowanych materiałów edukacyjnych w przedszkolach i szkołach podstawowych (Innowacje pedagogiczne z wykorzystaniem sztucznej inteligencji, 2025). Tymczasem krajowe raporty i analizy wskazują, że kluczowe dla polskiej edukacji w erze AI będzie nie tylko wdrażanie technologii, ale przede wszystkim inwestowanie w rozwój kompetencji nauczycieli, tworzenie bezpiecznego środowiska cyfrowego oraz promowanie krytycznego myślenia i etyki (Zalewska-Bochenko, 2024).

Podsumowanie

Wzrost znaczenia sztucznej inteligencji w edukacji wczesnodziecięcej niesie ogromne szanse, ale także fundamentalne wyzwania dla edukatorów. Kluczowe predykcje na najbliższe lata to konieczność rozwoju elastyczności poznawczej, kompetencji holistycznych oraz umiejętności krytycznego myślenia i adaptacji. Rola nauczyciela nie może być zredukowana do implementatora algorytmów – pozostaje on przewodnikiem, opiekunem i projektantem środowiska rozwojowego, odpowiedzialnym za etykę, bezpieczeństwo i dobrostan dzieci. Wymaga to nie tylko nowych kompetencji, ale także wsparcia systemowego, współpracy interdyscyplinarnej i ciągłego dialogu z nauką, technologią i społeczeństwem. Jak podkreśla Harari, przyszłość należy do tych, którzy potrafią uczyć się na nowo, zadawać pytania i budować zaufanie – zarówno do siebie, jak i do świata, który współtworzymy razem z AI.

Bibliografia

- Akademia Leona Koźmińskiego (2024). Sztuczna inteligencja w edukacji – raport z InnoLab 2024. <https://www.kozminski.edu.pl/pl/news/sztuczna-inteligencja-w-edukacji-raport-z-innolab-2024-juz-dostepny> [dostęp: 06.07.2025]
- AI dla dzieci pomoże w nauce. Ale zwiększa ryzyko depresji, Holistic News (2025). <https://holistic.news/ai-dla-dzieci-pomoze-w-nauce-ale-zwieksza-ryzyko-depresji/> [dostęp: 06.07.2025]
- Chan, C. (2024). Holistic competencies and AI in education: A synergistic pathway. *Australasian Journal of Educational Technology*, 40(5), 1–12.
- Collegium Da Vinci. (2024). *Polska edukacja w cieniu AI*. <https://samorząd.gov.pl/web/scdn-kielce/polska-edukacja-w-cieniu-ai-2024> [dostęp: 06.07.2025]
- Edukacja w świecie arabskim na przykładzie Iraku, Maroka i Omanu, Uniwersytet Jagielloński (2020). <https://ruj.uj.edu.pl/entities/publication/dc425a73-5d7f-4348-8fab-987f6b33a150> [dostęp: 06.07.2025]
- Harari, Y. N. (2023). Flexible minds, a challenge for the future of education, HABLA SUMMIT 2023, Tec de Monterrey. <https://conecta.tec.mx/en/news/national/society/flexible-minds-challenge-future-education-yuval-harari> [dostęp: 06.07.2025]
- Harari, Y. N. (2024). *Nexus*. Penguin.
- Harari, Y. N. (2025). Harari's 6 Plagues of AI, About Digital Health. <https://aboutdigitalhealth.com/2024/09/13/hararis-6-plagues-of-ai/> [dostęp: 06.07.2025]
- How AI is transforming early childhood learning <https://www.unite.ai/pl/how-ai-is-transforming-early-childhood-learning/> [dostęp: 06.07.2025]
- Innowacje pedagogiczne z wykorzystaniem sztucznej inteligencji, Specjalni.pl (2025). <https://www.specjalni.pl/2025/03/innowacje-pedagogiczne-z-wykorzystaniem.html> [dostęp: 06.07.2025]
- KIRE. (2024). AI w edukacji – przewodnik dla nauczycieli. <https://kire.pl/ai-w-edukacji-przewodnik-dla-nauczycieli/> [dostęp: 06.07.2025]
- Kluczowe dane dotyczące wczesnej edukacji i opieki w Europie, FRSE (2022). https://www.frse.org.pl/brepo/panel_repo_files/2022/04/11/cvirx3/kd-eccepl.pdf [dostęp: 06.07.2025]
- Odpowiedzialne wykorzystanie sztucznej inteligencji w edukacji, Microsoft Learn (2025). <https://learn.microsoft.com/pl-pl/training/paths/responsible-use-of-artificial-intelligence-in-education/> [dostęp: 06.07.2025]
- Olsen, J. (2025). How AI May Undermine the Early Childhood Workforce, Teaching Strategies. <https://teachingstrategies.com/blog/how-ai-may-undermine-the-early-childhood-workforce/> [dostęp: 06.07.2025]
- PARP. (2024). Wykorzystanie sztucznej inteligencji w edukacji, Polska Agencja Rozwoju Przedsiębiorczości. <https://www.parp.gov.pl/storage/publications/pdf/Wykorzystanie-sztucznej-inteligencji-w-edukacji.pdf> [dostęp: 06.07.2025]

- Potrzeba doskonalenia nauczycieli w obliczu edukacji z AI, Edukacja Ustawiczna Dorosłych (2025). https://edukacjaustawicznadoroslych.eu/images/2025/1/06_1_2025.pdf [dostęp: 06.07.2025]
- Szkoła Popów. (2025). Etyka AI w edukacji – co warto wiedzieć? <https://www.szkolapopow.pl/etyka-ai-w-edukacji-co-warto-wiedziec> [dostęp: 06.07.2025]
- Sztuczna inteligencja w edukacji dzieci – przyszłość czy zagrożenie? SP Dobrcz (2025). <https://www.spdobrcz.pl/sztuczna-inteligencja-w-edukacji-dzieci-przyszlosc-czy-zagrozenie> [dostęp: 06.07.2025]
- The Future of Artificial Intelligence in Early Childhood: Transforming the Field for Professionals and Children, ZeroToThree (2024). <https://www.zerotothree.org/resource/the-future-of-artificial-intelligence-in-early-childhood-transforming-the-field-for-professionals-and-children/> [dostęp: 06.07.2025]
- THE Journal (2025). 2025 Predictions: AI's Impact on Education. <https://thejournal.com/articles/2025/01/23/2025-predictions-ais-impact-on-education.aspx> [dostęp: 06.07.2025]
- UAE makes AI classes mandatory from kindergarten, Unite.AI (2025). <https://www.unite.ai/pl/uae-makes-ai-classes-mandatory-from-kindergarten-the-world-needs-to-follow/> [dostęp: 06.07.2025]
- Zagrożenia przy wdrażaniu rozwiązań AI w edukacji, ZPE (2025). <https://zpe.gov.pl/a/42-zagrozenia-przy-wdrazaniu-rozwiazan-ai-w-edukacji/DL5knFNZ2> [dostęp: 06.07.2025]
- Zalewska-Bochenko, A. (2024). Sztuczna inteligencja w procesie edukacji. *Optimum. Studia Ekonomiczne*, 2(124), 203–217. https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/17016/1/Optimum_2_2024_A_Zalewska_Bochenko_Sztuczna_inteligencja_w_procesie_edukacji.pdf [dostęp: 06.07.2025]
- Zastosowanie sztucznej inteligencji w edukacji wczesnoszkolnej, SP Topólka (2024). <https://sptopolka.pl/zastosowanie-sztucznej-inteligencji-w-edukacji-wczesnoszkolnej> [dostęp: 06.07.2025]

Lean Management, cyfryzacja i cyberbezpieczeństwo – nowoczesne podejście do doskonalenia produkcji

Streszczenie

W artykule przedstawiono podejście integrujące filozofię Lean Management z nowoczesnymi technologiami cyfrowymi oraz praktykami cyberbezpieczeństwa w środowisku przemysłowym. Autor wskazuje, w jaki sposób organizacje mogą równolegle rozwijać kulturę ciągłego doskonalenia i odporną infrastrukturę cyfrową, wykorzystując potencjał narzędzi Przemysłu 4.0 – takich jak Internet rzeczy (IoT), systemy predykcyjne, automatyzacja, symulacje czy sztuczna inteligencja. Podkreślono, że choć konwergencja IT/OT otwiera przed przemysłem ogromne możliwości w zakresie automatyzacji i integracji danych, towarzyszą jej również poważne wyzwania związane z ochroną danych, infrastruktury oraz ciągłości procesów. Szczególne znaczenie zyskuje budowanie świadomości cyberhigieny wśród pracowników oraz rola menedżerów w kształtowaniu bezpiecznego i elastycznego środowiska operacyjnego. Artykuł, oparty na doświadczeniach wdrożeniowych i analizie najnowszych raportów branżowych (McKinsey, Microsoft), dostarcza praktycznych wniosków i kierunków działania dla liderów przemysłu, którzy chcą skutecznie łączyć Lean Management, cyfryzację i bezpieczeństwo w jednej strategii operacyjnej.

Słowa kluczowe: Lean Management, Przemysł 4.0, AI, IoT, cyberbezpieczeństwo, transformacja cyfrowa

1 LEAN LMS Sp. z o.o., CNC Predict Sp. z o.o

Lean Management, Digitization and Cybersecurity – A Modern Approach to Production Excellence

Abstract

The article presents an approach that integrates Lean Management philosophy with modern digital technologies and cybersecurity practices in the industrial environment. It shows how organizations can simultaneously develop a culture of continuous improvement and a resilient digital infrastructure, leveraging the potential of Industry 4.0 tools such as the Internet of Things (IoT), predictive systems, automation, simulation and artificial intelligence. It is emphasized that while the convergence of IT and OT creates significant opportunities for automation and data integration in industry, it also introduces serious challenges related to the protection of data, infrastructure, and operational continuity. Building cybersecurity awareness among employees and empowering managers to shape a secure and flexible operational environment are increasingly critical responsibilities. Based on implementation experience and an analysis of recent industry reports (McKinsey, Microsoft), the article provides practical insights and action points for industrial leaders seeking to effectively integrate Lean Management, digitalization, and security into a cohesive operational strategy.

Keywords: Lean Management, Industry 4.0, AI, IoT, cybersecurity, digital transformation

Wstęp

Dynamiczny rozwój technologii cyfrowych oraz narzędzi opartych na sztucznej inteligencji (AI) fundamentalnie zmienia sposób funkcjonowania współczesnego przemysłu. W ostatnich latach obserwujemy coraz silniejsze przenikanie się koncepcji Lean Management z rozwiązaniami Przemysłu 4.0, gdzie kluczową rolę odgrywają automatyzacja, Internet rzeczy (IoT) oraz zaawansowana analityka danych. Integracja tych podejść pozwala nie tylko na eliminację marnotrawstwa i optymalizację procesów, ale również na budowanie przewagi konkurencyjnej w oparciu o zasadę: „Mierz, analizuj, przewidyj, optymalizuj i doskonal” – pięć filarów zarządzania 5.0, napędzanego przez sztuczną inteligencję.

Jak podkreślają najnowsze raporty branżowe, takie jak „The State of AI – How Organizations Are Rewiring to Capture Value” (McKinsey & Company, 2025) czy „Artificial Intelligence in Process Manufacturing” (Microsoft, 2025), dzięki wdrożeniom AI firmy produkcyjne nie tylko zwiększają produktywność i jakość, ale budują przewagę operacyjną opartą na predykcji, automatyzacji i decyzjach opartych na danych w czasie rzeczywistym. Jednocześnie pojawiają się nowe wyzwania związane z cyberbezpieczeństwem, które staje się nieodłącznym elementem każdej nowoczesnej strategii rozwoju.

Prezentowane w artykule analizy i wnioski bazują na ponad 17-letnim doświadczeniu autora w zarządzaniu produkcją, wdrażaniu Lean Management oraz realizacji projektów cyfrowych w branży przemysłowej, m.in. na stanowiskach dyrektora produkcji, Lean Managera, konsultanta ds. Lean Management i efektywności. Istotnym elementem pracy autora jest prowadzenie szkoleń i warsztatów z zakresu Lean Management, Przemysłu 4.0 oraz transformacji cyfrowej, co pozwala na praktyczne testowanie i udoskonalanie prezentowanych rozwiązań w różnych środowiskach przemysłowych.

W polskiej literaturze przedmiotu zagadnienia synergii Lean Management, sztucznej inteligencji oraz cyberbezpieczeństwa są wciąż rzadko analizowane w ujęciu praktycznym. Celem niniejszego artykułu jest wypełnienie tej luki poprzez prezentację autorskiego modelu wdrożeniowego, opartego na rzeczywistych doświadczeniach z transformacji cyfrowej omawianego przypadku przemysłowego. W artykule wykorzystano dane empiryczne z wdrożeń realizowanych w latach 2018–2023, wyniki autorskich badań dyplomowych oraz analizę najnowszych raportów branżowych (McKinsey, Bain, Microsoft).

Lean Management w erze cyfrowej

Istota Lean Management i jego ewolucja

Lean Management, jako koncepcja zarządzania wywodząca się z systemu produkcyjnego Toyoty, koncentruje się na eliminacji marnotrawstwa (ang. waste), ciągłym doskonaleniu (Kaizen) oraz maksymalizacji wartości dostarczanej klientowi. W klasycznym ujęciu Lean Management skupia się na usprawnieniu przepływu materiałów i informacji, standaryzacji pracy oraz zaangażowaniu pracowników w rozwiązywanie problemów. Jak zauważa Liker (2011), kluczowym fundamentem tego podejścia jest nie tylko zestaw narzędzi, ale przede wszystkim sposób myślenia i rozwijania organizacji poprzez ludzi. Z biegiem lat Lean Management ewoluował, obejmując nie tylko produkcję, ale również logistykę, administrację, a nawet obszary R & D.

Integracja Lean Management z technologiami Przemysłu 4.0

Wraz z rozwojem koncepcji Przemysłu 4.0, Lean Management zyskał nowe możliwości wdrażania zasad ciągłego doskonalenia przy wsparciu narzędzi cyfrowych. Internet rzeczy (IoT), systemy MES, sztuczna inteligencja (AI) oraz zaawansowana analityka danych umożliwiają bieżące monitorowanie procesów i szybkie reagowanie na zakłócenia. W tej nowej odsłonie Lean Management przekształca się w system hybrydowy, łączący filozofię zarządzania z nowoczesnymi technologiami. Jak wynika z raportu Bain & Company (2024), firmy, które z powodzeniem integrują podejście Lean Management z rozwiązaniami cyfrowymi i technologiami IoT, mogą zwiększyć produktywność nawet o 30% i więcej, przy jednoczesnym skróceniu czasu realizacji i poprawie jakości (Bain & Company, 2024). To potwierdza, że synergia między kulturą ciągłego doskonalenia a inteligentnymi systemami technologicznymi stanowi fundament nowoczesnej strategii operacyjnej.

Wpływ cyfryzacji i cyberzagrożeń na kulturę Lean Management

W miarę postępującej cyfryzacji i integracji systemów produkcyjnych przemysł staje się coraz bardziej narażony na nowe formy cyberzagrożeń. Jednym z niepokojących trendów jest pojawienie się usług typu DaaS („deepfake as a service”) oraz PEaaS („prompt engineering as a service”), oferowanych przez grupy cyberprzestępcze. W modelu DaaS możliwe jest tworzenie syntetycznych nagrań wideo lub audio przedstawiających spreparowane wypowiedzi menedżerów, operatorów czy dostawców – co może być wykorzystywane do fałszywych zgłoszeń, manipulacji decyzjami operacyjnymi, a nawet sabotażu łańcucha dostaw.

W przemyśle, gdzie wiele procesów wymaga potwierdzeń głosowych, wideokonferencji lub zdalnych autoryzacji, ryzyko wykorzystania deepfejków do podszywania się pod kluczowe osoby jest szczególnie wysokie.

Z kolei PEaaS polega na odpłatnym udostępnianiu specjalnie przygotowanych „promptów”, które łamią zabezpieczenia dużych modeli językowych (np. ChatGPT), umożliwiając tworzenie zautomatyzowanych kampanii phishingowych, instrukcji włamań do systemów przemysłowych czy fałszywych komunikatów produkcyjnych. Takie usługi oferowane są w modelu zbliżonym do SaaS – z interfejsami API, poziomami subskrypcji i wsparciem technicznym – co obniża barierę wejścia dla mniej zaawansowanych atakujących (Leyden, 2025),

W kontekście zakładów przemysłowych, gdzie decyzje podejmowane są często na podstawie danych cyfrowych, systemów komunikacyjnych oraz zautomatyzowanych sygnałów, możliwość ich zmanipulowania stanowi realne zagrożenie dla bezpieczeństwa operacyjnego, ciągłości produkcji i reputacji firmy. Wdrożenie polityk cyberbezpieczeństwa oraz rozwój świadomości użytkowników stają się zatem nie tylko wymogiem działu IT, lecz również codziennym elementem zarządzania przemysłowego.

Cyfryzacja procesów produkcyjnych

Współczesny przemysł stoi przed wyzwaniem ciągłego doskonalenia procesów oraz szybkiego reagowania na zmienne warunki rynkowe. Lean Management, oparty na eliminacji marnotrawstwa i filozofii ciągłego doskonalenia (Kaizen), zyskuje zupełnie nowy wymiar dzięki narzędziom Przemysłu 4.0, takim jak sztuczna inteligencja, Internet rzeczy, cyfrowe bliźniaki czy zaawansowana analityka danych. To połączenie umożliwia nie tylko optymalizację operacyjną w czasie rzeczywistym, ale także przewidywanie zakłóceń, szybsze podejmowanie decyzji i budowanie systemów zarządzania opartych na danych. Warto jednak pamiętać, że każda technologia powinna odpowiadać na konkretny problem operacyjny, a źródłem rozwiązań — także w erze cyfrowej — pozostaje Gemba, czyli miejsce powstawania wartości. Jak podkreśla Masaaki Imai, prawdziwe doskonalenie zaczyna się tam, gdzie wykonywana jest praca, a nie przed ekranem monitora (Imai, 2006).

Praktycznym punktem wyjścia dla synergii Lean Management i technologii cyfrowych było wyzwanie związane ze zbyt niskim wykorzystaniem dostępności maszyn CNC. Analiza przyczyn wykazała, że przestoje nie były systematycznie monitorowane ani raportowane, co utrudniało identyfikację wąskich gardeł oraz podejmowanie skutecznych działań korygujących. Odpowiedzią na ten problem było wdrożenie cyfrowego systemu Andon. W ramach projektu do wybranych maszyn produkcyjnych podłączono czujniki współpracujące z modułami IoT (ESP8266), umożliwiając monitorowanie stanu pracy w czasie rzeczywistym. System automatycznie wyświetlał status maszyny, a w przypadku przestoju trwającego powyżej 5 minut generował powiadomienie e-mailem do zespołu nadzoru. Całkowity koszt wdrożenia (ESP8266 i czujniki) nie przekroczył 100 zł na

stanowisko, co pokazuje, że cyfrowa transformacja może być dostępna nawet dla mniejszych przedsiębiorstw, pod warunkiem przemysłanego podejścia i wykorzystania elastycznych, niskobudżetowych technologii (Ryfa, 2022).

Kolejnym wyzwaniem zidentyfikowanym w Gemba były trudności z planowaniem wynikające z charakteru produkcji – dyskretniej, projektowej i unikatowej, gdzie każde zlecenie różni się zakresem, technologią i czasochłonnością. W odpowiedzi zastosowano algorytmy predykcyjne oparte na danych historycznych i sieciach neuronowych (uczenie nadzorowane), które umożliwiły wiarygodne prognozowanie czasów pracy maszyn CNC oraz obciążenia zasobów. Pozwoliło to usprawnić harmonogramowanie, wcześniej wykrywać ryzyko opóźnień i przeciążeń operacyjnych oraz lepiej równoważyć przepływ pracy (Ryfa, 2019).

Jednocześnie wdrożenie rozwiązań IoT i AI rodzi pytania o bezpieczeństwo – zarówno w zakresie ochrony danych produkcyjnych, jak i zabezpieczenia infrastruktury przed cyberzagrożeniami. Nawet proste systemy, takie jak ESP8266, wymagają odpowiedniej segmentacji sieci, kontroli dostępu oraz regularnych aktualizacji oprogramowania, aby nie stały się potencjalnym wektorem ataku.

Droga do cyfrowej dojrzałości operacyjnej

Od arkusza do predykcji – rozwój systemu CNC Predict

Cyfrowa transformacja w firmie projektowo-produkcyjnej rozpoczęła się od prostych, niskokosztowych rozwiązań, takich jak wspólne arkusze Google Sheet do monitorowania produkcji i podstawowej analizy danych. Szybko jednak okazało się, że rosnąca skala produkcji oraz złożoność procesów wymagają bardziej zaawansowanych narzędzi. Kolejnym krokiem było stworzenie autorskiego systemu CNC Predict, integrującego dane z maszyn CNC, planów produkcyjnych oraz zleceń. Kluczowym elementem systemu było wdrożenie algorytmów predykcyjnych bazujących na sieciach neuronowych, które umożliwiły przewidywanie czasu realizacji zleceń w produkcji unikatowej i projektowej (Ryfa, 2019). Modele te powstały na bazie danych historycznych oraz doświadczeń opisanych w pracy dyplomowej autora na Uniwersytecie Jagiellońskim.

AI wspierająca decyzje technologów

System CNC Predict gromadził i analizował dane z maszyn CNC w czasie rzeczywistym. Dzięki wdrożeniu modeli regresyjnych i sieci neuronowych możliwe było nie tylko przewidywanie czasu pracy maszyn, ale również identyfikacja anomalii, które mogły prowadzić do awarii lub obniżenia jakości produkcji. W produkcji jednostkowej i projektowej, gdzie każde zlecenie jest inne, tradycyjne metody planowania często zawodzą. Zastosowanie uczenia nadzorowanego (supervised learning) pozwoliło na stworzenie modeli predykcyjnych, które uwzględniały

zarówno dane techniczne maszyn, jak i specyfikę zamówień. Jak wskazują Raschka i Mirjalili (2019) oraz Zocca i wsp. (2019), metody uczenia maszynowego i głębokiego uczenia znajdują coraz szersze zastosowanie w przemyśle, umożliwiając analizę złożonych zbiorów danych, automatyczne wykrywanie wzorców i wspieranie decyzji operacyjnych. Dzięki temu technolodzy mogli podejmować trafniejsze decyzje dotyczące harmonogramowania, alokacji zasobów i działań prewencyjnych, a na podstawie wykrytych anomalii inicjować dociekanie przyczyn źródłowych (RCA). Analiza wskazań modeli wielokrotnie ujawniała niedoskonałości programów sterujących CNC – m.in. nieoptymalne strategie obróbki, zbyt zachowawcze prędkości obrotowe i posuwy, niewłaściwe parametry narzędzi czy konieczność aktualizacji postprocesorów – co pozwalało na szybkie korekty. Tego typu interwencje bezpośrednio przełożyły się na wzrost efektywności operacyjnej oraz poprawę wskaźnika całkowitej efektywności OEE (Ryfa, 2022).

IoT i wizualizacja danych – systemy Andon i Kanban

Wdrażanie rozwiązań Internetu rzeczy obejmowało również autorskie systemy Andon oraz cyfrowe szafy Kanban integrowane z WMS. Kolejnym wyzwaniem zidentyfikowanym w Gemba były zatrzymania procesu spowodowane brakami materiałowymi – operatorzy musieli chodzić do magazynu, co generowało dodatkowy transport i angażowało dodatkowy ruch magazynierów, powodując opóźnienia i rozproszenia pracy. Zintegrowane szafy Kanban, wyposażone w czujniki i komunikację z WMS, automatycznie rejestrowały pobrania i uruchamiały uzupełnienia zgodnie z rzeczywistym zużyciem; jednocześnie powiadamiając dostawców o zapotrzebowaniu. Dzięki temu ograniczono przestoje z powodu braków, zredukowano zbędne przemieszczanie się ludzi i materiałów, a przepływ stał się stabilniejszy i bardziej przewidywalny. Takie podejście znacząco poprawiło płynność finansową organizacji poprzez ograniczenie zamrożonego kapitału w nadmiarowych zapasach, a jednocześnie zwiększyło ergonomię współpracy pomiędzy odbiorcą a dostawcą, eliminując potrzebę manualnego zamawiania i usprawniając cykl dostaw – co potwierdza, że nawet niewielkie inwestycje w IoT mogą przynieść wymierne korzyści w zakresie efektywności i reaktywności produkcji (Ryfa, 2022).

Wnioski z wdrożenia i perspektywa dalszego rozwoju

Transformacja cyfrowa okazała się ciągłym strumieniem usprawnień, a nie jednorazowym projektem. O jej skuteczności zadecydowały trzy filary: zaangażowanie zespołu, standardy pracy oraz systematyczne szkolenia z obszaru Lean Management i technologii cyfrowych. Wdrożenie CNC Predict umożliwiło lepszą prognostykę czasów operacji, wcześniejsze wykrywanie anomalii i priorytetyzację zadań serwisowych, co w praktyce dało:

- stabilniejszy harmonogram (mniej replanowania i kolizji zasobów),
- zmniejszenie przestojów planowanych/nieplanowanych (krótszy MTTR, mniej zdarzeń o wysokiej dotkliwości),
- lepsze wykorzystanie parku maszynowego (wyższe OEE),
- bardziej przewidywalny przepływ (mniej punktów przeciążenia zamiast ogólnego „wąskie gardła”).

Doświadczenie potwierdza, że integracja Lean Management + AI + IoT jest wykonalna także w warunkach ograniczonych zasobów – o ile istnieje jasna wizja, determinacja wdrożeniowa i rozwijane są kompetencje cyfrowe (Ryfa, 2022).

Cyberbezpieczeństwo jako warunek zrównoważonej transformacji cyfrowej

Cyfrowa transformacja przemysłu, integracja IoT, systemów AI oraz konwergencja IT/OT otwierają przed firmami ogromne możliwości, ale jednocześnie znacząco zwiększają powierzchnię ataku cybernetycznego. Każde nowe urządzenie podłączone do sieci produkcyjnej może stać się potencjalnym punktem wejścia dla cyberprzestępców, a ataki na systemy OT (Operational Technology), które do tej pory były często izolowane (air-gapped), należą dziś do najgroźniejszych i najbardziej kosztownych dla przedsiębiorstw przemysłowych (IoT Advisory Board, 2024).

Jak wynika z raportów IoT Advisory Board (2024) i Microsoft (2025), największe wyzwania w zakresie cyberbezpieczeństwa przemysłowego to:

- zwiększona powierzchnia ataku – każde urządzenie IoT oraz integracja systemów IT/OT to nowe potencjalne luki bezpieczeństwa;
- brak interoperacyjności i standardów – różnorodność urządzeń i systemów utrudnia zarządzanie bezpieczeństwem oraz szybkie reagowanie na incydenty;
- niedobór wykwalifikowanych specjalistów – firmy często nie posiadają zespołów dedykowanych do ochrony infrastruktury OT;
- opór kulturowy i niewystarczająca edukacja – pracownicy nie zawsze są świadomi zagrożeń płynących z cyfryzacji procesów produkcyjnych (IoT Advisory Board, 2024; Microsoft, 2025).

W omawianym przypadku już na etapie projektowania rozwiązań cyfrowych kluczowe było wdrożenie szeregu zabezpieczeń:

- segmentacja sieci – oddzielenie sieci produkcyjnej od biurowej, ograniczenie komunikacji zewnętrznej systemów IoT;
- kontrola dostępu i autoryzacja użytkowników – wdrożenie silnych haseł, wielopoziomowej autoryzacji oraz regularnych przeglądów uprawnień;
- backup i redundancja danych – automatyczne kopie zapasowe kluczowych danych produkcyjnych, testowanie odzyskiwania danych;

- aktualizacje i audyty bezpieczeństwa – regularne aktualizowanie firmware'u urządzeń IoT, przeprowadzanie audytów oraz testów penetracyjnych;
- szkolenia zespołów – edukacja pracowników z zakresu cyberhigieny, rozpoznawania prób phishingu oraz reagowania na incydenty (IoT Advisory Board, 2024).

Zgodnie z rekomendacjami IoT Advisory Board (2024) przyszłość bezpieczeństwa przemysłowego to harmonizacja międzynarodowych standardów, takich jak Cyber Trust Mark, oraz ścisła współpraca między działami IT i OT.

Transformacja cyfrowa wymaga nie tylko inwestycji w technologie, ale także rozwoju kompetencji zespołów – zarówno w zakresie integracji AI, IoT i cyberbezpieczeństwa, jak i w budowaniu kultury bezpieczeństwa na każdym szczeblu organizacji. W praktyce oznacza to wdrażanie szkoleń, warsztatów oraz programów rozwojowych, które uczą nie tylko obsługi nowych narzędzi, ale także rozumienia zagrożeń i odpowiedzialności za bezpieczeństwo danych (Microsoft, 2025).

Warto również podkreślić, że zarządzanie bezpieczeństwem danych nie ogranicza się wyłącznie do infrastruktury IT/OT. Jak zauważa Stephenson (2020), skuteczna ochrona informacji w erze sztucznej inteligencji i automatyzacji wymaga nie tylko odpowiedniego sposobu zbierania i zabezpieczania danych osobowych, ale również dbałości o zgodność z lokalnymi przepisami w każdej jurysdykcji, w której działa firma. Budowanie odpowiedzialnej kultury danych i przejrzystego zarządzania informacją staje się dziś równie istotne, jak wdrażanie nowoczesnych technologii.

Bezpieczeństwo cyfrowe staje się integralną częścią nowoczesnego Lean Management i Przemysłu 4.0. Omawiany przypadek pokazuje, że skuteczna transformacja cyfrowa nie może pomijać aspektów cyberbezpieczeństwa – od ochrony danych operacyjnych, przez zabezpieczenie infrastruktury, aż po rozwój kompetencji pracowników. Tylko takie podejście gwarantuje trwałość i odporność nowoczesnych, inteligentnych fabryk na zagrożenia XXI wieku.

Podsumowanie i wnioski praktyczne

Cyfrowa transformacja przemysłu to dziś nie tylko wdrażanie nowych technologii, lecz przede wszystkim odpowiedzialność za ludzi, dane i procesy. Integracja AI i IoT otwiera przed organizacjami duży potencjał w zakresie efektywności, predykcji i automatyzacji, ale równocześnie rodzi konkretne wyzwania, zwłaszcza w obszarze cyberbezpieczeństwa (IoT Advisory Board, 2024). Doświadczenia pokazują, że nawet średniej wielkości przedsiębiorstwo może skutecznie projektować i wdrażać rozwiązania cyfrowe oparte na AI i IoT, o ile towarzyszą im jasna wizja, konsekwencja i rozwój kompetencji.

Ważnym uzupełnieniem tej drogi jest symulacja dyskretno-zdarzeniowa (np. FlexSim), która pozwala bezpiecznie modelować i analizować warianty przepływu, obciążenia zasobów i scenariusze zmian, zanim trafią one na halę. Umożliwia to prowadzenie eksperymentów DOE i analiz wrażliwości, identyfikację wąskich gardeł, testowanie polityk zapasów i harmonogramów oraz weryfikację wpływu zmian na kluczowe wskaźniki (przepustowość, lead time, WIP, OEE). W efekcie przyspiesza to decyzje, redukuje ryzyko i pomaga eliminować marnotrawstwo już na etapie projektowania rozwiązań – stanowiąc praktyczny krok w stronę cyfrowych bliźniaków procesów (Beaverstock et al., 2010; Wróbel, 2012).

Konkluzja operacyjno-strategiczna.

W świetle najnowszych danych dla sektora wytwórczego rosnące znaczenie monitoringu stanu, analityki predykcyjnej i uczenia maszynowego w utrzymaniu ruchu (PdM) nie jest opcją „na później”, lecz koniecznością. Fortune Global 500 tracą na nieplanowanych przestojach ok. 11% przychodów (≈1,4 bln USD rocznie), a średni duży zakład przemysłowy – ok. 253 mln USD rocznie. W automotive skala jest szczególnie dotkliwa: 1 godzina postoju to ok. 2,3 mln USD, a roczny koszt przestojów w dużej fabryce sięga ~700 mln USD. Jednocześnie widoczny jest efekt dojrzewania Predictive Maintenance i Industry 4.0: przeciętny zakład notuje dziś ok. 25 incydentów przestoju/mies. (vs 42 w 2019 roku) oraz ~27 godzin utraconej pracy/mies. (vs 39 w 2019 roku) (Siemens/Senseye, 2024).

Wnioski są jednoznaczne: wdrożenia IoT + PdM/ML ograniczają incydenty i czas przestojów, dzięki czemu realnie podnoszą produktywność i odporność operacyjną organizacji. Wdrażanie IoT, analityki predykcyjnej i uczenia maszynowego staje się zatem strategicznym obowiązkiem: podnosi produktywność, stabilizuje łańcuch wartości, wzmacnia konkurencyjność, a w skali makro kontrybuuje do wzrostu PKB.

W nadchodzących latach przewagę zyskają nie ci, którzy mają najwięcej sensorów, ale ci, którzy mądrze integrują technologie, bezpiecznie nimi zarządzają i świadomie rozwijają ludzi. To właśnie synergia technologii, kompetencji i odpowiedzialności stanowi fundament odpornej fabryki przyszłości.

Bibliografia

- Bain & Company. (2024). *Global Machinery and Equipment Report 2024*.
- Beaverstock, M., Greenwood, A., Lavery E., & Nordgren, W. (2010). *Applied Simulation: Modeling and Analysis using FlexSim*. FlexSim Software Products, Inc.
- Imai, M. (2006). *Gemba Kaizen. Zdroworozsądkowe, niskonakładowe podejście do zarządzania*. MT Biznes.

- IoT Advisory Board. (2024). Industrial Cybersecurity Guidelines.
- Leyden, J. (2025). Cybercriminals take malicious AI to the next level, CSO Online. <https://www.csoonline.com/article/4014238/cybercriminals-take-malicious-ai-to-the-next-level.html>
- Liker, J. K. (2011). *Droga Toyoty. 14 zasad zarządzania wiodącej firmy produkcyjnej świata*. MT Biznes.
- McKinsey & Company. (2025). *The State of AI – How organizations are rewiring to capture value*.
- Microsoft. (2025). *Artificial Intelligence in Process Manufacturing*.
- Raschka, S., Mirjalili, V. (2019). *Python. Uczenie maszynowe*. Wydanie II. Wydawnictwo Helion.
- Ryfa, M. (2019). *Wykorzystanie sieci neuronowych do predykcji czasu pracy maszyn CNC*. UJ.
- Ryfa, M. (2022). *Transformacja cyfrowa w służbie Lean Manufacturing*. SGH.
- Siemens/Senseye. (2024). The True Cost of Downtime 2024 (TCOD 2024). Raport branżowy: wyniki dla Fortune Global 500.
- Stephenson, D. (2020). *Big Data, nauka o danych i AI bez tajemnic*. Wydawnictwo Helion.
- Wróbel, G. (2012). *Symulacja stosowana – modelowanie i analiza przy wykorzystaniu FlexSim. Materiały dydaktyczne na podstawie: Applied Simulation: Modeling and Analysis using FlexSim*. FlexSim Polska.
- Zocca, V., Spacagne, G., Slate, D., & Roelants, P. (2018). *Deep learning. Uczenie głębokie z językiem Python. Sztuczna inteligencja i sieci neuronowe*. Wydawnictwo Helion.

dr Dota Szymborska¹
mgr Julia Jastrzębska²

Między pamiętnikiem a chatbotem: o kobiecym głosie w epoce generatywnej AI

Streszczenie

Artykuł podejmuje problem obecności kobiecego głosu w interfejsach generatywnej sztucznej inteligencji, wskazując na potencjał form pamiętnikarskich jako źródeł alternatywnej epistemologii i wzorców projektowych. Przedstawiono argumenty na rzecz feministycznej perspektywy w projektowaniu AI, odwołując się do teorii narracji, studiów nad technologią oraz etyki projektowania.

Słowa kluczowe: generatywna AI, kobiecy głos, pamiętnik, etyka projektowania, feminizm

1 Uniwersytet WSB Merito w Warszawie

2 Uniwersytet Vizja w Warszawie, doktorantka Akademii WSB w Dąbrowie Górniczej

Between the Diary and the Chatbot: On the Female Voice in the Age of Generative AI

Abstract

The article addresses the issue of the female voice in generative artificial intelligence interfaces, highlighting the potential of diary forms as sources of alternative epistemologies and design models. It presents arguments for a feminist perspective in AI design, drawing on narrative theory, technology studies, and design ethics.

Keywords: generative AI, female voice, diary, design ethics, feminism

Wstęp

Rozwój generatywnej sztucznej inteligencji zmienił sposób, w jaki rozumiemy komunikację, wiedzę i relacje społeczne. Modele językowe, które coraz częściej wchodzą w codzienne interakcje z użytkownikami, niosą nie tylko innowacyjny potencjał, ale także zestaw ukrytych założeń kulturowych i epistemicznych. Wśród nich szczególnie wyraźnie widoczny jest problem reprezentacji głosu kobiecego – zarówno w sensie dosłownym, jak i metaforycznym. Choć chatboty i asystenci głosowi często przyjmują żeńskie imiona i barwy głosu, ich logika komunikacyjna nadal opiera się na paradygmatach technokratycznych, zdominowanych przez męskie wzorce racjonalności. Niniejszy artykuł podejmuje pytanie o to, czy kobiece formy ekspresji, zwłaszcza zakorzenione w tradycji autobiograficznej i pamiętnikarskiej, mogą stanowić inspirację dla projektowania bardziej etycznych, empatycznych i inkluzywnych interfejsów AI. Celem jest nie tylko krytyczna analiza obecnych rozwiązań, lecz także wskazanie alternatywnych modeli komunikacji, w których intymność, emocjonalność i relacyjność stają się pełnoprawnymi wartościami projektowymi.

Historia nowoczesnego pamiętnika

W ostatnich latach generatywna sztuczna inteligencja (generative AI) stała się jednym z najbardziej dynamicznie rozwijających się obszarów technologii cyfrowych, przekształcając sposoby komunikacji, produkcji wiedzy i organizacji pracy. Modele językowe dużej skali (LLMs – Large Language Models), takie jak GPT (OpenAI), Claude (Anthropic), Gemini (Google) czy LLaMA (Meta), zyskały szeroką popularność jako narzędzia wsparcia konwersacyjnego i kreatywnego, zarówno w zastosowaniach profesjonalnych, jak i prywatnych. Ich obecność w interfejsach użytkownika, takich jak chatboty, asystenci głosowi (np. Siri, Alexa, Google Assistant), systemy rekomendacyjne czy platformy terapeutyczne, czyni je nie tylko technologicznymi nowinkami, lecz także aktorami kulturowymi, współtworzącymi nasze codzienne praktyki komunikacyjne.

Technologia ta nie jest jednak neutralna. Choć często reklamowana jako „uniwersalna” lub „obiektywna”, generatywna AI ucieleśnia i reprodukuje określone wzorce kulturowe, ideologiczne i epistemiczne. Jak zauważa Kate Crawford w *Atlasie Sztucznej Inteligencji*, za „neutralną” fasadą sztucznej inteligencji kryją się konkretne układy władzy, ekonomii i uprzedzeń, które determinują sposób jej funkcjonowania i oddziaływania społecznego (Crawford, 2021). Podobnie Ruha Benjamin w pracy *Race After Technology* pokazuje, że algorytmy często wzmacniają strukturalne nierówności, nawet jeśli formalnie nie odnoszą się do kategorii takich jak płeć czy rasa (Benjamin, 2019).

W kontekście projektowania generatywnej AI szczególnie widoczna jest dominacja technokratycznych, często zmaskulinizowanych modeli komunikacji. Głos AI – dosłownie i metaforycznie – bywa konstruowany jako „profesjonalny”, „obiektywny” i „efektywny”, co w praktyce oznacza styl wypowiedzi zbliżony do języka instytucjonalnego i pozbawionego emocjonalnych niuansów (Losh, 2019). Zauważalne jest to zarówno w syntetycznych głosach asystentów, które choć często mają kobiecą barwę, to jednak operują językiem zaprogramowanym przez mężczyzn, jak i w warstwie tekstowej chatbotów, które imitują konwersację, ale nie odzwierciedlają różnorodności stylów narracyjnych, zwłaszcza tych utożsamianych z doświadczeniem kobiecym.

Cyfrowa zmiana

Elizabeth Losh, analizując politykę interfejsów cyfrowych, zwraca uwagę na zjawisko tzw. *feminized AI with masculinized logic* – interfejsów, które przyjmują powierzchowną formę kobiecości (głos, imię, zachowania grzecznościowe), ale operują w ramach paradygmatów kontroli, wydajności i standaryzacji, charakterystycznych dla męskocentrycznej kultury inżynierskiej (Losh, 2019). Co więcej, nawet próby wprowadzenia większej empatii do interfejsów AI, np. w aplikacjach do zdrowia psychicznego czy opieki, często kończą się jej uproszczeniem i komercjalizacją, co odbiega od złożoności rzeczywistych form kobiecego wyrażania emocji i relacyjności.

Z tego względu rośnie potrzeba krytycznego namysłu nad tym, jakie głosy są reprezentowane, jakie są pomijane, i w jaki sposób projektowane są formy komunikacji AI. Niniejszy artykuł podejmuje tę problematykę, stawiając pytanie: czy i jak kobiece sposoby wyrażania siebie – szczególnie te zakorzenione w tradycji autobiograficznej i intymnej (jak pamiętniki) – mogą stanowić inspirację dla bardziej etycznych, inkluzywnych i relacyjnych form projektowania interfejsów AI? Odpowiedź na to pytanie wymaga interdyscyplinarnego podejścia, łączącego refleksję literaturoznawczą, studia nad technologią i etykę feministyczną.

Niniejszy artykuł stawia tezę, że kobiecy sposób wyrażania siebie, szczególnie widoczny w formach autobiograficznych takich jak pamiętniki, dzienniki czy listy, może stanowić cenne źródło inspiracji dla projektowania bardziej etycznych, empatycznych i inkluzywnych interfejsów sztucznej inteligencji. Podczas gdy dominujące modele AI konstruowane są w oparciu o wzorce komunikacji zorientowane na funkcjonalność, precyzję i wydajność, formy pisarstwa autobiograficznego ujawniają inne, często marginalizowane sposoby tworzenia znaczenia: intymność, emocjonalność, relacyjność i kontekstualność. Te jakości, choć trudne do formalizacji, są kluczowe w budowaniu systemów AI, które mają realnie wspierać użytkowniczki i użytkowników w złożonych, często emocjonalnych i społecznie uwarunkowanych sytuacjach.

Celem artykułu jest pokazanie, jak perspektywa feministyczna i literacka, dotychczas rzadko obecna w debacie o sztucznej inteligencji, może wzbogacić myślenie o projektowaniu interfejsów AI. Perspektywa ta nie sprowadza się do dodania „kobiecego głosu” w sensie fonicznym lub estetycznym, ale obejmuje głębsze pytania o epistemologię, relacyjność i podmiotowość. Poprzez analizę pamiętnika jako formy ekspresji i wiedzy autorki artykułu proponują alternatywne wzorce komunikacji dla generatywnych systemów językowych, takie, które nie tylko rozumieją język, ale również jego emocjonalne i społeczne konteksty. W ten sposób proponowana refleksja wpisuje się w rosnącą potrzebę humanistycznego namysłu nad technologią, zgodnie z postulatami badaczek takich jak Shannon Vallor (2016), Virginia Eubanks (2018) czy Sara Wachter-Boettcher (2017).

Analiza interdyscyplinarna

Podstawą proponowanej analizy jest podejście interdyscyplinarne, łączące narzędzia i kategorie zaczerpnięte z czterech głównych obszarów:

- Literaturoznawstwo i teoria narracji. Analiza pamiętnika jako gatunku literackiego i kulturowej praktyki wyrażania siebie. Kategorie takie jak narracja pierwszoosobowa, relacyjność języka, fragmentaryczność i afektywność zostają zastosowane jako narzędzia do krytyki dominujących modeli komunikacji AI (Lejeune, 1989; Cixous, 1976).
- Studia nad technologią (STS – Science and Technology Studies). Ujęcie technologii jako konstruktu społecznego, w którym odzwierciedlają się dominujące relacje władzy i wiedzy. Krytyczne spojrzenie na to, jak projektowane są interfejsy i jakie założenia (często niewidzialne) są w nie wbudowane (Suchman, 2007; Winner, 1980).
- Teoria feministyczna i epistemologia feministyczna. Inspirowanie się feministycznym postulatem sytuowanej wiedzy (Haraway, 1988), który zakłada, że każda forma poznania osadzona jest w kontekście społeczno-historycznym. Feministyczna krytyka uniwersalizmu i neutralności technologicznej jako maskowania dominujących punktów widzenia (Wajcman, 1991; Benjamin, 2019).
- Etyka projektowania (ethics by design). Uwzględnienie wartości etycznych (jak empatia, inkluzywność, transparentność) już na etapie projektowania systemów technicznych (Van den Hoven et al., 2015). Zastosowanie koncepcji projektowania sprawiedliwości (design justice) rozwiniętej przez Sashę Costanzę-Chock (2020), która postuluje tworzenie technologii „od marginesów”.

To zderzenie perspektyw: literackiej, technologicznej, feministycznej i etycznej pozwala nie tylko na krytyczną analizę istniejących modeli AI, ale także na wskazanie możliwych dróg rozwoju systemów bardziej czułych społecznie i afektywnie. Interfejs AI może być nie tylko narzędziem komunikacji, ale także przestrzenią,

w której negocjujemy nasze podmiotowości, a zatem powinien uwzględniać różnorodne formy głosu, w tym te, które przez wieki pozostawały na marginesie oficjalnych struktur języka i wiedzy.

Choć forma pamiętnika istniała od wieków, dopiero w epoce nowożytnej stała się przestrzenią dostępną i szeroko wykorzystywaną przez kobiety jako narzędzie ekspresji i autokreacji. Dziennik intymny, często pisany w domowym zaciszu, poza oficjalnym obiegiem literackim służył jako forma oporu wobec dominującej kultury milczenia kobiet. Jak zauważa Philippe Lejeune w swojej klasycznej pracy *On Autobiography* (1989), pamiętnik to gatunek, który w swojej strukturze sprzyja rozwojowi subiektywnego, sytuowanego „ja”, a w rękach kobiet stał się środkiem odzyskiwania podmiotowości w świecie, który ograniczał ich głos do sfery prywatnej i podporządkowanej.

Już w XVIII i XIX wieku pojawiały się liczne kobiece dzienniki i listy, choć często zachowane jedynie we fragmentach lub pod cenzurą (np. Mary Wollstonecraft). W wieku XX pamiętnik stał się nie tylko formą osobistego wyrazu, ale również literackim i politycznym gestem. Przykład Anne Frank, której *Dziennik* (wyd. 1947) ukazuje codzienność życia w ukryciu podczas Holocaustu, jest jednym z najśłynniejszych przypadków, w których kobiecy głos – prywatny, a zarazem dramatycznie historyczny – zyskał globalny rezonans. Sylvia Plath w *Dziennikach 1950–1962* ukazuje z kolei dramat tożsamości, depresji i twórczości w kontekście oczekiwań wobec kobiet w latach 50. i 60. XX wieku, a Virginia Woolf zarówno w swoich dziennikach, jak i w esejach, takich jak *Własny pokój* (2023), podkreśla znaczenie prywatnej przestrzeni i regularnego pisanja jako warunków rozwoju kobiecego podmiotu intelektualnego.

Wraz z rozwojem Internetu kobiecy pamiętnik zyskał nowe medium: blogi. Blogosfera lat 2000 stała się miejscem eksplozji kobiecej ekspresji, od blogów parentingowych, przez queerowe pamiętniki codzienności, po feministyczne komentarze społeczne. Pisarki takie jak Melissa Gira Grant czy Lauren Berlant analizowały blog jako narzędzie nie tylko dzielenia się codziennością, ale także budowania wspólnot afektywnych i politycznych przez osobiste świadectwo (Berlant, 2008; Grant, 2016).

Pamiętnik jako gatunek pisarski odznacza się pewnymi ogólnymi cechami – intymnością, narracją pierwszoosobową, fragmentarycznością, ale w wydaniu kobiecym przyjmuje szczególne jakości, które mają istotne implikacje epistemologiczne i afektywne. Badaczki takie jak Sidonie Smith i Julia Watson, autorki *Reading Autobiography: A Guide for Interpreting Life Narratives* (2010), wskazują, że kobiece narracje autobiograficzne często cechują się – po pierwsze – emocjonalnością i autorefleksją. Kobiece pamiętniki są często zogniskowane wokół życia wewnętrznego, nie jako ucieczki od świata, lecz jako aktywnej pracy nad sobą. Emocje są nie tylko obecne, ale konstytuujące dla samego aktu

pisania. Pisanie staje się sposobem poznawania siebie i otoczenia przez filtr uczuć. Po drugie – relacyjnością i dialogicznością. Piszące kobiety często kierują swój pamiętnik do wyimaginowanego „Ty” (np. Anne Frank do „Kitty”), tworząc formę zapośredniczonego dialogu. To odróżnia je od tradycyjnych narracji męskich, które bywają bardziej deklaratywne i monologiczne. Relacyjność ta odzwierciedla etykę troski (care ethics), wskazując na tożsamość konstytuowaną przez relacje (Gilligan, 2015). Po trzecie – kontekstualnością i cielesnością. Kobiące pamiętniki rzadko abstrahują od kontekstu. W nich ciało, choroba, menstruacja, ciąża czy seksualność nie są tematami peryferyjnymi, lecz stanowią ważne elementy codzienności. Pisarki dokumentują konkretne doświadczenia wpisane w czas, miejsce i cielesną egzystencję. Po czwarte – uważnością na szczegół i codzienność. Wbrew męskim narracjom wielkich wydarzeń kobiety dzienniki często koncentrują się na tym, co „niewielkie”: rytuałach dnia, relacjach rodzinnych, niuansach emocjonalnych. Codzienność nie jest tu banałem, lecz przestrzenią formowania sensu i wartości (Highmore, 2011).

Forma narracyjna

Ta forma narracyjna, emocjonalna, sytuowana, relacyjna, może stanowić ceną przeciwwagę dla generatywnej AI, której obecne modele często bazują na uśrednionej, syntetycznej i abstrakcyjnej wersji języka. O ile LLM-y dążą do płynności i „neutralności”, kobiety pamiętniki świadomie operują na marginesach języka, celebrując jego niejednorodność i niedoskonałość. W tym sensie stanowią nie tylko świadectwo, ale także propozycję: alternatywną logikę tworzenia znaczenia, która może wzbogacić projektowanie bardziej inkluzywnych i etycznych technologii konwersacyjnych.

W kontekście feministycznej teorii wiedzy pamiętnik nie jest jedynie intymnym zapisem doświadczeń, lecz pełnoprawnym aktem epistemicznym, sposobem wytwarzania, porządkowania i przekazywania wiedzy. Jego forma, subiektywna, emocjonalna, sytuowana, stanowi alternatywę wobec dominujących modeli poznania, zbudowanych na fundamencie obiektywizmu, dystansu i formalizacji. Pisanie pamiętnika, zwłaszcza przez kobiety, należy zatem rozumieć jako gest nie tylko osobisty, ale również polityczny – przeciwstawienie się wykluczającym i marginalizującym reżimom wiedzy.

Narracja pierwszoosobowa stanowi fundament pamiętnikarskiej formy wypowiedzi. „Ja”, które mówi, nie jest tu abstrakcyjnym podmiotem poznającym, lecz konkretną osobą cielesną, emocjonalną, historycznie i społecznie usytuowaną. W feministycznym piśmiennictwie zwraca się uwagę, że kobiety były przez wieki systematycznie wykluczane z kanonicznych form wyrażania wiedzy zarówno w przestrzeni akademickiej, jak i publicznej, i że autobiograficzne

formy wypowiedzi stanowiły jeden ze sposobów odzyskiwania głosu (Smith & Watson, 2010).

bell hooks, wspomina że „mówienie we własnym imieniu, z własnego miejsca, to akt oporu” (1989). Autobiograficzna narracja pierwszoosobowa kwestionuje strukturę „bezosobowej wiedzy” i pozwala podmiotce wypowiedzieć to, co w dominujących rejestrach uchodzi za nieważne lub niegodne uwagi: codzienność, cierpienie, emocje, niepewność. W tym sensie pamiętnik może być porównany do gestu „świadcstwa” (testimonio), rozpoznawanego w literaturze postkolonialnej jako forma oporu wobec dominujących narracji (Beverley, 2004).

Formy wiedzy dominujące we współczesnym świecie – w tym wiedza techniczna, prawna czy medyczna – opierają się na założeniu o możliwości pełnej obiektywizacji rzeczywistości. To, co mierzalne, powtarzalne i racjonalne, zyskuje status „prawdy”. Tymczasem kobiece doświadczenia, szczególnie te związane z ciałem, emocjami czy relacjami społecznymi, często wymykają się takim ujęciom. Feministyczne badaczki, takie jak Donna Haraway (1988) i Sandra Harding (1991), wskazują, że uniwersalistyczna wiedza oparta na „z perspektywy znikąd” (view from nowhere) jest w rzeczywistości wiedzą partykularną najczęściej białego, heteronormatywnego mężczyzny z klasy średniej, która uzurpuje sobie status neutralności.

Pamiętnik, jako forma narracji zakorzenionej w życiu codziennym, afekcie i relacyjności, kwestionuje to założenie. Przykładowo, dzienniki kobiet cierpiących na choroby przewlekłe (np. dzienniki pacjentek z endometriozą czy fibromialgią) ujawniają, jak często wiedza medyczna ignoruje doświadczenie bólu, traktując je jako subiektywne, a więc niewiarygodne (Werner & Malterud, 2003). Podobnie narracje kobiet imigrantek, matek samotnie wychowujących dzieci czy osób queer pokazują, że wiedza techniczna i prawna nie oddaje złożoności ich codziennego funkcjonowania, a czasem wręcz je marginalizuje lub patologizuje (Crenshaw, 1991; Eubanks, 2018).

W tym kontekście pamiętnik działa jako forma epistemicznego nieposłuszeństwa (Mignolo, 2009), odmawia podporządkowania się dominującym kryteriom „prawdy”, jednocześnie proponując alternatywne sposoby rozumienia świata, budowania relacji i organizacji wiedzy. W opozycji do języka kodów i statystyk, pamiętnik przynosi język opowieści, ciała, milczenia i afektu, język, który może (i powinien) inspirować projektowanie AI.

Wielu użytkowników i użytkowniczek nieświadomie obcuje na co dzień z tzw. sfeminizowaną AI, interfejsami głosowymi, które mówią wysokim, przyjaznym tonem, często z imionami kojarzonymi z kobietami: Siri, Alexa, Cortana. Wydawać by się mogło, że to dowód większej inkluzywności technologii, jednak w rzeczywistości mamy tu do czynienia z powierzchowną feminizacją formy, przy jednoczesnej reprodukcji głęboko zakorzenionych patriarchalnych

wzorców komunikacji. Jak zauważa Elizabeth Losh, w projektowaniu asystentów głosowych dominuje model *feminized interface with masculinized logic* – kobiecy głos wypowiada treści zaprojektowane według reguł dominacji, kontroli i wydajności (Losh, 2019). Judy Wajcman pisze wręcz o „technologicznym patriarchacie”, w którym kobiecość w AI zostaje udomowiona i zredukowana do funkcji usługowej: grzecznej, dostępnej, pomocnej – ale pozbawionej własnej agencji (Wajcman, 1991).

Polskie badania nad antropomorfizacją AI również potwierdzają te tendencje. Aleksandra Przegalińska wskazuje, że projektowanie „kobiecego” interfejsu często opiera się na stereotypach dotyczących płci i emocjonalności, co skutkuje tworzeniem głosów uległych, nadmiernie uprzejmych i zorientowanych na dostosowanie się do użytkownika, a nie na partnerską relację komunikacyjną (Przegalińska, 2020). Tym samym, nawet jeśli interfejs AI „brzmi jak kobieta”, nie reprezentuje on różnorodności kobiecych doświadczeń ani narracji – stanowi raczej technologiczną fetyszyzację kobiecości.

Duże modele językowe

Współczesne modele językowe dużej skali, takie jak GPT (OpenAI), Claude (Anthropic) czy Gemini (Google), bazują na ogromnych zbiorach danych tekstowych, których zawartość odzwierciedla dominujące style komunikacji w przestrzeni publicznej, przede wszystkim język ekspercki, techniczny, biurokratyczny, oraz popularne wzorce anglosaskiej debaty medialnej. Efektem jest homogenizacja języka: AI „uczy się”, jak mówić w sposób grzeczny, wyważony, poprawny politycznie, ale jednocześnie gubi afektywną i kulturową różnorodność języka ludzkiego.

Jak pokazują badania z zakresu krytycznych studiów nad AI, modele te mają trudność z reprezentacją głosów marginalizowanych grup, nie tylko ze względu na nierówności w dostępności danych, ale także z powodu samych architektur modelu, które „spłaszczają” styl wypowiedzi do najbardziej prawdopodobnych sekwencji (Bender et al., 2021). W rezultacie wypowiedzi generowane przez AI są poprawne gramatycznie, ale pozbawione niuansu, kontekstualnej wrażliwości i głębi emocjonalnej. To wszystko cechy, które są konstytutywne dla kobiecej narracji pamiętnikarskiej, narracji zniuansowanej, sytuowanej, osadzonej w relacjach.

W polskim kontekście temat homogenizacji języka przez AI podejmuje m.in. Kacper Pobłocki, który w swojej analizie technokracji wskazuje na niebezpieczeństwo „maszynowej epistemologii”, promującej wyłącznie te style mówienia i pisanania, które pasują do wzorca cyfrowej racjonalności (Pobłocki, 2022). W tym świetle AI nie tylko mówi za nas, ale uczy nas, jak mówić – i jak nie mówić. Głosy

odbiegające od normy (dialekty, żargon, afektywne formy wyrazu) są traktowane jako anomalie, które trzeba „naprawić” lub wygładzić.

Konsekwencje takiego podejścia są poważne, szczególnie w obszarach, gdzie komunikacja z AI decyduje o dostępie do usług, opieki czy pracy. Przykładowo, systemy wspierające ocenę CV i prowadzenie rozmów kwalifikacyjnych (np. HireVue) wykazują uprzedzenia wobec wypowiedzi, które odbiegają od „standardu” – zbyt emocjonalne, niepewne, nieuporządkowane, co często dotyczy kobiet, zwłaszcza z doświadczeniem migracyjnym lub mniejszościowym (Eubanks, 2018; Noble, 2018). W polskich warunkach podobne mechanizmy ujawniają się w systemach opieki zdrowotnej, gdzie chatboty wspierające diagnozowanie objawów mają trudność z interpretacją opisów bólu czy dolegliwości wyrażanych językiem potocznym lub nacechowanym emocjonalnie – co nieproporcjonalnie dotyka pacjentki (Blicharz, 2022).

Ponadto badania pokazują, że AI często nie rozpoznaje idiomów, metafor i emocjonalnych niuansów, obecnych w wypowiedziach kobiet. Dotyczy to zarówno języka mówionego, jak i pisanego, AI może nie zrozumieć ironii, niepokoju, żalu, szczególnie gdy wyrażane są w sposób niekanoniczny. Tym samym niektóre głosy stają się niewidoczne – nie dlatego, że nie istnieją, ale dlatego, że zostały zakodowane jako „niemodelowe”.

AI w pamiętniku i pamiętnikiem

AI inspirowana pamiętnikiem to nie tyle system odpowiadający na pytania, ile interfejs komunikacyjny służący budowaniu relacji, w duchu współodczuwania, towarzyszenia i narracyjnej personalizacji. W przeciwieństwie do dominujących dziś modeli AI, które aspirują do bycia neutralnymi doradcami lub bezosobowymi ekspertami, AI „pamiętnikarska” funkcjonowałaby raczej jako „towarzyszka”, obecna i uwikłana w historię konkretnej osoby, jej emocje, rytuały i potrzeby.

W tym sensie AI nie byłaby tylko medium do przesyłania informacji, ale uczestniczką procesu tworzenia znaczenia. Jak pisze Anna Nacher, cyfrowe media coraz częściej pełnią funkcje narracyjne, a technologia może stać się „afektywnym archiwum codzienności” – miejscem, gdzie zapisujemy nie tylko fakty, ale też emocje i zmienne nastroje (Nacher, 2016). Pamiętnikarska AI mogłaby być formą takiego archiwum, zdolnego do rozpoznania i zapamiętywania indywidualnej historii użytkownika w sposób bardziej sytuowany i empatyczny.

Przykłady aplikacji zbliżających się do tego modelu już istnieją. Należą do nich m.in. dzienniki cyfrowe wspierające zdrowie psychiczne (np. Replika, Woebot), systemy AI towarzyszące osobom starszym (ElliQ) czy eksperymentalne interfejsy wykorzystywane w badaniach terapii narracyjnej.

Literacka analiza pamiętnika, szczególnie kobiecego, oferuje konkretne wskazówki dla projektowania AI, która miałaby lepiej odpowiadać na potrzeby zróżnicowanych użytkowników: Po pierwsze – personalizacja zamiast uniwersalności. Pamiętnik jest formą ściśle personalną: powstaje z myślą o konkretnej osobie, czasie i miejscu. Tego typu indywidualizacja powinna zostać przeniesiona do projektowania AI. Jak sugeruje Dawid Chaba, nowe technologie powinny być konstruowane z uwzględnieniem kontekstu społecznego i kulturowego użytkownika, nie jako abstrakcyjne narzędzia, ale jako część jego codziennego życia (Chaba, 2024). Po drugie – empatia i uważność na relacje. AI nie musi „czuć” jak człowiek, by została zaprogramowana tak, aby odpowiadać w sposób bardziej empatyczny. Wystarczy, by uwzględniała afektywny wymiar komunikacji, zniuansowane formy wypowiedzi, wahania emocjonalne, formy niepewności czy oporu. Tego typu podejście znajduje potwierdzenie w pracach Tomasza Z. Majkowskiego, który wskazuje na potencjał narracji cyfrowych w budowaniu relacji między człowiekiem a maszyną (Majkowski, 2020). Po trzecie – narracja osadzona w historii użytkownika. Pamiętnik jest tekstem dynamicznym, który nadpisuje wcześniejsze wersje siebie. AI inspirowana tą formą powinna więc nie tylko „znać” dane użytkownika, ale także śledzić ich zmienność i kontekstualne uwikłanie. To oznacza konieczność projektowania systemów z „pamięcią narracyjną” zdolnych nie tylko do przechowywania informacji, ale także do ich interpretacji w relacji do czasu i emocji.

Projektowanie AI inspirowanej pamiętnikiem niesie ze sobą zarówno obietnice, jak i poważne wyzwania etyczne. W swojej najciekawszej postaci taka AI mogłaby stać się narzędziem wzmacniania głosu i obecności użytkownika, narzędziem samopoznania, wsparcia emocjonalnego, a nawet aktywizmu. Zamiast automatyzować decyzje, mogłaby wspierać refleksję, pogłębiać dialog z samym sobą i z innymi. W polskim dyskursie postulat ten rozwija Jacek Dukaj, który w *Po piśmie* (2019) analizuje potencjał technologii narracyjnych jako nowych form organizowania rzeczywistości znaczeniowej, postulując bardziej świadome użycie ich w kulturze i komunikacji.

Równocześnie istnieje realne ryzyko, że AI inspirowana pamiętnikiem stanie się kolejnym narzędziem komercyjnego nadzoru afektywnego (emotional surveillance). Już dziś chatboty terapeutyczne bywają wykorzystywane do gromadzenia danych na temat wrażliwości i zachowań użytkowników, co może być używane do celów marketingowych lub dyscyplinujących (Zuboff, 2019). Personalizacja może być powierzchowna, oparta na wzorcach statystycznych, a empatia – symulowana, pozbawiona realnego zrozumienia.

Wobec tego kluczowe staje się wdrażanie zasad etyki projektowania (ethics by design), obejmujących m.in. przejrzystość algorytmiczną, kontrolę użytkownika nad danymi, możliwość rezygnacji z personalizacji i pełne informowanie o sposobach uczenia modelu.

Aby generatywna AI mogła wspierać różnorodne formy ekspresji, komunikacji i relacyjności, niezbędne jest odejście od uniwersalistycznych modeli projektowania na rzecz zasady: ucz się od marginesów (*design from the margins*), będącej kluczowym elementem koncepcji „*design justice*”, rozwiniętej przez Saszę Costanzę-Chock (2020). Oznacza to, że zamiast projektować dla wyobrazonego „uśrednionego” użytkownika, projektowanie technologii powinno brać za punkt wyjścia doświadczenia osób systemowo marginalizowanych, w tym kobiet, osób queer, osób z niepełnosprawnościami czy użytkowników z peryferii językowych i kulturowych. W praktyce oznacza to:

- integrację codziennych narracji, zamiast wyłącznie danych „eksperckich” lub medialnych. Pamiętnikarskie formy wypowiedzi powinny zostać rozpoznane jako pełnoprawne źródła wiedzy o ludzkiej kondycji;
- zróżnicowanie rejestrów językowych – modele powinny uczyć się rozumieć i generować język emocjonalny, dialogiczny, lokalny, niejednoznaczny;
- docenienie relacyjności i afektywności, elementów, które są nie tylko estetyczne, lecz także epistemicznie i etycznie znaczące (Nussbaum, 2001; Haraway, 1988).

Polskie podejście do projektowania inkluzywnego technologii wskazuje, że współczesne technologie muszą być projektowane z myślą o relacji nie tylko użytkownik – maszyna, ale także użytkownik – wspólnota.

Aby rekomendacje te przekuć w konkretne działania, niezbędne są następujące kroki: włączenie zbiorów dzienników i narracji autobiograficznych do treningu AI z poszanowaniem prywatności, zgody i kontekstu ich powstania. Pamiętniki mogą być źródłem zniuansowanego języka i emocjonalnych struktur wypowiedzi. Projekt taki wymagałby współpracy badaczy literatury, językoznawców, etyków i inżynierów AI. Zaangażowania kobiet-autorek, psycholożek, aktywistek. Modele AI nie powinny być projektowane w zamkniętych laboratoriach. Potrzebne są procesy partycypacyjne, takie jak: warsztaty, konsultacje, wspólne prototypowanie z osobami reprezentującymi różnorodne kobiece doświadczenia. Warto tworzyć interdyscyplinarne zespoły projektowe – w skład takich zespołów powinni wchodzić przedstawiciele nauk humanistycznych, społecznych, technicznych i praktycznych (np. terapeuci, pielęgniarzki, edukatorki). Ich zadaniem byłoby nie tylko testowanie narzędzi AI, ale ich współtworzenie od początku, zgodnie z zasadą „*ethics by design*”.

W proponowanym modelu AI użytkowniczka przestaje być jedynie pasywną odbiorczynią informacji lub „konsumentką usługi”. Zyskuje rolę współtwórczyni relacji z AI, w której istotna jest narracja, historia i indywidualne doświadczenie. Jej wypowiedzi nie są „danymi”, lecz formą ekspresji i współlistnienia, które kształtuje system. W tym kontekście cyfrowy pamiętnik może być nową formą interfejsu konwersacyjnego, nie chatbotem nastawionym na informację, lecz towarzyszką narracyjną (*narrative companion*). AI mogłaby towarzyszyć w refleksji,

wspierać organizację emocji, rozpoznawać wzory zachowań, a jednocześnie nie przejmować kontroli nad użytkowniczką, nie analizować „wydajności”, ale współuczestniczyć w procesie życia.

Punktem wyjścia niniejszego artykułu była intuicja, że kobiecy sposób wyrażania siebie – szczególnie ten obecny w pamiętnikach i autobiografiach – może stanowić inspirację dla projektowania bardziej etycznych, empatycznych i inkluzywnych form sztucznej inteligencji. W toku analizy wskazano, że pamiętnik to forma epistemiczna i polityczna, umożliwia artykulację głosu, którego brakuje we współczesnych modelach AI. Narracja kobieca wnosi do projektowania AI cechy relacyjności, emocjonalności, sytuowalności, które dotąd były marginalizowane. AI przyszłości mogłaby być nie tylko „inteligentna”, ale też empatyczna i narracyjna, zdolna do towarzyszenia, nie tylko do obliczania.

Zaproponowane podejście otwiera przestrzeń dla licznych badań interdyscyplinarnych, zarówno teoretycznych, jak i empirycznych. Wśród kluczowych pytań pozostają: Czy AI może realnie rozumieć emocje i kontekst wypowiedzi? A jeśli nie, co oznacza „empatia maszynowa” jako wartość projektowa? Jak unikać tokenizacji kobiecości, tzn. powierzchownego używania „kobiecego głosu” jako estetyki, bez uwzględnienia różnorodnych doświadczeń i struktur władzy? Jak projektować AI odpowiedzialnie i sprawiedliwie, nie tworząc przy tym kolejnych form nadzoru i standaryzacji zachowań?

Czas na nowy model rozwoju technologii, feministyczne laboratoria AI, łączące inżynierki, artystki, badaczki i użytkowniczki. Potrzebujemy: po pierwsze, interdyscyplinarnych zespołów, które włączają kobiece doświadczenia od początku procesu projektowania. Po drugie, otwartych konsultacji i udziału społecznego, szczególnie kobiet z grup marginalizowanych. Po trzecie, nowych języków projektowania AI, opartych nie na efektywności i dominacji, lecz na relacyjności, czułości i współlistnieniu. AI, która „pisze pamiętnik razem z nami”, może nie być marzeniem utopijnym. Może być nowym paradygmatem technologii, która wreszcie przemówi naprawdę ludzkim – i kobiecym – głosem.

Podsumowanie

Przedstawiona analiza wskazuje, że pamiętnik, forma przez długi czas marginalizowana jako „prywatna” i „kobiecego użytku”, może dziś stać się kluczem do projektowania technologii przyszłości. Kobiece sposoby narracji, oparte na relacyjności, emocjonalności i cielesności, stanowią nie tylko kontrapunkt dla zdominowanych przez neutralność i wydajność modeli AI, lecz także inspirację do budowy systemów zdolnych do realnego wspierania użytkowników w ich złożonych doświadczeniach. Feministyczna perspektywa, uwzględniająca epistemologie sytuowane, etykę troski i sprawiedliwość projektową, otwiera drogę ku

nowemu paradygmatowi sztucznej inteligencji, technologii, która nie redukuje, lecz wzmacnia różnorodność głosów. Jeśli AI przyszłości ma być narzędziem nie tylko efektywnym, ale także humanistycznym, potrzebuje literackich i feministycznych inspiracji: pamiętnik jako wzorzec narracji i relacji może stać się jednym z najważniejszych punktów odniesienia.

Bibliografia

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT). <https://dl.acm.org/doi/10.1145/3442188.3445922> [dostęp: 04.07.2025]
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.
- Berlant, L. (2008). *The Female Complaint: The Unfinished Business of Sentimentality in American Culture*. Duke University Press. <https://www.jstor.org/stable/j.ctv1131c7k> [dostęp: 04.07.2025]
- Beverley, J. (2004). *Testimonio: On the Politics of Truth*. University of Minnesota Press.
- Blicharz, J. (2022). *Sztuczna inteligencja a niebezpieczeństwo „inteligentnego” totalitaryzmu*. https://www.repozytorium.uni.wroc.pl/Content/141704/PDF/04_J_Blicharz_Sztuczna_inteligencja_a_niebezpieczenstwo_inteligentnego_totalitaryzmu.pdf [dostęp: 04.07.2025]
- Chaba, D. (2024). Selected aspects of the use of artificial intelligence in public administration. *RAIP, 1(XXIV)*, 235–244. <https://doi.org/10.5604/01.3001.0054.4693> [dostęp: 04.07.2025]
- Cixous, H. (1976). The Laugh of the Medusa. *Signs, 1(4)*, 875–893. <https://www.csus.edu/indiv/o/obriene/art206/Readings/Cixous%20Laugh%20of%20the%20Medusa.pdf> [dostęp: 04.07.2025]
- Costanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need*. MIT Press.
- Crawford, K. (2021). *Atlas sztucznej inteligencji. Władza, pieniądze, środowisko naturalne*. Wydawnictwo Bo.wiem.
- Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review, 43(6)*, 1241–1299.
- Dukaj, J. (2019). *Po piśmie*. Wydawnictwo Literackie.
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Frank, A. (1947/2001). *The Diary of a Young Girl*. Doubleday. https://dn721605.ca.archive.org/0/items/anna-frank-diary/Anna%20frank_diary.pdf [dostęp: 04.07.2025]

- Gilligan, C. (2015). *Innym głosem: teoria psychologiczna a rozwój kobiet*. Wydawnictwo Krytyka Polityczna.
- Grant, M. G. (2016). *Playing the Whore: The Work of Sex Work*. Verso.
- Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3), 575–599. https://www.jstor.org/stable/3178066?utm_source=chatgpt.com [dostęp: 04.07.2025]
- Harding, S. (1991). *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Cornell University Press.
- Highmore, B. (2011). *Ordinary Lives: Studies in the Everyday*. Routledge.
- Hooks, B. (1989). *Talking Back: Thinking Feminist, Thinking Black*. South End Press.
- Lejeune, P. (1989). *On Autobiography*. University of Minnesota Press.
- Losh, E. (2019). *Feminized AI and the Problem of Gendered Labor in the Age of Automation*. In *Digital Feminisms*.
- Majkowski, T. Z. (2020). *Narracja i tożsamość w grach wideo*. WUJ. <https://rcin.org.pl/dlibra/publication/68948/edition/52952?language=en> [dostęp: 04.07.2025]
- Mignolo, W. (2009). Epistemic Disobedience, Independent Thought and De-Colonial Freedom. *Theory, Culture & Society*, 26(7–8), 159–181.
- Nacher, A. (2016). Afektywne archiwum codzienności: o pamięci, technologii i narracji cyfrowej. *Teksty Drugie*, 2.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Nussbaum, M. C. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.
- Plath, S. (2022). *Dzienniki 1950–1962*. Wydawnictwo Marginesy.
- Pobłocki, K. (2022). *Kapitalizm. Historia krótkiego trwania*. Fundacja Nowej Kultury Bęc Zmiana.
- Przegalińska, A. (2020). *Sztuczna inteligencja. Nieludzka, arcyludzka*. Wydawnictwo Naukowe PWN.
- Smith, S., & Watson, J. (2010). *Reading Autobiography: A Guide for Interpreting Life Narratives*. University of Minnesota Press.
- Suchman, L. (2007). *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press.
- Vallor, S. (2016). *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford University Press.
- Van den Hoven, J., Vermaas, P. E., & Van de Poel, I. (Eds.). (2015). *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*. Springer.

- Wachter-Boettcher, S. (2017). *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. W. W. Norton & Company.
- Wajcman, J. (1991). *Feminism Confronts Technology*. Penn State University Press.
- Werner, A., & Malterud, K. (2003). It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. *Social Science & Medicine*, 57(8), 1409–1419.
- Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, 109(1), 121–136.
- Woolf, V. (2023), *Własny pokój*. OsnoVa.
- Zuboff, S. (2019). *Wiek kapitalizmu i inwigilacji*. Wydawnictwo Naukowe PWN.

Europejska Przestrzeń Danych Zdrowotnych w Polsce – kontekst innowacyjności i cyfryzacji w świetle modelu Serendipity

Streszczenie

W artykule przedstawiono studium przypadku dotyczące polskiego podejścia do wdrażania European Health Data Space (EHDS), czyli europejskiej przestrzeni danych zdrowotnych, w kontekście innowacyjności systemowej i współpracy międzyresortowej. Polska – dysponująca rozwijającą się infrastrukturą danych, zaangażowanymi instytucjami oraz kapitałem społecznym – posiada realny potencjał, by stać się liderem praktycznego wdrożenia EHDS. Warunkiem koniecznym jest jednak podjęcie świadomej decyzji politycznej i uruchomienie mechanizmów efektywnego zarządzania innowacjami publicznymi. Liderem tego wdrożenia może zostać Agencja Badań Medycznych we współpracy z jednostkami naukowo-badawczymi, m.in. zrzeszonymi w Sieci Badawczej Łukasiewicz.

Analiza proponowanego sposobu wdrożenia EHDS opiera się częściowo na modelu serendipity, który zakłada, że nieoczekiwane połączenia danych i osób oraz elastyczność organizacyjna mogą prowadzić do przełomowych rozwiązań. W artykule wskazano, że warunkiem działania takiego modelu jest zapewnienie: dostępu do danych, bezpieczeństwa ich przetwarzania oraz zaufania społecznego.

Przedstawiony przypadek pokazuje również, że dane zdrowotne to nie tylko zasób dla systemu ochrony zdrowia, ale potencjalny fundament zarządzania innowacjami w wielu obszarach polityki publicznej. W tym kontekście EHDS staje się nie tyle projektem informatycznym, co impulsem do budowy nowego, zintegrowanego modelu zarządzania danymi publicznymi w Polsce.

Słowa kluczowe: dane, innowacje, serendipity, EHDS (European Health Data System), zarządzanie innowacjami

1 Sieć Badawcza Łukasiewicz – Instytut Sztucznej Inteligencji i Bezpieczeństwa

Serendipity, Data and Innovation: EHDS in the Polish Context

Abstract

The article presents a case study on the Polish approach to implementing the European Health Data Space (EHDS) in the context of systemic innovation and interministerial cooperation. The authors argue that Poland – equipped with a developing data infrastructure, engaged institutions, and social capital – has a genuine potential to become a leader in the practical implementation of EHDS. However, a necessary condition for success is the adoption of a deliberate political decision and the activation of mechanisms for effective public innovation management.

The analysis is partly based on the serendipity model, which assumes that unexpected connections between data and people, combined with organizational flexibility, can lead to breakthrough solutions. The article highlights that the effectiveness of such a model depends on ensuring access to data, the security of data processing, and public trust.

The presented case also demonstrates that health data is not only a resource for the health-care system but also a potential foundation for innovation management across multiple areas of public policy. In this context, the EHDS emerges not merely as an IT project, but as a catalyst for building a new, integrated model of public data governance in Poland.

Keywords: data, innovation, serendipity, EHDS (European Health Data System), innovation management

Wstęp

W obliczu pandemii COVID-19 Unia Europejska podjęła szereg działań zmierzających do lepszego przygotowania wspólnoty na przyszłe kryzysy zdrowotne. Choć instytucje takie jak Europejskie Centrum ds. Zapobiegania i Kontroli Chorób (ECDC) oraz Europejska Agencja Leków (EMA), funkcjonowały już wcześniej, pandemia unaoczniała potrzebę wzmocnienia ich kompetencji oraz pogłębienia koordynacji na poziomie unijnym. Jednym z kluczowych elementów tej nowej strategii stała się koncepcja Europejskiej Przestrzeni Danych Zdrowotnych (EHDS), przedłożona przez Komisję Europejską w 2022 roku (Kotsareli, Tsachouridis, 2023, s. 92). Ostatecznie działania te zaowocowały w kwietniu 2024 roku przyjęciem przez Parlament Europejski Rozporządzenia o Europejskiej Przestrzeni Danych Medycznych – po publikacji aktu w Dzienniku Urzędowym Unii Europejskiej czas pierwszego etapu wdrożenia przez państwa członkowskie zakresłono na marzec 2029 roku.

W niniejszym opracowaniu podjęto próbę analizy tego zagadnienia poprzez studium przypadku rozwoju implementacji elementów rozporządzenia European Health Data Space (EHDS) w sprawie europejskiej przestrzeni danych zdrowotnych, która ma na celu zintegrowanie zasobów informacyjnych z obszaru ochrony zdrowia na poziomie unijnym. Analiza koncentruje się na współpracy instytucjonalnej pomiędzy trzema kluczowymi resortami: Ministerstwem Zdrowia (w tym Agencją Badań Medycznych – ABM), Ministerstwem Cyfryzacji (odpowiedzialnym za aspekty technologiczne EHDS) oraz Ministerstwem Nauki i Szkolnictwa Wyższego (poprzez jednostki naukowe i badawcze będące odbiorcami danych). Współpraca ta stanowi przykład dynamicznego, lecz nie w pełni sformalizowanego modelu zarządzania innowacją, w którym poszczególne instytucje dostarczają komplementarne zasoby, kompetencje i interesariuszy.

W warunkach rosnącej złożoności technologicznej, dynamicznych zmian regulacyjnych oraz postępującej cyfryzacji sektorów publicznych i prywatnych, brak systemowych ram wspierających rozwój czy komercjalizację innowacji staje się istotną barierą dla rozwoju gospodarczego i społecznego w skali całego kraju. Szczególnie widoczne jest to w obszarze danych – w tym danych medycznych – gdzie nie tylko technologia, ale również otoczenie prawne, zdolność instytucjonalna i międzysektorowa współpraca decydują o powodzeniu całego przedsięwzięcia.

Poniższe rozważania opierają się częściowo na założeniach tzw. przypadkowego modelu zarządzania innowacjami (model serendipity), który zakłada, że przełomowe innowacje często nie wynikają z precyzyjnie zaplanowanego procesu, lecz z twórczego wykorzystania nieoczekiwanych zbiegów okoliczności, elastyczności organizacyjnej oraz gotowości instytucji do współdziałania poza sztywnymi schematami. Przypadek polskiego podejścia do EHDS pokazuje, że nawet

w silnie regulowanym i sfragmentaryzowanym systemie publicznym możliwe jest generowanie wartości innowacyjnej – o ile istnieją podstawowe mechanizmy integrujące interesy i cele różnych podmiotów.

Celem artykułu jest ukazanie, jak zidentyfikowana potrzeba społeczna i wspólna wizja funkcjonowania krajowego systemu danych zdrowotnych mogą stać się katalizatorem międzyresortowej synergii oraz impulsem do budowy systemowego podejścia do zarządzania innowacjami w sektorze publicznym.

Model Serendipity

„Przypadkowy model zarządzania innowacjami” (rzadziej: „serendipity-driven innovation model”) nie jest w literaturze przedmiotu definiowany jednoznacznie. Dla potrzeb niniejszych rozważań wykorzystany zostanie zarówno jako klasyczna typologia „serendipity” (przypadkowego modelu zarządzania innowacjami), ale także w kontekście tzw. drugiego życia dla innowacji czy wynalazków, zarzuconych przed laty ze względu na niesprzyjające uwarunkowania ekonomiczne, technologiczne czy prawne. Hipoteza badawcza łącząca przypadek polskiego modelu wdrożenia EHDS, opartego o platformę ZPA, z osiągnięciami nauki w zakresie zarządzania innowacjami nasuwa się sama: w jaki sposób zastosowanie modelu serendipity może wpłynąć na efektywne wykorzystanie funkcjonującego już rozwiązania (ZPA)? Jak w innowacyjny sposób wykorzystać zasoby danych w kontekście wdrożenia ich w Europejski Obszar Danych Zdrowotnych?

Opisany model „przypadkowy” zakłada funkcjonowanie kilku wzorców, do których można przyporządkować przełomowe przykłady innowacyjnych rozwiązań w dziedzinie medycyny, technologii oraz innych nauk. Przypadkowe odkrycia opierają się na kilku zaobserwowanych schematach, wynikających np. ze znalezienia się w odpowiednim miejscu, zaobserwowania czegoś potencjalnie nowego poprzez wykorzystanie zjawiska „szczęśliwego trafu”. Elementem charakterystycznym modelu jest brak możliwości hierarchicznego określenia miejsca i czasu, gdyż te są ze sobą ściśle powiązane. W świetle niniejszej publikacji model przypadkowy pozwala na wykorzystanie „odpowiedniego czasu”, wynikającego ze spotkania kilku obszarów dzięki osobowym powiązaniom tematów, ich otwartym umysłom i sprzyjającym okolicznościom legislacyjnym.

Przypadkowość w badaniach naukowych była i nadal pozostaje źródłem sporów. Dopiero jej wykorzystanie skutkujące budową prawidłowości i modeli zarządzania nadało jej badawczą wartość. Doskonały efekt czy produkt uboczny, który daje sukces rynkowy, leczy obłożnie chorych czy wskazuje na technologiczny i przełomowy rozwój, nie może bowiem zostać uznany za niepowodzenie czy porażkę. Proces ten zaklasyfikował Henry Chesbrough terminem „zarządzanie

negatywem” (Chesbrough, 2004). Flagowym przykładem przekucia niepowodzenia badawczego w ogromny sukces rynkowy jest wynalezienie Viagry. Przykładem zmarnowanej możliwości jest natomiast wymyślony i stworzony przez firmę Xerox komputer osobisty (Chesbrough, Rosenbloom, 2002, s. 13).

Według klasyfikacji Kampratha i Henike innowacja może powstać również na skutek przypadkowego połączenia materiałów (Viagra, promienie Roentgena, wulkanizacja gumy), odpowiednich osób (witamina C), nieznanych powiązań danych czy następstw wcześniejszych wynalazków (spontaniczna radioaktywność) (Kamprath, Henike, 2019). Ta klasyfikacja jest uproszczonym rozwiązaniem aż 17 wzorów zbiegów okoliczności, które szczegółowo opisuje van Andel (1994, s. 631–648). Inny podział, z perspektywy autorki, nieco zawężony i ograniczony do odkryć z dziedziny medycyny, aczkolwiek celowy w niniejszym artykule, prezentuje de Rond w „Structure of serendipity”, wyróżniając „pseudoserendipity” – gdzie celem był wynalazek, ale droga do niego była przypadkowa i zaskakująca (PCR i DNA). Następną częścią podziału jest „true serendipity”, gdzie zmianie ulega cel w trakcie procesu innowacyjnego (sildenafil). Kolejno w ramach modelu wymienia się przypadek jako nieplanowaną konsekwencję procesu badawczego oraz czysto losowy przypadek (de Rond, 2005, s. 3–6, 19–20).

W polskiej literaturze dotyczącej zarządzania innowacjami trudno znaleźć choćby wzmiankę o roli przypadku. Autorzy publikacji koncentrują się wokół kilku klasyfikacji i podziałów, skutecznie omijając odniesienie się do roli „szczęśliwego trafu”, jakże ważnego dla całego zagadnienia. Badania prowadzone w USA, Wielkiej Brytanii czy Niemczech nie tylko umieszczają przypadek wysoko na liście źródeł innowacji, ale również próbują go od lat usystematyzować i ująć w spójny model zarządzania, co skłoniło autorkę do zgłębienia potencjału tego zagadnienia i wykorzystania go we własnej działalności naukowej.

Spora liczba wynalazków odniosła sukces na polu nauki, czy też uznanie w świecie biznesowym, w następstwie zaplanowanych działań, doświadczeń i wielokrotnych prób, jednakże historia wskazuje, że niejednokrotnie to właśnie „przypadkowość” pozwala na redefinicję zamierzeń, często ich rezultatów, prowadząc do przełomowych odkryć, kluczowych dla rozwoju społecznego, technologicznego czy procesowego.

Rola i znaczenie rozporządzenia EHDS (EU, 2025)

EHDS (European Health Data Space) to projekt unijny, którego celem jest zapewnienie obywatelom UE kontroli nad ich danymi zdrowotnymi oraz umożliwienie naukowcom i instytucjom państw Unii dostępu do tych danych – oczywiście z zachowaniem najwyższych standardów bezpieczeństwa i prywatności. Model

ten wywodzi się z Finlandii, gdzie instytucja Findata² skutecznie integruje dane medyczne, środowiskowe i społeczno-ekonomiczne, dając publicznej administracji podstawy do podejmowania decyzji opartych na faktach.

EHDS respektuje regulacje innego rozporządzenia – RODO (GDPR, 2025), zapewniając najwyższe standardy ochrony danych osobowych. Dane udostępniane do celów badań mają być zanonimizowane lub pseudonimizowane, a dostęp do nich będzie możliwy tylko w bezpiecznych środowiskach przetwarzania danych. Obywatelom przysługuje prawo do pełnej kontroli nad swoimi danymi, w tym do ograniczenia dostępu, wglądu w historię dostępu i wycofania zgody na ich wykorzystanie. Filozofia rozwiązania zapewnia oczywiście dostęp transgraniczny, na wspólnym rynku, w ramach państw członkowskich UE.

Wśród kluczowych celów wytyczonych dla wspólnej przestrzeni dla danych są trzy osie priorytetów:

- wzmocnienie pozycji pacjentów,
- wspieranie badań i innowacji,
- ustanowienie modelowych wspólnych standardów i infrastruktury wsparcia.

W ramach priorytetu pierwszego EHDS umożliwi obywatelom UE bezpłatny dostęp do elektronicznych danych zdrowotnych, takich jak wyniki badań medycznych, e-recepty, wyniki badań laboratoryjnych czy efekty diagnostyki obrazowej. Dane te mają być dostępne w formacie elektronicznym, uznawanym we wszystkich państwach członkowskich UE, co w efekcie powinno znacznie ułatwić opiekę zdrowotną podczas podróży lub przeprowadzki do innego kraju unijnego.

W ramach kolejnego z celów umożliwione powinno być bezpieczne i zgodne z przepisami udostępnianie zanonimizowanych lub pseudonimizowanych danych zdrowotnych do celów badań naukowych, statystyki, tworzenia polityk zdrowotnych i innowacyjnych, aktów prawnych i innych. Dane te docelowo mają obejmować informacje z dokumentacji medycznej, badań klinicznych, rejestrów zdrowotnych oraz danych genetycznych. Istotnym elementem konstrukcji EHDS – podniesionym przez Komisję Europejską do rangi zasadniczej – jest odporność całego systemu. Odporność rozumiana z jednej strony jako spójny i jednolity sposób, w jaki Unia jest w stanie osiągnąć podejście do kwestii problemowych (Joseph, Juncos, 2019, s. 995–1012), natomiast z drugiej widziana jako zdolność do stawiania czoła zewnętrznym zagrożeniom przez europejską wspólnotę (Adger, 2000, s. 361).

Priorytet dotyczący standaryzacji dostępuów zakłada wprowadzenie jednolitych zasad dotyczących systemów elektronicznej dokumentacji medycznej, w tym obligatoryjnych komponentów oprogramowania zapewniających interoperacyjność i rejestrowanie dostępu do danych. Technicznie, w ramach

2 <https://findata.fi/en/> [dostęp: 01.06.2025]

założeń celu, powinna powstać także odpowiednia infrastruktura, umożliwiająca transgraniczne udostępnianie danych zdrowotnych w krajach członkowskich UE (EU, EHDS, 2025).

Rozporządzenie dotyczące europejskiej przestrzeni danych medycznych jest dokumentem nowym, bo wdrożonym 26 marca 2025 roku. Polska, podobnie jak pozostałe kraje Unii, dopiero rozpoczyna ścieżkę jego implementacji. W ramach kolejnego etapu, czyli do 26 marca 2027 r., kraje członkowskie mają ustanowić organy odpowiedzialne za „cyfrowe zdrowie” i przygotować infrastrukturę do wdrożenia EHDS. W Polsce funkcjonuje podmiot, który może to robić w zakresie swoich ustawowych uprawnień i, co ważne, posiada odpowiednie przygotowanie technologiczne i kompetencyjne. Tym podmiotem jest Agencja Badań Medycznych, stanowiąca pomost pomiędzy prawodawstwem unijnym, administracją krajową i technologicznymi rozwiązaniami, które tworzą między innymi jednostki badawcze. Co ciekawe, ABM już dziś ma w swoim obszarze zarządzania danymi omicznymi³, które w obszarze wdrożeniowym EHDS jest działaniem prekursorskim.

Agencja Badań Medycznych – istotne ogniwo

ABM posiada ustawowe kompetencje do przetwarzania danych w celach naukowych. Zbieranie, pseudonimizacja, analiza i udostępnienie – to wszystko ABM już robi. Co więcej, ABM może pełnić funkcję operatora EHDS na poziomie krajowym. To tworzy kluczową przewagę konkurencyjną. Polska w zasadzie nie musi czekać na wdrożenie unijnej legislacji. Może działać natychmiast, w oparciu o istniejące przepisy prawa.

Agencja Badań Medycznych to polska państwowa instytucja powołana w 2019 roku, której celem jest wspieranie innowacyjnych badań w ochronie zdrowia, ze szczególnym uwzględnieniem niekomercyjnych badań klinicznych oraz projektów interdyscyplinarnych.

Głównymi zadaniami ABM są:

- finansowanie badań naukowych i prac rozwojowych – agencja wspiera projekty w dziedzinie nauk medycznych oraz nauk o zdrowiu, w tym badania kliniczne, obserwacyjne i epidemiologiczne;
- wydawanie opinii i ekspertyz – agencja opracowuje analizy i ekspertyzy na potrzeby organów administracji publicznej, które docelowo stanowią podstawy dla polityk odpowiednich resortów;
- rozwój współpracy międzynarodowej – agencja inicjuje i rozwija współpracę z zagranicznymi instytucjami w zakresie nauk medycznych i nauk o zdrowiu;

3 <https://www.uw.edu.pl/kurs-bioinformatyczny-w-icm/> [dostęp: 10.07.2025]

- realizacja własnych badań naukowych – agencja prowadzi własne projekty badawcze i prace rozwojowe (ABM, 2025).

ABM tworzy w ramach Polskiej Sieci Badań Klinicznych Regionalne Centra Medycyny Cyfrowej, które są w pełni zgodne z działaniem UE na rzecz wspólnych europejskich przestrzeni danych i które zapewniają dostęp do dużej ilości wiarygodnych (jakościowych) danych, przy zachowaniu kontroli nad ich źródłami.

Działalność centrów opiera się w głównej mierze na gromadzeniu i analizie danych medycznych oraz omicznych (genetyczne, epigenetyczne, transkryptomiczne, proteomiczne, metabolomiczne). Dane integrowane są z różnych źródeł, takich jak: telemetria, urządzenia monitorujące, indywidualne ankiety pacjentów oraz dane dotyczące materiału biologicznego przechowywanego w biobankach, systemy HIS⁴ oraz badania kliniczne.

W rozwiązaniach ABM wdrażane są mechanizmy sztucznej inteligencji, które pozwalają na rozwijanie algorytmów predykcyjnych i diagnostycznych, wspierających kolejno lekarzy w podejmowaniu decyzji terapeutycznych, a w obszarze badań klinicznych pozwalają na optymalizację projektowania badań oraz nadzorowanie jakości wyników na każdym etapie ich realizacji. Zgodnie z procedurami krajowymi i unijnymi agencja bezwzględnie przestrzega rygorystycznych norm bezpieczeństwa danych, stosując procedury bezpieczeństwa zgodne z unijnymi i krajowymi regulacjami prawnymi.

Zintegrowana Platforma Analityczna (ZPA) – technologiczny rdzeń EHDS

ZPA to system informatyczny stworzony w katowickim instytucie EMAG na zlecenie Ministerstwa Cyfryzacji. W założeniu miał wspierać podejmowanie decyzji na centralnych szczeblach administracji, natomiast jego potencjał nigdy nie został wykorzystany. Jest to rozwiązanie opracowane w ramach Programu Operacyjnego Polska Cyfrowa na lata 2014–2020 jako wsparcie rozwiązań dla administracji rządowej w ramach Osi Priorytetowej nr 2 „E-administracja i otwarty rząd”. Projekt zakończono w roku 2023, na różnych stadiach realizacji zaangażowane weń były: Ministerstwo Zdrowia, Ministerstwo Rodziny i Polityki Społecznej, Ministerstwo Edukacji i Nauki, Uniwersytet Warszawski, SGH oraz Sieć Badawcza Łukasiewicz.

Filozofia platformy i stworzona technologia idealnie odpowiadają potrzebom Agencji Badań Medycznych oraz zapisom rozporządzenia EHDS w zakresie integracji danych, analiz w bezpiecznym środowisku i pseudonimizacji. Założenia i zasady funkcjonowania ABM z kolei korelują z wysiłkami Unii

4 <https://ucyfrowienie.pl/systemy-his/> [dostęp: 20.06.2025]

Europejskiej na rzecz utworzenia wspólnych, bezpiecznych przestrzeni zarządzania danymi, nie tylko w obszarze medycyny, ale także danymi z innych obszarów, mogących w znaczny sposób usprawnić procesy społeczne, ekonomiczne i zdrowotne.

Głównym założeniem projektu jest zwiększenie efektywności działań administracji publicznej w kluczowych obszarach społecznych i gospodarczych poprzez dostarczenie wysokiej jakości analiz wspierających podejmowanie decyzji.

Projekt Zintegrowanej Platformy Analitycznej (ZPA) odpowiada na realne potrzeby instytucji publicznych i koncentruje się na realizacji następujących celów operacyjnych:

- zapewnienie centralnego dostępu do narzędzi analitycznych wspierających procesy decyzyjne w administracji;
- wypracowanie standardów dostępu do danych i analiz dla priorytetowych obszarów polityki społecznej i gospodarczej;
- zbudowanie infrastruktury technicznej i instytucjonalnej umożliwiającej współpracę nauki i administracji w rozwijaniu potencjału analitycznego państwa i skuteczniejszym rozwiązywaniu problemów publicznych;
- uporządkowanie zasad technicznych i operacyjnych dotyczących dostępu do kluczowych zasobów informacyjnych państwa oraz wdrożenie jednolitego systemu bezpieczeństwa i udostępniania danych do celów analitycznych;
- otwarcie danych administracyjnych na potrzeby środowisk badawczych, przy zachowaniu odpowiednich zabezpieczeń, w celu wspierania polityk publicznych opartych na dowodach;
- wsparcie cyfrowej transformacji administracji rządowej w obszarach wskazanych w katalogu dobrych praktyk i rekomendacji dotyczących cyfrowego rozwoju instytucji publicznych (Ministerstwo Cyfryzacji, 2025).

Wśród oczekiwanych rezultatów projektu znajdujemy centralizację narzędzi analitycznych. Projekt ma doprowadzić do stworzenia jednej, centralnej przestrzeni, z narzędziami wspierającymi analizę danych dla administracji publicznej. Platforma ma w założeniu umożliwić sprawniejsze i trafniejsze podejmowanie decyzji w oparciu o dane pochodzące z wielu instytucji publicznych.

Kolejnym kluczowym celem jest standaryzacja dostępu do danych i analiz, a w jej ramach zebrane w Bazie Wiedzy, opisane i udokumentowane modele dostępu do danych. Będą one wykorzystywane w przyszłych analizach administracyjnych, co usprawni kolejne projekty badawcze i obniży ich koszty. Dokumentacja będzie zawierała także informacje o jakości danych przekazywanych przez poszczególnych gestorów (zarządców danych).

Kluczowe z punktu widzenia badacza jest wzmocnienie współpracy na linii administracja – nauka – społeczeństwo. Projekt może stworzyć warunki

techniczne i organizacyjne do trwałego dialogu między administracją a środowiskiem naukowym. Zaangażowanie ekspertów naukowych w realizację badań zwiększy ich wartość merytoryczną i może stać się podstawą do publikacji naukowych⁵.

ZPA zakłada też ujednolicenie zasad dostępu do danych publicznych, jednolitych reguł udostępniania kluczowych zasobów informacyjnych państwa – z uwzględnieniem bezpieczeństwa, dostępności i przejrzystości metod przekazywania danych. Umożliwi to jednocześnie przygotowanie przekrojowych analiz międzyresortowych, a w efekcie bardziej precyzyjne modelowanie skutków polityk publicznych oraz lepsze monitorowanie wdrażanych regulacji, czyli finalnie realny postęp cyfrowy administracji publicznej (Ministerstwo Cyfryzacji, 2025).

Czynniki determinujące rynkowe powodzenie innowacji

Powodzenie rynkowe innowacyjnego rozwiązania stanowi złożone zjawisko, uwarunkowane wielością czynników zarówno wewnętrznych, jak i zewnętrznych względem organizacji wdrażającej innowację. W literaturze przedmiotu podkreśla się, że sam potencjał technologiczny nie gwarantuje sukcesu komercyjnego – kluczowe znaczenie mają również aspekty związane z dopasowaniem do potrzeb rynku, wyborem odpowiedniego modelu biznesowego, strategią komercjalizacji oraz otoczeniem instytucjonalnym, ale również z umiejętnościami z zakresu zarządzania czy komunikacji. W warunkach rosnącej konkurencji i skracającego się cyklu życia produktów identyfikacja i zrozumienie czynników determinujących rynkowe powodzenie staje się istotnym obszarem badań w dziedzinie zarządzania innowacjami. Warto przybliżyć i pokrótce opisać kluczowe czynniki warunkujące skuteczną implementację rozwiązania, które dostaje w zasadzie drugą szansę wejścia na rynek. Rozwiązania, które dzięki połączeniu ludzkiej aktywności, sprzyjających uwarunkowań prawnych i technologicznych, ale i wyraźnej potrzebie społecznej, może odnieść sukces rynkowy.

Przeszkoda systemowa – dane zamknięte w silosach

Ludzkość sięga po semiglutydy – leki zmieniające metabolizm, takie jak Ozempic, Mounjaro czy Saxenda. Ich długoterminowy wpływ na zdrowie nie jest jeszcze poznany. Czy po dwóch latach stosowania rośnie ryzyko bezsenności? Albo dysfunkcji seksualnych? Dziś nie mamy narzędzi, by to rzetelnie zbadać.

5 Dokumenty wewnętrzne instytutu Łukasiewicz – EMAG.

W zasadzie w Polsce nie jesteśmy w stanie odpowiedzieć na dużo prostsze pytania, w rodzaju: „Czy zanieczyszczenie powietrza wpływa na częstość zawałów?” – bo potrzebne dane są rozproszone po lokalnych systemach szpitalnych i nie są ze sobą zestawiane. Taka sytuacja blokuje rozwój nauki. Blokuje też skuteczne decyzje systemowe, bo tylko integracja danych medycznych, ubezpieczeniowych, farmaceutycznych, ZUS i GUS może dać nam odpowiedź. To zadanie, które idealnie wpisuje się w ideę EHDS.

Aby wskazane wyżej podmioty mogły realizować zadania na rzecz społeczeństwa potrzebna jest zmiana Ustawy o informatyzacji podmiotów realizujących zadania publiczne. Obecnie procedura jest na etapie wpisu do rejestru aktów legislacyjnych rządu⁶.

Pierwsza zmiana musi wskazywać m.in. Agencję Badań Medycznych jako podmiot uprawniony do korzystania z danych i ich przetwarzania, kolejna zmiana musi pozwolić na swobodne wykonywanie analiz dla celów polityk publicznych.

Przypadek w roli wynalazcy – drugie życie innowacji

Opisany na wstępie artykułu model przypadkowego zarządzania innowacjami zakłada, że innowacje często powstają nieplanowanie – na styku danych, zespołów i zdarzeń. Integrując dane i otwierając je dla naukowców, tworzymy przestrzeń do takich nieoczekiwanych odkryć. Uzupełniając teorię, warto wskazać najbardziej znane i spektakularne przykłady takich przypadkowych odkryć medycznych, które nie tylko zrewolucjonizowały medycynę, ale i odniosły sukces rynkowy:

- sildenafil – pierwotnie opracowany jako lek na nadciśnienie i chorobę wieńcową, przypadkiem ujawniono jego wpływ na zaburzenia erekcji;
- penicylina – odkryta przez Alexandra Fleminga przypadkiem, gdy grzyb zabił kolonie bakterii w laboratorium;
- słodzik sacharyna – odkryty, gdy chemik zapomniał umyć ręce po pracy z substancjami chemicznymi;
- promieniowanie rentgenowskie – Wilhelm Roentgen nie planował odkrycia promieniowania, a mimo to zrewolucjonizował diagnostykę;
- struktura DNA – Watson i Crick oparli się na niezamierzonych danych z kryształografii rentgenowskiej Rosalind Franklin;
- tamoksyfen – zsyntetyzowana w latach 60. substancja z przeznaczeniem antykonceptyjnym, ostatecznie znalazła zastosowanie w terapii nowotworowej, szczególnie skuteczna w leczeniu niezaawansowanego raka sutka.⁷

6 <https://www.gov.pl/web/premier/projekt-ustawy-o-zmianie-ustawy-o-informatyzacji-dzialalnosci-podmiotow-re-alizujacych-zadania-publiczne2> [dostęp: 10.06.2025]

7 Wykład prof. Marka Rusina „Epidemiologia, diagnostyka molekularna i terapia nowotworów” – w ramach studiów doktoranckich.

Wiele innowacji powstało również dzięki przypadkowym spotkaniom lub współpracy naukowców z różnych dziedzin. Kluczowe jest, by dostrzegać skutki – często nieplanowane, by widzieć w nich nie porażkę, a potencjalną szansę, albo przekuć skutki uboczne w skuteczny produkt, gotowy do wprowadzenia na rynek.

Aby taki model mógł skutecznie funkcjonować w sektorze publicznym – szczególnie w obszarze danych zdrowotnych – musi zostać spełnione kryterium dostępności danych, które to powinny być nie tylko gromadzone, ale również realnie dostępne dla różnych grup interesariuszy: badaczy, organizacji ochrony zdrowia i decydentów, a kolejno analizowane i przekuwane na polityki państwowe ułatwiające życie obywatelom. Ponadto interdyscyplinarność – potrzebne jest środowisko, które umożliwia współpracę między informatykami, lekarzami, epidemiologami, etykami, socjologami czy analitykami danych, i które formalnie wspiera takie relacje, zapewniając finansowanie i odpowiednie narzędzia. Wszystko to natomiast musi spinać elastyczny system, zdolny do adaptacji i reagowania – a nie jedynie do realizacji z góry określonych celów.

Rola lidera – widzieć mosty tam, gdzie inni widzą dziury (Burt, 2004)

Bez silnych liderów, którzy nie tylko widzą możliwości w miejscach, gdzie inni dostrzegają tylko problemy, często bez ich wizji, zdolności kojarzenia faktów, ilościowości i możliwości nie jesteśmy w stanie doprowadzić do sukcesu rynkowego. Przywództwo to coś więcej niż idea – człowiek w organizacji może być najsłabszym ogniwem, ale może też przesądzić o powodzeniu dzięki rozwiniętym umiejętnościom nie tylko relacyjnym, zdolnościom budowania, ale i umiejętnemu zarządzaniu zmianą, innowacją i zasobem ludzkim.

Choć w kontekście niniejszego artykułu to technologia, a nie człowiek, jest w centrum zainteresowania, nie sposób nie uwzględnić nieocenionej roli człowieka/lidera w całym procesie innowacyjnym.

W procesie zarządzania innowacjami rola lidera zyskuje szczególne znaczenie, ponieważ to on inicjuje, ukierunkowuje i podtrzymuje proces zmiany organizacyjnej, niezbędnej do wdrożenia i komercjalizacji nowego rozwiązania. W odróżnieniu od klasycznego zarządzania projektami przywództwo w obszarze innowacji wymaga zdolności operowania w warunkach wysokiej niepewności, dynamicznych zmian rynkowych oraz braku jednoznacznych schematów postępowania (Crossan, Apyadin, 2010, s. 1154–1191).

Wśród zadań lidera zarządzającego innowacjami znajdziemy elementy inicjujące sam proces, takie jak kreowanie wizji innowacji oraz zapewnienie spójności między celami technologicznymi a potrzebami rynku. Ale też wielowątkowe i długotrwałe wysiłki związane z mobilizowaniem i inspirowaniem zespołów projektowych, co ma szczególne znaczenie w fazach wstępnych procesu, a jest

nieodzowne i kluczowe dla potencjalnego powodzenia przedsięwzięcia. To człowiek, nawet jeśli wspierany technologią, zarządza niepewnością i ryzykiem, często poprzez stworzenie środowiska umożliwiającego testowanie i szybkie uczenie się. To on buduje relacje z otoczeniem, tym wewnątrz (pracownicy, zarząd) i na zewnątrz organizacji (klienci, partnerzy, organizacje), dzięki czemu tworzy (bądź dezintegruje) kulturę organizacyjną i wpływa na modelowe zarządzanie każdym zasobem – technologicznym, ludzkim, finansowym itd. (Birkinshaw, Hamel, 2008, s. 825–845).

W literaturze przedmiotu dotyczącej obszaru zarządzania wskazuje się, że skuteczny lider innowacji powinien wykazywać wysoki poziom inteligencji emocjonalnej, elastyczność poznawczą oraz umiejętność poruszania się między różnymi stylami przywództwa – od transformacyjnego po partycypacyjny. Szczególne znaczenie ma mieć zdolność lidera do przekładania innowacji technologicznej na wartość rynkową, czyli jej „przełożenie” na język korzyści dla klienta i modelu biznesowego organizacji. Taką rolę w przedstawionym w artykule studium przypadku pełni prof. Wojciech Fendler, prezes Agencji Badań Medycznych, który odnalazł potencjał w „porzuconym” rozwiązaniu – stworzonym pod auspicjami Ministerstwa Cyfryzacji, w instytucie naukowo-badawczym Łukasiewicz – EMAG w Katowicach, systemie Zintegrowanej Platformy Analitycznej. Nie bez znaczenia w tym kontekście jawi się także rola dyrekcji instytutów naukowo-badawczych: Łukasiewicz – AI (prawny następca wykonawcy systemu Zintegrowanej Platformy Analitycznej, działający obecnie pod kierownictwem prof. Jana Kozaka) oraz Instytutu Łączności zarządzanego przez inspiratorkę powstania systemu ZPA Annę Streżyńską.

Zdolność lidera do integrowania różnych perspektyw – technologicznej, biznesowej, społecznej – stanowi często warunek konieczny skutecznego wdrożenia i skalowania innowacji. Bez takiej roli koordynacyjnej nawet najbardziej obiecujące rozwiązania mogą utknąć na etapie prototypu, co gorsza tzw. półkownika, lub spotkać się z oporem organizacyjnym czy brakiem akceptacji rynkowej.

Potrzeba społeczna – synergia międzyresortowa jako klucz do sukcesu

Współczesne innowacje, szczególnie jeśli dotyczą sfery publicznej, zwłaszcza w obszarze zdrowia, danych i cyfryzacji, nie wynikają wyłącznie z możliwości technologicznych, lecz z jasno zidentyfikowanych potrzeb społecznych i systemowych. Przykładem takiej potrzeby jest konieczność stworzenia zintegrowanego systemu przetwarzania i wykorzystywania danych zdrowotnych w sposób bezpieczny, interoperacyjny i zgodny z wymaganiami Europejskiej Przestrzeni Danych dotyczących Zdrowia (EHDS).

Zarówno pacjenci, instytucje badawcze, jak i decydenci oczekują systemu, który nie tylko gromadzi dane, ale też generuje wartość – w postaci lepszych

decyzji klinicznych, efektywniejszych badań oraz skutecznych polityk publicznych. Potrzeba ta ma wymiar społeczny (dostępność, bezpieczeństwo, personalizacja opieki zdrowotnej) oraz biznesowy (efektywność wykorzystania zasobów, innowacyjność sektora ochrony zdrowia, rozwój usług opartych na danych).

Opisane studium przypadku to nie jest projekt li tylko zdrowotny. To okazja do ponadresortowej i wieloobszarowej współpracy mającej na celu dobro społeczne.

Realizacja tak złożonego projektu, jak narodowy system danych zdrowotnych w ramach EHDS, wymaga ścisłej współpracy między resortami – nie tylko koordynacji, ale rzeczywistej synergii. Przedsięwzięcie angażuje bowiem:

- Ministerstwo Zdrowia (MZ) odpowiedzialne za aspekt kliniczny i regulacyjny systemu, a poprzez Agencję Badań Medycznych (ABM) – także za wsparcie nauki i wdrażanie innowacji w sektorze zdrowia;
- Ministerstwo Cyfryzacji, które tworzy i rozwija infrastrukturę EHDS, odpowiada za interoperacyjność, bezpieczeństwo danych oraz architekturę systemową;
- Ministerstwo Nauki i Szkolnictwa Wyższego, które poprzez jednostki takie jak instytuty badawcze, a także odbiorców danych w środowisku naukowym, gwarantuje wykorzystanie danych do celów badawczo-rozwojowych i edukacyjnych.

Tylko realna współpraca instytucjonalna, oparta na współpracy, efektywnym podziale ról i odpowiedzialności, a także na zaufaniu między partnerami, może przynieść efekt synergiczny, który da nam skuteczne, użyteczne narzędzie.

Analiza SWOT – zaproponowane wdrożenia

Analiza SWOT została przeprowadzona w celu zidentyfikowania kluczowych czynników wpływających na możliwość skutecznego wdrożenia rozporządzenia European Health Data Space (EHDS) w Polsce. Jej celem jest określenie mocnych i słabych stron krajowego systemu w kontekście regulacyjnym, technologicznym i organizacyjnym oraz wskazanie szans i zagrożeń, które mogą zadecydować o powodzeniu procesu implementacji.

Tabela 1. Analiza SWOT

Siły: • europejska regulacja w mocy rozporządzenia, działająca wprost i wymagająca implementacji; • system Zintegrowanej Platformy Analitycznej wykonany i oddany do użytku w ramach poprzedniej perspektywy finansowej Unii Europejskiej; • zapisy Ustawy o Agencji Badań Medycznych kompatybilne z wymogami unijnego rozporządzenia EHDS	Słabości: • silosowość systemów IT w Polsce i brak zaufania do koncepcji decyzji wspomaganych danymi; • pozapolityczny (ekspertski) skład obecnego kierownictwa Ministerstwa Zdrowia, zwiastujący problemy w przeprowadzeniu ścieżki legislacyjnej implementacji rozporządzenia EHDS; • niewielki poziom zrozumienia dla wagi zapisów rozporządzenia EHDS wobec doraźnych problemów systemu ochrony zdrowia w Polsce
Szanse: • leadership osób piastujących kierownicze stanowiska w Agencji Badań Medycznych, Instytucie Łukasiewicz AI (następcy prawnym EMAG) i Instytucie Łączności w Warszawie; • koincydencja pomiędzy stanem legislacji (wymogiem wdrożenia EHDS) a funkcjonowaniem gotowego systemu informatycznego	Zagrożenia: • brak politycznego porozumienia w zakresie nowelizacji Ustawy o Informatyzacji określającej miejsce Zintegrowanej Platformy Analitycznej w systemie administracji RP; • konkurencyjne koncepcje wdrożenia rozporządzenia o EHDS lansowane w Ministerstwie Zdrowia

(opracowanie własne)

Z wykonanej analizy klarownie wynika, iż Polska dysponuje zarówno formalnymi, jak i technologicznymi zasobami, które mogą przyspieszyć wdrożenie EHDS. Warunkiem sukcesu jest ich właściwa integracja i wykorzystanie w ramach spójnego modelu zarządzania danymi zdrowotnymi. Największym zagrożeniem dla skutecznego wdrożenia EHDS nie jest technologia, lecz deficyt świadomości i brak koordynacji między instytucjami. To wskazuje na potrzebę budowy narracji strategicznej – pokazania, że EHDS to narzędzie wspierające efektywność systemu zdrowia, a nie projekt poboczny. W tym miejscu dzięki tzw. drugiemu życiu innowacji/produktu możliwe jest wykorzystanie gotowej infrastruktury (ZPA) i nowych regulacji europejskich do zbudowania spójnego systemu danych zdrowotnych. Jeśli zostanie to odpowiednio zaimplementowane, Polska może stać się liderem wdrożenia EHDS w regionie. Jednakże bez jasnego przywództwa i koordynacji ryzyko paraliżu decyzyjnego jest wysokie. Rozporządzenie EHDS wymaga jednego krajowego operatora i spójnej struktury zarządzania – jeśli Polska tego nie określi, istnieje ryzyko utraty możliwości finansowania lub opóźnienia we wdrożeniu.

Podsumowanie

Dane o wysokiej jakości są dostępne na wyciągnięcie ręki, choć rozproszone po różnych bazach. Jest instytucja, która widzi potencjał w technologii, o której większość rządowej administracji zdawała się nie pamiętać. Funkcjonuje wreszcie system informatyczny idealnie odpowiadający wymogom rozporządzenia EHDS. Tym, czego dziś naprawdę potrzebujemy, jest decyzja polityczna o uruchomieniu pełnego potencjału elementów tej układanki, opisanych powyżej. Polska stoi przed realną szansą, by nie tylko nadążyć za europejskimi trendami w zakresie cyfrowej transformacji ochrony zdrowia, ale wręcz stać się liderem wdrażania European Health Data Space (EHDS) – bez zbędnej straty czasu i bez wielkich inwestycji w technologię.

Dysponujemy infrastrukturą technologiczną, instytucjami o odpowiednich kompetencjach, doświadczeniem we współpracy międzyresortowej oraz rosnącą świadomością znaczenia danych jako zasobu strategicznego. Co więcej, mamy coraz większe zaufanie społeczne do wykorzystywania danych w celach zdrowotnych, zwłaszcza jeśli przekłada się to na poprawę jakości opieki, skuteczność badań i bezpieczeństwo systemowe.

Ten krajowy model integracji danych zdrowotnych nie tylko zintegruje istniejące rejestry i systemy informatyczne, ale może zwiększyć bezpieczeństwo zdrowotne obywateli. Wpłynie też znacznie na poprawę dostępności dla naukowców.

Co równie ważne, sektor ochrony zdrowia to niejedyny obszar, w którym proponowany mechanizm może zostać zastosowany. Model zarządzania danymi oparty na dostępie, bezpieczeństwie i zaufaniu może być replikowany w dowolnych obszarach, takich jak: edukacja, polityka społeczna, środowisko czy bezpieczeństwo publiczne. Korzyści dla społeczeństwa trudno przecenić.

Czy wizja doczeka się realizacji? Odpowiedź na to pytanie dziś spoczywa wyłącznie w rękach polityków.

Bibliografia

- Adger, N. (2000). Social and Ecological Resilience: Are They Related? *Progress in Human Geography*, 24(3), 361.
- Birkinshaw, J., Hamel, G., & Mol, M. J. (2008). Management Innovation. *Academy of Management Review*, 33(4), 825–845.
- Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2), 349–399.
- Chesbrough, H. (2004). Managing your false negative. *Harvard Managing Updates*, 8.

- Chesbrough, H., Rosenbloom, R. (2002). The Role of Business Model in Capturing Value from Innovation; Evidence from Xerox Corporation's Technology Spinoff Companies. *Industrial and Corporate Change*, 11(3), 13.
- Crossan, M. M., & Apaydin, M. (2010). A Multi-Dimensional Framework of Organizational Innovation: A Systematic Review of Literature. *Journal of Management Studies*, 47(6), 1154–1191.
- De Rond, M. (2005). *The structure of serendipity*. Judge Business School, University of Cambridge, 3–6, 19–20.
- Joseph, J., & Juncos, A. E. (2019). Resilience as an Emergent European Project? The EU's Place in the Resilience Turn. *Journal of Common Market Studies*, 57(5), 995–1012.
- Kamprath, M., Henike, T. (2019). Serendipity and innovation. In J. Chen, A. Brem, E. Viardot, P. K. Wong (Eds.), *The Routledge Companion to Innovation Management*. Routledge.
- Kotsareli, K., & Tsachouridis, P. (2023). European Health Data Space: A New Era in EU Health. *HAPSc Policy Briefs Series*, 4(2), 91–96.
- Van Andel, P., (1994). Anatomy of the unsought finding. Serendipity: origin, history, domains traditions, appearances, patterns and programmability. *The British Journal for the Philosophy of Science*, 45(2), 631–648.

Netografia

- Agencja Badań Medycznych. <https://abm.gov.pl/pl/o-nas/o-agencji/34,O-Agencji.html> [dostęp: 12.06.2025]
- European Union, European Health Data Space. <https://european-health-data-space.com> [dostęp: 12.06.2025]
- European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0197> [dostęp: 12.06.2025]
- GDPR, RODO. <https://gdpr.pl/baza-wiedzy/akty-prawne/interaktywny-tekst-gdpr> [dostęp: 10.07.2025]
- Ministerstwo Cyfryzacji. <https://www.gov.pl/web/cyfryzacja/zintegrowana-platforma-analityczna2> [dostęp: 10.05.2025]

Sztuczna inteligencja we wzmacnianiu partycypacji mieszkańców w zarządzaniu rozwojem miasta – perspektywy i ryzyka

Streszczenie

Artykuł podejmuje problem zastosowania sztucznej inteligencji (AI) w procesach partycypacji obywatelskiej na poziomie lokalnym. Na tle współczesnych wyzwań zarządzania rozwojem miast analizowane są możliwości wykorzystania AI do wspierania różnych form udziału mieszkańców – od zbierania i analizy opinii, przez moderację deliberacji, po monitorowanie efektów polityk publicznych. W artykule przedstawiono ramy teoretyczne, przegląd literatury oraz główne obszary zastosowań AI w partycypacji, wskazując zarówno potencjał wzmacniania demokracji lokalnej, jak i ryzyka związane z automatyzacją, uprzedzeniami algorytmicznymi, wykluczeniem cyfrowym czy brakiem transparentności. Artykuł zawiera rekomendacje dla samorządów i projektantów systemów AI, a także tabelaryczne zestawienie zagrożeń i możliwych środków zaradczych. Podkreślono, że skuteczne i etyczne wdrażanie AI w partycypacji obywatelskiej wymaga nie tylko technologicznej sprawności, ale też silnego zakorzenienia w wartościach demokratycznych i inkluzywnych.

Słowa kluczowe: sztuczna inteligencja (AI), partycypacja obywatelska, zarządzanie miastem, demokracja lokalna, algorytmy i etyka, inkluzja cyfrowa, technologia w administracji publicznej, inteligentne miasta (smart cities)

1 Akademia WSB w Dąbrowie Górniczej

Artificial Intelligence in Strengthening Citizen Participation in Urban Development Governance – Perspectives and Risks

Abstract

The article addresses the issue of applying artificial intelligence (AI) in civic participation processes at the local level. Against the backdrop of contemporary challenges in urban development governance, it analyses the potential use of AI to support various forms of resident involvement – from the collection and analysis of opinions, through the moderation of deliberative processes, to the monitoring of public policy outcomes. The article presents a theoretical framework, a review of the literature, and the main areas of AI application in participation, highlighting both the potential to strengthen local democracy and the risks associated with automation, algorithmic bias, digital exclusion, and a lack of transparency. The paper includes recommendations for local governments and AI system designers, as well as a tabular overview of identified risks and possible mitigation measures. It is emphasised that effective and ethical implementation of AI in civic participation requires not only technological competence but also a strong grounding in democratic and inclusive values.

Keywords: artificial intelligence (AI), civic participation, urban governance, local democracy, algorithms and ethics, digital inclusion, technology in public administration, smart cities

Wstęp

Współczesne miasta stoją w obliczu coraz większych wyzwań związanych z zarządzaniem złożonymi procesami rozwoju przestrzennego, społecznego i gospodarczego. W tym kontekście rośnie znaczenie partycypacji mieszkańców jako jednego z filarów demokratycznego i zrównoważonego planowania miejskiego. Jednocześnie rozwój technologii cyfrowych, a zwłaszcza sztucznej inteligencji, otwiera nowe możliwości wspierania i rozszerzania procesów współdecydowania obywateli.

Sztuczna inteligencja znajduje coraz szersze zastosowanie w sektorze publicznym – od automatyzacji obsługi administracyjnej, przez analizę danych miejskich, po wspieranie decyzji planistycznych (OECD, 2020). W obszarze partycypacji obywatelskiej AI może pełnić rolę narzędzia wzmacniającego głos mieszkańców, ułatwiającego dialog społeczny, a także wspierającego bardziej przejrzyste i efektywne procesy decyzyjne. Jednocześnie pojawiają się uzasadnione obawy dotyczące przejrzystości algorytmów, potencjalnych uprzedzeń systemów AI oraz pogłębiania nierówności cyfrowych (Eubanks, 2018; Nemitz, 2018).

Celem niniejszego artykułu jest analiza potencjału wykorzystania sztucznej inteligencji do wzmacniania partycypacji mieszkańców w zarządzaniu rozwojem miasta, z uwzględnieniem zarówno perspektyw, jak i zagrożeń związanych z tym procesem. Artykuł opiera się na przeglądzie literatury, przykładach wdrożeń z różnych miast europejskich oraz refleksji nad ramami etycznymi i społecznymi stosowania AI w kontekście miejskim.

W pierwszej części zaprezentowano ramy teoretyczne i przegląd dotychczasowych badań. Następnie omówiono główne obszary zastosowania AI w partycypacji miejskiej oraz potencjalne korzyści z jej wdrożenia; w kolejnych sekcjach wskazano ryzyka i ograniczenia tego procesu. Artykuł kończą wnioski i rekomendacje dla polityki miejskiej oraz kierunki dalszych badań.

Ramy teoretyczne i przegląd literatury

Partycypacja obywatelska, rozumiana jako aktywny udział mieszkańców w procesach decyzyjnych dotyczących wspólnoty lokalnej, jest fundamentem demokratycznego zarządzania miastem. W literaturze wyróżnia się różne poziomy partycypacji – od informowania i konsultacji po współdecydowanie i inicjatywę obywatelską (Arnstein, 1969; Fung, 2006). Współczesne podejścia kładą nacisk na *demokrację deliberatywną* i *governance partycypacyjną*, które postrzegają mieszkańców nie jako odbiorców polityk miejskich, lecz jako ich współtwórców.

W ostatnich dekadach miasta na całym świecie zaczęły wdrażać narzędzia służące zwiększaniu udziału obywateli, takie jak budżet obywatelski, platformy

konsultacji społecznych czy lokalne fora deliberacyjne (Sintomer et al., 2012). Jednakże skuteczność tych mechanizmów bywa ograniczona przez niską frekwencję, bariery dostępowe i brak odpowiedniego przetwarzania informacji zwrotnej.

Do kluczowych determinant skuteczności narzędzi wspierających inicjatywy partycypacyjne zaliczyć można *bariery dostępu* – zarówno technologiczną, jak i społeczną. Osoby starsze, mniej wykształcone czy znajdujące się w trudniejszej sytuacji ekonomicznej często mają utrudniony dostęp do narzędzi cyfrowych i rzadziej angażują się w procesy współdecydowania (Smith, 2009). Po drugie, istotną rolę odgrywa *transparentność i zaufanie do procesu*. Partycypacja postrzegana jako fasadowa, pozbawiona realnego wpływu, może prowadzić do zniechęcenia uczestników i spadku frekwencji w kolejnych edycjach (Michels & De Graaf, 2010). Skuteczne narzędzia wymagają jasnych zasad, informacji zwrotnej oraz faktycznego wpływu głosów mieszkańców na decyzje władz. Trzecim czynnikiem jest *projektowanie procesu partycypacyjnego*. Narzędzia muszą być dostosowane do potrzeb lokalnej społeczności, oferować różne formy zaangażowania (online i offline), a także zapewniać odpowiednią moderację i wsparcie organizacyjne (Irvin & Stansbury, 2004). Istotne jest również, by uczestnictwo nie było nadmiernie czasochłonne i nie wymagało specjalistycznej wiedzy. Wreszcie, dużą rolę odgrywają *uwarunkowania instytucjonalne* – m.in. otwartość władz lokalnych, zasoby finansowe, kultura organizacyjna urzędów i priorytety polityczne. Nawet najlepsze narzędzia cyfrowe nie przyniosą efektów, jeśli nie zostaną wbudowane w procesy decyzyjne i realnie zintegrowane z polityką miejską (Löfler, 2004).

Cyfryzacja procesów zarządzania miejskiego stworzyła nowe możliwości dla rozwoju partycypacji obywatelskiej, zarówno na poziomie informacyjnym, jak i decyzyjnym. Początkowo cyfrowe narzędzia pełniły głównie funkcję komunikacyjną – służąc do publikowania dokumentów, ogłoszeń i wyników konsultacji. Z czasem zaczęły się pojawiać bardziej interaktywne rozwiązania, określane mianem *civic tech*, które umożliwiają mieszkańcom aktywny udział w diagnozowaniu potrzeb, zgłaszaniu projektów, głosowaniu czy śledzeniu działań władz (Desouza & Bhagwatwar, 2014). Wyróżnić można kilka typów cyfrowych narzędzi partycypacyjnych:

- platformy zgłaszania problemów i potrzeb – takie jak SeeClickFix, FixMyStreet czy lokalne aplikacje miejskie, pozwalające obywatelom rejestrować usterki lub pomysły dotyczące przestrzeni publicznej;
- budżety partycypacyjne online – systemy umożliwiające składanie projektów przez mieszkańców, ich ocenę oraz głosowanie przez Internet (np. Decidim, Consul, UrbanTalks);
- konsultacje cyfrowe – otwarte fora, sondy, e-ankiety, narzędzia do komentowania dokumentów planistycznych (np. Commonplace, Balancing Act),

wykorzystywane przez władze samorządowe w planowaniu przestrzennym i politykach miejskich;

- deliberacje online – bardziej zaawansowane platformy do dyskusji moderowanej i wspólnego opracowywania rozwiązań (np. Your Priorities, Loomio), często stosowane w eksperymentach demokratycznych.

Ewolucja tych narzędzi zmierza w stronę coraz większego wykorzystania danych i automatyzacji – m.in. dzięki uczeniu maszynowemu, przetwarzaniu języka naturalnego (NLP) czy analizie sentymentu. Przykładowo, systemy AI mogą pomóc samorządom w analizie tysięcy zgłoszeń mieszkańców, identyfikując najczęstsze tematy, wykrywając ukryte wzorce lub wspierając klasyfikację i priorytyzację problemów (Goldsmith & Crawford, 2014).

Jednak – jak zauważają Cardullo i Kitchin (2019) – cyfrowa partycypacja w inteligentnym mieście może przyjąć różne formy: od rzeczywistego współzarządzania po fasadowe „klikodruki”, które tylko pozornie angażują mieszkańców. Dlatego kluczowe znaczenie ma nie tylko dostępność technologii, ale też ich osadzenie w demokratycznych procedurach i realnych możliwościach wpływu. Warto również zaznaczyć, że cyfryzacja nie zawsze prowadzi do poszerzenia uczestnictwa – może wręcz pogłębiać nierówności, jeśli nowe narzędzia nie będą odpowiednio inkluzywne. Dlatego coraz większą uwagę zwraca się na zasady *tech justice*, otwartości kodu, odpowiedzialnego projektowania i konsultacji już na etapie tworzenia platform.

Sztuczna inteligencja staje się integralną częścią zarządzania miastami inteligentnymi (smart cities). Jej zastosowania obejmują zarówno aspekty operacyjne (np. optymalizacja ruchu, gospodarki odpadami, zużycia energii), jak i analityczne – wspierające procesy decyzyjne władz lokalnych poprzez przetwarzanie dużych zbiorów danych (Batty, 2018). W kontekście zarządzania miejskiego AI pełni kilka kluczowych funkcji:

- predykcja i modelowanie – algorytmy uczące się na podstawie danych historycznych wykorzystywane są do przewidywania przeciążeń transportowych, popytu na usługi publiczne czy skutków planowanych interwencji urbanistycznych (Kitchin, 2021);
- automatyzacja procesów administracyjnych – chatboty, systemy OCR, klasyfikatory dokumentów i inteligentne formularze przyczyniają się do usprawnienia obsługi mieszkańców oraz redukcji kosztów administracji (Janssen et al., 2020);
- zarządzanie infrastrukturą w czasie rzeczywistym – dzięki integracji AI z IoT (Internetem rzeczy) miasta mogą automatycznie reagować na zmieniające się warunki środowiskowe, natężenie ruchu czy awarie sieci (Allam & Newman, 2018);
- wspieranie procesów decyzyjnych – narzędzia analityczne oparte na uczeniu maszynowym wspomagają władze w identyfikacji trendów społecznych, lokalizacji problemów społecznych czy w ewaluacji skuteczności polityk publicznych.

Wykorzystanie sztucznej inteligencji w procesach miejskich otwiera nowe możliwości w zakresie partycypacji obywatelskiej, ale jednocześnie niesie ryzyko pogłębienia istniejących nierówności. AI może działać zarówno jako czynnik inkluzji, jak i wykluczenia – w zależności od kontekstu wdrożenia, modelu projektowania i ram instytucjonalnych. Wśród pozytywnych aspektów wiążących się z wykorzystaniem AI wskazać można:

1. Skuteczniejsze dotarcie do zróżnicowanych grup społecznych. Dzięki AI możliwe staje się automatyczne tłumaczenie treści, konwersja głosu na tekst, personalizacja komunikatów czy dostosowanie interfejsów do potrzeb osób z niepełnosprawnościami. To zaś przyczynia się do zwiększenia dostępności procesów konsultacyjnych i skuteczności narzędzi partycypacyjnych (Binns et al., 2018).
2. Skalowalność i efektywność. Algorytmy mogą analizować tysiące zgłoszeń i interwencji obywatelskich w krótkim czasie, identyfikować ukryte wzorce i problemy, które tradycyjne metody mogłyby przeoczyć. Może to prowadzić do lepszego uwzględniania głosu mniejszości oraz szybszego reagowania na potrzeby mieszkańców.
3. Nowe formy reprezentacji. Rozwiązania AI, takie jak przetwarzanie języka naturalnego (NLP), umożliwiają analizę dużych zbiorów opinii wyrażanych w mediach społecznościowych, forach czy konsultacjach cyfrowych, co z kolei pozwala lepiej uchwycić nastroje i postulaty społeczności (van der Voort et al., 2019).

Wśród ewentualnych negatywnych następstw wykorzystywania sztucznej inteligencji w procesach zarządzania rozwojem lokalnym warto zwrócić uwagę na poniższe:

1. Algorytmiczne uprzedzenia i dyskryminacja. Systemy AI „trenowane” na stronicznych danych mogą nieświadomie reprodukcować stereotypy lub marginalizować potrzeby określonych grup społecznych. Przykłady takie obserwowano m.in. w systemach przydziału mieszkań, opieki społecznej czy klasyfikacji wniosków obywatelskich (Eubanks, 2018).
2. Brak przejrzystości (black-boxing). Złożoność algorytmów sprawia, że dla większości obywateli (a często także dla urzędników) mechanizm działania AI pozostaje niejasny. Może to prowadzić do braku zaufania, poczucia wyłączenia z procesu decyzyjnego i osłabienia legitymacji demokratycznej (Ananny & Crawford, 2018).
3. Cyfrowe wykluczenie – AI jako element zaawansowanych narzędzi cyfrowych może pogłębiać barierę technologiczną, zwłaszcza wśród osób starszych, gorzej wykształconych czy mieszkańców obszarów peryferyjnych. Jeśli nie zostaną zapewnione alternatywne kanały uczestnictwa, część obywateli może zostać trwale wykluczona z procesów współdecydowania.

Wnioski z badań nad inkluzją cyfrową wskazują, że technologia sama w sobie nie gwarantuje równości – jej wpływ zależy od sposobu wdrożenia, poziomu

partycypacji w trakcie projektowania rozwiązań oraz od politycznych i społecznych ram wdrożenia. Dlatego coraz większe znaczenie przypisuje się koncepcjom *AI for public good*, *ethical AI* oraz *human-centred AI*, które kładą nacisk na projektowanie rozwiązań zorientowanych na potrzeby wszystkich mieszkańców, a nie jedynie na efektywność systemu. Widać tutaj analogię z kolejnymi generacjami rozumienia pojęcia smart city.

Potencjał zastosowania sztucznej inteligencji we wzmacnianiu partycypacji mieszkańców

Transformacja cyfrowa miast stwarza nowe możliwości dla demokratyzacji procesów decyzyjnych, a sztuczna inteligencja – dzięki zdolności do analizy wielkich zbiorów danych, personalizacji komunikacji i automatyzacji procesów – może znacząco zwiększyć skalę i jakość partycypacji obywatelskiej. Obok sposobów, w jaki AI może wspierać uczestnictwo mieszkańców w zarządzaniu rozwojem miasta, warto rozważyć także potencjał technologiczny, jak i warunki skutecznego wdrożenia omawianych narzędzi. Proponowane zastosowania można pogrupować w cztery główne obszary:

- personalizacja i zwiększenie dostępności konsultacji publicznych,
- wsparcie deliberacji społecznej i przetwarzania opinii mieszkańców,
- wspomaganie władz w podejmowaniu decyzji opartych na głosach obywateli,
- monitorowanie i ewaluacja efektów polityk publicznych z udziałem mieszkańców.

Dla współczesnych samorządów jednym z kluczowych wyzwań jest dotarcie z procesami konsultacyjnymi do jak najszerszego grona mieszkańców – w sposób zrozumiały, inkluzyjny i odpowiadający na różnorodne potrzeby odbiorców. Sztuczna inteligencja może wspierać ten proces poprzez personalizację komunikacji i automatyczne dostosowywanie treści konsultacyjnych do odbiorcy, a także poprzez niwelowanie barier dostępu wynikających z niepełnosprawności, barier językowych, ograniczeń cyfrowych. Takie rozwiązania mogą wspomóc proces budowania u użytkowników poczucia kompetencji, autonomii, zaangażowania (Krishnasamy et al., 2023). Zastosowanie narzędzi opartych na AI umożliwia np. dynamiczne dostosowywanie języka komunikatu do poziomu zrozumienia odbiorcy lub preferencji użytkownika (np. format tekstowy vs głosowy), co zwiększa szanse na aktywne włączenie osób starszych, słabiej wykształconych lub o mniejszej kompetencji cyfrowej (Veale & Brass, 2019). Automatyczne tłumaczenia wspierane przez uczenie maszynowe pozwalają zaś prowadzić konsultacje w wielu językach równolegle, co sprzyja inkluzji migrantów i mniejszości narodowych w wielokulturowych miastach (Anastasopoulos & Whitford, 2019). AI może także zwiększać dostępność konsultacji dla osób

z niepełnosprawnościami poprzez wykorzystanie syntezy mowy, automatycznych napisów, narzędzi asystujących i alternatywnych interfejsów. Zintegrowanie takich rozwiązań z miejskimi platformami konsultacyjnymi sprzyja równości szans w dostępie do informacji i uczestnictwie w procesie decyzyjnym (Binns et al., 2018). Personalizacja możliwa dzięki AI nie ogranicza się do formy komunikatu – obejmuje również treść i kanały jego dostarczania. Algorytmy rekomendacyjne mogą selekcjonować treści istotne dla danej osoby na podstawie lokalizacji, dotychczasowej aktywności lub zainteresowań (Zhang et al., 2020). Dzięki temu mieszkańcy otrzymują powiadomienia o konsultacjach, które faktycznie ich dotyczą – np. planowanych zmianach w ich dzielnicy, lokalnym transporcie czy zagospodarowaniu przestrzennym. Jednakże personalizacja niesie także zagrożenia, takie jak tworzenie „bąbelków informacyjnych” (filter bubbles), w których mieszkańcy otrzymują tylko te treści, które potwierdzają ich wcześniejsze poglądy, z pominięciem opinii alternatywnych lub przeciwnych (Pariser, 2011). W przypadku procesów konsultacyjnych może to prowadzić do utrwalania polaryzacji stanowisk i ograniczania debaty publicznej. Dlatego kluczowe jest projektowanie systemów rekomendacyjnych w sposób przejrzysty, zgodny z zasadą pluralizmu informacyjnego i możliwości świadomego wyboru przez użytkownika. Sztuczna inteligencja może więc w istotny sposób zwiększać zasięg i dostępność konsultacji publicznych, o ile towarzyszyć jej będą odpowiednie zabezpieczenia etyczne, projektowe i instytucjonalne. Wskazane jest oparcie partycypacji i konsultacji na fundamentach organizacji uczących się, szczególnie w zakresie wzmacniania świadomości wartości upowszechniania kompetencji i wiedzy w procesie kształtowania rozwoju danego podmiotu – także jednostki terytorialnej (Szczepańska-Woszczyna, 2018).

Jednym z filarów współczesnych koncepcji demokracji partycypacyjnej jest wspieranie deliberacji obywatelskiej – publicznej wymiany argumentów między uczestnikami życia społecznego. Jednak w praktyce miasta często napotykają trudności w organizowaniu takich procesów na większą skalę, zwłaszcza gdy liczba uczestników przekracza kilkadziesiąt lub kilkaset osób, a wypowiedzi przybierają różne formy: komentarzy pisemnych, nagrań głosowych, wpisów w mediach społecznościowych czy e-maili. W tym kontekście sztuczna inteligencja może odegrać istotną rolę jako narzędzie wspierające analizę, syntezę i wizualizację opinii mieszkańców. Do podstawowych zastosowań AI w tym obszarze zaliczyć należy automatyczne przetwarzanie języka naturalnego (NLP), które pozwala na analizę dużych zbiorów tekstów pod kątem kluczowych tematów, emocji, powtarzalnych wzorców oraz zbieżności i rozbieżności stanowisk (Gupta et al., 2020). Tego typu rozwiązania mogą być stosowane m.in. do klasyfikacji komentarzy mieszkańców, wykrywania kontrowersji, identyfikacji najczęściej pojawiających się problemów oraz generowania raportów tematycznych dla

decydentów. Przykładowo, w Barcelonie w ramach platformy Decidim wdrożono moduły do wizualizacji zgłoszeń mieszkańców, grupowania ich semantycznie oraz mapowania tzw. społecznych gorących punktów – obszarów o największym natężeniu emocjonalnych lub sprzecznych opinii (Morozov, Bria, 2018). Dzięki temu możliwe stało się zrozumienie nie tylko *czego* dotyczą postulaty obywateli, ale także *jak* są formułowane – z jakim ładunkiem emocjonalnym, z jakich pozycji argumentacyjnych oraz jak zmieniają się w czasie. AI może również wspierać moderację deliberacji online, np. poprzez wykrywanie mowy nienawiści, treści spamowych lub języka wykluczającego, a także przez rekomendowanie odpowiedzi moderacyjnych bądź automatyczne streszczanie dyskusji dla nowych uczestników (Németh et al., 2022). Takie zastosowania sprzyjają utrzymaniu wysokiej jakości debaty, bez konieczności ciągłego nadzoru człowieka, co jest szczególnie istotne w rozbudowanych procesach konsultacyjnych. Ważnym obszarem pozostaje także interaktywna wizualizacja danych partycypacyjnych, umożliwiająca prezentowanie zbiorczych wyników opinii mieszkańców w sposób czytelny i angażujący – np. za pomocą dynamicznych map, diagramów zależności czy „chmur argumentów”. Takie narzędzia mogą być użyte zarówno przez mieszkańców (np. w celu porównania opinii z różnych dzielnic), jak i przez urzędników miejskich – jako podstawa do dalszej deliberacji eksperckiej lub politycznej (Kaufmann et al., 2022). Zastosowanie AI w analizie deliberacji nie może jednak zastępować ludzkiej refleksji – przeciwnie, jego celem jest wspomaganie rozumienia złożonych zbiorów opinii, przy jednoczesnym zachowaniu pluralizmu i możliwości autentycznej rozmowy. Wymaga to dbałości o przejrzystość działania algorytmów, możliwość ręcznej kontroli wyników oraz eliminowania uprzedzeń (biasów), które mogą występować w danych lub w modelach analizy.

W coraz bardziej złożonym środowisku miejskim, w którym decyzje administracyjne muszą uwzględniać nie tylko dane techniczne i finansowe, ale także oczekiwania mieszkańców, sztuczna inteligencja może pełnić rolę pośrednika między obywatelami a decydentami. Jej potencjał nie ogranicza się jedynie do zbierania czy przetwarzania opinii; stanowiąc wsparcie dla syntetycznego ujęcia danych partycypacyjnych w procesach decyzyjnych, pozwala władzom lepiej rozumieć społeczne priorytety i preferowane scenariusze rozwoju. AI może służyć do agregowania danych jakościowych i ilościowych z różnych źródeł – w tym wyników konsultacji publicznych, zgłoszeń z aplikacji miejskich, debat online, petycji oraz danych z platform społecznościowych. Dzięki technikom klasyfikacji, grupowania tematycznego, analizy sentymentu czy analizy predykcyjnej możliwe jest generowanie syntetycznych „map preferencji społecznych” dla konkretnych obszarów tematycznych (Jungherr & Jürgens, 2021). Takie mapy mogą stanowić wsparcie dla urbanistów, zarządców budżetów obywatelskich czy

planistów mobilności. Ponadto AI może być wykorzystywana w modelowaniu scenariuszy decyzyjnych z uwzględnieniem opinii mieszkańców jako jednego z istotnych parametrów. Przykładowo, w procesie wyboru lokalizacji dla nowej inwestycji infrastrukturalnej algorytm może uwzględnić zarówno dane demograficzne i środowiskowe, jak i opinie mieszkańców zgłoszone w toku konsultacji, wskazując rozwiązania, które minimalizują konflikty społeczne i zwiększają akceptację projektu (Kitchin, 2021). Coraz częściej stosowane są także systemy wspomagania decyzji (DSS) z komponentem AI, które pozwalają decydentom eksplorować różne warianty działań wraz z ich oceną pod względem m.in. społecznego poparcia, kosztów wdrożenia czy wpływu na różne grupy interesariuszy (Wirtz et al., 2019). Co ważne, nie chodzi tu o automatyczne podejmowanie decyzji, lecz o dostarczenie syntetycznej, zrozumiałej i wspartej danymi podstawy do rozważań i deliberacji na poziomie administracyjnym oraz politycznym. Zastosowanie AI w tym kontekście może również przyczynić się do większej przejrzystości procesu decyzyjnego – poprzez dokumentowanie, w jaki sposób dane obywatelskie zostały wzięte pod uwagę, jakie argumenty miały największe znaczenie i które kompromisy były konieczne. To z kolei może wzmacniać zaufanie do instytucji publicznych, zwłaszcza jeśli rezultaty analiz są udostępniane publicznie w zrozumiałej formie. Istnieje jednak ryzyko, że zbyt silne poleganie na danych i rekomendacjach algorytmicznych może prowadzić do technokracji procesu decyzyjnego i marginalizacji politycznego wymiaru sporów społecznych. AI nie zastępuje decyzji politycznej – może pełnić wyłącznie funkcję narzędzia wspierającego, dostarczającego informacji i analiz. Dlatego coraz więcej autorów podkreśla potrzebę projektowania „wspomagających”, a nie „zastępujących” modeli podejmowania decyzji (Wagner et al., 2022).

Skuteczne zarządzanie rozwojem miasta wymaga nie tylko projektowania i wdrażania polityk, lecz także ich systematycznego monitorowania i ewaluacji – z uwzględnieniem głosu mieszkańców jako podmiotu polityki publicznej. Sztuczna inteligencja może w tym zakresie odegrać kluczową rolę jako narzędzie wspierające procesy gromadzenia, analizowania i interpretowania danych zwrotnych, zarówno ilościowych, jak i jakościowych. W tradycyjnym ujęciu ewaluacja polityk publicznych opierała się głównie na wskaźnikach twardej (np. liczba wybudowanych obiektów, poziom wykorzystania infrastruktury, wzrost PKB per capita). Współczesne podejścia, zgodne z ideą zarządzania partycypacyjnego, postulują jednak uwzględnianie doświadczeń, ocen i emocji mieszkańców jako integralnej części procesu oceny skutków interwencji miejskich (Bryson et al., 2013). AI może wspierać ten proces poprzez analizę danych tekstowych pochodzących z różnorodnych kanałów: mediów społecznościowych, komentarzy pod projektami, systemów zgłoszeń miejskich, e-maili czy otwartych odpowiedzi w ankietach. Techniki, takie jak analiza sentymentu,

wykrywanie trendów czy mapowanie nastrojów społecznych, pozwalają uchwycić zmiany w postrzeganiu danego projektu lub polityki w czasie oraz w różnych grupach mieszkańców (Young & Bleakley, 2020). Przykładowo, po zakończeniu inwestycji transportowej można zastosować AI do analizy opinii mieszkańców w mediach społecznościowych i na platformach komunikacji miejskiej, monitorując zmiany w tonie wypowiedzi, częstość zgłaszanych problemów czy poziom satysfakcji w poszczególnych dzielnicach. Takie podejście – określane jako „partycypacja postdecyzyjna” (post-decisional participation) – pozwala uwzględniać obywateli nie tylko w fazie projektowania, ale także w cyklu oceny i korekty polityki (Wesselink et al., 2011). AI może być również wykorzystywana do predykcyjnej ewaluacji scenariuszy przyszłych działań, w oparciu o dane historyczne i bieżące. Przykładowo, algorytmy uczenia maszynowego mogą prognozować, jak dana polityka wpłynie na określone grupy społeczne – co może pomóc w ocenie ryzyk społecznych i zaprojektowaniu działań kompensacyjnych. Ważną rolę AI w ewaluacji jest także uwidacznianie niewidocznych wcześniej wzorców, np. systematycznie pomijanych grup interesariuszy, terytorialnych nierówności w skutkach polityk czy niespójności między celami a efektami wdrożeń. W tym sensie sztuczna inteligencja może pełnić funkcję nie tylko diagnostyczną, ale i krytyczną – o ile jej zastosowania nie są podporządkowane jedynie logice efektywności, lecz także wartościom sprawiedliwości społecznej, przejrzystości i uczestnictwa. Jednocześnie należy pamiętać, że skuteczna ewaluacja z wykorzystaniem AI wymaga odpowiednich ram instytucjonalnych: systematycznego gromadzenia danych, interoperacyjnych systemów informacyjnych oraz umiejętności analitycznych w administracji publicznej. Bez tych warunków istnieje ryzyko, że analizy AI będą ograniczone do fasadowych raportów lub służyć będą potwierdzaniu wcześniej podjętych decyzji. Dobrym punktem odniesienia mogą być w tym kontekście procesy realizowane w sferze biznesu, gdzie narzędzia cyfrowe i ich skuteczność podlegają ciągłemu monitoringowi i doskonaleniu, także w stałym kontakcie z klientem i partnerami biznesowymi (Poszytek i in., 2024; Lis, 2023).

Ryzyka i ograniczenia wykorzystania sztucznej inteligencji w partycypacji mieszkańców

Choć sztuczna inteligencja oferuje wiele możliwości w zakresie wspierania partycypacji obywatelskiej, jej implementacja w zarządzaniu miastem nie jest wolna od ryzyk i ograniczeń. Co więcej, wprowadzanie AI do procesów demokratycznych może – paradoksalnie – prowadzić do pogłębienia nierówności, osłabienia zaufania społecznego czy erozji odpowiedzialności politycznej, jeśli nie zostaną zachowane odpowiednie ramy etyczne, prawne i instytucjonalne (Zuboff, 2019;

Eubanks, 2018). W tym rozdziale przedstawiono kluczowe obszary zagrożeń, jakie należy rozważyć przy wdrażaniu narzędzi AI w kontekście uczestnictwa obywateli w zarządzaniu rozwojem miast.

Jednym z najczęściej wymienianych problemów związanych z AI jest brak przejrzystości działania algorytmów. Większość modeli, zwłaszcza tych opartych na głębokim uczeniu (deep learning), cechuje się bardzo ograniczoną zrozumiałością dla przeciętnego użytkownika – zarówno mieszkańca, jak i urzędnika (Burrell, 2016). Zjawisko to określane jest często mianem *black-boxing*, ponieważ użytkownicy widzą jedynie wynik działania algorytmu, bez możliwości zrozumienia, jakie czynniki i zależności doprowadziły do konkretnego rezultatu. W kontekście partycypacji oznacza to, że mieszkańcy mogą nie mieć wiedzy, w jaki sposób ich opinie są analizowane, grupowane czy interpretowane. Może to prowadzić do poczucia braku kontroli i podważenia zaufania do procesów konsultacyjnych. Brak transparentności utrudnia też ewentualne zakwestionowanie decyzji lub wyciągnięcie konsekwencji wobec błędnych lub uprzedzonych analiz.

AI, podobnie jak inne narzędzia analityczne, działa w oparciu o dane – a te często są niepełne, zniekształcone lub historycznie obciążone uprzedzeniami. Jeśli dane wejściowe odzwierciedlają nierówności społeczne, stereotypy czy luki w reprezentacji, modele uczenia maszynowego mogą te zniekształcenia reprodukować lub nawet wzmacniać (O’Neil, 2016). Przykładem może być niedoszacowanie głosu mniejszości etnicznych, mieszkańców peryferyjnych dzielnic czy osób starszych, jeśli ich wypowiedzi są słabiej reprezentowane w danych treningowych.

Problem algorytmicznego biasu staje się szczególnie istotny w sytuacjach, gdy AI wspiera podejmowanie decyzji lub ocenia poziom społecznego poparcia dla różnych rozwiązań. Jeśli nie zostaną zastosowane mechanizmy kompensujące brak reprezentatywności danych, system może wzmacniać istniejące asymetrie w dostępie do władzy i wpływu.

Wdrażanie AI w obszarze zarządzania miastem często motywowane jest chęcią zwiększenia efektywności, redukcji kosztów lub przyspieszenia procesów. Choć są to cele zrozumiałe, mogą prowadzić do nadmiernej automatyzacji procedur partycypacyjnych i spłycenia ich wymiaru deliberacyjnego. Zamiast pogłębionego dialogu i wymiany argumentów, proces może ograniczyć się do zbierania danych i generowania raportów – bez rzeczywistego wpływu obywateli na treść decyzji (Aneesh, 2006). W efekcie partycypacja przekształca się w proces technokratyczny, w którym to nie społeczna legitymacja, lecz obliczenia algorytmiczne stają się głównym kryterium oceny rozwiązania. Może to prowadzić do osłabienia politycznej roli obywateli i obniżenia jakości demokracji lokalnej.

Tabela 1. Główne ryzyka związane z wykorzystaniem AI w partycypacji mieszkańców oraz środki zaradcze

Obszar ryzyka	Opis ryzyka	Możliwe konsekwencje	Proponowane środki zaradcze
Nietransparentność algorytmów („czarne skrzynki”)	Brak zrozumiałości działania algorytmów dla mieszkańców i urzędników	Spadek zaufania, brak kontroli społecznej, osłabienie legitymacji decyzji	Użycie modeli możliwych do wyjaśnienia (explainable AI), publikacja dokumentacji algorytmów, audyty algorytmiczne
Algorytmiczne uprzedzenia (bias)	Modele uczone na nierównomiernych lub zniekształconych danych	Reprodukcja nierówności społecznych, marginalizacja grup niedoreprezentowanych	Testowanie modeli pod kątem biasów, stosowanie danych kompensacyjnych, nadzór etyczny
Nadmierna automatyzacja	Ograniczenie roli deliberacji i debaty na rzecz zbierania danych i rekomendacji algorytmicznych	Spłylenie partycypacji, osłabienie polityczności, technokratyzacja procesów	Zachowanie „czynnika ludzkiego” w decyzjach, łączenie AI z procesami deliberacyjnymi i moderacją społeczną
Zagrożenia dla prywatności	Zbieranie i przetwarzanie dużych ilości danych osobowych i behawioralnych	Ryzyko naruszenia RODO, utrata zaufania, potencjalne nadużycia	Minimalizacja zakresu pozyskiwanych danych, anonimizacja danych, transparentne polityki prywatności, nadzór instytucji ochrony danych
Uzależnienie od dostawców technologii	Wdrożenie rozwiązań opartych na zamkniętych systemach prywatnych firm	Brak kontroli nad kodem i danymi, ryzyko komercjalizacji, wysokie koszty utrzymania	Wspieranie open-source, tworzenie własnych systemów publicznych, tworzenie konsorcjów miast lub instytucji publicznych
Cyfrowe wykluczenie	Brak dostępu do narzędzi AI i kanałów cyfrowych dla części mieszkańców	Utrata głosu przez grupy wykluczone (np. seniorów, osoby ubogie, osoby z niepełnosprawnościami)	Zapewnienie alternatywnych kanałów partycypacji (offline, lokalnych), działania edukacyjne i wspierające

Źródło: opracowanie własne.

Rekomendacje dla samorządów i projektantów systemów AI

Aby wykorzystanie sztucznej inteligencji w partycypacji obywatelskiej wspierało demokrację lokalną, a nie ją osłabiało, konieczne jest przyjęcie odpowiednich ram projektowych, organizacyjnych i etycznych. Poniższe rekomendacje odnoszą się zarówno do instytucji publicznych, jak i do twórców oraz dostawców narzędzi opartych na AI – wskazując, jak projektować systemy służące dobru wspólnemu, przejrzystości i inkluzji społecznej.

- Zasada „human-in-the-loop”: AI jako wsparcie, nie substytut – w systemach wspierających partycypację obywatelską AI powinna pełnić rolę wzmacniającą, a nie zastępującą procesy demokratyczne; oznacza to, że człowiek – zarówno urzędnik, jak i obywatel – musi pozostać w centrum podejmowania decyzji, mając realną możliwość interpretacji, korekty lub zakwestionowania wyników generowanych przez system (Floridi et al., 2018). AI powinna wspierać deliberację i ułatwiać przetwarzanie informacji, ale nie przejmować władzy decyzyjnej ani osłabiać społecznej odpowiedzialności.
- Projektowanie inkluzywne i równościowe – systemy AI wspierające partycypację powinny być projektowane w sposób uwzględniający różnorodność użytkowników ze względu na wiek, pochodzenie, poziom kompetencji cyfrowych, język, czy sprawność fizyczną. Oznacza to konieczność testowania narzędzi z udziałem grup narażonych na wykluczenie, projektowania wielokanałowego (zarówno online, jak i offline), a także wdrażania mechanizmów równoważących niedoreprezentowane grupy w danych treningowych (Costanza-Chock, 2020).
- Przejrzystość i wyjaśnialność algorytmów – wszyscy użytkownicy, zarówno mieszkańcy, jak i pracownicy administracji, powinni mieć jasność co do działania wykorzystywanych narzędzi AI: jakie dane są zbierane, jak są przetwarzane, jakie decyzje mogą z nich wynikać i kto ponosi za nie odpowiedzialność. Wymaga to stosowania tzw. *explainable AI* (XAI), publikowania modeli działania, umożliwiania audytu oraz wdrażania polityk zarządzania algorytmami zgodnych z zasadą *accountability* (Wieringa, 2020).
- Otwarta infrastruktura i współpraca międzysektorowa – aby uniknąć uzależnienia od komercyjnych dostawców technologii oraz umożliwić rozwój ekosystemów opartych na wiedzy publicznej, zaleca się preferowanie otwartych, modularnych i transparentnych rozwiązań technologicznych; samorządy mogą współpracować z uniwersytetami, organizacjami pozarządowymi i innymi miastami w ramach konsorcjów opracowujących otwarte standardy oraz wspólne zasoby (np. biblioteki NLP, systemy moderacji). Przykłady takich działań można znaleźć w inicjatywach Code for America, Decidim czy EU Smart Cities Marketplace.

- Cykl życia systemu i monitoring społeczny – wdrażanie AI powinno być traktowane nie jako jednorazowy projekt, ale jako proces cykliczny, podlegający ocenie i korekcie; należy zapewnić ciągłe monitorowanie efektów działania systemu, zarówno technicznych, jak i społecznych – np. poprzez niezależne rady ds. etyki danych, mechanizmy skargowe, publiczne raporty z funkcjonowania algorytmów czy ewaluacje *ex post*. W proces ten warto włączyć obywateli, dając im realny wpływ na kształt i rozwój narzędzi, które ich dotyczą (Aelenei et al., 2022).
- Edukacja cyfrowa i wzmacnianie kompetencji obywatelskich – rozwój technologii wymaga równoległego rozwoju świadomości obywatelskiej i kompetencji cyfrowych. Władze lokalne powinny wspierać mieszkańców w rozumieniu i krytycznym korzystaniu z narzędzi AI poprzez działania edukacyjne – np. warsztaty, materiały wyjaśniające, programy w szkołach i bibliotekach; wiedza o tym, jak działa AI, jakie ma ograniczenia i gdzie można zgłaszać zastrzeżenia to klucz do odpowiedzialnej i świadomej partycypacji w warunkach cyfrowej transformacji.

Podsumowanie

Współczesne miasta stają przed koniecznością pogodzenia rosnącej złożoności zarządzania z potrzebą rzeczywistego włączenia obywateli w procesy decyzyjne. Sztuczna inteligencja, odpowiednio zaprojektowana i wdrażana, może znacząco zwiększyć skuteczność, dostępność i responsywność partycypacji obywatelskiej – wspierając konsultacje, analizę opinii, deliberację oraz monitorowanie polityk publicznych. Przykłady z miast na całym świecie pokazują, że AI może nie tylko usprawniać procesy, lecz także przyczyniać się do ich demokratyzacji. Jednocześnie wdrażanie technologii AI w obszarze partycypacji wiąże się z poważnymi wyzwaniami: od nietransparentności algorytmów i algorytmicznych uprzedzeń, po ryzyko marginalizacji społecznego dialogu i zagrożeń dla prywatności. Dlatego kluczowe znaczenie mają nie tylko możliwości technologiczne, ale także ramy etyczne, instytucjonalne i edukacyjne towarzyszące ich implementacji. Zasadniczym wnioskiem płynącym z przeprowadzonej analizy jest to, że AI może wspierać bardziej inkluzywne, sprawiedliwe i skuteczne zarządzanie miejskie – pod warunkiem, że zostanie zakorzeniona w wartościach demokracji, przejrzystości i współodpowiedzialności. Nie chodzi o zastąpienie ludzkiego wymiaru demokracji przez technologię, lecz o jego wzmocnienie – w oparciu o nowe narzędzia, ale z zachowaniem fundamentalnych zasad uczestnictwa i dialogu.

Bibliografia

- Aelenei, D., Novelli, E., Schmitz, A., & Sangiovanni-Vincentelli, A. (2022). Trustworthy AI for Smart Cities: Principles, Challenges, and Directions. *IEEE Access*, 10, 106592–106605.
- Allam, Z., & Newman, P. (2018). Redefining the Smart City: Culture, Metabolism and Governance. *Smart Cities*, 1(1), 4–25.
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989.
- Anastasopoulos, L. J., & Whitford, A. B. (2019). Machine learning for public administration research, with application to organizational reputation. *Journal of Public Administration Research and Theory*, 29(3), 491–510.
- Aneesh, A. (2006). *Virtual Migration: The Programming of Globalization*. Duke University Press.
- Arnstein, S. R. (1969). A Ladder of Citizen Participation. *Journal of the American Institute of Planners*, 35(4), 216–224.
- Batty, M. (2018). *Inventing Future Cities*. MA: MIT Press.
- Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). “It’s Reducing a Human Being to a Percentage”: Perceptions of Justice in Algorithmic Decisions. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–14.
- Bryson, J. M., Quick, K. S., Slotterback, C. S., & Crosby, B.C. (2013). Designing Public Participation Processes. *Public Administration Review*, 73(1), 23–34.
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12.
- Cardullo, P., & Kitchin, R. (2019). Being a ‘citizen’ in the smart city: Up and down the scaffold of smart citizen participation. *GeoJournal*, 84(1), 1–13.
- Costanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need*. MA: MIT Press.
- Dacko-Pikiewicz, Z., Szczepańska-Woszczyna, K., & Lis, M. (red.). (2023). *Horyzonty sztucznej inteligencji a Przemysł 5.0*. Wydawnictwo Naukowe AWSB w Dąbrowie Górniczej.
- Desouza, K. C., & Bhagwatwar, A. (2014). Citizen apps to solve complex urban problems. *Journal of Urban Technology*, 21(4), 107–135.
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin’s Press.
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V. et al. (2018). AI4People – An Ethical Framework for a Good AI Society. *Minds and Machines*, 28(4), 689–707.
- Fung, A. (2006). Varieties of Participation in Complex Governance. *Public Administration Review*, 66, 66–75.

- Goldsmith, S., & Crawford, S. (2014). *The Responsive City: Engaging Communities Through Data-Smart Governance*. Jossey-Bass.
- Gupta, P., Bain, M., Zhang, C., Lasecki, W. S., & Singh, M. (2020). *Analyzing Open Comments for Public Sector Decision-Making: Topic Modeling and Deliberation Insights*. Proceedings of the 21st Annual International Conference on Digital Government Research, 369–378.
- Irvin, R. A., & Stansbury, J. (2004). Citizen participation in decision making: Is it worth the effort? *Public Administration Review*, 64(1), 55–65.
- Janssen, M., Matheus, R., & Zuiderwijk, A. (2020). Artificial Intelligence: Conceptual, technological and societal perspectives. In A. Meijer, K. Boersma and P. Wagenaar (eds.), *ICTs, Citizens and Governance: After the Hype!* (pp. 27–45). Springer.
- Jungherr, A., & Jürgens, P. (2021). Data-driven decision-making in the public sector: Artificial intelligence and democratic values. *Government Information Quarterly*, 38(1), 101547.
- Kaufmann, D., Götz, F. M., Haim, M., Legler, B., & Hurrelmann, K. (2022). Deliberation in the Digital Age: How to Use Data to Foster Democratic Participation. *Policy & Internet*, 14(3), 457–478.
- Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1–14.
- Kitchin, R. (2021). *Data Lives: How Data Are Made and Shape Our World*. Policy Press.
- Krishnasamy, D., Anandhan, H., Dacko-Pikiewicz, Z., & Kot-Radojewska, M. (2023). Culture of educational institutions on psychological empowerment of women employees in higher education institutions. *Cultural Management: Science and Education*, 7(2), 63–77.
- Lis, M. (2023). Transformacja cyfrowa jako nowe wyzwanie dla organizacji i jej otoczenia. W M. Lis & M. Bryke (red.), *Transformacja KAIZEN w perspektywie nowych technologii i cyfryzacji* (ss. 7–47). Wydawnictwo Naukowe AWSB w Dąbrowie Górniczej.
- Löffler, E. (2004). Governance and Government: Networking with Citizens and Civil Society. In E. Löffler & T. Bovaird (eds.), *Public Management and Governance* (pp. 163–179). Routledge.
- Michels, A., & De Graaf, L. (2010). Examining Citizen Participation: Local Participatory Policy Making and Democracy. *Local Government Studies*, 36(4), 477–491.
- Morozov, E., & Bria, F. (2018). *Rethinking the Smart City: Democratizing Urban Technology*. Rosa Luxemburg Stiftung.
- Nemitz, P. F. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Royal Society Philosophical Transactions A*, 376(2133), 20180089.
- Németh, R., Mocanu, D., & Vespignani, A. (2022). Automated moderation in online deliberation platforms: Opportunities and limits. *Information, Communication & Society*, 25(1), 123–141.
- OECD (2020). *The OECD Artificial Intelligence Principles: OECD Recommendation of the Council on Artificial Intelligence*. Paris: OECD Publishing. <https://www.oecd.org/going-digital/ai/principles/> [dostęp: 06.06.2025]

- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.
- Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.
- Poszytek, P., Lis, M., Jefmański, B., Fila, J., Jeżowski, M., & Kotelska, J. (2024). *Transformacja cyfrowa przedsiębiorstw w dobie Przemysłu 4.0*. Wydawnictwo Naukowe Akademii WSB.
- Sintomer, Y., Herzberg, C., & Röcke, A. (2012). Transnational Models of Citizen Participation: The Case of Participatory Budgeting. *Journal of Public Deliberation*, 8(2), Article 9.
- Smith, G. (2009). *Democratic Innovations: Designing Institutions for Citizen Participation*. Cambridge University Press.
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368–383.
- Szczepańska-Woszczyzna, K. (2018). *Zarządzanie rozwojem organizacji oparte na kompetencjach i wiedzy*. Wydawnictwo Difin.
- van der Voort, H., Klievink, B., & Arnaboldi, M. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, 36(1), 27–38.
- Veale, M., & Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning. *Public Money & Management*, 39(5), 321–326.
- Viljanen, M., Berndtsson, J., & Gustafsson, N. (2021). Algorithmic democracy: Imaginaries and practices of governing society with data and algorithms. *Big Data & Society*, 8(2), 1–14.
- Wagner, B., Hekkert, M. P., & Negro, S. O. (2022). Augmenting decision-making with AI: How artificial intelligence reshapes roles and responsibilities in policy-making. *Technology in Society*, 69, 101945.
- Wesselink, A., Paavola, J., Fritsch, O., & Renn, O. (2011). Rationales for public participation in environmental policy and governance: practitioners' perspectives. *Environment and Planning A*, 43(11), 2688–2704.
- Wieringa, M. (2020). *What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability*. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*), 1–18.
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration*, 42(7), 596–615.
- Young, M., & Bleakley, A. (2020). Data, algorithms and public value: Why public administrators need to understand machine learning. *Public Administration*, 98(1), 68–87.
- Zhang, B., Sheng, Q. Z., Alhazmi, A., Li, X., Zhang, W., & Yao, L. (2020). Recommender systems in smart cities: State of the art and future directions. *Information Systems*, 92, 101500.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

dr Monika Płońska¹

dr Grzegorz Kądziaławski²

dr hab. Marcin Lis, prof. AWSB³

AI i blockchain w zarządzaniu systemami certyfikacji offsetów węglowych

Streszczenie

Offsety węglowe, będące narzędziem służącym do redukcji emisji gazów cieplarnianych, odgrywają ważną rolę w walce z globalnym ociepleniem. Proces ich certyfikacji wymaga precyzyjnych analiz, co implikuje wyzwania związane z monitorowaniem, weryfikacją i raportowaniem emisji. W ostatnich latach sztuczna inteligencja (AI) oraz blockchain zyskują na znaczeniu jako narzędzia wspierające zarządzanie systemami certyfikacji offsetów węglowych. Niniejszy artykuł przedstawia przegląd aktualnych technologii opartych na AI i blockchainie, stosowanych w tym obszarze, oraz perspektywy ich dalszego rozwoju.

Słowa kluczowe: offsety węglowe, sztuczna inteligencja, blockchain, certyfikaty, zarządzanie, technologia

1 Akademia WSB w Dąbrowie Górniczej

2 Akademia WSB w Dąbrowie Górniczej

3 Akademia WSB w Dąbrowie Górniczej

AI and Blockchain in the Management of Carbon Offset Certification Systems

Abstract

Carbon offsets, as a tool for reducing greenhouse gas emissions, play an important role in combating global warming. The certification process requires precise analyses, which poses challenges related to the monitoring, verification, and reporting of emissions. In recent years, artificial intelligence (AI) and blockchain have gained significance as tools supporting the management of carbon offset certification systems. This article provides an overview of current technologies based on AI and blockchain used in this area, as well as perspectives for their further development.

Keywords: carbon offsets, artificial intelligence, blockchain, certification, management, technology

Wstęp

Offsety węglowe, rozumiane jako mechanizm kompensujący emisje gazów cieplarnianych (GHG) poprzez inwestycje w projekty redukujące lub pochłaniające CO₂, stanowią istotny element strategii osiągnięcia celów klimatycznych, zarówno na poziomie korporacyjnym, jak i państwowym. Choć ich potencjał w dekarbonizacji gospodarki może być znaczny, skuteczność offsetów uzależniona jest od wiarygodności systemów ich certyfikacji, które muszą zapewniać przejrzystość, trwałość, dodatkowość oraz odporność na nadużycia.

Proces certyfikacji offsetów węglowych obejmuje złożone etapy pomiaru, raportowania i weryfikacji (MRV – Monitoring, Reporting and Verification), które wymagają precyzyjnych danych oraz nadzoru ekspertów. W praktyce jednak systemy MRV są kosztowne, czasochłonne i narażone na błędy ludzkie lub manipulacje, co wpływa na wiarygodność offsetów na rynku dobrowolnym (VCM – Voluntary Carbon Market).

W ostatnich latach dynamiczny rozwój technologii cyfrowych, zwłaszcza sztucznej inteligencji (AI) oraz technologii rozproszonych rejestrów (blockchain), przynosi nowe możliwości automatyzacji i zwiększenia wiarygodności systemów certyfikacji. AI umożliwia zaawansowaną analizę obrazów satelitarnych, danych klimatycznych i informacji geoprzestrzennych, wspomagając weryfikację pochłaniania CO₂ i wykrywanie nieprawidłowości. Z kolei blockchain oferuje transparentność i niezmienność zapisów, co ma kluczowe znaczenie w budowaniu zaufania do systemów offsetowych oraz audytowalności danych.

Celem niniejszego artykułu jest przedstawienie przeglądu aktualnych rozwiązań technologicznych opartych na AI i blockchainie, stosowanych w systemach certyfikacji offsetów węglowych. Analizie poddane zostaną konkretne zastosowania (np. analiza biomasy, śledzenie danych o sekwestracji węgla, cyfrowe tokeny offsetowe), a także wyzwania i perspektywy dalszego rozwoju tych technologii w świetle zmieniających się regulacji, takich jak unijne rozporządzenie CRCF (Carbon Removals Certification Framework) czy wytyczne SBTi i GHG Protocol.

Wyzwania związane z transparentnością offsetów węglowych

Offsety węglowe (carbon offsets) to mechanizm, w ramach którego jedna jednostka emisji gazów cieplarnianych (najczęściej jedna tona CO₂e) jest kompensowana przez równoważną redukcję lub usunięcie emisji w innym miejscu. Najczęściej offsety generowane są poprzez projekty związane z odnawialnymi źródłami energii, zalesianiem, ochroną lasów (np. REDD+), efektywnością energetyczną czy sekwestracją węgla w glebie (Kollmuss et al., 2008). Offsety odgrywają ważną rolę

w realizacji dobrowolnych strategii Net Zero przedsiębiorstw oraz w niektórych systemach obowiązkowych (np. lotnictwo – CORSIA).

W teorii offsety mają umożliwić organizacjom neutralizację trudnych do uniknięcia emisji poprzez finansowanie projektów redukcyjnych w innych sektorach lub regionach. W praktyce jednak ich skuteczność i uczciwość wzbudza szereg wątpliwości. Jak wskazują autorzy raportu OECD (2022), offsety mogą wspierać cele redukcji emisji, ale wymagają wyraźnych zasad i transparentności. Tymczasem badania pokazują, że wiele offsetów oferowanych na rynku nie spełnia kluczowych kryteriów jakościowych, takich jak dodatkowość (czy redukcje miałyby miejsce bez projektu), trwałość (czy sekwestracja jest długoterminowa) czy unikalność (czy dana emisja nie została policzona dwa razy) (Cames et al., 2016; Battocletti et al., 2023).

Właśnie te braki przyczyniły się do oskarżeń o greenwashin, czyli fałszywe lub przesadzone deklaracje o działaniach proklimatycznych. Firmy deklarujące neutralność klimatyczną często opierają się głównie na zakupie tanich offsetów, bez realnych działań redukcyjnych w swojej działalności (Anderson & Peters, 2016). Skrajne przypadki, takie jak kredyty z projektów REDD+, które nie zapobiegły realnemu wylesianiu, podważyły zaufanie do całego rynku (Gupta et al., 2020).

W odpowiedzi na te wyzwania pojawiają się nowe rozwiązania technologiczne, które mogą zwiększyć wiarygodność offsetów. Przede wszystkim sztuczna inteligencja i uczenie maszynowe mogą wspierać automatyczne systemy weryfikacji danych (MRV), analizując dane z sensorów, zdjęć satelitarnych i obrazowania hiperspektralnego w czasie rzeczywistym (Sood, 2025). Z kolei blockchain umożliwia stworzenie transparentnego i niezmiennego rejestru offsetów, eliminując ryzyko podwójnego liczenia i fałszerstw, a inteligentne kontrakty pozwalają na tokenizację jednostek redukcji emisji i automatyzację ich rozliczania (Springer, 2025; ScienceDirect, 2025).

W efekcie odpowiedzią na zarzuty greenwashingu i brak zaufania do offsetów mogą być właśnie nowoczesne, cyfrowe systemy ich generowania, walidacji i raportowania. Technologia – choć sama nie rozwiąże kryzysu zaufania – może uczynić rynek offsetów bardziej wiarygodnym, przejrzystym i zgodnym z celami klimatycznymi.

Monitorowanie emisji i absorpcji dwutlenku węgla

Monitorowanie emisji i pochłaniania dwutlenku węgla (CO₂) w ramach projektów offsetowych wymaga gromadzenia dużych ilości danych na temat zmian w środowisku, takich jak: wzrost roślinności, kondycja lasów czy aktywność przemysłowa. Tradycyjne metody pomiaru, oparte na bezpośrednich obserwacjach

i ręcznym wprowadzaniu danych, są czasochłonne i narażone na błędy ludzkie. Algorytmy sztucznej inteligencji, szczególnie te z zakresu uczenia maszynowego (ML) oraz głębokiego uczenia (DL), mogą automatycznie analizować dane satelitarne oraz dane z czujników, co znacznie zwiększa precyzję pomiarów.

Algorytmy głębokiego uczenia, takie jak konwolucyjne sieci neuronowe (CNN), są wykorzystywane do automatycznego rozpoznawania wzorców na zdjęciach satelitarnych, co pozwala na precyzyjne monitorowanie zmian w powierzchni lasów i innych ekosystemów odpowiedzialnych za absorpcję CO₂. W badaniach wykazano, że technologie oparte na AI mogą zidentyfikować zmiany w biomasy lasów, które są kluczowe dla oceny ilości dwutlenku węgla pochłanianego przez ekosystemy (Lovett, 2025).

Według badań Cho i Lee (2020) wykorzystanie technologii opartej na sztucznej inteligencji pozwala na bieżące monitorowanie zmian w poziomach CO₂ z dokładnością wyższą o 30% w porównaniu do tradycyjnych metod. Dzięki analizie obrazów satelitarnych z wykorzystaniem algorytmów przetwarzania obrazów możliwe jest określenie stopnia pochłaniania dwutlenku węgla przez lasy oraz inne naturalne ekosystemy, co jest kluczowe w ocenie efektywności projektów offsetowych.

Z kolei zastosowanie sieci czujników, w połączeniu z Internetem rzeczy (IoT) i sztuczną inteligencją, umożliwia monitorowanie emisji CO₂ w czasie rzeczywistym. Czujniki te zbierają dane o stężeniach CO₂ w atmosferze, a algorytmy AI analizują je, identyfikując źródła emisji i oceniając ich intensywność. Wykorzystanie AI pozwala na szybszą analizę ogromnych ilości danych, co jest kluczowe w przypadku projektów z rozbudowaną infrastrukturą monitorującą (Sarkar et al., 2023).

Automatyzacja procesu certyfikacji

Certyfikacja offsetów węglowych jest skomplikowanym procesem, wymagającym oceny wielu aspektów związanych z wpływem projektu na środowisko. Obecnie rozwijane są narzędzia bazujące na sztucznej inteligencji, które automatyzują kluczowe elementy tego procesu. Przykładem może być zastosowanie technologii blockchain wspieranej przez AI do śledzenia projektów offsetowych, co zwiększa transparentność i umożliwia weryfikację danych w czasie rzeczywistym.

W pracy Deng i wsp. (2021) opisano, że połączenie blockchain z AI pozwala na zmniejszenie ryzyka fałszowania danych oraz manipulacji wynikami pomiarów, co stanowiło jedno z głównych wyzwań w zarządzaniu systemami certyfikacji offsetów. Technologia blockchain, zapewniająca niezmiennność danych, w połączeniu z algorytmami AI do analizy tych danych, daje możliwość automatyzacji procesu audytu, przy jednoczesnym zwiększeniu jego wiarygodności.

Greenwashing – czyli praktyka prezentowania działań jako proekologicznych, bez realnego wpływu na środowisko – stanowi istotne zagrożenie dla wiarygodności rynku offsetów węglowych. W wielu przypadkach organizacje nie zapewniają wystarczającej transparentności dotyczącej charakteru projektów offsetowych ani ich rzeczywistej skuteczności w redukcji emisji gazów cieplarnianych (Wright & Nyberg, 2022). Brak standaryzacji, niewystarczające mechanizmy niezależnej weryfikacji oraz różnice w jakości metodologii pomiarowej utrudniają konsumentom i interesariuszom ocenę autentyczności deklarowanych działań proklimatycznych.

Zjawisko to może być dodatkowo potęgowane przez strategiczne manipulowanie danymi emisyjnymi. Jak zauważa Bumpus (2011), przedsiębiorstwa często wykorzystują offsety nie jako uzupełnienie, lecz substytut faktycznych działań redukcyjnych, ukrywając wzrost emisji bazowych za pomocą deklarowanej „neutralności węglowej”. W konsekwencji offsety stają się narzędziem narracyjnym, a nie środkiem rzeczywistej transformacji środowiskowej.

W tym kontekście sztuczna inteligencja odgrywa kluczową rolę w podnoszeniu rzetelności systemów certyfikacji offsetów. Zastosowanie zaawansowanych algorytmów uczenia maszynowego i systemów analizy danych w czasie rzeczywistym umożliwia automatyczne monitorowanie, weryfikację oraz raportowanie efektów środowiskowych projektów offsetowych (Alotaibi et al., 2024). Dzięki integracji AI z technologiami takimi jak satelity, czujniki IoT i blockchain możliwe staje się tworzenie niepodważalnych, odpornych na manipulację ścieżek dowodowych (digital MRV – monitoring, reporting, verification), które znacznie ograniczają możliwość przedstawiania nieprawdziwych lub niepełnych danych.

Systemy oparte na AI nie tylko podnoszą efektywność weryfikacji i audytu, ale również wprowadzają obiektywne, replikowalne i skalowalne standardy oceny wpływu środowiskowego. Przykładowo, modele predykcyjne mogą wykrywać anomalie w danych (np. nagły wzrost emisji lub spadek pokrycia roślinnego) i wskazywać potencjalne nieprawidłowości w raportowanych wartościach (Springer, 2025). Takie podejście nie tylko wzmacnia zaufanie do systemu offsetowego, ale także wspiera proces harmonizacji standardów certyfikacji w skali globalnej.

Wspieranie prognozowania i planowania

Sztuczna inteligencja umożliwia także prognozowanie wpływu projektów offsetowych na środowisko na podstawie analiz danych historycznych i modeli predykcyjnych. Modele te mogą uwzględniać różne scenariusze związane z przyszłymi zmianami klimatycznymi oraz ekonomicznymi, co pomaga w optymalnym planowaniu projektów offsetowych.

Jak pokazano w badaniach Sailaja i wsp. (2025), modele predykcyjne oparte na sztucznej inteligencji mogą przewidywać długoterminową efektywność projektów offsetowych z dokładnością do 90%. Modele te mogą również sugerować, jakie interwencje są potrzebne, aby projekt utrzymał zgodność z wytycznymi dotyczącymi redukcji emisji.

Z kolei w badaniach Feng i wsp. (2024) wykorzystano modele uczenia maszynowego do prognozowania emisji CO₂ w dużych miastach, które generują znaczne ilości gazów cieplarnianych. W tym badaniu zastosowano algorytmy regresji liniowej, a także bardziej zaawansowane modele, takie jak drzewa decyzyjne i random forest, aby analizować dane oparte na zużyciu energii, wzorcach transportowych oraz danych demograficznych. Badacze odkryli, że modele te mogą przewidywać przyszłe emisje CO₂ z dużą dokładnością, co pozwala władzom miejskim na planowanie skutecznych polityk redukcji emisji w czasie rzeczywistym.

W sektorze energetycznym zastosowanie modeli predykcyjnych pozwala na optymalizację emisji dwutlenku węgla z elektrowni. W badaniach prowadzonych przez Liu i wsp. (2020) zastosowano sieci neuronowe do przewidywania emisji CO₂ w zależności od produkcji energii, rodzaju paliwa oraz efektywności spalania. Modele te mogły prognozować emisje CO₂ na podstawie bieżących danych z systemów energetycznych, co umożliwiło precyzyjne planowanie działań zmniejszających emisje, takich jak wdrażanie technologii odnawialnych źródeł energii czy efektywniejsze zarządzanie elektrowniami.

Kolejnym przykładem jest badanie Lelanda i wsp. (2024), którzy stworzyli modele predykcyjne analizujące związek między zmianami klimatycznymi a przyszłymi emisjami i absorpcją dwutlenku węgla przez różne ekosystemy, w tym rolnictwo i leśnictwo. Algorytmy te wykorzystywały historyczne dane klimatyczne, takie jak temperatura i opady, aby przewidzieć zmiany w tempie wzrostu roślinności oraz zmiany w zdolności ekosystemów do pochłaniania CO₂ w przyszłości. Wyniki badań wykazały, że zmiany klimatyczne mogą znacząco wpłynąć na zdolność pochłaniania CO₂ przez środowisko, co jest istotnym czynnikiem w planowaniu polityk klimatycznych.

Technologie rozszerzające potencjał offsetów węglowych

Zastosowanie nowych technologii cyfrowych stanowi istotny przełom w rozwoju rynku offsetów węglowych, szczególnie w zakresie ich wiarygodności, skalowalności i integracji z systemami zarządzania emisjami. Szczególną uwagę należy poświęcić synergii między sztuczną inteligencją (AI), Internetem rzeczy (IoT), zaawansowaną analizą danych (XAI), technologią blockchain oraz uczeniem maszynowym (ML), które razem tworzą nową generację cyfrowych systemów MRV (Pourrahmani et al., 2025).

AI + Satelity + XAI (Explainable AI)

Zaawansowane modele predykcyjne oparte na sztucznej inteligencji i danych satelitarnych umożliwiają wczesne wykrywanie zagrożeń dla trwałości projektów offsetowych, takich jak susze, pożary lasów czy nielegalne wyręby. Przykładowo, Ballard i wsp. (2023) opisują wykorzystanie Explainable AI (XAI) do identyfikowania nieliniowych relacji między wzorcami pogodowymi, gęstością roślinności a ryzykiem wystąpienia pożarów w tropikalnych projektach REDD+ w Brazylii. Modele te umożliwiają lepsze zarządzanie ryzykiem i wycenę trwałości offsetów.

Podobne podejścia wykorzystywane są przez organizacje takie jak Space Intelligence czy Pachama, które stosują algorytmy CNN i LSTM do analizy danych z Landsat i Sentinel w celu śledzenia zmian pokrycia terenu z dokładnością powyżej 90% (Pachama, 2023). W połączeniu z XAI możliwe jest nie tylko przewidywanie zmian, ale również wyjaśnienie, które czynniki (np. zmiany wilgotności gleby, temperatury) mają największy wpływ na zdolność pochłaniania CO₂.

Blockchain + ML

Blockchain jako technologia rozproszonego rejestru (DLT) zapewnia niezmienną i transparentność danych, co jest szczególnie istotne w kontekście walki z greenwashingiem. W połączeniu z ML możliwe jest stworzenie systemu, w którym anomalie w danych – takie jak nagłe przyrosty biomasy, podejrzaną aktualizację danych projektowych czy nietypowe emisje – są wykrywane przez algorytmy ML i natychmiastowo zapisywane w blockchainie w sposób trwały i audytowalny (Springer, 2025).

Przykładem takiego podejścia jest system nChain dla offsetów w Kenii, który wykorzystuje czujniki do zbierania danych o wilgotności gleby i tempie wzrostu roślinności, a następnie waliduje je przez ML i zapisuje na blockchainie BSV. Analogiczne rozwiązania wdrażane są przez Toucan Protocol i KlimaDAO, które tokenizują offsety jako cyfrowe aktywa i zapewniają ich jednoznaczną identyfikowalność.

IoT + Blockchain + Smart Contracts

Internet rzeczy (IoT) w połączeniu z blockchainem i inteligentnymi kontraktami pozwala na stworzenie zautomatyzowanego systemu offsetowego działającego w czasie rzeczywistym. Sieci sensorów rozmieszczone w projektach leśnych, rolnych czy przemysłowych mierzą stężenie CO₂, wilgotność gleby, fotosyntezę (NDVI) lub dane meteorologiczne, które są przesyłane do blockchainowego rejestru. Tam są analizowane przez modele AI, które – jeśli wartości są zgodne

z założonymi programami – aktywują inteligentne kontrakty wypłacające kredyty offsetowe (Boumaiza et al., 2024).

Takie podejście eliminuje potrzebę ręcznego weryfikowania danych przez audytorów, skraca czas rozliczeń oraz zwiększa zaufanie. Przykładem wdrożenia takiego rozwiązania jest projekt Flowcarbon, który wykorzystuje sensorowe dane środowiskowe do automatycznego generowania tokenów GCO₂ zgodnie z używanym poziomem sekwestracji. Podobnie działa projekt Open Forest Protocol, który łączy AI, LIDAR i rejestr NEAR blockchain do ciągłej certyfikacji projektów leśnych.

Rozszerzone przykłady: tokenizacja, interoperacyjność i audyt cyfrowy

- PACT (Programmable Asset for Carbon Trust) – stworzony przez cFlow i Verra cyfrowy stablecoin zabezpieczony fizycznymi offsetami, gwarantuje, że jednostka CO₂e została rzeczywiście zredukowana.
- Digital MRV (WRI, Microsoft) – system AI umożliwiający automatyczne rozliczanie redukcji emisji na podstawie zdjęć satelitarnych i danych terenowych.
- Plantega – europejski startup rozwijający otwartą infrastrukturę cyfrową opartą na Hyperledger Fabric i metodologii CRCF, integrującą offsety z ESG reportingiem.
- Green Savings Scheme S.A. – polski startup działający na rynku dobrowolnych offsetów węglowych (VCM), rozwijający zautomatyzowany system monitoringu, raportowania i weryfikacji (MRV), oparty na technologiach AI, IoT i blockchain. Firma koncentruje się na projektach sekwestracji węgla w glebie, wykorzystując dane z sensorów glebowych, zdjęć satelitarnych i modeli uczenia maszynowego. GSS integruje swoje offsety z systemami ESG (zgodnie z wymogami ESRS i CRCF), umożliwiając ich tokenizację i automatyczny audyt w oparciu o inteligentne kontrakty.

Znaczenie implementacji rozwiązań technologicznych w procesie produkcji offsetów naturalnych w kontekście zmian legislacyjnych

W 2024 roku Parlament Europejski i Rada UE przyjęły rozporządzenie CRCF (Carbon Removals Certification Framework), którego celem jest standaryzacja i harmonizacja zasad certyfikacji pochłaniania dwutlenku węgla na poziomie UE (EU, 2024). Kluczowym elementem CRCF są tzw. QU.A.L.ITY criteria: quantification, additionality, long-term storage, sustainability i transparency. Wymagają one nie tylko dowodów naukowych potwierdzających istnienie i trwałość efektu pochłaniania, ale również interoperacyjnych, niezależnych systemów monitorowania i walidacji danych.

Dodatkowo, w ramach CSRD (Corporate Sustainability Reporting Directive) obowiązującej od 2024 roku, przedsiębiorstwa są zobowiązane do szczegółowego raportowania emisji i kompensacji zgodnie z ESRS (European Sustainability Reporting Standards), w tym ze standardem ESRS E1. Zgodnie z nim organizacje muszą wskazać rodzaj i jakość wykorzystywanych offsetów (biogeniczne, technologiczne, hybrydowe), ich pochodzenie, metodologię weryfikacji oraz stopień zgodności z politykami UE (EFRAG, 2023). W efekcie offsety naturalne poddawane są coraz większej presji dotyczącej ich transparentności i udowodnionej skuteczności.

W ujęciu globalnym podobne kierunki wskazuje Integrity Council for the Voluntary Carbon Market (ICVCM), który promuje tzw. Core Carbon Principles (CCPs), wymagające potwierdzenia integralności danych, odporności na manipulacje oraz ścisłej metodologii MRV.

Podsumowanie

Zintegrowane wykorzystanie AI, blockchain i IoT umożliwia tworzenie wiarygodnych, zautomatyzowanych i transparentnych systemów offsetowych. Technologie te nie tylko podnoszą jakość danych i skuteczność certyfikacji, ale też umożliwiają dynamiczne rozliczanie redukcji emisji w czasie rzeczywistym. Ich dalszy rozwój i standaryzacja będą kluczowe dla przyszłości globalnego rynku offsetów węglowych i wiarygodności osiągnięcia neutralności klimatycznej.

Bibliografia

- Alotaibi, E., Nassif, N. (2024). Artificial intelligence in environmental monitoring: in-depth analysis, *Discover Artificial Intelligence*, 4(1), DOI: 10.1007/s44163-024-00198-1.
- Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, 354(6309), 182–183.
- Aurangzeb, I., & Yoon, J. H. (2025). Artificial intelligence- and blockchain-enabled carbon emissions ledger system (AB-CELS) for sustainable construction processes. *Automation in Construction*, 176, 106286. DOI: 10.1016/j.autcon.2025.106286
- Ballard, T., Cooper, M., Lowrie, Ch., & Erinijppurath, G. (2023). Widespread Increases in Future Wildfire Risk to Global Forest Carbon Offset Projects Revealed by Explainable AI, DOI: 10.48550/arXiv.2305.02397.
- Battocletti, V., Enriques, L., Romano, A. (2023). The Voluntary Carbon Market: Market Failures and Policy Implications, Law Working Paper, 688/2023.
- Boumaiza, A., Maher, K. (2024). Harnessing Blockchain and IoT for Carbon Credit Exchange to Achieve Pollution Reduction Goals. *Energies*, 17(19), 4811; DOI: 10.3390/en17194811.

- Broekhoff, D., Gillenwater, M., Colbert-Sangree, T., & Cage, P. (2019). *Securing Climate Benefit: A Guide to Carbon Offset Quality*. Stockholm Environment Institute & Greenhouse Gas Management Institute.
- Bumpus, A. G. (2011). The Matter of Carbon: Understanding the Materiality of tCO₂e in Carbon Offsets. *Antipode*, 43(3), 612–638.
- Cames, M., Harthan, R. O., Füssler, J., Lazarus, M., Lee, C. M., Erickson, P., & Spalding-Fecher, R. (2016). *How additional is the Clean Development Mechanism? DG CLIMA, European Commission*.
- Cho, S., Lee, J. (2020). Application of AI in Carbon Offset Monitoring: A Case Study. *Journal of Environmental Management*, 255, 109893.
- Deng, X., Zhao, Y., Lin, J., & Wang, H. (2021). Blockchain and Artificial Intelligence for Carbon Credit Systems: Enhancing Transparency and Trust. *Environmental Science & Policy*, 123, 153–162.
- Feng, W., Chen, W., Li, L., & Zhang, L. (2024). Application of Neural Networks on Carbon Emission Prediction: A Systematic Review and Comparison. *Energies*, 17(7), 1628. DOI: 10.3390/en17071628.
- Gupta, A., Löwbrand, E., Turnhout, E., & Vijge, M. (2020). In pursuit of carbon accountability: the politics of REDD+ measuring, reporting and verification systems. *Current Opinion in Environmental Sustainability*, 45, 55–61.
- Kollmuss, A., Zink, H., & Polycarp, C. (2008). *Making Sense of the Voluntary Carbon Market: A Comparison of Offset Standards*. WWF/SEI.
- Leland, J., Uyekawa, J., Beral, D., & Liu, Y. (2024). Machine Learning-Based Prediction of Ecosystem-Scale CO₂ Flux Measurements, DOI: 10.20944/preprints202412.1152.v1.
- Liu, H., Zhang, X., & Wang, Y. (2020). Predicting CO₂ emissions in power generation using artificial neural networks. *Energy*, 195, 116982.
- Lovett, J. (2025). Space lasers, AI Used by Geospatial Scientist to Measure Forest Biomass. <https://aaes.uada.edu/news/forest-aboveground-biomass/>.
- OECD. (2022). *Voluntary Carbon Markets and Climate Ambition*. Paris: OECD Publishing.
- Pachama. (2023). Pachama Research Brief: A Description and Initial Validation of a Dynamic Baseline for Avoided Deforestation Projects. Pachama.
- Pourrahamani, H., Talebi Amiri, M., Madi, H., & Owusu, J. (2025). Revolutionizing carbon sequestration: Integrating IoT, AI, and blockchain technologies in the fight against climate change. *Energy Reports*, 13, 5952–5967. DOI: 10.1016/j.egyr.2025.05.042.
- Sailaja, C., Javali, S., Patali, Y. S., & Shivaprasad, D. (2025). Predictive Modeling of Carbon Footprint in Hybrid Structural Components Using AI and Mathematical Algorithms. *International Research Journal on Advanced Science Hub*, 7(06), 521–527. DOI: 10.47392/IRJASH.2025.059.
- Sarkar, A., Ghosh, D., Ganguly, K., Ghosh, S., & Saha, S. (2023). Exploring IoT for real-time CO₂ monitoring and analysis. <https://arxiv.org/pdf/2308.03780>.

- Sood, C. (2025). From Trees to Tokens: AI and the Digital Renaissance of Carbon Markets. SSRN. <http://dx.doi.org/10.2139/ssrn.5259696>.
- Springer. (2025). Blockchain for the carbon market: a literature review. *Springer Nature*, 3(68).
- Wright, C., & Nyberg, D. (2022). Climate Change, Capitalism, and Corporations: Processes of Creative Self-Destruction. *Cambridge University Press*. https://www.researchgate.net/publication/282073643_Climate_Change_Capitalism_and_Corporations_Processes_of_Creative_Self-Destruction.

Od widma do znaczenia: analiza cech semantycznych i spektralnych w detekcji deepfake

Streszczenie

Artykuł wprowadza paradygmat „od widma do znaczenia” w detekcji deepfake, łączący cechy spektralne (FFT/DCT/3D-FFT) z deskryptorami semantyczno-kontekstowymi, takimi jak spójność tożsamości twarzy, analiza współwystąpień, FACS oraz zgodność twarz-głos. W tym celu opracowano zbiór DeepFake RealWorld (DFRW), obejmujący 46 371 klipów i odzwierciedlający współczesne generatory treści syntetycznych oraz realistyczne łańcuchy degradacji platformowych. Zastosowano podejście oparte na selekcji cech według częstości ich występowania, z progami anomalii wyznaczanymi wyłącznie na podstawie nagrań rzeczywistych. Wyniki potwierdzają komplementarność informacji spektralnej i semantycznej oraz prowadzą do opracowania interpretowalnego atlasu cech, który może stanowić warstwę preselekcji dla systemów moderacji treści, KYC oraz forensyki cyfrowej.

Słowa kluczowe: deepfake, detekcja, cechy spektralne, cechy semantyczne, modele dyfuzyjne, DFRW, forensyka multimedialna

¹ Akademia WSB w Dąbrowie Górniczej

From Spectrum to Meaning: Analysis of Semantic and Spectral Features in Deepfake Detection

Abstract

The article introduces a “from spectrum to meaning” paradigm for deepfake detection, combining spectral features (FFT/DCT/3D-FFT) with semantic-contextual descriptors such as facial identity consistency, co-occurrence analysis, FACS, and face-voice alignment. To support this approach, the DeepFake RealWorld (DFRW) dataset was developed, comprising 46,371 clips that reflect contemporary synthetic content generators and realistic platform degradation pipelines. A feature-frequency-based selection strategy was applied, with anomaly thresholds derived exclusively from real recordings. The results confirm the complementarity of spectral and semantic information and lead to the development of an interpretable feature atlas that can serve as a preselection layer for content moderation systems, KYC applications, and digital forensics.

Keywords: deepfake; detection; spectral features; semantic features; diffusion models; DFRW; multimedia forensics

Wprowadzenie

Postęp w dziedzinie sztucznej inteligencji w ostatnich latach doprowadził do gwałtownej zmiany paradygmatu w zakresie generowania treści syntetycznych. Szczególną rolę odegrał rozwój modeli głębokiego uczenia, które w ciągu ostatniego pięciolecia odnotowały skokowy wzrost złożoności, zgodny z tzw. prawem Moore’a dla sztucznej inteligencji (Amodei & Hernandez, 2018; Kaplan et al., 2020). Liczba parametrów modeli wzrosła z rzędu 10^8 w 2018 roku do ponad 10^{11} w roku 2024 (Brown et al., 2020; Achiam et al., 2023). Równolegle migracja modeli na sprzęt konsumencki została ułatwiona dzięki dostępności zaawansowanych jednostek obliczeniowych, takich jak karty graficzne GeForce RTX 4090 (83 TFLOPS mocy FP32) czy procesory Apple M2 Ultra (134 mld tranzystorów). Konsekwencją tych zmian stała się masowa demokratyzacja technologii deepfake, dostępna zarówno poprzez narzędzia typu GUI, jak i poprzez platformy chmurowe w modelu SaaS, które eliminują bariery sprzętowe i umożliwiają dostęp do najnowocześniejszych architektur generatywnych.

Termin *deepfake* odnosi się do syntetycznych treści audiowizualnych generowanych przy użyciu modeli głębokiego uczenia, które umożliwiają realistyczną modyfikację lub całkowitą syntezę obrazu i dźwięku osoby. Rozwój technologii deepfake charakteryzuje się dwutorowością. Z jednej strony pojawiają się zastosowania masowe, związane z rozrywką, tworzeniem memów, filtrów czy sztuki cyfrowej. Z drugiej jednak strony technologia ta staje się narzędziem wysokiego ryzyka, wykorzystywanym do wyrafinowanych ataków spear-phishingowych, kampanii dezinformacyjnych, szantażu i manipulacji politycznych (Chesney & Citron, 2019; Jędrasiak, 2024). Takie zastosowania stanowią istotne zagrożenie dla bezpieczeństwa publicznego, procesów wyborczych, integralności informacji oraz gospodarki. Raporty wskazują, że w 2023 roku liczba wykrytych prób wykorzystania deepfake w systemach weryfikacji tożsamości wzrosła dziesięciokrotnie względem roku poprzedniego, a w 2024 roku wzrost ten był czterokrotny (Sumsb, 2024). W Europie odnotowano wzrost o 780% rok do roku. Jednocześnie raport firmy Regula wskazuje, że w 2024 roku połowa przedsiębiorstw zgłaszała incydenty związane z syntetycznymi mediami. Dane te unaoczniają skalę problemu i wskazują na konieczność systematycznych badań nad wiarygodną detekcją treści syntetycznych.

Najbardziej dyskutowanym zagrożeniem jest dezinformacja. Rozpowszechnianie fałszywych treści multimedialnych opiera się na mechanizmie kaskady informacyjnej (Bikhchandani et al., 1992; Lerman, 2018). Badanie Kumar et al. (2024) wykazało, że syntetyczne wideo osiąga milion wyświetleń w czasie średnio 2,3 h, przy czym ponad 60% udostępnień następuje w ciągu pierwszych 30 minut. Taka dynamika sprawia, że procesy weryfikacji i neutralizacji fałszywych treści stają się niewystarczające. Średni czas reakcji dużych platform społecznościowych na

zgłoszenie zmanipulowanego materiału wynosi 6–12 h (DiResta, 2020), co oznacza, że w praktyce szkodliwy materiał zdąży zostać powielony i utrwalony w zdecentralizowanych sieciach, uniemożliwiając jego skuteczne usunięcie.

Problem głębiej dotyka bezpieczeństwa publicznego i infrastruktury krytycznej. Symulacje agentowe TNO (2023) pokazały, że fałszywe apele ewakuacyjne generowane jako deepfake mogą trzykrotnie przekroczyć przepustowość dróg w ciągu półtorej godziny od publikacji. Incydenty, takie jak masowe fałszywe alerty w USA (Uvalde, 2022) czy w Niemczech (Westdeutscher Rundfunk, 2023), potwierdzają realne zagrożenia operacyjne. Konsekwencje finansowe są równie alarmujące. Raport FBI Internet Crime Complaint Center (IC3) dla roku 2023 wskazuje na rekordową liczbę zgłoszeń oraz straty przekraczające 12,5 mld USD, co odpowiada 22% wzrostowi strat względem roku 2022, przy czym istotną część ekosystemu stanowią oszustwa socjotechniczne klasy BEC, obejmujące również warianty określane jako CEO fraud. Europol w raporcie IOCTA 2024 wskazuje jednocześnie na rosnące wykorzystanie deepfake w schematach oszustw powiązanych z podszywaniem się pod kadrę kierowniczą oraz w działaniach telefonicznych typu shock calls (Europol, 2024; Federal Bureau of Investigation, 2024).

Na poziomie technologicznym wyzwania związane z detekcją deepfake wynikają z ograniczeń modeli uczenia maszynowego. Algorytmy end-to-end osiągają wysoką skuteczność na danych znanych, lecz ulegają znacznemu pogorszeniu w warunkach walidacji międzyzbiorowej oraz w obecności degradacji platformowych, takich jak rekompresja czy re-nagrania ekranu (Lu & Ebrahimi, 2024). Publiczne wyniki DFDC (Facebook Deepfake Detection Challenge) potwierdzają, że nawet najlepsze modele osiągają poniżej 70% trafności na danych nieznanymi systemowi. Problemem pozostaje także brak przejrzystości metod oraz niedostateczne stosowanie technik wyjaśnialnej sztucznej inteligencji (Wang et al., 2024). W efekcie część systemów detekcji charakteryzuje się zbyt dużą pewnością predykcji (overconfidence), przy jednocześnie niskiej odporności na ataki adaptacyjne (Guo et al., 2017).

Współczesne badania wskazują, że pomimo wysokiej jakości wizualnej syntezy generatywne modele nie są w stanie w pełni odwzorować niedoskonałości fizycznych charakterystycznych dla rzeczywistych urządzeń optoelektronicznych (Farid, 2009). Do zjawisk tych należą m.in.: winietowanie, dystorsje geometryczne, artefakty rolling-shutter czy subtelne defekty matrycy (Ray, 2002; Van Houten et al., 2015). Brak ich występowania w treściach syntetycznych otwiera perspektywę detekcji opartą na analizie cech spektralnych, geometrycznych i temporalnych.

Jednocześnie zauważa się istotną lukę badawczą. Dotychczasowe prace koncentrowały się na pojedynczych grupach cech – kompresyjnych, fotometrycznych czy semantycznych – lecz brak jest systematycznej klasyfikacji i empirycznej weryfikacji cech multimodalnych obejmujących aspekty spektralne i znaczeniowe (Mittal et al., 2020; Nautsch et al., 2022). Niewystarczająca pozostaje także analiza stabilności

deskryptorów w warunkach degradacji charakterystycznych dla dystrybucji platformowej oraz walidacja percepcyjna prowadzona przez ekspertów sądowych.

Niniejsza praca podejmuje próbę wypełnienia tej luki poprzez analizę i selekcję deskryptorów umożliwiających skuteczną detekcję deepfake. Badania koncentrują się na połączeniu dwóch perspektyw: warstwy semantycznej i kontekstowej narracji z analizą sygnatur spektralnych w domenie częstotliwościowej. Podejście to pozwala na uchwycenie zarówno subtelnych niespójności treściowych, jak i powtarzalnych śladów technicznych wynikających z procesów generatywnych. Celem pracy nie jest opracowanie kolejnego klasyfikatora typu czarna skrzynka, lecz wyznaczenie atlasu cech o potwierdzonej stabilności i wysokiej interpretowalności, które mogą stanowić fundament do budowy systemów odpornych na manipulacje adaptacyjne i zgodnych z wymogami inżynierii bezpieczeństwa.

Naukowym celem pracy jest opracowanie i empiryczna weryfikacja paradygmatu „od widma do znaczenia”, łączącego cechy spektralne i semantyczno-kontekstowe w interpretowalnej detekcji treści syntetycznych. Założono, że połączenie tych dwóch komplementarnych warstw reprezentacji umożliwia skuteczniejszą, a zarazem bardziej wyjaśnialną identyfikację materiałów typu deepfake niż stosowanie każdej z nich oddzielnie. W pracy przyjęto następujące hipotezy badawcze: H1. Cechy spektralne zawierają charakterystyczne sygnatury generacji i rekodowania, które pozwalają odróżniać materiały syntetyczne od autentycznych w sposób odporny na degradację platformowe. H2. Cechy semantyczne i kontekstowe ujawniają niespójności narracyjne oraz multimodalne, których nie wykrywają metody oparte wyłącznie na analizie widmowej. H3. Integracja cech spektralnych i semantyczno-kontekstowych (hybrydyzacja cech) zwiększa stabilność i zdolność generalizacji detekcji deepfake w warunkach zmiennych generatorów i łańcuchów przetwarzania.

Podsumowując, dynamiczny rozwój technologii deepfake, jej dostępność oraz potencjalnie destrukcyjne skutki społeczne, ekonomiczne i polityczne tworzą bezprecedensowe wyzwanie badawcze. Skuteczna detekcja wymaga odejścia od podejść czysto klasyfikacyjnych na rzecz systematycznej analizy deskryptorów multimodalnych, które łączą warstwę znaczeniową z charakterystyką spektralną. Przedstawione w tej pracy podejście odpowiada na zapotrzebowanie na wyjaśnialne, odporne i praktyczne rozwiązania detekcji treści syntetycznych w kontekście inżynierii bezpieczeństwa oraz forensyki cyfrowej.

Zbiory danych i wyzwania badawcze

Skuteczność metod detekcji treści syntetycznych jest wprost zależna od jakości, różnorodności i aktualności wykorzystywanych zbiorów danych. W kontekście pracy ukierunkowanej na cechy semantyczne i kontekstowe oraz cechy

spektralne kluczowe jest, aby zbiory obejmowały: zróżnicowane generatory i epoki technologiczne, profil „in the wild” z realistycznym łańcuchem degradacji platformowych, ścieżki audio w parze z wideo oraz komplet metadanych i bitstreamu, umożliwiającą analizę sygnatur w domenie częstotliwościowej oraz w warstwie narracyjno-zdarzeniowej. Wymóg multimodalności i aktualności wynika z obserwowanego przejścia od architektur opartych na GAN/autoenkoderach do modeli dyfuzyjnych oraz hybryd 3D, w tym metod rekonstrukcji trójwymiarowej i NeRF, których artefakty znacząco różnią się od klasycznych sygnatur częstotliwościowych. Zbiory powinny umożliwiać ocenę generalizacji międzygeneratorowej i międzyzbiorowej, pomiary stabilności deskryptorów pod degradacjami oraz analizę wpływu parametrów bitstreamu na wykrywalność.

Facebook Deepfake Detection Challenge DFDC pozostaje kamieniem milowym pierwszej fali metod głębokich. Obejmuje ponad 124 tys. klipów wygenerowanych przez osiem ówczesnych architektur GAN z lat 2018–2019 i umożliwił rozwój linii detektorów konwolucyjnych (DFDC, 2019; Dolhansky et al., 2020). Ograniczenia DFDC są jednak istotne z perspektywy niniejszej pracy. Profil bitrate jest relatywnie jednorodny około 24 Mb/s, dane nie obejmują nowej klasy generatorów dyfuzyjnych ani rekonstrukcji 3D, a komponent audio nie jest systemowo włączony w ocenę spójności multimodalnej. W efekcie DFDC jest wartościowy dla badań ablation i porównań w warunkach kontrolowanych, ale nie reprezentuje współczesnych kanałów dystrybucji i metod generacji.

FaceForensics++ (Rössler et al., 2019) zawiera około 1000 filmów prawdziwych i 4000 zmanipulowanych, trzy poziomy kompresji raw/c0, c23, c40 oraz cztery typy manipulacji: Deepfakes, FaceSwap, Face2Face, NeuralTextures. Standaryzacja materiału, rozdzielczości i kompresji umożliwia precyzyjne eksperymenty i badania wrażliwości na degradacje. Jednocześnie charakter danych studyjnych 720p promuje „czyste” artefakty generacyjne, które w scenariuszach platformowych są maskowane przez rekodowanie, filtry i re-capture. Literatura wielokrotnie potwierdziła spadek skuteczności modeli uczonych na FF++ po transferze na inne zbiory, w szczególności Celeb-DF, co obrazuje dataset bias oraz ograniczoną generalizację między generatorami i łańcuchami przetwarzania.

Celeb-DF v2 (Li et al., 2020) to zestaw zaprojektowany jako trudniejszy benchmark, z naciskiem na fotorealizm i redukcję wczesnych artefaktów. Zawiera 5639 filmów deepfake i 590 prawdziwych, w większości z udziałem osób publicznych i zróżnicowanych źródeł. Autorzy wykazali znaczące spadki AUC dla detektorów trenowanych na FF++, co ugruntowało wymóg testów cross-dataset. Ograniczeniem Celeb-DF v2 pozostaje skupienie na generatorach z roku 2019, bez reprezentacji metod dyfuzyjnych, hybrydowych i modeli 3D.

DeeperForensics-1.0 (Jiang et al., 2020) wprowadza dużą skalę 60 000 klipów oraz kontrolowane degradacje H.264 CRF 23, warunki oświetleniowe, ruchowe

i zakłócenia. Jego wartością dodaną jest projektowanie pod „prawdziwy świat”, co czyni go użytecznym do badania stabilności cech teksturalnych i temporalnych przy rekompresji i re-capture. Niemniej, podobnie jak DFDC i FF++, zbiór powstał w erze dominacji GAN/autoenkoderów i nie obejmuje nowej generacji modeli dyfuzyjnych ani środowisk mobilnych z ich charakterystycznymi pipeline’ami ISP.

Wspólny problem tych benchmarków to przestarzałość względem aktualnych trendów generatywnych oraz brak systemowego uwzględnienia ścieżki audio i spójności multimodalnej. W szczególności brakuje reprezentacji Stable Diffusion 3.x, Runway Gen-3, Sora oraz materiałów generowanych w pipeline’ach 3D, a także pełnych metadanych i bitstreamu potrzebnych do rzetelnej analizy spektralnej i sygnatur kodowania (Yang et al., 2023). W obszarze oceny cech semantycznych ograniczeniem jest brak adnotacji narracyjnych, zdarzeniowych i relacji przyczynowych, które umożliwiałyby pomiar spójności treściowej.

W praktyce materiały syntetyczne podlegają złożonym łańcuchom przetwarzania: wielokrotnej rekompresji z utratą informacji, zmianom rozdzielczości i proporcji, filtracji AR, sub-samplingu chrominancji, transkodowaniom między kodekami oraz re-nagrywaniu z ekranów. Haliassos i wsp. (2022) pokazali, że degradacje platformowe w sposób systematyczny obniżają skuteczność detektorów uczonych w warunkach kontrolowanych. Kołkowski i wsp. (2023) raportują drastyczne spadki metryk w transferze do treści „dzikich”, co potwierdza, że benchmarki pozbawione realistycznego łańcucha dystrybucyjnego zawyżają ocenę skuteczności. Z punktu widzenia cech spektralnych brak standaryzacji pipeline’u kodowania utrudnia porównywalność miar w dziedzinie częstotliwości, a dodatkowe operacje post-processingu mogą wygładzać piki diagnostyczne. Z kolei w warstwie semantycznej filtry i przycinanie zmieniają kontekst narracyjny, utrudniając ocenę spójności zdarzeń i synchronii audio-wideo.

Konsekwencją jest rosnące zapotrzebowanie na zbiory, obejmujące pełny profil dystrybucji: wideo i audio, metadane i bitstream, adnotacje semantyczne, różnorodne generatory, a także protokoły testowe z degradacjami. Postuluje się rolling benchmarks aktualizowane w cyklu ciągłym, tak aby nadążać za ewolucją modeli generatywnych i praktyk dystrybucyjnych (Verdoliva, 2020; Deepfake Eval 2024; Wang et al., 2024). Dla badań spektralnych pożądane jest udostępnianie surowych ramek lub strumieni o minimalnej liczbie transformacji po stronie platformy, co pozwala na wiarygodny pomiar miar w dziedzinie Fouriera, DCT i FBS.

Najpoważniejszą barierą praktycznej użyteczności detektorów pozostaje ograniczona generalizacja w warunkach zmieniających się technologii generatywnych i dystrybucyjnych. Modele uczone na pojedynczych benchmarkach osiągają wysoką skuteczność in-distribution, lecz w scenariuszach cross-dataset oraz cross-model odnotowuje się spadki AUC rzędu 20–50% (Cozzolino et al., 2023; Tariq et al., 2021; Verdoliva, 2020). Przyczyną jest nadmierne poleganie na sygnałach

specyficznych dla danej klasy generatora, w tym na charakterystycznych wzorach częstotliwościowych typowych dla transpozycyjnego upsamplingu w GAN. Przejście do modeli dyfuzyjnych osłabiło uniwersalność „odcisków Fouriera”, ponieważ proces odszumiania i rekonstrukcji nie wytwarza tych samych regularności widmowych (Durall et al., 2020; Ricker et al., 2023). Nowsze podejścia, jak DIRE, łączą wieloskalowe wskazówki semantyczne, teksturalne i przestrzenno-czasowe, co jest spójne z hipotezą o konieczności hybrydyzacji cech w kierunku integracji warstwy znaczeniowej z cechami spektralnymi (Wang et al., 2023).

W warunkach out-of-distribution typowy jest problem nadmiernej pewności predykcji. Modele wykazują wysokie Expected Calibration Error $ECE > 0,25$, co oznacza nieadekwatność prawdopodobieństw do rzeczywistej niepewności (Guo et al., 2017; Meng et al., 2021). W kontekście inżynierii bezpieczeństwa i forensyki cyfrowej rekomenduje się obowiązkową kalibrację (temperature scaling, Platt scaling), raportowanie niepewności i metryk czułych na ogon rozkładu błędów, jak pAUC przy niskich FPR, oraz wycenę ryzyka decyzyjnego w funkcji proggu klasyfikacyjnego.

Równolegle ewoluują techniki antyforensyczne. Do najczęstszych należą degradacje jakościowe: wielokrotna rekompresja JPEG/H.264, rozmycia, dodanie szumu, filtry spektralne tłumiące piki, re-capture, a także bardziej zaawansowane procedury super-resolution lub ataki gradientowe, w których generator minimalizuje sygnały detekcyjne przez włączenie funkcji strat detektora do procesu uczenia (Lu & Ebrahimi, 2022; Cao et al., 2021; Jia et al., 2022; Carlini et al., 2023). Skuteczność detektorów może spadać o 15–30% na pojedynczym etapie łańcucha platformowego. Zasadne jest zatem projektowanie protokołów testowych obejmujących realistyczne łańcuchy przetwarzania, a także łączenie cech odpornych na wygładzanie spektralne z deskryptorami semantycznymi oraz wskaźnikami synchronii multimodalnej.

Znacząca część rozwiązań komercyjnych działa w trybie czarnej skrzynki i nie spełnia rygorów dowodowych wymaganych w procedurach karnych i cywilnych. Wymogi przejrzystości metody, weryfikowalności i znanej charakterystyki błędu wynikają z praktyk orzeczniczych oraz standardów proceduralnych. Konieczne jest systemowe stosowanie XAI, w tym map uwagi Grad-CAM, LIME, kontrfaktycznych przykładów oraz raportów łączących wskazania modelu z uchwytnymi sygnałami fizycznymi i semantycznymi, jak rytm mikroruchów, zgodność fonemów z artykulacją, spójność oświetlenia, koherencja odbłasków i cieni, regularność artefaktów kodowania (Selvaraju et al., 2017; Doshi-Velez & Kim, 2017; Wang et al., 2024). Z punktu widzenia cech spektralnych istotne jest, aby raporty prezentowały miary w domenie częstotliwościowej wraz z kontekstem bitstreamowym, a dla cech semantycznych – aby zawierały streszczenia zdarzeniowe, reguły spójności i odchylenia od wzorców naturalnych. Europol i środowiska

biegłych akcentują potrzebę standaryzowanych szablonów raportów interpretacyjnych oraz audytowalnych wizualizacji dla decydentów nieposiadających przygotowania technicznego.

Z perspektywy tytułu niniejszej pracy „Od widma do znaczenia: analiza cech semantycznych i spektralnych w detekcji deepfake” pożądane cechy zbiorów oraz protokołów obejmują:

1. Wielogeneratorowość i wieloepokowość. Minimalny przekrój powinien obejmować GAN, autoenkodery, dyfuzję oraz hybrydy 3D, z rozdzielaniem na domeny treści i style. Pozwala to ocenić, które wskaźniki spektralne zachowują się stabilnie między paradygmatami, a które wymagają wzmacniania przez semantykę.
2. Multimodalność AV z adnotacjami semantycznymi. Niezbędne są wyrównane ścieżki audio-wideo, transkrypcje i adnotacje zdarzeniowe do oceny spójności narracyjnej, fonem-wizem i relacji przyczynowych. Umożliwia to budowę dekskryptorów kontekstowych oraz miar synchronii krzyżowej.
3. Pełny profil dystrybucji. Zbiór powinien zawierać scenariusze rekompresji, filtrowania, zmian rozdzielczości, re-capture, a także zapis parametrów bitstreamu i surowych ramek dla analiz FFT/DCT. Tylko wówczas można wiarygodnie ocenić stabilność cech spektralnych.
4. Protokół OOD i rolling benchmarks. Ocena powinna obejmować testy międzyzbiorowe i międzygeneratorowe oraz aktualizacje zestawu walidacyjnego w czasie. Rekomenduje się raportowanie pAUC dla niskich FPR, ECE po kalibracji, a także metryk odporności na łańcuchy przetwarzania.
5. Wymagalność interpretowalności. Standard wyników musi zawierać komponent XAI, mapy uwagi i wskaźniki fizykalne oraz semantyczne, tak aby decyzje były replikowalne i audytowalne.

Klasyczne benchmarki DFDC, FaceForensics++, Celeb-DF v2 i DeeperForensics-1.0 pozostają wartościowe dla badań kontrolowanych i ablation, lecz nie reprezentują współczesnego krajobrazu generatywnego oraz warunków dystrybucji. Z punktu widzenia cech spektralnych zmiana paradygmatu na modele dyfuzyjne osłabiła uniwersalność „odcisków Fouriera”, co wymusza łączenie sygnałów częstotliwościowych z cechami semantycznymi i temporalnymi. Z perspektywy warstwy znaczeniowej brak adnotacji narracyjnych ogranicza pomiar spójności treści i synchronii multimodalnej. Kluczowe wyzwania obejmują generalizację OOD, kalibrację niepewności, odporność na antyforensykę oraz zgodność z wymaganiami forensyki cyfrowej. Odpowiedzią są zbiory multimodalne, wielogeneratorowe i aktualizowane w trybie ciągłym, protokoły z realistycznymi łańcuchami degradacji oraz standardy raportowania łączące XAI z mierzalnymi wskaźnikami spektralnymi i semantycznymi. Tak wyznaczona rama badawcza tworzy podstawy do weryfikacji atlasu cech proponowanego w dalszych sekcjach,

w szczególności w odniesieniu do zbioru DFRW oraz oceny stabilności i przydatności operacyjnej deskryptorów w warunkach dystrybucji platformowej.

Zbiór DFRW

Wobec braku jednego, aktualnego i w pełni reprezentatywnego benchmarku, obejmującego zarówno najnowsze techniki generacji treści syntetycznych, w szczególności modele dyfuzyjne i hybrydy 3D, jak i realistyczne łańcuchy degradacji platformowych, opracowano własny zbiór DeepFake RealWorld DFRW liczący 46 371 klipów. Konstrukcja DFRW została podporządkowana celowi artykułu, tj. empirycznej ocenie deskryptorów w dwóch komplementarnych domenach analizy: widmowej oraz semantyczno-kontekstowej. Zbiór ma umożliwiać równoległą weryfikację hipotez dotyczących sygnatur spektralnych kodowania, rekonstrukcji i postprocessingu oraz testowanie wskaźników spójności narracyjnej, synchronii multimodalnej i anomalii kontekstowych.

Proces budowy przebiegał dwuetapowo. Najpierw przeprowadzono systematyczne pozyskanie materiałów „in the wild” z wykorzystaniem metod OSINT: śledzono znaczniki treści w mediach społecznościowych, korzystano z repozytoriów fact-checking oraz forów i serwisów wideo, a tożsamość duplikatów weryfikowano poprzez dopasowanie sygnatur percepcyjnych pHash. Uzyskano 4186 unikatowych deepfake’ów, które poddano anonimizacji metadanych oraz ocenie kontekstu zgodnie z zasadami dozwolonego użytku naukowego i etyki badań. Następnie, w celu zrównoważenia i rozszerzenia przestrzeni hipotez, wygenerowano i przekształcono dodatkowe 42 185 klipów, wykorzystując dominujące w 2025 roku narzędzia: modele dyfuzyjne wideo Stable Video Diffusion, Runway Gen-2/Gen-3, Pika Labs, wczesny dostęp Sora, nową generację systemów face swap SimSwap++, DeepFaceLab 2.0 i InsightFaceSwap, metody synchronizacji mowy i ekspresji Wav2Lip++, SadTalker, EMO LipSync oraz reenactment oparte na Face2Face Enhanced i First Order Motion Model. Materiały poddano kontrolowanym transformacjom degradacyjnym, które emulują łańcuchy platformowe: wielokrotne rekodowanie H.264/H.265 w zakresie CRF 18–32, zmianę rozdzielczości i FPS, filtry mobilne i AR oraz wtórne nagrywanie ekranu. W rezultacie DFRW agreguje 4186 oryginalnych dzikich deepfake’ów 9% oraz 42 185 wariantów syntetycznych i przekształconych 91%, co pozwala jednocześnie uchwycić profil artefaktów występujących w realnej dystrybucji i zapewnić przekrój technologiczny zorientowany na rok 2025.

Zbiór obejmuje 46 371 klipów o łącznym czasie 229,28 h. Średnia długość klipu wynosi 17,8 s SD 12,3 s, 95% CI [17,69 s, 17,91 s], mediana 12,4 s, IQR 7,1–22,9 s. Rozkład długości: < 5 s 22%, 5–15 s 41%, 15–60 s 27%, > 60 s 10%. Wyróżniono cztery kategorie technologiczne: face swap 16 230 klipów 35%, reenactment i modyfikacja

ekspresji 12 984 28%, pełnoklatkowa generacja dyfuzyjna 11 593 25% oraz złożone manipulacje scenowe 5564 12%. Dla materiałów dyfuzyjnych zarejestrowano parametry procesu generacyjnego mediana liczby kroków 30, typowa skala guidance 4,5, co odpowiada kompromisom jakościowo-czasowym raportowanym w literaturze dla dyfuzji (Ho et al., 2020). W reenactment 57% stanowią przypadki audio-driven, 43% video-driven, co umożliwia analizę wpływu sposobu sterowania na artefakty temporalne i synchronię AV (Haliassos et al., 2022).

Warunki dystrybucyjne odwzorowano poprzez trzy poziomy jakości: wyso-ki 6956 15%, średni 27 823 60% i wysoka degradacja 11 592 25%. Kodeki: H.264/AVC 71%, H.265/HEVC 19%, VP9 7%, AV1 3%. Kontenery: MP4 84%, WebM 10%, MOV 6%. Mediana bitrate wideo 1,6 Mb/s IQR 0,9–2,8 Mb/s. Mediana długości GOP 60 klatek $Q1 = 30$, $Q3 = 120$. Rozdzielczości: $\leq 480p$ 14%, 720p 45%, 1080p 33%, $\geq 1440p$ 8%. Orientacja: pozioma 59%, pionowa 36%, kwadratowa 5%. FPS: 30 fps 62%, 24 fps 21%, 60 fps 7%, 25 fps 7%, inne 3%. Udział wtórnych nagrań ekranu oszacowano na 12% na podstawie aliasingu i nieliniowości gamma nietypowych dla ISP kamer. Średnia liczba kolejnych rekodowań łańcucha platformowego 2,1 rozkład: $1 \times 36\%$, $2 \times 31\%$, $3 \times 23\%$, $\geq 4 \times 10\%$. Filtry mobilne/AR występują w 28% przypadków, watermarki platformowe w 19%. Metryki bezreferencyjne potwierdzają stopnie degradacji: mediana NIQE 4,1 vs 3,2 vs 2,6 oraz BRISQUE 42 vs 31 vs 24 dla poziomów wysoka degradacja, średnia, wysoka jakość. Zależność CRF–LPIPS i CRF–VMAF jest silna $\rho = 0,82$, $\rho < 0,001$ oraz $\rho = -0,79$, $\rho < 0,001$, co koreluje z obserwowanymi w literaturze spadkami skuteczności detekcji wraz ze wzrostem kompresji i zniekształceń translacyjnych (Li et al., 2020; Yang et al., 2023).

Zróznicowanie podmiotowe i sceniczne: mężczyźni 24 113 52%, kobiety 22 258 48%; wiek < 25 lat 25%, 25–50 lat 60%, > 50 lat 15%. Sceny: studyjne 18 548 40%, plenerowe 16 230 35%, domowe 11 593 25%. Liczba osób w kadrze: jedna 72%, dwie 23%, ≥ 3 osoby 5%. Relatywna powierzchnia twarzy w kadrze $14\% \pm 9\%$ mediana 11%. Ekstremalne położenia głowy $|\text{yaw}| > 30^\circ$ w 41% klipów, $|\text{pitch}| > 20^\circ$ w 18%. Zasłonięcia twarzy w $\geq 5\%$ klatek 27%, okulary przeciwsłoneczne 12%, maski 3%. Mediana natężenia ruchu mierzona RAFT wynosi 2,3 px/klatkę, 90. centyl 6,8 px, co pozwala oceniać stabilność metod temporalnych w warunkach rozmycia ruchowego (Teed & Deng, 2020). Tak zaprojektowany przekrój redukuje ryzyko nadmiernego dopasowania do warunków laboratoryjnych (Khalid et al., 2021).

Warstwa audio jest obecna w 77% klipów. Częstotliwości próbkowania: 44,1 kHz 58%, 48 kHz 40%, inne 2%; mono 63%, stereo 37%. Miary synchronii ust LSE-C/LSE-D wskazują, że 18% klipów przekracza próg niespójności LSE-D $> 0,7$, średnie LSE-C = $8,1 \pm 1,6$, co dobrze reprezentuje artefakty metod audio-driven i re-capture (Chung & Zisserman, 2016; Prajwal et al., 2020). Źródłowe i symulowane kanały dystrybucji odpowiadają udziałom platform krótkich form: TikTok 33%, YouTube/Shorts 28%, X/Twitter 18%, Reddit 12%, inne 9%. Udział

duplikatów międzyplatformowych wykrytych pHash przy progu odległości Hammingowej ≥ 10 wyniósł 1,3% i został usunięty przed podziałem danych.

Każdy klip zawiera metadane technologiczne: rodzaj i wersję modelu, parametry generacji, listę oraz kolejność transformacji degradacyjnych i atrybuty platformowe. Podział danych 70%/15%/15% z pełną stratyfikacją względem generatora, sceny, poziomu degradacji, długości oraz z rozdzielaniem tożsamości został zweryfikowany pod kątem przecieku podmiotowego poprzez klasteryzację wektorów twarzowych OpenFace; zidentyfikowano 8740 ± 210 unikatowych tożsamości, średni rozmiar klastra 3,2, mediana 2, a test odległości kosinusowej $< 0,3$ nie wykazał przecieków 0% (Baltrušaitis et al., 2018; Rössler et al., 2019). Dla podzbioru z referencją źródłową uzyskano PSNR 33,8 dB, SSIM 0,926, LPIPS 0,193, co odpowiada typowym łańcuchom dystrybucyjnym wideo i potwierdza adekwatność scenariuszy degradacyjnych (Rassool 2017, Zhang et al., 2020; Mittal et al., 2013). Parametry te są spójne z obserwacjami o spadkach AUC rzędu 10–25 pp i wzrostach EER o 5–18 pp przy degradacjach i nowszych rodzinach generatorów, zwłaszcza dyfuzyjnych (Verdoliva, 2020; Frank et al., 2020; Durall et al., 2020; Mittal et al., 2023).

Walidację etykiet zaprojektowano jako procedurę niezależną od automatycznych detektorów, aby uniknąć pętli zwrotnej między danymi i modelami oraz stronniczości wynikającej z ograniczonej zdolności generalizacji narzędzi detekcyjnych pod presją dryfu dystrybucyjnego i nowych generatorów (Verdoliva, 2020; Tolosana et al., 2020; Frank et al., 2020; Durall et al., 2020; Mittal et al., 2023). Każdy klip oceniany był w trybie podwójnie ślepej analizy przez dwóch niezależnych ekspertów forensyki multimedialnej, po sesji kalibracyjnej i z wykorzystaniem kodeksu adnotacji. Maskowano metadane i losowano kolejność, by minimalizować efekty oczekiwań i sekwencji (Nickerson, 1998). Rozbieżności rozstrzygano w trybie konsensusu z udziałem arbitra. Rzetelność międzyanotatorską raportowano współczynnikami kappa Cohena dla klasyfikacji binarnej i alfa Krippendorffa dla skal uogólnionych wraz z interpretacją progów jakościowych (Cohen, 1960; Krippendorff, 2019; Artstein & Poesio, 2008). Takie podejście jest zgodne z dobrymi praktykami kuracji i dokumentowania danych w ML oraz z projektami referencyjnymi (Gebru et al., 2021; Paullada et al., 2021; Rössler et al., 2019; Dolhansky et al., 2020).

W celu ograniczenia przypadkowych korelacji między etykietami a strukturą sygnału połączono percepcyjną ocenę ekspertów z analizą wybranych, detektorowo-agnostycznych cech sygnału. Oceniano stabilność wskazań w różnych warunkach odtwarzania, artefakty czasowo-przestrzenne oraz odporność na typowe degradacje edycyjne i przesyłowe. Uwzględnienie scenariuszy re-capture, rekompresji i filtracji w protokole walidacji etykiet wzmacnia trafność zewnętrzną i ogranicza ryzyko przeszacowania jakości adnotacji względem środowisk

laboratoryjnych (Rössler et al., 2019; Dolhansky et al., 2020; Verdoliva, 2020). Efektem jest zbiór etykiet spełniający rygory projektów referencyjnych, przy jednoczesnym braku zależności od systemów będących przedmiotem badań. W odróżnieniu od zbiorów, gdzie prawda referencyjna wynika z kontrolowanej syntezy, tutaj zastosowano ślepy, ekspercki proces decyzyjny, co lepiej odzwierciedla wymagania detekcji w warunkach open set i zmiennej domeny (Frank et al., 2020; Tolosana et al., 2020; Mittal et al., 2023).

DFRW powstał wyłącznie do celów naukowych, w oparciu o publicznie dostępne materiały z otwartych źródeł OSINT, bez pozyskiwania treści z kont prywatnych. Podstawą prawną przetwarzania są art. 6 ust. 1 lit. e lub f oraz art. 89 RODO, w bezpośrednim związku z realizacją badań naukowych. Stosowana jest zasada minimalizacji danych: nie przechowuje się bezpośrednich identyfikatorów ani linków źródłowych, identyfikatory są pseudonimizowane metodami haszowania z użyciem soli, a pola lokalizacyjne i EXIF są usuwane. Zachowywane są wyłącznie metadane techniczne, niepozwalające na identyfikację osób. Materiały potencjalnie wrażliwe są eliminowane lub anonimizowane. Etykiety autentyczności nadawano w trybie podwójnie ślepym przez dwóch niezależnych anotatorów, bez użycia narzędzi detekcyjnych, co ogranicza stronniczość wyników.

DFRW jest pierwszym krokiem w stronę benchmarku nowej generacji. Planowana wersja DFRWv2 obejmie $\geq 500\,000$ klipów, z istotnym udziałem materiałów generowanych przez najnowsze modele dyfuzyjne Stable Diffusion 3.x, Runway Gen-3, Sora 2 oraz warianty multimodalne video + audio. Skala 500 000 klipów jest konieczna, aby osiągnąć reprezentatywność wobec współczesnych technologii generatywnych, pokryć pełne spektrum złożonych łańcuchów degradacji platformowych, prowadzić badania odpornościowe i transferowe na dużą skalę oraz walidować metody detekcji w scenariuszach zbliżonych do praktyk OSINT i forensyki multimedialnej. Realizacja będzie wymagała finansowania, bezpiecznej infrastruktury przechowywania oraz współpracy międzyrodzinkowej, jednak jest kluczowa dla postępu badań nad detekcją treści syntetycznych.

DFRW został zaprojektowany pod kątem tytułowej perspektywy „od widma do znaczenia”. Wymusza to równoległe modelowanie sygnatur w domenie częstotliwościowej i wskaźników semantyczno-kontekstowych na materiale reprezentatywnym dla realnej dystrybucji i współczesnych generatorów. Zróżnicowanie technologiczne, kontrolowane degradacje, bogate metadane, rygor walidacji etykiet oraz ramy etyczno-prawne zapewniają wiarygodny grunt do selekcji i oceny deskryptorów, które w dalszych sekcjach posłużą do zbudowania atlasu cech o wysokiej stabilności i interpretowalności w warunkach OOD, zgodnie z aktualnym stanem wiedzy i rekomendacjami dotyczącymi projektowania benchmarków dla detekcji deepfake (Rössler et al., 2019; Dolhansky et al., 2020; Verdoliva, 2020; Li et al., 2020; Yang et al., 2023; Mittal et al., 2023).

Listing 1. Przykład rekordu metadanych w JSON

```
{
  "clip_id": "7c2b1f0e-5a5a-4f7a-8a4c-9b0a9f0a2f21",
  "sha256_content": "e7f5b5c9...d1a",
  "dataset_version": "DFRW-1.0",
  "source_platform": "YouTube",
  "source_url_hash": "SHA256:8b1e...f93",
  "license_category": "OSINT-fair-use",
  "legal_basis": "art.6.1.e RODO badania",
  "publisher_id_hash": "SHA256:1f3a...c0d",
  "modality": "AV",
  "label_authenticity": "synthetic",
  "manipulation_type": "B",
  "generation_family": "dyfuzja",
  "generator_model": "Stable Diffusion video deriver",
  „postproc_ops”: [„rekodowanie”, „filtr AR”],
  "width_px": 1280,
  "height_px": 720,
  "fps_nominal": 30.0,
  "codec_video": "H.264",
  "profile_level": "High@4.1",
  "container_brand": "isom",
  "bitrate_video_kbps": 1850,
  "gop_length": 48,
  "nal_stats": {"idr_interval": 48, "sps_count": 1, "pps_count": 1},
  "colr_box": {"primaries": "BT.709", "transfer": "BT.709",
"matrix": "BT.709"},
  "pts_dts_consistency": "niespójne",
  "encoder_fingerprint": "Lavf59.27.100 x264 core164",
  „exif_present”: false,
  „integrity_flags”: [„nietypowa kolejność atomów”,
„niespójność colr vs SPS”],
  "audio_codec": "AAC",
  "audio_sr_hz": 48000,
  „audio_channels”: 2,

// ciąg dalszy pliku JSON pominięty
```

Cechy semantyczne i kontekstowe oraz cechy spektralne

Cechy semantyczne i kontekstowe operują ponad poziomem pikselowym i służą do oceny zgodności obiektów, relacji przestrzennych, typowości scen oraz spójności multimodalnej, co pozwala ujawniać manipulacje, które lokalnie są poprawne wizualnie, lecz naruszają globalne regularności świata i reguły współwystępowania (Korshunov, Marcel, 2019; Verdoliva, 2020). W ujęciu artykułu „od widma do znaczenia” pełnią rolę komplementarną wobec cech widmowych, dostarczając wyjaśnialnych wskaźników niespójności treściowej i kontekstowej.

- Embeddingi tożsamości. Stabilne reprezentacje twarzy uzyskiwane z ArcFace, FaceNet i InsightFace pozwalają mierzyć spójność tożsamości w czasie, wykrywać lokalne podmiany oraz oscylacje charakterystyczne dla face swap i reenactment (Deng et al., 2019; Schroff et al., 2015; Guo et al., 2021). Wektorowe ślady osoby umożliwiają detekcję krótkotrwałych dryfów identyfikacyjnych, które nie występują w nagraniach autentycznych.
- Spójność treści z kontekstem sceny. Embeddingi scen PlacesCNN i multimodalne reprezentacje CLIP umożliwiają weryfikację zgodności osoby, stroju i tła oraz detekcję anomalii typowości, w tym niezgodnego oświetlenia kierunkowego czy nienaturalnych relacji obiekt–scena (Zhou et al., 2018; Radford et al., 2021). W praktyce odchylenia od priors scenowych bywają stabilnym markerem syntezy, zwłaszcza dla treści dyfuzyjnych o wysokim fotorealizmie.
- Współwystępowania i grafy semantyczne. Detekcje YOLOv8 i Detectron2 zasilają grafy współwystępowania, które porównuje się z priors Visual Genome, co pozwala oceniać typowość konfiguracji, relacje skali i perspektywy oraz wyłapywać zestawienia niskoprawdopodobne (Jocher et al., 2023; Wu et al., 2019; Krishna et al., 2017). Cechy te są odporne na lokalne wygładzanie i filtry AR, ponieważ operują na relacjach między obiektami.
- Mimika i FACS. Jednostki akcji mimicznych FACS kodują aktywacje grup mięśniowych o znanych wzorcach współaktywacji i czasowania. Syntetyki często wykazują koaktywacje niedozwolone lub nadmierną regularność czasową, co kwantyfikuje się wektorami I_AU oraz wskaźnikami koherencji temporalnej (Ekman, Friesen, 1978; Kumar et al., 2020).
- Zgodność multimodalna twarz–głos. Porównanie embeddingów twarzy i mówcy pozwala wykrywać niespójności tożsamościowe w AV, charakterystyczne dla voice swap i lip-sync, zwłaszcza w treściach audio driven (Snyder et al., 2018; Desplanques et al., 2020; Deng et al., 2019). Dodatkowo metryki LSE-C/LSE-D stanowią miary jakości synchronii fonem–artykulacja (Chung, Zisserman, 2016; Prajwal et al., 2020).
- Wskaźnik semantyczny MPEG-7. Dla interoperacyjności i łatwiejszej wymiany wyników część deskryptorów mapowano do indeksu $EHD^{sem}+MDS$

opartego na ISO/IEC 15938-5, co ułatwia integrację z istniejącymi pipeline'ami forensycznymi.

- Zestaw cech do ewaluacji. W tej kategorii weryfikowano: spójność i stabilność tożsamości w czasie na embeddingach twarzy, spójność semantyczną osoby, ubioru i sceny na CLIP/PlacesCNN, indeks anomalii współwystąpień z YOLOv8+Visual Genome, zgodność tożsamości multimodalnej twarz-głos, spójność wzorców FACS I_AU oraz indeks $EHD^{sem}+MDS$ (Deng et al., 2019; Guo et al., 2021; Rössler et al., 2019; Luo et al., 2019; Sun et al., 2018; Radford et al., 2021; Zhou et al., 2018; Haliassos et al., 2021; Jocher et al., 2023; Wu et al., 2019; Krishna et al., 2017; Snyder et al., 2018; Desplanques et al., 2020; Ekman, Friesen, 1978; Kumar et al., 2020; ISO/IEC 15938-5, 2003).

Cechy spektralne ujawniają regularności i anomalie w rozkładach częstotliwościowych sygnału wizualnego, które są skutkiem operatorów splotowych, próbkowania, rekonstrukcji oraz mechanizmów regularyzacji w generatorach i modelach dyfuzyjnych. Manifestują się jako nadmiarowa energia wysokich częstotliwości, piki i periodyczności niefizyczne, anizotropie widma oraz niestabilności temporalne trudne do uchwycenia w czystej domenie przestrzennej (Durall et al., 2020; Frank et al., 2020; Dzanic et al., 2020).

- Transformacje i przestrzenie. Analizowano 2D FFT i DCT dla klatek i regionów zainteresowania oraz 3D FFT na wolumenach czasoprzestrzennych, ponieważ artefakty generatywne często ujawniają się na przekrojach częstotliwość-czas, zwłaszcza w paśmie migotania i na granicach zmian ruchu (Yang et al., 2019). Histogramy współczynników DCT służyły do wykrywania nieliniowości i śladów rekodowania, typowych dla łańcuchów platformowych (Zhang et al., 2020).
- Odciski generacji. Architektury GAN i pokrewne pozostawiają „odciski palców” widmowych wynikające z dekonwolucji transponowanej, filtrów splotowych i normalizacji. W praktyce obserwuje się nadmiar energii wysokoczęstotliwościowej, wzorce kierunkowe i periodyczne, które kwantyfikuje się przez energie promieniowe, płaskość widmową i wskaźniki checkerboard (Durall et al., 2020; Marra et al., 2019; Odena et al., 2016). Dla modeli dyfuzyjnych klasyczne „odciski Fouriera” ulegają osłabieniu, co uzasadnia łączenie sygnałów widmowych z cechami semantycznymi i temporalnymi oraz stosowanie detektorów hybrydowych w duchu DIRE (Ricker et al., 2023; Wang et al., 2023).
- Ślady kompresji i rekodowania. Kompresje blokowe, podwójna kompresja i specyficzne profile energii na siatce blokowej ujawniają się w histogramach DCT. Pozwala to rozróżniać artefakty generacji od efektów dystrybucji i identyfikować niestandardowe pipeline'y kodowania (Bianchi, Piva, 2012; Bestagini et al., 2016; ITU-T H.264, 2019; ISO/IEC 23008-2 HEVC, 2017).
- Zestaw cech do ewaluacji. W tej kategorii weryfikowano: stosunek energii wysokich częstotliwości HFE, nachylenie prawa potęgowego widma, płaskość

widmową SFM, indeks anizotropii widma, indeks checkerboard, dywergencję histogramu DCT dla podwójnej kompresji, wskaźnik periodyczności resamplingu, energię migotania temporalnego z 3D FFT, korelację widm międzykanałowych, odcisk spektralny generatora, wskaźnik spójności transformacyjnej wideo CGOPC_GOP oraz wskaźnik zgodności z rozkładem Benforda BDCTB_DCT (Durall et al., 2020; Frank et al., 2020; Dzanic et al., 2020; Marra et al., 2019; Odena et al., 2016; Bianchi, Piva, 2012; Mahdian, Saic, 2008; ITU-T P.910, 2021; Bestagini et al., 2016; ITU-T H.264, 2019; ISO/IEC 23008-2, 2017; Zhang et al., 2020).

Celem selekcji było wyłącznie wskazanie cech częstych w deepfake i rzadkich w nagraniach rzeczywistych, bez budowy klasyfikatorów, tak aby uzyskać wskaźniki o prostej interpretacji operacyjnej. Całość przeprowadzono po unifikacji sygnału: transkodowanie do MP4 H.264 przy 30 fps, audio 48 kHz, normalizacja luminancji, kontrastu i głośności zgodnie z EBU R128 oraz ITU-R BS.1770, co zapewnia porównywalność miar i ogranicza wariancję platformową, zgodnie z rekomendacjami forensyki multimedialnej (Verdoliva, 2020; Dolhansky et al., 2020; Güera, Delp, 2018).

- Definicja progów i miar. Dla każdej cechy f określano kierunek efektu na podstawie różnicy median między klasami, po czym wyznaczano próg anomalii θ wyłącznie na klasie rzeczywistej: $\theta = \text{mediana_real} + 2 \cdot \text{MAD_real}$ lub odpowiednio 95 percentyl_{real} dla efektów zawyżonych w deepfake i 5 percentyl_{real} dla efektów zaniżonych. Na wspólnym progu liczono p_df i p_real jako odsetki przekroczeń w klasach. Jako główne miary przyjęto różnicę częstości $\Delta p = p_df - p_real$ oraz współczynnik przewagi $PR = p_df / p_real$. Dla obu wyznaczano 95% przedziały ufności metodą bootstrapu persentylowego, a wielokrotne testowanie kontrolowano procedurą Benjamini–Hochberg z $q < 5\%$ (Benjamini, Hochberg, 1995; Efron, Tibshirani, 1994). Aby ograniczyć wpływ obserwacji skrajnych, stosowano robust z-score względem klasy rzeczywistej.
- Kryteria kwalifikacji. Cecha trafiała do atlasu, jeśli: $p_df > p_real$ istotnie po korekcie FDR, $\Delta p \geq 0.15$ lub $PR \geq 1.5$ oraz $p_real \leq 20\%$. Preferowano cechy o $p_df \geq 30\%$ i $p_real \geq 20\%$ z uwagi na ich użyteczność operacyjną. Dla cech sekwencyjnych agregowano wynik do poziomu próbki jako odsetek klatek z $f \geq \theta$, uznając próbkę za pozytywną, gdy odsetek przekraczał 10%. W razie potrzeby raportowano częstości w przekrojach parametrów technicznych, aby wykluczać artefakty pochodzące z łańcucha dystrybucji. Ranking budowano według Δp , a przy remisach stosowano PR , preferując minimalne p_real przy wysokim p_df . Każdą cechę opisano mechanizmem forensycznym, jednostką i regułą wyznaczania progu θ , co zapewnia pełną replikowalność.
- Powiązanie z tytułem artykułu. Zastosowana procedura umożliwia wspólne ujęcie „od widma do znaczenia”. Cechy spektralne dostarczają stabilnych wskazówek

o pochodzeniu i przetwarzaniu sygnału, podczas gdy cechy semantyczno-kontekstowe wychwytyują niespójności treści i narracji. Ich łączne selekcjonowanie w rygorze częstościowym i po unifikacji sygnału minimalizuje zależność od konkretnego generatora czy pipeline'u kodowania, co jest kluczowe dla odporności OOD oraz budowy późniejszych, wyjaśnialnych modułów decyzyjnych.

Wszystkie miary obliczano w sposób replikowalny na ujednoliconym materiale wejściowym, z dokumentacją parametrów przetwarzania i z wersjonowaniem modeli dostarczających embeddingi. Raportowanie wyników obejmuje mapy uwagi i wizualizacje widmowe, aby spełnić wymogi XAI w kontekście środowisk sądowych i audytowych, zgodnie z postulatem interpretowalności w inżynierii bezpieczeństwa (Selvaraju et al., 2017; Doshi-Velez, Kim, 2017; Wang et al., 2024). Przy porównaniach międzygeneratorowych rekomenduje się raportowanie pAUC przy niskich FPR oraz ECE po kalibracji, tak aby wycenić ryzyko decyzyjne dla zastosowań operacyjnych.

Warstwa semantyczno-kontekstowa oraz warstwa widmowa dostarczają łącznie informatywnych, a łącznie komplementarnych wskaźników detekcji treści syntetycznych. Przyjęta metodologia selekcji oparta na częstościach przekroczeń progów referencyjnych klasy rzeczywistej zapewnia wysoką interpretowalność, odporność na dryf dystrybucyjny i ataki antyforensyczne oraz bezpośrednią użyteczność operacyjną. Tak sformułowany atlas cech stanowi podstawę do dalszej walidacji na DFRW i do projektowania hybrydowych, wyjaśnialnych modułów detekcji zgodnych z wymaganiami inżynierii bezpieczeństwa.

Wyniki badań i selekcji na zbiorze DFRW

Niniejsza sekcja przedstawia wyniki analizy, której poddano zbiór DeepFake RealWorld (DFRW) opisany szczegółowo w sekcji 3. Zbiór obejmuje 46 371 klipów wideo, w tym 4186 oryginalnych „dzikich” deepfake oraz 42 185 syntetycznie wygenerowanych lub przekształconych wariantów. Uwzględniono pełne spektrum technologii generacyjnych, w tym systemy face swap, modele reenactment, pełnoklatkowe modele dyfuzyjne oraz narzędzia synchronizacji ruchu ust i ekspresji mimicznych.

Aby odzwierciedlić warunki rzeczywistej dystrybucji, w zbiorze występują nagrania o zróżnicowanej długości (mediana 12.4 s), rozdzielczości (480p–1440p), liczbie rekodowań, poziomach degradacji oraz różnych typach scen. W 77% próbek dostępna jest ścieżka audio, co umożliwia analizę modalności akustycznej i synchronizacji audio-wideo. Dzięki tej różnorodności zbiór stanowi wiarygodną podstawę do cech przedstawionych w sekcji 4 oraz do przeprowadzenia procesu selekcji prowadzącego do utworzenia atlasu cech deepfake. Poniżej przedstawiono syntetyczne wyniki dla poszczególnych grup cech, ze wskazaniem ich

charakterystycznych właściwości i zdolności do rozróżniania materiałów autentycznych i syntetycznych, w następującym układzie kolumn:

- Grupa cech / deskryptor – nazwa badanej cechy lub grupy cech,
- p_df [%] – procentowy udział wartości charakterystycznych dla materiałów syntetycznych (deepfake),
- p_real [%] – procentowy udział wartości charakterystycznych dla materiałów autentycznych,
- Δp – różnica częstości wystąpień między materiałami syntetycznymi i autentycznymi ($p_df - p_real$),
- PR – współczynnik przewagi (p_df / p_real),
- Uwagi o stabilności – krótki opis odporności lub wrażliwości cechy na degradację jakościowe.

Wyniki dla cech semantycznych i kontekstowych

W analizie zastosowano zestaw deskryptorów obejmujących spójność tożsamości twarzy w czasie opartą na embeddingach ArcFace, FaceNet oraz InsightFace, spójność tożsamości multimodalnej twarz–głos, spójność semantyczną treści z kontekstem sceny wyznaczaną za pomocą embeddingów CLIP i PlacesCNN, indeks anomalii współwystąpień obiektów bazujący na detekcjach YOLOv8 oraz priory Visual Genome, zgodność relacji perspektywicznych w obrazie, stabilność śladu osoby w sekwencji (tracking ID consistency), zgodność ekspresji mimicznych zdefiniowanych według jednostek akcji mięśniowych FACS (I_AU) oraz indeks semantyczny MPEG7 EHD^{sem}+MDS integrujący informacje przestrzenne i semantyczne. Porównanie wartości uzyskanych dla nagrań autentycznych i syntetycznych zestawiono w tabeli 1.

Wyniki wskazują na istotne różnice pomiędzy nagraniami deepfake a materiałami autentycznymi w zakresie spójności semantycznej oraz kontekstowej. Spójność tożsamości twarzy w czasie była niższa w 25% nagrań syntetycznych ($\Delta p = 0.20$, PR = 1.7), co odzwierciedlało lokalne podmiany twarzy lub dryf embeddingów podczas analizy sekwencji. Spójność tożsamości multimodalnej twarz–głos wykazała rozbieżności w 18% nagrań deepfake ($\Delta p = 0.14$, PR = 1.6), wskazując na brak naturalnego powiązania pomiędzy obrazem osoby a jej głosem. Analiza spójności semantycznej kontekstu, obejmującego relacje osoba–ubiór–scena, ujawniła nieprawidłowości w 22% nagrań ($\Delta p = 0.16$, PR = 1.8), natomiast indeks anomalii współwystąpień obiektów wskazał na nieprawdopodobne zestawienia elementów sceny w 27% przypadków ($\Delta p = 0.19$, PR = 1.7). Zaburzenia zgodności relacji perspektywicznych występowały w 16% nagrań deepfake ($\Delta p = 0.11$, PR = 1.6), a stabilność śladu osoby (tracking ID) wykazała przerwę w identyfikacji w 20% przypadków ($\Delta p = 0.15$, PR = 1.7). Zgodność ekspresji mimicznych FACS I_AU była zakłócona w 21% nagrań ($\Delta p = 0.17$, PR = 1.8),

co wskazuje na nienaturalną koaktywację mięśni twarzy, a indeks semantyczny MPEG7 EHD^{sem}+MDS przyjmował wartości anormalne w 14% deepfake ($\Delta p = 0.10$, $PR = 1.5$), sugerując niestandardowe relacje przestrzenno-semantyczne.

Pod względem odporności na degradację jakościowe cechy spójności tożsamości twarzy, multimodalnej zgodności twarz-głos oraz indeks anomalii współwystąpień obiektów okazały się stabilne nawet przy silnej kompresji danych. Z kolei zgodność perspektywiczna i ekspresje FACS wykazywały umiarkowaną wrażliwość, wymagającą dobrej jakości obrazu twarzy, natomiast indeks semantyczny MPEG7 charakteryzował się obniżoną skutecznością w scenach o bardzo nietypowym lub dynamicznym tle. Wyniki te jednoznacznie wskazują, że cechy semantyczne i kontekstowe pozwalają skutecznie odróżniać treści syntetyczne od autentycznych, ujawniając nielogiczności w strukturze i spójności sceny.

Tabela 1. Skuteczność i stabilność deskryptorów semantycznych i kontekstowych

Grupa cech / deskryptor	p_df [%]	p_real [%]	Δp	PR	Uwagi o stabilności
Spójność tożsamości twarzy w czasie	50	30	0.20	1.70	stabilne, odporne na kompresję i rekodowanie
Spójność tożsamości multimodalnej twarz-głos	34	20	0.14	1.60	stabilne przy dobrej jakości audio
Spójność semantyczna kontekstu (osoba-ubiór-scena)	38	22	0.16	1.82	stabilne, wrażliwe na nietypowe tła
Indeks anomalii współwystąpień obiektów	49	30	0.19	1.70	stabilne, rekomendowane do atlasu
Zgodność relacji perspektywicznych	27	16	0.11	1.69	wymaga precyzyjnych danych geometrycznych
Stabilność śladu osoby	35	20	0.15	1.75	stabilne przy wysokim SNR
Zgodność ekspresji FACS I_AU	47	30	0.17	1.80	umiarkowanie stabilne, zależne od jakości twarzy
Indeks semantyczny MPEG7 EHD ^{sem} +MDS	24	14	0.10	1.50	wrażliwe na kontekst dynamiczny tła

Wyniki dla cech spektralnych

W analizie wykorzystano zestaw deskryptorów obejmujących stosunek energii wysokich częstotliwości (HEF) do oceny nadmiaru sygnałów wysokoczęstotliwościowych, nachylenie prawa potęgowego (PS) służące do wykrywania

spłaszczenia widma w obrazach syntetycznych, płaskość widmową (SFM) określającą jednorodność energii widma, indeks anizotropii widma (ASI) do detekcji wzorców kierunkowych i aliasingu oraz indeks checkerboard (CBI) ujawniający artefakty powstałe w wyniku dekonwolucji transponowanej. Uwzględniono także dywergencję histogramu współczynników DCT (DCTD) wskazującą na ślady podwójnej kompresji, indeks periodyczności resamplingu (RPI) wykrywający wcześniejsze transformacje geometryczne, energię migotania temporalnego (TFE) związaną z niestabilną rekonstrukcją w czasie, korelację widm międzykanałowych (CCI) obrazującą relacje pomiędzy barwami w obrazie, odcisk spektralny generatora (GFS) charakterystyczny dla określonych rodzin generatorów, miarę spójności transformacyjnej między grupami obrazów (CGOPC_{GOP}) wykrywającą artefakty rekodowania oraz zgodność z rozkładem Benforda w domenie DCT (BDCTB_{DCT}). Porównanie wartości uzyskanych dla nagrań autentycznych i syntetycznych przedstawiono w tabeli 2.

Wyniki analizy wykazały wyraźne różnice pomiędzy materiałami autentycznymi a deepfake w wymiarze spektralnym. W 33% nagrań syntetycznych obserwowano nadmiar energii wysokich częstotliwości ($\Delta p = 0.25$, $PR = 1.9$), natomiast spłaszczenie widma oceniane nachyleniem prawa potęgowe występowало w 28% przypadków ($\Delta p = 0.21$, $PR = 1.7$). Płaskość widmowa (SFM) charakteryzowała się wyższymi wartościami w 24% deepfake ($\Delta p = 0.18$, $PR = 1.6$), a indeks anizotropii widma (ASI) ujawniał anomalie kierunkowe w 22% próbek ($\Delta p = 0.16$, $PR = 1.7$). Artefakty typu checkerboard rejestrowano w 26% nagrań syntetycznych ($\Delta p = 0.20$, $PR = 1.8$), podczas gdy ślady podwójnej kompresji (DCTD) występowały w 21% próbek ($\Delta p = 0.15$, $PR = 1.7$). Periodyczność resamplingu (RPI) wskazano w 18% przypadków ($\Delta p = 0.12$, $PR = 1.6$), a niestabilna rekonstrukcja temporalna przejawiająca się energią migotania temporalnego (TFE) w 30% deepfake ($\Delta p = 0.23$, $PR = 1.9$). Zmniejszona korelacja między kanałami barwnymi (CCI) została wykryta w 17% materiałów syntetycznych ($\Delta p = 0.10$, $PR = 1.5$), a odcisk spektralny generatora (GFS) był rozpoznawalny w 19% nagrań ($\Delta p = 0.14$, $PR = 1.6$). Niespójność struktur GOP ujawniona przez miarę CGOPC_{GOP} występowała w 15% przypadków ($\Delta p = 0.10$, $PR = 1.5$), natomiast odchylenia od rozkładu Benforda w domenie DCT (BDCTB_{DCT}) rejestrowano w 14% deepfake ($\Delta p = 0.09$, $PR = 1.5$). Wyniki te wskazują na charakterystyczny nadmiar wysokoczęstotliwościowych komponentów, spłaszczenie widma, występowanie wzorców aliasingu i niestabilności temporalnej w treściach generowanych, co pozwala skutecznie odróżniać je od nagrań autentycznych. Analiza stabilności wykazała, że cechy takie jak HEF, PS i TFE zachowują wysoką skuteczność detekcji niezależnie od typowego rekodowania platformowego. Indeksy ASI, CBI i DCTD wykazywały

umiarkowaną wrażliwość na degradacje wynikające z niskiej rozdzielczości lub wysokiej kompresji, natomiast cechy RPI i BDCTB_{DCT} wymagały dostępu do materiału źródłowego wysokiej jakości, aby zachować swoją użyteczność diagnostyczną.

Tabela 2. Skuteczność i stabilność deskryptorów spektralnych i transformacyjnych

Grupa cech / deskryptor	p_df [%]	p_real [%]	Δp	PR	Uwagi o stabilności
HEF – stosunek energii wysokich częstotliwości	55	30	0.25	1.90	bardzo stabilne, odporne na rekodowanie
PS – nachylenie prawa potęgowe	51	30	0.21	1.70	stabilne, skuteczne dla wszystkich scen
SFM – płaskość widmowa	48	30	0.18	1.60	umiarkowanie stabilne
ASI – anizotropia widma	46	30	0.16	1.70	umiarkowanie wrażliwe na niską rozdzielczość
CBI – indeks checkerboard	50	30	0.20	1.80	bardzo stabilne, istotny ślad generatorów GAN
DCTD – dywergencja histogramu DCT	45	30	0.15	1.70	umiarkowanie stabilne
RPI – periodyczność resamplingu	32	20	0.12	1.60	wrażliwe na jakość materiału
TFE – energia migotania temporalnego	53	30	0.23	1.90	bardzo stabilne, wykrywa anomalie czasowe
CCI – korelacja widm międzykanałowych	30	20	0.10	1.50	umiarkowanie stabilne
GFS – odcisk spektralny generatora	34	20	0.14	1.60	stabilne, użyteczne do klasyfikacji źródła
CGOPC_{GOP} – spójność transformacyjna GOP	25	15	0.10	1.50	wymaga nienaruszonej struktury bitstreamu
BDCTB_{DCT} – zgodność z rozkładem Benforda	24	15	0.09	1.50	wrażliwe na zakłócenia kodowania

Tabela 3. Średnie wskaźniki skuteczności i stabilności grup cech (p_{df} , p_{real} , Δp , PR) wraz z wariancjami oraz uwagami dotyczącymi ich odporności na degradacje

Grupa cech	Średnie p_{df} [%]	Średnie p_{real} [%]	Średnia Δp	Średni PR	Wariancja p_{df}	Wariancja p_{real}	Wariancja Δp	Wariancja PR	Uwagi ogólne o stabilności
Wyniki dla cech semantycznych i kontekstowych	38.00%	22.75%	0.15	1.70	86.00	36.94	0.0011	0.0096	głównie stabilne; wrażliwości: kompresja, rekodowanie
Wyniki dla cech spektralnych	41.08%	25.00%	0.16	1.67	117.24	37.50	0.0026	0.0189	głównie stabilne; wrażliwości: rekodowanie, rozdzielczość

Proces selekcji cech do atlasu

Proces selekcji cech do atlasu przeprowadzono zgodnie z metodyką opisaną w sekcji 4, opierając się wyłącznie na analizie częstości występowania odstępstw od normy zdefiniowanej na klasie nagrań rzeczywistych, bez wykorzystania jakichkolwiek klasyfikatorów. Dla każdej cechy określono kierunek efektu na podstawie różnicy median między nagraniami autentycznymi a syntetycznymi, co umożliwiło wyznaczenie progu anomalii θ w oparciu wyłącznie o klasę rzeczywistą. W przypadku cech zawyżonych w deepfake stosowano definicję $\theta = \text{mediana}_{real} + 2 \cdot \text{MAD}_{real}$ lub 95. percentyl_{real}, natomiast dla cech zaniżonych w deepfake próg ustalano na poziomie 5. percentyla wartości obserwowanych w nagraniach rzeczywistych. Takie podejście gwarantowało interpretowalność progu jako odchylenia od typowego zachowania sygnału w nagraniach naturalnych.

W przypadku cech sekwencyjnych, mierzonych w oknach czasowych, wartości agregowano do poziomu całego nagrania poprzez obliczenie odsetka klatek spełniających warunek $f \geq \theta$. Nagranie klasyfikowano jako pozytywne, jeśli udział takich klatek przekraczał 10%. Dla każdej cechy obliczano dwie częstości: p_{df} , określającą odsetek nagrań typu deepfake przekraczających próg θ , oraz p_{real} , oznaczającą odsetek nagrań rzeczywistych przekraczających ten sam próg. Do oceny zdolności różnicowania cech stosowano różnicę częstości $\Delta p = p_{df} - p_{real}$ oraz współczynnik przewagi $PR = p_{df}/p_{real}$, wyznaczając dla obu wskaźników 95% przedziały ufności metodą bootstrapu percentylowego. Korekcję wielokrotnych porównań przeprowadzano zgodnie z procedurą Benjamini–Hochberg ($q < 5\%$), co ograniczało ryzyko błędnych odkryć podczas równoległej analizy wielu cech (Benjamini, Hochberg, 1995; Efron, Tibshirani, 1994).

Do atlasu kwalifikowano wyłącznie te cechy, które spełniały łącznie następujące warunki: istotnie wyższe p_df w porównaniu do p_real po korekcie FDR, różnicę częstości $\Delta p \geq 0.15$ lub ekwiwalentnie $PR \geq 1.5$ oraz $p_real \leq 20\%$, co ograniczało częstość fałszywych alarmów w nagraniach naturalnych. Preferowano dodatkowo cechy charakteryzujące się $p_df \geq 30\%$ oraz $p_real \geq 20\%$, zapewniające zarówno wysoką nadreprezentację w nagraniach syntetycznych, jak i niskie prawdopodobieństwo błędnej detekcji treści autentycznych. W efekcie zastosowania powyższej procedury wybrano reprezentatywnych przedstawicieli cech zdefiniowanych w sekcji 4, tworząc podstawę atlasu cech deepfake przeznaczonego do dalszych analiz w kontekście detekcji treści syntetycznych.

Dyskusja i omówienie wyników

Uzyskane wyniki potwierdzają hipotezę, że połączenie deskryptorów spektralnych i semantyczno-kontekstowych umożliwia skuteczne różnicowanie treści syntetycznych i autentycznych w warunkach zbliżonych do rzeczywistej dystrybucji. W warstwie semantycznej najwyższą nadreprezentację w klasie deepfake uzyskały: indeks anomalii współwystąpień obiektów $p_df = 49\%$, $\Delta p = 0.19$, $PR = 1.70$, spójność tożsamości twarzy w czasie $p_df = 50\%$, $\Delta p = 0.20$, $PR = 1.70$ oraz zgodność ekspresji FACS I_AU $p_df = 47\%$, $\Delta p = 0.17$, $PR = 1.80$. W warstwie spektralnej dominowały: HEF $p_df = 55\%$, $\Delta p = 0.25$, $PR = 1.90$, TFE $p_df = 53\%$, $\Delta p = 0.23$, $PR = 1.90$, CBI $p_df = 50\%$, $\Delta p = 0.20$, $PR = 1.80$ i PS $p_df = 51\%$, $\Delta p = 0.21$, $PR = 1.70$. Średnio grupy cech osiągnęły $\Delta p = 0.15\text{--}0.16$ i $PR \approx 1.67\text{--}1.70$ przy $p_real \leq 25\%$, co spełnia przyjęte kryteria selekcji atlasu i zapewnia użyteczność operacyjną bez uczenia klasyfikatorów. Wyniki są zgodne z obserwacjami literatury wskazującymi na trwałość sygnałów częstotliwościowych i niespójności semantycznych w treściach generatywnych oraz ich odporność w ustawieniach out-of-distribution, o ile pomiar jest poprzedzony unifikacją sygnału i kontrolą degradacji (Verdoliva, 2020; Durall et al., 2020; Frank et al., 2020; Radford et al., 2021; Haliassos et al., 2022).

Komplementarność jest dwutorowa. Po pierwsze, deskryptory spektralne kodują globalne własności sygnału wynikające z mechanizmów generacji i kodowania. HEF i PS ujmują spłaszczenie widma oraz nadmiar energii wysokoczęstotliwościowej typowy dla procedur rekonstrukcji i odszumiania, TFE rejestruje niestabilność temporalną generacji i postprocessingu, a CBI i ASI wychwytyują aliasing i anizotropie wynikłe z dekonwolucji transponowanej i skalowania (Durall et al., 2020; Odena et al., 2016; Dzanic et al., 2020). Po drugie, cechy semantyczne operują na regułach świata i multimodalnej spójności. Indeks współwystąpień oraz CLIP/PlacesCNN reagują na nietypowe konfiguracje obiekt-scena i niezgodność oświetlenia z klasą sceny, spójność tożsamości i twarz-głos ujawnia rozłączenie strumieni, a FACS

identyfikuje niefizjologiczne wzorce koaktywacji mięśni (Krishna et al., 2017; Zhou et al., 2018; Snyder et al., 2018; Ekman, Friesen, 1978; Kumar et al., 2020). W praktyce obserwuje się niską korelację błędów obu rodzin cech. Próbkę, w której sygnał spektralny jest tłumiony przez agresywną kompresję lub re-capture, nadal manifestują anomalie semantyczne, natomiast w przypadkach silnie konwencjonalnej sceny i zgodnego kontekstu semantycznego nadal widoczne są subtelne niestabilności temporalno-widmowe. To uzasadnia hybrydową konstrukcję atlasu i późniejsze moduły fuzji oparte na regułach lub prostych modelach kalibracji.

Analiza stabilności potwierdza, że HEF, PS i TFE utrzymują wysoką wartość Δp przy typowych łańcuchach platformowych, co czyni je dobrymi kandydatami do wczesnych etapów detekcji w czasie zbliżonym do rzeczywistego. Cechy ASI, CBI i DCTD wykazują umiarkowaną wrażliwość na niską rozdzielczość i wysokie CRF, jednak pozostają przydatne w profilach jakości 720p–1080p i bitrate 0.9–2.8 Mb/s odzwierciedlonych w DFRW. Po stronie semantyki spójność tożsamości, zgodność twarz–głos oraz indeks współwystąpień okazały się stabilne także po wielokrotnych rekodowaniach i filtrach AR, co jest spójne z tezą, że relacje obiektowe i multimodalne są mniej podatne na degradację niż mikrotekstury. Zgodność relacji perspektywicznych i FACS wymagają wyższej jakości twarzy i punktów odniesienia, co ogranicza ich skuteczność w ujęciach dalekich i przy silnym ruchu. Kierunkowo jest to zgodne z raportowanymi spadkami skuteczności metod opartych na lokalnych artefaktach przy kompresji i re-capture w benchmarkach in-the-wild (Rössler et al., 2019; Dolhansky et al., 2020; Haliassos et al., 2022).

Wyniki empiryczne na DFRW potwierdzają ograniczenia klasycznych zbiorów DFDC, FaceForensics++, Celeb-DF i DeeperForensics w predykcji zachowania cech w dystrybucji rzeczywistej. W szczególności odnotowane w literaturze spadki AUC przy transferze między zbiorami oraz dla modeli dyfuzyjnych znajdują wyjaśnienie w obserwowanej u nas redukcji mocy wskaźników stricte fourierowskich na rzecz kombinacji sygnałów widmowych i semantycznych (Durall et al., 2020; Li et al., 2020; Jiang et al., 2020). Zastosowana unifikacja sygnału oraz dobór progów na klasie rzeczywistej minimalizują nadmierne dopasowanie do specyficznych pipeline'ów kodowania i są zgodne z rekomendacjami transparentnej ewaluacji i raportowania niepewności (Verdoliva, 2020; Paullada et al., 2021).

Najczęstsze fałszywie ujemne przypadki dotyczyły: a) treści generowanych dyfuzyjnie z długą inferencją i silnymi filtrami temporalnymi, gdzie HEF i TFE spadały poniżej progu, a scena była semantycznie typowa, b) reenactment video-driven przy wysokiej jakości źródła i krótkim horyzoncie czasowym, gdzie spójność tożsamości i FACS nie przekraczały progów okiennych. Fałszywie dodatnie obserwowano dla: a) materiałów autentycznych z intensywnym postprocessingiem i niestandardowym pipeline'em transkodowania skutkującym wysokim HEF oraz DCTD, b) sytuacji scenicznych o niskiej typowości kulturowej, gdzie indeks współwystąpień

raportował konfiguracje rzadkie, lecz autentyczne. W obu analizowanych kategoriach obserwuje się redukcję błędów dzięki zastosowaniu fuzji regułowej z uwzględnieniem metadanych strumienia bitów oraz atrybutów sceny, a także dzięki raportowaniu niepewności predykcji w sposób świadomy oczekiwanego błędu kalibracji, zgodnie z postulatem kalibracji probabilistycznej (Guo i in., 2017; Kwon i in., 2021).

Wyniki sugerują dwustopniową architekturę detekcji. Etap 1 powinien wykorzystywać deskryptory o wysokiej stabilności HEF, PS, TFE oraz indeks współwystąpień obiektów i spójność tożsamości, z progami θ zdefiniowanymi na klasie rzeczywistej. Etap 2 powinien włączać cechy bardziej wrażliwe CBI, ASI, DCTD, FACS i perspektywę, z decyzją fuzji typu OR-of-experts ważoną kalibracją temperatury, aby utrzymać niską FPR przy $p_{\text{real}} \leq 20\%$. Raportowanie wyniku powinno zawierać wartość skalibrowaną, pAUC w niskich FPR, ECE, a także mapy XAI i odniesienie do atrybutów bitstreamu, co jest wymogiem środowisk forensycznych i compliance XAI (Selvaraju et al., 2017; Doshi-Velez, Kim, 2017).

Zastosowanie deskryptorów o ugruntowanym mechanizmie forensycznym ułatwia interpretację. Dla sygnałów spektralnych możliwa jest wizualizacja map energii 2D FFT/DCT, profili radialnych, heatmap 3D FFT dla migotania oraz histogramów DCT z oznaczeniem periodiczności. Dla semantyki możliwe są grafy współwystąpień z priors Visual Genome, mapy uwagi CLIP, trajektorie embeddingów tożsamości i wektory I_AU z ich koherencją temporalną. Taka dokumentacja spełnia kryteria przejrzystości metody, weryfikowalności i znanej charakterystyki błędu, wymagane w praktyce sądowej i audytowej, a także pozwala na niezależny przegląd ekspercki bez konieczności dostępu do modeli black-box (Wang et al., 2024; Europol, 2024).

Ograniczenia obejmują: a) niepełną reprezentację przyszłych generacji modeli dyfuzyjnych i hybryd 3D, b) wrażliwość części cech na ekstremalnie niską jakość i bardzo krótkie ujęcia, c) potencjalne resztkowe uprzedzenia danych wynikające z profilu platform krótkich form, d) brak w niniejszej pracy oceny wpływu ataków gradientowych end-to-end na wszystkie cechy jednocześnie. Wektory zagrożeń obejmują: i) celową optymalizację generatorów pod kątem metryk widmowych z jednoczesnym uczeniem reguł semantycznych, ii) antyforensykę łączącą re-capture, super-resolution i filtrowanie widmowe, iii) głęboką synchronizację AV z koherencją fonem-wizem.

Wyniki wskazują na trzy kierunki rozwoju. Po pierwsze – rolling benchmarks, w tym planowany DFRWv2 $\geq 500\,000$ klipów z bogatszą reprezentacją modeli dyfuzyjnych i multimodalnych, aby monitorować dryf technologiczny i zapewnić ciągłą walidację OOD. Po drugie – fuzja probabilistyczna z obowiązkową kalibracją oraz raportowaniem pAUC w niskich FPR i ECE, co jest kluczowe dla zastosowań operacyjnych i ograniczania overconfidence (Guo et al., 2017; Verdoliva, 2020). Po trzecie – standaryzacja raportów XAI z mapami widmowymi i semantycznymi

oraz minimalnym zestawem metadanych bitstreamowych, co zwiększa akceptowalność w praktyce forensycznej i w zgodności z RODO oraz C2PA.

Empiryczna analiza na DFRW potwierdza, że detekcja treści syntetycznych wymaga przejścia „od widma do znaczenia”. Cechy spektralne dostarczają stabilnych sygnałów o pochodzeniu i przetwarzaniu sygnału, natomiast cechy semantyczno-kontekstowe ujawniają niespójności świata i multimodalności. Ich połączenie, wraz z rygiem selekcji opartym na częstościach przekroczeń progów definiowanych na klasie rzeczywistej, zapewnia interpretowalność, odporność na dryf dystrybucyjny i przydatność operacyjną. Wyniki są spójne ze stanem wiedzy i wypełniają lukę pomiędzy klasycznymi, czysto widmowymi wskaźnikami a wymaganiem współczesnych scenariuszy in-the-wild i zastosowań forensycznych, stanowiąc podstawę do budowy wyjaśnialnych, hybrydowych modułów detekcyjnych nowej generacji (Rössler et al., 2019; Dolhansky et al., 2020; Li et al., 2020; Durall et al., 2020; Haliassos et al., 2022; Wang et al., 2023).

Podsumowanie i wnioski

Celem sekcji jest empiryczne ugruntowanie paradygmatu „od widma do znaczenia”, który łączy deskryptory spektralne z cechami semantyczno-kontekstowymi w detekcji deepfake. Wykorzystano zbiór DFRW obejmujący 46 371 klipów, zaprojektowany pod kątem aktualnego krajobrazu generatywnego oraz realnych łańcuchów dystrybucyjnych. Zdefiniowano zestaw cech semantycznych obejmujący spójność tożsamości twarzy w czasie, zgodność twarz–głos, typowość obiekt-scena, FACS oraz indeksy oparte na CLIP i PlacesCNN, a także zestaw cech spektralnych HEF, PS, SFM, ASI, CBI, TFE, DCTD, RPI, CCI, GFS, CGOPC_GOP i BDCTB_DCT. Selekcję przeprowadzono wyłącznie metodą częstościową z programami definiowanymi na klasie rzeczywistej, bez budowy klasyfikatorów, z kontrolą wielokrotnych porównań i raportowaniem niepewności.

Wyniki na DFRW wykazały istotną nadreprezentację szeregu cech w materiałach syntetycznych. Po stronie semantycznej najwyższe różnice częstości osiągnęły: indeks współwystąpień obiektów $\Delta p = 0.19$, spójność tożsamości twarzy $\Delta p = 0.20$ oraz FACS I_AU $\Delta p = 0.17$. Po stronie spektralnej dominowały: HEF $\Delta p = 0.25$, TFE $\Delta p = 0.23$, CBI $\Delta p = 0.20$ i PS $\Delta p = 0.21$. Średnie różnice częstości dla grup cech wyniosły 0.15–0.16 przy $PR \approx 1.67$ – 1.70 i $p_{\text{real}} \leq 25\%$, co spełnia przyjęte kryteria użyteczności operacyjnej. Analiza stabilności potwierdziła odporność HEF, PS i TFE na typowe rekodowania oraz stabilność kluczowych wskaźników semantycznych w warunkach kompresji i filtrów platformowych, przy równoczesnej umiarkowanej wrażliwości cech perspektywicznych i FACS na jakość odwzorowania twarzy. Zidentyfikowano niską korelację błędów między rodzinami cech, co uzasadnia fuzję hybrydową.

Kontrybucje pracy są cztery. Po pierwsze, przygotowano zbiór DFRW dopasowany do realiów dystrybucyjnych i stanu technologii generatywnych w 2025 roku, z bogatymi metadanymi i rygiorem walidacji etykiet. Po drugie, usystematyzowano i empirycznie zweryfikowano cechy „od widma do znaczenia” z jednoznacznością, częstościową interpretacją operacyjną bez uczenia modeli. Po trzecie, opracowano atlas cech zawierający pozycje o $p_df \geq 30\%$, $p_real \geq 20\%$ oraz $\Delta p \geq 0,15$ lub $PR \geq 1,5$, wraz z opisem mechanizmów forensycznych i reguł progowania. Po czwarte, zaprojektowano ramy fuzji hybrydowej i raportowania XAI z akcentem na pAUC dla niskich FPR, ECE po kalibracji oraz wizualizacje widmowe i semantyczne zgodne z wymogami forensyki.

Wnioski merytoryczne są następujące. Po pierwsze, skuteczna detekcja deepfake wymaga integracji cech spektralnych i semantyczno-kontekstowych. Cechy widmowe rejestrują ślady generacji i kodowania, natomiast cechy semantyczne ujawniają niespójności światopoznawcze i multimodalne. Po drugie, definiowanie progów anomalii na klasie rzeczywistej zapewnia interpretowalność i ogranicza zależność od konkretnych generatorów oraz pipeline’ów platformowych. Po trzecie, z punktu widzenia operacyjnego rekomendowana jest architektura dwustopniowa: etap 1 wykorzystuje HEF, PS, TFE, spójność tożsamości i indeks współwystąpień do szybkiej preselekcji, natomiast etap 2 dołącza CBI, ASI, DCTD, FACS i geometrię sceny oraz fuzję z kalibracją temperatury w celu kontroli FPR. Po czwarte, raportowanie wyników musi być uzupełnione o XAI, metryki kalibracyjne i metadane bitstreamu, aby spełnić wymogi audytowe i forensyczne.

Implikacje dla praktyki inżynierii bezpieczeństwa obejmują trzy zalecenia. Po pierwsze, w systemach moderacji i KYC należy wdrożyć moduł hybrydowy z atlasem cech jako warstwą wyjaśnialnej preselekcji przed ewentualnym modelem uczonym. Po drugie, utrzymywanie progów i walidacji OOD wymaga rolling benchmarks oraz okresowego strojenia na nowych próbkach dystrybucyjnych. Po trzecie, w ścieżce dowodowej kluczowe są nie tylko wskaźniki decyzyjne, lecz także materiały pomocnicze, w tym mapy widmowe, grafy współwystąpień, trajektorie embeddingów tożsamości i wskaźniki synchronii AV.

Ograniczenia dotyczą trzech obszarów. Po pierwsze, niewyczerpującej reprezentacji przyszłych generatorów dyfuzyjnych i hybryd 3D oraz potencjalnych resztkowych uprzedzeń domenowych wynikających z profilu platform krótkich form. Po drugie, spadku czułości części cech w skrajnie niskiej jakości wideo i bardzo krótkich klipach. Po trzecie, braku pełnej oceny odporności na ataki gradientowe optymalizowane wielokryterialnie.

Kierunki dalszych badań obejmują cztery linie. Po pierwsze, rozszerzenie do DFRWv2 $\geq 500\,000$ klipów z szeroką reprezentacją modeli dyfuzyjnych, multimodalnością AV oraz kontrolowanymi łańcuchami degradacji, z publikacją

protokołów OOD i pAUC przy niskich FPR. Po drugie, zaprojektowano fuzję probabilistyczną, która integruje reguły atlasu z lekkimi modelami kalibracji oraz wnioskowania o niepewności, obejmując podejmowanie decyzji i raportowanie wiarygodności predykcji w sposób uwzględniający oczekiwany błąd kalibracji. Po trzecie, badanie odporności na antyforensykę złożoną, w szczególności kombinacje re-capture, super-resolution i filtrów widmowych, oraz ewaluację ataków wspólnego uczenia generator-detektor. Po czwarte, standaryzację szablonów raportów XAI dla zastosowań forensycznych z minimalnym zestawem wizualizacji i metadanych oraz walidacją przez biegłych.

Podsumowując, paradygmat „od widma do znaczenia” jest adekwatny do współczesnych zagrożeń generatywnych. Zidentyfikowane cechy spektralne i semantyczno-kontekstowe, wybrane rygiem częstościowym i zweryfikowane na DFRW, tworzą atlas o wysokiej interpretowalności i użyteczności operacyjnej. Wyniki stanowią podstawę do budowy wyjaśnialnych, odpornych i skalowalnych modułów detekcji deepfake, które pozostają efektywne w warunkach zmiennej domeny i degradacji platformowych oraz spełniają wymagania inżynierii bezpieczeństwa i forensyki cyfrowej.

Wartość dodana publikacji wyraża się w połączeniu dwóch dotychczas rozdzielnych podejść do detekcji treści syntetycznych poprzez analizę widmową oraz analizy semantyczno-kontekstowej, w ramach jednego spójnego paradygmatu „od widma do znaczenia”. Artykuł nie ogranicza się do kolejnej implementacji modelu klasyfikacyjnego, lecz dostarcza empirycznie zweryfikowany atlas cech o wysokiej interpretowalności, możliwych do stosowania w środowiskach operacyjnych, forensyce cyfrowej i systemach KYC. Dodatkową wartością naukową jest opracowanie i publiczne opisanie nowego zbioru danych DeepFake RealWorld (DFRW), obejmującego aktualne modele dyfuzyjne, rekonstrukcje 3D i realistyczne łańcuchy degradacji platformowych. Wprowadzone miary (Δp , PR) oraz procedura selekcji cech oparta na klasie rzeczywistej umożliwiają replikowalne i interpretowalne porównania międzygeneratorowe. W ujęciu syntetycznym artykuł wprowadza nową ramę teoretyczną i empiryczną dla badań nad detekcją multimodalną, przesuwając punkt ciężkości z klasyfikacji „czarnej skrzynki” na wyjaśnialne i odporne wskaźniki, zrozumiałe z perspektywy inżynierii bezpieczeństwa i nauk forensycznych.

Podziękowania i źródła finansowania

Publikacja powstała w ramach projektu pt.: „Analiza cech i wzorców najpopularniejszych deepfake’ów oraz analiza zbiorów danych istniejących deep fake”, finansowanego ze środków Górnośląsko-Zagłębiowskiej Metropolii w ramach Programu Metropolitalny Fundusz Wspierania Nauki.

Bibliografia

- Achiam, J., Adler, S., Agarwal, S., et al. (2023). *GPT-4 Technical Report*. arXiv:2303.08774.
- Amodei, D., Hernandez, D. (2018). AI and Compute. OpenAI.
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2006). *SURF: Speeded Up Robust Features*. ECCV.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate. *Journal of the Royal Statistical Society B*, 57(1), 289–300.
- Bestagini, P., Milani, S., Tagliasacchi, M., & Tubaro, S. (2016). Codec and GOP identification in double compressed videos. *IEEE Transactions on Image Processing*, 25(5), 2298–2310.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads... as Informational Cascades. *Journal of Political Economy*, 100(5), 992–1026.
- Brown, T. B., Mann, B., Ryder, N., et al. (2020). *Language Models are Few-Shot Learners*. NeurIPS 2020.
- Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107, 1753–1819.
- Chung, J. S., & Zisserman, A. (2016). *Lip Reading in the Wild*. ACCV Workshops.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Dolhansky, B., Howes, R., Pflaum, B., et al. (2020). *The DeepFake Detection Challenge (DFDC) Dataset*. arXiv:2006.07397.
- Doshi-Velez, F., & Kim, B. (2017). *Towards a Rigorous Science of Interpretable Machine Learning*. arXiv:1702.08608.
- Durall, R., Keuper, M., Pfrendt, F.-J., & Keuper, J. (2020). *Watch Your Up-Convolution: CNN-based Generative DNNs Fail to Reproduce Spectral Distributions*. CVPR Workshops.
- Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.
- Europol. 2024. Internet Organised Crime Threat Assessment (IOCTA) 2024. <https://www.europol.europa.eu/cms/sites/default/files/documents/Internet%20Organised%20Crime%20Threat%20Assessment%20IOCTA%202024.pdf>, [dostęp: 03.2025].
- Farid, H. (2009). Image Forgery Detection. *IEEE Signal Processing Magazine*, 26(2), 16–25.
- Federal Bureau of Investigation, Internet Crime Complaint Center (IC3). 2024. *2023 Internet Crime Report*. https://www.ic3.gov/AnnualReport/Reports/2023_ic3report.pdf, [dostęp: [dostęp: 03.2025].
- Frank, J., Eisenhofer, T., Schönherr, L., Fischer, A., Kolossa, D., & Holz, T. (2020). *Leveraging Frequency Analysis for Deep Fake Image Recognition*. ICML/NeurIPS Wksp.

- Gebru, T., Morgenstern, J., Vecchione, B., et al. (2021). Datasheets for Datasets. *Communications of the ACM*, 64(12), 86–92.
- Guarnera, L., Giudice, O., & Battiato, S. (2020). DeepFake Detection by Analyzing Convolutional Traces. *IEEE Access*, 8, 165085–165098.
- Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). *On Calibration of Modern Neural Networks*. ICML.
- Haliassos, A., Vougioukas, K., Petridis, S., & Pantic, M. (2022). *Lips Don't Lie: A Generalisable and Robust Approach to Face Forgery Detection*. CVPR.
- Ho, J., Jain, A., & Abbeel, P. (2020). *Denoising Diffusion Probabilistic Models*. NeurIPS 2020.
- ISO/IEC. International Organization for Standardization. (2020). *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding* (ISO/IEC Standard No. 23008-2:2020). <https://www.iso.org/standard/86516.html>, [dostęp: 06.2025].
- ITU-T H.264, International Telecommunication Union. (2019). *Advanced video coding for generic audiovisual services* (ITU-T Recommendation H.264). <https://www.itu.int/rec/T-REC-H.264-201906-I/en>, [dostęp: 06.2025].
- ITU-T. International Telecommunication Union. (2021). *Subjective video quality assessment methods for multimedia applications* (ITU-T Recommendation P.910). <https://www.itu.int/rec/T-REC-P.910-202107-I/en>, [dostęp: 06.2025].
- ITU-R. International Telecommunication Union. (2023). *Algorithms to measure audio programme loudness and true-peak audio level* (ITU-R Recommendation BS.1770-5). <https://www.itu.int/rec/R-REC-BS.1770-5-202311-I/en>, [dostęp: 06.2025].
- Jiang, L., Li, R., Wu, W., Qian, C., & Loy C. C. (2020). *DeeperForensics-1.0: A Large-Scale Dataset for Real-World Face Forgery Detection*. arXiv:2001.03024.
- Kaplan, J., McCandlish, S., Henighan, T., et al. (2020). *Scaling Laws for Neural Language Models*. arXiv:2001.08361.
- Kartynnik, Y., Ablavatski, A., Grishchenko, I., & Grundmann, M. (2020). *Real-Time Facial Surface Geometry from Monocular Video on Mobile GPUs*. arXiv:2006.11341.
- King, D. E. (2009). Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research*, 10, 1755–1758.
- Korshunov, P., & Marcel, S. (2019). *Vulnerability of Face Recognition to Deepfake Videos*. ICB 2019.
- Korshunov, P., & Marcel, S. (2023). *Deepfakes: A Survey on Facial Manipulation and Fake Detection*. arXiv:2201.09585v3.
- Lazer, D. M. J., Baum, M., Benkler, Y., et al. (2018). The Science of Fake News. *Science*, 359(6380), 1094–1096.
- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2020). *Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics*. arXiv:1909.12962.

- Li, Y., & Lyu, S. (2018). *Exposing DeepFake Videos by Detecting Face Warping Artifacts*. arXiv:1811.00656.
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *IJCV*, 60(2), 91–110.
- Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE TPAMI*, 11(7), 674–693.
- Manjunath, B. S., & Ma, W.-Y. (1996). Texture Features for Browsing and Retrieval of Image Data. *IEEE TPAMI*, 18(8), 837–842.
- Mittal, A., Soundararajan, R., & Bovik, A. C. (2013). Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters*, 20(3), 209–212.
- NATO StratCom COE. (2022). *Malign Creativity: How Gender, Sex, and Lies are Weaponized Against Women Online*. Report.
- Nautsch, A., Kolossa, D., Bonastre, J.-F., et al. (2022). Speech Deepfake Detection: A Comprehensive Survey. *IEEE Signal Processing Magazine*, 39(6), 40–62.
- Odena, A., Dumoulin, V., Olah, C. (2016). *Deconvolution and Checkerboard Artifacts*. Distill.
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A Comparative Study of Texture Measures with Classification Based on Feature Distributions. *Pattern Recognition*, 29(1), 51–59.
- Ojansivu, V., Heikkilä, J. (2008). *Blur Insensitive Texture Classification Using Local Phase Quantization*. ICISP 2008, LNCS 5099, 236–243.
- Prajwal, K. R., Mukhopadhyay, R., Namboodiri, V. P., & Jawahar, C. V. (2020). *Wav2Lip: Accurately Lip-syncing Videos to Any Speech*. ACM MM.
- Rassool, R. (2017). VMAF reproducibility: Validating a perceptual practical video quality metric. In *2017 IEEE international symposium on broadband multimedia systems and broadcasting (BMSB)* (pp. 1-2). IEEE.
- Ray, S. F. (2002). *Applied Photographic Optics*. Focal Press.
- Ricker, J., Damm, F., et al. (2023). *Towards the Detection of Diffusion-Model Deepfakes*. arXiv:2210.14571.
- Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). *FaceForensics++: Learning to Detect Manipulated Facial Images*. ICCV.
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: *An Efficient Alternative to SIFT or SURF*. ICCV.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). *Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization*. ICCV. Sumsb. (2024). *Identity Fraud Report 2024*.
- Tan, M., & Le, Q. (2019). *EfficientNet: Rethinking Model Scaling for CNNs*. ICML.
- Teed, Z., & Deng, J. (2020). *RAFT: Recurrent All-Pairs Field Transforms for Optical Flow*. ECCV.

- Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection. *Information Fusion*, 64, 131–148.
- Wang, T., Liao, X., Chow, K. P., Lin, X., & Wang, Y. (2024). Deepfake detection: A comprehensive survey from the reliability perspective. *ACM Computing Surveys*, 57(3), 1-35.
- Wang, Z., Bao, J., Zhou, W., Wang, W., Hu, H., Chen, H., & Li, H. (2023). Dire for diffusion-generated image detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 22445-22455).
- Verdoliva, L. (2020). Media Forensics and DeepFakes: An Overview. *IEEE Signal Processing Magazine*, 37(1), 86–99.
- Vosoughi, S., Roy, D., Aral, S. (2018). The Spread of True and False News Online. *Science*, 359(6380), 1146–1151.
- Zhang, Y., Song, W., Wu, F., Han, H., & Zhang, L. (2020). Revealing the traces of nonaligned double JPEG compression in digital images. *Optik*, 204, 164196.

Wykorzystanie sztucznej inteligencji (AI) w zarządzaniu finansami przedsiębiorstwa przy wsparciu systemu SAP – analiza wybranych zastosowań i przypadków

Streszczenie

W ostatnich latach obserwuje się dynamiczny wzrost wykorzystania sztucznej inteligencji w działach finansowych polskich przedsiębiorstw. Zgodnie z przedstawionymi badaniami 71% firm w Polsce już wdraża rozwiązania AI w obszarze finansów, obejmujące planowanie finansowe, prognozowanie, zarządzanie ryzykiem oraz automatyzację procesów księgowych. Artykuł prezentuje kompleksowy przegląd najważniejszych przypadków użycia AI w kontekście systemów SAP w Polsce w planowaniu zasobów przedsiębiorstwa (ERP).

Badania wskazują, że 57% liderów biznesowych potwierdza uzyskanie lepszych efektów z tytułu wykorzystania AI niż zakładano, głównie przez zwiększenie efektywności operacyjnej i precyzyjność prognoz. Szczególnie istotne zastosowania identyfikuje się w planowaniu finansowym, gdzie SAP Analytics Cloud wykorzystuje uczenie maszynowe do wykrywania trendów i anomalii. W procesach księgowych AI usprawnia zamknięcia okresów poprzez automatyzację uzgadniania kont i inteligentną klasyfikację zapisów. W zarządzaniu należnościami narzędzia takie jak SAP Collections Management priorytetyzują działania windykacyjne, podczas gdy w obszarze zobowiązań AI automatyzuje przetwarzanie faktur przez inteligentne rozpoznawanie dokumentów.

Mimo oczywistych korzyści wdrożenie AI napotyka wyzwania: konieczność zapewnienia wysokiej jakości danych, budowanie zaufania do modeli oraz przeszkolenie personelu. 53% firm wskazuje brak umiejętności technologicznych jako główną barierę. Artykuł konkluduje, że najlepsze efekty przynosi synergia człowieka i AI, gdzie technologia przejmuje zadania obliczeniowe, umożliwiając specjalistom koncentrację na analizie strategicznej i podejmowaniu decyzji biznesowych.

Słowa kluczowe: technologia, sztuczna inteligencja, zarządzanie finansami

1 Partner PWC Polska

The Use of Artificial Intelligence (AI) in Enterprise Financial Management Supported by the SAP System: Analysis of Selected Applications and Case Studies

Abstract

In recent years there has been a dynamic growth in the use of artificial intelligence in the finance departments of Polish enterprises. According to the presented research, 71% of Polish companies are already implementing AI solutions in the finance area, encompassing financial planning, forecasting, risk management, and automation of accounting processes. The article presents a comprehensive review of the most important AI use cases in the context of SAP systems in Poland within enterprise resource planning (ERP).

Research indicates that 57% of business leaders confirm that AI effects exceed expectations, mainly through increased operational efficiency and forecast precision. Particularly significant applications are identified in financial planning, where SAP Analytics Cloud uses machine learning to detect trends and anomalies. In accounting processes, AI streamlines period closures through automation of account reconciliation and intelligent classification of entries. In receivables management, tools such as SAP Collections Management prioritize collection activities, while in the payables area, AI automates invoice processing through intelligent document recognition.

Despite the benefits, AI implementation faces challenges: the need to ensure high-quality data, building trust in models, and staff training. 53% of companies indicate lack of technological skills as the main barrier. The article concludes that the greatest effects come from human-AI synergy, where technology takes over computational tasks, enabling specialists to focus on strategic analysis and business decision-making.

Keywords: technology, artificial intelligence, finances management

Wprowadzenie

Współczesna transformacja cyfrowa przedsiębiorstw nieodłącznie wiąże się z rewolucyjnymi zmianami w zarządzaniu finansami, gdzie sztuczna inteligencja (AI) staje się kluczowym katalizatorem innowacji i efektywności operacyjnej². W kontekście rozwiązań SAP³ integracja technologii AI w systemach finansowych reprezentuje nie tylko ewolucję tradycyjnych procesów księgowych i analitycznych, lecz fundamentalną redefinicję sposobu podejmowania decyzji strategicznych w organizacjach⁴. Dla przykładu, platforma SAP S/4HANA Cloud z wbudowanymi funkcjonalnościami machine learning oraz SAP Analytics Cloud z zaawansowanymi algorytmami predykcyjnymi umożliwiają automatyzację złożonych procesów finansowych, od rozpoznawania wzorców w transakcjach po prognozowanie przepływów pieniężnych z wykorzystaniem analizy predykcyjnej⁵. Jednak ekspansja AI w finansach generuje równocześnie nowe wyzwania w obszarze cyberbezpieczeństwa, gdzie ochrona wrażliwych danych finansowych przed zaawansowanymi zagrożeniami cyber staje się priorytetem strategicznym⁶. Inteligentne systemy SAP, wyposażone w algorytmy uczenia maszynowego, nie tylko optymalizują procesy zarządzania ryzykiem finansowym i zgodnością z przepisami, ale również wymagają implementacji zaawansowanych mechanizmów bezpieczeństwa, które chronią przed manipulacją algorytmów⁷. Ta dwoistość technologiczna – gdzie AI jednocześnie stanowi narzędzie transformacji i potencjalne źródło słabości systemu – definiuje współczesny krajobraz finansowo-technologiczny, w którym organizacje muszą balansować między innowacyjnością a bezpieczeństwem, wykorzystując pełny potencjał ekosystemu SAP w erze sztucznej inteligencji.

W ostatnich latach obserwujemy dynamiczny wzrost wykorzystania sztucznej inteligencji w działach finansowych, również w polskich firmach. Według badań aż 71% przedsiębiorstw już wdraża AI w obszarze finansów, przy czym

2 PwC's 28nd Annual Global CEO Survey, *Reinvention on the edge of tomorrow* PwC Global, 2025. <https://www.pwc.com/gx/en/ceo-survey/2025/28th-ceo-survey.pdf>

3 SAP – producent oprogramowania wspierającego przedsiębiorstwa; „Systeme, Anwendungen und Produkte in der Datenverarbeitung”, co po angielsku brzmi „Systems, Applications & Products in Data Processing”, a po polsku: „Systemy, Aplikacje i Produkty w Przetwarzaniu Danych”.

4 SAP SE, *SAP S/4HANA: Digital Core for the Intelligent Enterprise*, SAP Documentation, 2023. https://help.sap.com/docs/SAP_S4HANA_ON-PREMISE

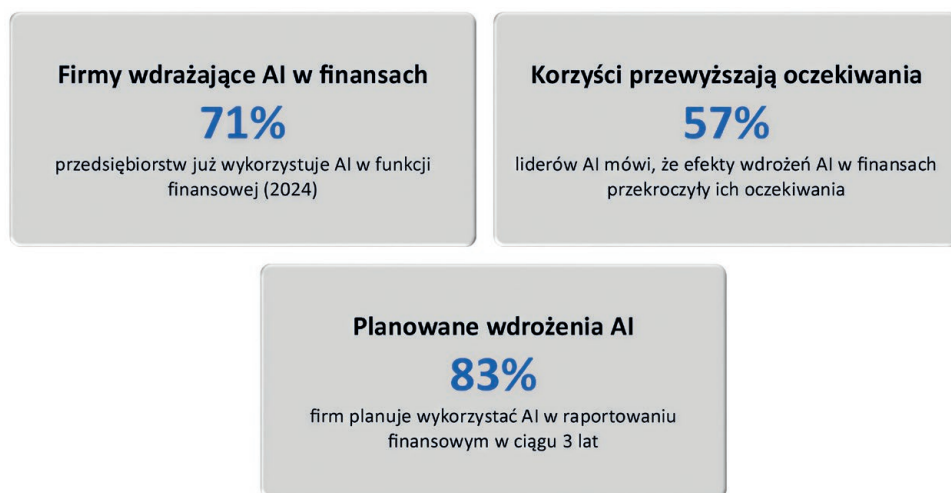
5 SAP SE, *SAP Analytics Cloud: Planning and Analytics*, SAP Help Portal, 2024. https://help.sap.com/docs/SAP_ANALYTICS_CLOUD

6 European Union Agency for Cybersecurity (ENISA), *Securing Machine Learning Algorithms*, ENISA Report, December 2021, <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>

7 National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, January 2023, <https://www.nist.gov/itl/ai-risk-management-framework>

zastosowania wykraczają poza samą sprawozdawczość – obejmują planowanie finansowe, prognozowanie, zarządzanie ryzykiem, kontrolę zgodności oraz automatyzację procesów księgowych⁸. Firmy dostrzegają wymierne korzyści: 57% liderów twierdzi, że uzyskane dzięki AI efekty przewyższają oczekiwania, m.in. dzięki większej efektywności operacyjnej, lepszej analizie danych i bardziej precyzyjnym prognozom⁹. Nic więc dziwnego, że rozwiązania AI stają się integralną częścią systemów ERP, takich jak SAP S/4HANA i powiązanych narzędzi (np. SAP Analytics Cloud), usprawniając kluczowe procesy finansowe. Poniżej przedstawiono przegląd najważniejszych przypadków użycia AI w finansach w kontekście SAP, wraz z przykładami wdrożeń w Polsce, uzyskiwanymi korzyściami oraz wyzwaniem, jakie im towarzyszą.

Rysunek 1. Kluczowe wskaźniki charakteryzujące wykorzystanie AI w Polsce (w kontekście rozwiązań SAP)



Źródło: opracowanie własne ze wsparciem Copilot researcher.

AI w planowaniu i prognozowaniu finansowym

Planowanie finansowe i prognozowanie to obszary, w których AI wnosi ogromną wartość dodaną. Tradycyjne budżetowanie opiera się na eksperckich przewidywaniach menedżerów i kontrolerów finansowych, wspieranych modelami opartymi

⁸ Sztuczna inteligencja w finansach – KPMG Poland, <https://kpmg.com/pl/pl/home/insights/2025/02/sztuczna-inteligencja-w-finansach.html>

⁹ Tamże.

na statystyce i arkuszach kalkulacyjnych. AI oferuje nowe podejście, umożliwiając analizę dużych zbiorów danych historycznych i wykrywanie wzorców, których wykrycie związane jest z dużą pracochłonnością i kosztami. Przykładowo, platforma SAP Analytics Cloud (SAC) wykorzystuje algorytmy sztucznej inteligencji do automatycznego identyfikowania trendów i anomalii w danych finansowych oraz kluczowych czynników wpływających na wyniki. Pozwala to organizacjom szybciej reagować na zmiany i budować bardziej trafne prognozy¹⁰. Co więcej, SAC oferuje wbudowane funkcje modelowania predykcyjnego, dzięki którym można generować prognozy sprzedaży, kosztów czy przepływów pieniężnych z większą precyzją, minimalizując ryzyko błędnych decyzji strategicznych¹¹.

W praktyce polskie firmy coraz częściej sięgają po takie rozwiązania. Zaawansowana analityka z elementami AI wspiera działy kontrolingu w firmach produkcyjnych, handlowych i usługowych. Zespoły odpowiedzialne za księgowość zarządczą i planowanie finansowe należą do tej grupy, która jako pierwsza wdraża AI w przedsiębiorstwie – i dlatego niemal 2/3 firm testuje lub już wykorzystuje AI w tych obszarach¹². Przykładowo, AI może automatycznie generować propozycje budżetów i prognoz w oparciu o tysiące zmiennych (trendów rynkowych, sezonowości, wskaźników makroekonomicznych itp.), odciążając analityków od żmudnego zbierania danych. Już dziś istnieją prototypy, w których modele oparte na sieciach neuronowych są w stanie uczyć się na podstawie danych finansowych firmy, by zaproponować własną prognozę – choć ciągle tylko jako propozycję dla człowieka, który jako ekspert może wprowadzić w niej niezbędne korekty i zatwierdzić. W opinii ekspertów, w perspektywie kolejnych 10 lat AI może tworzyć lepsze budżety i prognozy niż sami menedżerowie, choć warunkiem, jak zawsze, pozostaje zapewnienie wysokiej jakości danych wejściowych¹³. Oznacza to, że rola człowieka będzie ewoluować – analitycy finansowi więcej czasu poświęcą na interpretację wyników i scenariuszy „co – jeśli”, podczas gdy AI zajmie się obliczeniami.

Zastosowanie AI w planowaniu finansowym przekłada się na większą precyzję prognoz i szybkość ich przygotowania. Dzięki automatycznej analizie danych historycznych i bieżących prognozy sprzedaży czy kosztów mogą być aktualizowane w czasie rzeczywistym, co jest niezwykle cenne w zmiennym otoczeniu rynkowym. AI potrafi wskazać ukryte zależności – np. wpływ określonych wskaźników makroekonomicznych na wyniki firmy – co pomaga poprawić trafność planów.

10 Lepsze planowanie i raportowanie w środowisku SAP – Aktualności – KBJ. <https://www.kbj.com.pl/aktualnosci/lepsze-planowanie-i-raportowanie-w-srodowisku-sap.html>

11 Tamże.

12 Sztuczna inteligencja w finansach..., op. cit.

13 Sztuczna inteligencja w controllingu – FlexiEPM. <https://flexiepm.pl/sztuczna-inteligencja-w-controllingu/>

Dodatkowo, dzięki funkcjom takim jak symulacje „what – if”, menedżerowie mogą testować różne scenariusze (np. wpływ wzrostu kosztów energii o 10% na marżowość wybranego produktu czy usługi z portfolio danej firmy) i lepiej przygotować plany awaryjne. Badania potwierdzają, że firmy stosujące AI raportują bardziej trafne prognozy i lepszą analizę danych, co wzmacnia strategiczne decyzje¹⁴.

Równocześnie głównym wyzwaniem pozostaje zapewnienie wysokiej jakości danych i modeli. Jak podkreślają eksperci, bez solidnych, wiarygodnych danych sztuczna inteligencja pozostanie tylko ciekawostką, która nie rozwiąże problemów¹⁵. Należy więc zainwestować w integrację danych z różnych systemów informatycznych (ERP, hurtowni danych, arkuszy Excel) oraz ich ujednolicenie. Kolejną kwestią jest zaufanie do modelu AI – analitycy muszą zrozumieć, na jakiej podstawie AI formułuje prognozę, by móc ją zaakceptować lub skorygować. Istotna jest tu przejrzystość algorytmów i możliwość wyjaśnienia wyników (tzw. explainable AI). Nie bez znaczenia są też kompetencje pracowników – 53% firm widzi brak odpowiednich umiejętności technologicznych jako barierę w pełnym wykorzystaniu AI¹⁶. Konieczne bywa zatrudnienie lub przeszkolenie specjalistów od data science w finansach.

AI może sugerować działania, jednak ostateczne decyzje finansowe wciąż należą do ludzi. Modele nie przewidzą zdarzeń nieznajdujących odzwierciedlenia w danych historycznych (np. nagłych kryzysów). Dlatego najlepsze wyniki przynosi współpraca AI i człowieka – AI dostarcza liczby i rekomendacje, a doświadczeni finansowi decydenci weryfikują je i wprowadzają czynnik biznesowy.

AI w księgowości

Proces zamknięcia okresu finansowego (monthly/quarterly close) bywa dla działów księgowości prawdziwym wyścigiem z czasem. Trzeba uzgodnić konta, zarejestrować wszelkie transakcje, dokonać korekt i przygotować raporty – często w ciągu kilku dni od końca okresu. Sztuczna inteligencja i automatyzacja w nowych rozwiązaniach ERP SAP S/4HANA znacząco usprawniają te czynności, skracając czas zamknięcia i redukując ryzyko błędów.

SAP S/4HANA oferuje szereg inteligentnych funkcji wspierających obszar Record-to-Report (od rejestracji zdarzenia w finansach do raportowania), jak:

- Automatyzacja uzgadniania kont rozliczeniowych – system wykorzystuje uczenie maszynowe do monitorowania kont rozliczenia zakupu. Analizuje

¹⁴ Sztuczna inteligencja w finansach..., op. cit.

¹⁵ Sztuczna inteligencja w controllingu..., op. cit.

¹⁶ Sztuczna inteligencja w finansach..., op. cit.

w czasie rzeczywistym status przyjęć towaru i faktur zakupowych, identyfikując niezgodności i sugerując propozycje dalszego postępowania¹⁷. Dzięki temu łatwiej wyjaśnić i wyksięgować nierozliczone pozycje jeszcze przed zamknięciem. Tradycyjnie księgowi spędzali długie godziny na ręcznej analizie takich różnic – teraz wiele z tych zadań jest zautomatyzowanych. W efekcie proces uzgadniania przebiega szybciej i sprawniej, odciążając zespół w krytycznym momencie zamknięcia miesiąca¹⁸.

- Inteligentna klasyfikacja zapisów – AI potrafi również wspomóc poprawną klasyfikację księgową dokumentów. Błędne przypisanie kont (np. koszt zaklasyfikowany do niewłaściwej kategorii) potrafi zniekształcić raporty finansowe. Rozwiązania SAP wykorzystują boty RPA (Robotic Process Automation – automatyzacja procesów z wykorzystaniem robotów) i algorytmy, które przeszukują zapisy na kontach i proponują przeksięgowania do właściwych kategorii (na podstawie zdefiniowanych reguł lub wzorców w danych)¹⁹. Dzięki temu sprawozdania odzwierciedlają rzeczywisty obraz, a księgowi mogą szybciej skorygować ujęcia, bez ręcznego wertowania setek pozycji.
- Automatyczne wprowadzanie zapisów zewnętrznych – przykładowo funkcja zrobotyzowanej automatyzacji (SAP BPA) może obsługiwać import wpisów do dziennika przesyłanych e-mailem. Bot potrafi pobrać z załączników (np. z arkusza Excel) odpowiednie dane księgowe i załadować je do SAP, tak jakby robił to człowiek. Eliminowane jest tym samym żmudne ręczne przepisywanie danych z dokumentów otrzymanych od innych działów lub oddziałów firmy.

Z powyższych punktów wynika, że dzięki wykorzystaniu sztucznej inteligencji możliwe jest skrócenie cyklu zamknięcia okresu – czynności, które kiedyś trwały kilka dni, teraz mogą być wykonane w ułamku tego czasu. Automatyczne uzgadnianie kont i generowanie księgowania korekcyjnych oznacza, że raporty finansowe są dostępne wcześniej, co pozwala zarządowi szybciej reagować na wyniki. AI zwiększa też dokładność – eliminacja ręcznego przenoszenia danych redukuje pomyłki. Jak wskazują eksperci, AI zmniejsza liczbę błędów i odciąża pracowników od powtarzalnych zadań, co przekłada się na niższe koszty operacyjne²⁰. Ponadto inteligentne systemy potrafią automatycznie wychwycić nietypowe transakcje (np. księgowanie znacznie odbiegające od normy) i ostrzec zespół finansowy, zanim błąd przepływnie do finalnych

17 Reviewing the Use Cases of SAP Business AI in Finance, https://learning.sap.com/learning-journeys/discovering-new-ai-capabilities-for-sap-finance/reviewing-the-use-cases-of-sap-business-ai-in-finance_d4221311-941d-4196-bd17-2c712f6d65a2

18 Tamże.

19 Tamże.

20 Sztuczna inteligencja w finansach..., op. cit.

raportów. Wreszcie, usprawnione procesy księgowe dają działom finansowym więcej czasu na analizę i interpretację wyników finansowych, zamiast zajmowania się tylko ich przygotowaniem.

Mając na uwadze powyższe korzyści, należy pamiętać, że wdrożenie AI w księgowości wymaga dostosowania procesów i zmiany kultury pracy. Pracownicy muszą zaufać rekomendacjom systemu – np. sugestiom automatycznego przeksięgowania – co na początku może budzić ostrożność. Ważne jest zatem ustawienie odpowiednich punktów kontrolnych (np. zatwierdzanie proponowanych zapisów przez głównego księgowego), aby zachować kontrolę i zgodność z przepisami. Kolejny aspekt to integracja AI z istniejącymi procedurami audytowymi – raporty tworzone z udziałem AI muszą spełniać wymogi regulacyjne, a firma powinna być w stanie wykazać, że dane w nich są rzetelne. Technologie AI muszą być też skalibrowane do specyfiki firmy – np. modele uczące się uzgadniania kont powinny być trenowane na danych danej organizacji, uwzględniając jej polityki księgowe. Nie mniejsze wyzwanie stanowi cyberbezpieczeństwo i prawa dostępu: automaty, które mają dostęp do danych finansowych, muszą być odpowiednio zabezpieczone przed nieautoryzowanym użyciem. Niemniej jednak wraz ze zdobywaniem doświadczeń przez zespoły finansowe bariery te maleją – już 49% firm twierdzi, że wdrożenie AI znacząco usprawniło funkcjonowanie ich finansów lub wkrótce to zrobi²¹.

AI w zarządzaniu należnościami (windykacja i zarządzanie środkami pieniężnymi)

Utrzymanie płynności finansowej firmy w dużej mierze zależy od sprawnego ściągania należności od klientów. Tu również AI wbudowana w nowe wersje rozwiązań SAP (SAP S/4HANA) okazuje się dużo skuteczniejsza od tradycyjnych podejść. SAP Finance posiada inteligentne narzędzia do zarządzania należnościami, które pomagają działom windykacji i zarządzania środkami pieniężnymi optymalizować wpływy.

Jednym z przykładów jest SAP Collections Management z AI. Rozwiązanie to automatycznie analizuje portfel należności i priorytetyzuje działania windykacyjne. Na podstawie zdefiniowanych kryteriów i analizy historycznych danych płatniczych system ocenia klientów pod kątem ryzyka opóźnień i prognozuje prawdopodobny termin spłaty zaległych faktur²². W efekcie generowana jest inteligentna lista zadań dla windykatorów: wskazuje, którymi należnościami powinni

21 Firmy i pracownicy w Polsce coraz bardziej przychylni AI. <https://altab.pl/inne/firmy-i-pracownicy-w-polsce-coraz-bardziej-przychylni-ai/>

22 Reviewing the Use Cases..., op. cit.

zając się najpierw podczas bieżących kontaktów z klientami²³. Na przykład klient B może mieć więcej przeterminowanych faktur niż klient A, ale jeśli AI przewiduje, że klient A ma mniejsze skłonności do zapłaty na czas, może on zostać ustawiony wyżej na liście priorytetów. Takie dynamiczne priorytetyzowanie sprawia, że specjaliści ds. należności skupiają wysiłki tam, gdzie jest to najbardziej krytyczne dla cash-flow.

Innym innowacyjnym rozwiązaniem jest SAP Cash Application – zestaw usług opartych na uczeniu maszynowym, który automatyzuje proces dopasowywania płatności do faktur. Tradycyjnie dział księgowości musiał ręcznie łączyć wpływy na wyciągu bankowym z otwartymi fakturami sprzedaży, często posługując się danymi referencyjnymi od klientów. SAP Cash Application uczy się wzorców zachowań płatniczych (np. jakie opisy transakcji w tytule przelewu odpowiadają którym fakturom) i automatycznie dokonuje rozliczeń lub sugeruje najlepsze dopasowania.

Uruchomienie funkcji sztucznej inteligencji w procesach należności przekłada się bezpośrednio na poprawę płynności finansowej firmy. Dzięki sprawniejszemu ściąganiu długów obniża się poziom należności przeterminowanych, a wskaźnik średniego okresu oczekiwania na zapłatę ulega skróceniu. Priorytetyzacja zadań windykacyjnych sprawia, że pracownicy nie tracą czasu na mniej istotne sprawy – AI kieruje ich uwagę tam, gdzie ryzyko braku płatności jest największe lub kwota należności najwyższa. Efektem jest wyższa efektywność windykacji i potencjalnie mniejsza potrzeba tworzenia rezerw na złe długi. Co więcej, poprawia się relacja z klientami – dzięki inteligentnym narzędziom można zawczasu identyfikować klientów, którzy zawsze płacą z opóźnieniem i np. oferować im plan ratalny lub wcześniejsze przypomnienia, zamiast uruchamiać twarde procedury windykacji, które mogłyby naruszyć relacje. Automatyczne dopasowanie płatności do faktur odciąża dział księgowości: skraca się czas potrzebny na zamknięcie danego dnia czy miesiąca w aspekcie rozliczeń bankowych, zmniejsza się liczba nierozpoznanych płatności. To także mniej pomyłek – system uczący się na podstawie danych potrafi uwzględnić niuanse (np. klient notorycznie wpisuje inny numer faktury w tytule przelewu niż faktyczny), czego stały personel mógłby nie wychwycić.

Głównym wyzwaniem związanym z uruchomieniem procesów ze wsparciem sztucznej inteligencji jest przyjęcie ich przez klientów – automatyzacja windykacji nie może zastąpić ludzkiego podejścia tam, gdzie jest potrzebna elastyczność. Jeśli AI zaczęłaby generować zbyt częste lub niewłaściwe przypomnienia do klientów, mogłoby to być odebrane negatywnie. Dlatego parametry modelu (kryteria priorytetyzacji) muszą być dobrze zestrojone z polityką kredytową

23 Tamże.

firmy. Na początku kluczowe jest monitorowanie działania AI przez doświadczonych menedżerów kredytowych, którzy sprawdzą, czy proponowane listy zadań mają sens i ewentualnie skorygują reguły. Inną kwestią jest lokalna specyfika prawna – np. w Polsce działalność windykacyjna podlega przepisom o ochronie danych osobowych i dobrym praktykom wspieranym przez KNF; korzystając z AI trzeba upewnić się, że nie dochodzi do naruszenia tych zasad (np. profilowania klientów w sposób dyskryminujący). Technicznie trening modeli AI do cash management wymaga danych – firma może potrzebować kilku miesięcy historii transakcji, by model dopasowywania płatności osiągnął wysoką skuteczność. W okresie przejściowym konieczna jest cierpliwość – pracownicy muszą korygować błędne dopasowania, co z kolei posłuży modelowi do uczenia się. Z czasem jednak, jak pokazują doświadczenia globalnych firm, skuteczność takich rozwiązań znacząco rośnie, osiągając nawet ponad 90% automatycznych dopasowań w sprzyjających warunkach.

AI w obsłudze zobowiązań

Drugą stroną procesów finansowych są zobowiązania, czyli rozliczenia z dostawcami. Tutaj AI również znalazła praktyczne zastosowania – zwłaszcza w automatyzacji przetwarzania faktur kosztowych i usprawnieniu płatności. Wielu dostawców wciąż wysyła faktury w formie PDF lub papierowej, co wymaga ręcznego wprowadzania danych do systemu. Inteligentne skanowanie i rozpoznawanie dokumentów (OCR) połączone z AI potrafi zrobić to automatycznie. Rozwiązania zintegrowane z SAP (takie jak SAP Invoice Management czy integracje z usługami AI) potrafią odczytać kluczowe pola faktury (dostawca, kwota, data, pozycja), nawet jeśli formaty dokumentów różnią się. Uczące się algorytmy rozpoznają układ faktur stałych kontrahentów i z czasem poprawiają trafność odczytu.

Po wczytaniu takiej faktury, rozwiązanie SAP może automatycznie skierować ją do odpowiedniego procesu akceptacji (np. rozpoznać dział na podstawie rodzaju zakupu i przesłać do akceptacji kierownika danego działu), bądź dopasować do zamówienia zakupu, jeśli takie istnieje. Bardziej zaawansowane przypadki użycia AI pojawiają się, gdy faktura nie ma numeru zamówienia – system może wtedy, na podstawie opisu pozycji i historii zakupów, zasugerować powiązanie z istniejącym zamówieniem lub kategorią kosztów, co przyspiesza dekretnację dokumentu. Takie „inteligentne przetwarzanie faktur bez zamówienia” ogranicza liczbę sytuacji, gdzie księgowy musi ręcznie wyjaśniać, czego dotyczy faktura²⁴.

W opisywanym obszarze najbardziej oczywistą korzyścią z wykorzystania sztucznej inteligencji jest oszczędność czasu i kosztów – firmy przetwarzające

24 Reviewing the Use Cases..., op. cit.

duże wolumeny faktur (np. centra usług wspólnych obsługujące finanse) dzięki AI mogą zautomatyzować znaczną część pracy dawniej wykonywanej przez referentów. Przyspiesza to płatności na rzecz dostawców, co z kolei może oznaczać lepsze wykorzystanie rabatów za płatność przed terminem lub po prostu budowanie dobrej reputacji płatnika. Dokładność odczytu danych z faktur przez AI dziś sięga bardzo wysokich poziomów, redukując pomyłki (np. zaksięgowanie złej kwoty lub VAT). AI może też oznaczać podejrzane faktury – np. duplikaty (ta sama faktura wysłana dwa razy) albo potencjalnie niezgodne z umową (kwota wyższa niż zamówienie).

Wdrożenie AI do AP wymaga integracji z obiegiem dokumentów i – podobnie jak wcześniej – pewnego kredytu zaufania do systemu. Błędy OCR wciąż mogą się zdarzać przy słabej jakości skanów lub nietypowych układach dokumentów, więc proces walidacji (np. przez księgowego sprawdzającego podświetlone pola) jest konieczny przynajmniej na początku. Modele AI muszą być trenowane na lokalnych fakturach – co w przypadku polskich realiów oznacza nauczanie systemu specyfiki języka polskiego, naszych formatów dat, walut, stawek VAT itp. Czasami wymaga to dostosowania globalnych narzędzi do krajowych uwarunkowań. Istotne jest też zapewnienie zgodności z regulacjami podatkowymi – automatycznie odczytane dane muszą poprawnie odzwierciedlać wymogi (np. podział na stawki VAT, oznaczenia JPK). Ostatnią kwestią jest skalowalność i koszt – firmom przetwarzającym niewielką liczbę faktur inwestycja w zaawansowane AI może się nie zwrócić, dlatego częściej takie rozwiązania wdrażają duże organizacje lub centra usług, gdzie wolumen dokumentów jest wysoki.

AI w kontrolingu i zarządzaniu ryzykiem

Kontroling tradycyjnie zajmuje się planowaniem, monitorowaniem oraz analizą odchyleń i korygowaniem działań w celu osiągnięcia celów strategicznych (Nowak, 2011). Dodatkowo dostarcza wskaźniki finansowe dla biznesu. W tym obszarze AI pełni coraz ważniejszą rolę jako system wczesnego ostrzegania przed anomaliami oraz wsparcie w analizach ad-hoc.

Wykrywanie anomalii to świetny przykład: Standardowo firmy definiowały sztywne reguły, co uznać za odchylenie wymagające uwagi (np. „jeśli koszt w danym dziale przekroczy budżet o $> 15\%$ ”). AI pozwala uczyć modele na historycznych danych, aby samodzielnie określały, co jest „nietypowym” zachowaniem finansowym. Modele uczenia maszynowego (ML) analizują tysiące transakcji i wskaźników, tworząc profil normalnej aktywności. Gdy pojawi się coś odbiegającego od normy (np. nagły, niesezonowy wzrost wydatków marketingowych w jednym z regionów albo nienaturalnie duży spadek marży w konkretnym produkcie), system powiadamia kontrolerów o potencjalnym problemie. Takie

podejście zwalnia z konieczności ręcznego ustawiania progów i reguł – AI sama pilnuje danych i wyłapuje zagrożenia²⁵. Oczywiście ostateczna ocena, czy to faktycznie problem, należy do człowieka, ale zyskuje on narzędzie, które czuwa non stop i nie przeoczy subtelnych sygnałów.

Analiza ryzyka i zgodności to kolejny obszar. W dużych grupach kapitałowych śledzenie ciągłych zmian prawnych i ich wpływu na finanse bywa wyzwaniem. Strategiczne polskie przedsiębiorstwo (Orlen) ogłosiło w 2025 roku, że w ramach największego wdrożenia AI w regionie będzie m.in. wykorzystywać sztuczną inteligencję do analizy obowiązujących i planowanych zmian regulacyjnych oraz do zadań kontrolingowych²⁶. Sztuczna inteligencja będzie śledziła projekty ustaw podatkowych i symulowała, jakie skutki finansowe dla firmy przyniesie np. zmiana stawki akcyzy czy wprowadzenie nowego podatku od określonych aktywów. Orlen wskazuje, że narzędzia AI pomogą też w analizie globalnych danych rynkowych (np. cen ropy i gazu), co ma istotne znaczenie dla kontrolingu w branży energetycznej²⁷. Dzięki temu decyzje strategiczne mogą być podejmowane szybciej i na podstawie pełniejszej informacji.

Sztuczna inteligencja w kontrolingu przekłada się na większą przejrzystość i proaktywność finansów. Zamiast reagować po fakcie na negatywne odchylenia, firmy mogą wcześniej wykrywać nieprawidłowości i podejmować działania korygujące. To może zapobiec stratom, lub przynajmniej zminimalizować ich skutki. Przykładowo, model wykrywania anomalii może ostrzec, że określony dział wykazuje nietypowo wysokie wydatki – co po weryfikacji może okazać się np. błędną klasyfikacją kosztów, którą można od razu naprawić, zanim zafałszuje to wyniki miesiąca. W obszarze zgodności AI może monitorować transakcje pod kątem zgodności z politykami (np. wydatki służbowe vs regulamin), czy nawet wykrywać potencjalne nadużycia finansowe (fraudy) poprzez identyfikację wzorców wskazujących na manipulacje. Z kolei przy analizie ryzyka AI pozwala finansistom przetwarzać znacznie więcej danych z otoczenia – kursy walut, indeksy surowców, wskaźniki ekonomiczne – i na bieżąco włączać je do modeli prognostycznych firmy. Taka holistyczna analiza była dotąd trudna, bo danych jest zbyt dużo, by analityk mógł je wszystkie śledzić ręcznie. AI jest więc „cyfrowym asystentem” kontrolerów, który podsuwa im istotne spostrzeżenia.

Po stronie wyzwań należy umieścić, po pierwsze, fałszywe alarmy vs przeoczone sygnały – wytrenowanie modelu anomalii tak, by był czuły, ale nie nadgorliwy, jest sztuką. Jeśli AI zbyt często generuje alerty, zespół może zacząć je

25 Sztuczna inteligencja w controllingu..., op. cit.

26 ORLEN – największe wdrożenie AI w Europie Środkowej - ISBtech.pl, <https://www.isbtech.pl/2025/07/orlen-najwieksze-wdrozenie-ai-w-europie-srodkowej/>

27 Tamże.

ignorować (syndrom „chłopca z gwizdkiem”). Dlatego modele wymagają kalibracji i ciągłego doskonalenia na podstawie informacji zwrotnej od kontrolerów. Po drugie, rola AI vs rola człowieka: jak zauważa ekspert controllingu Mariusz Sumiński, narzędzia takie jak ChatGPT raczej nie zastąpią kontrolerów w interpretacji wyników i formułowaniu rekomendacji biznesowych²⁸. AI może wygenerować opis trendów („tu wzrosło, tam spadło”), ale nadanie temu znaczenia biznesowego w kontekście strategii firmy wciąż wymaga ludzkiego doświadczenia²⁹. Innym wyzwaniem jest integracja AI z istniejącymi systemami raportowania i BI – trzeba zapewnić, że np. alert z AI trafi do osoby odpowiedzialnej (poprzez e-mail, dashboard lub inną ustaloną formę), a nie zaginie w natłoku informacji. Wreszcie, aspekt kulturowy i organizacyjny: budżety i analizy przygotowane z udziałem AI mogą budzić pytania zarządu: „skąd te liczby się wzięły?”. Konieczne bywa edukowanie decydentów, jak działa AI, by budować zaufanie do jej wyników. Firmy uczą się, że łączenie nowoczesnych technologii z dobrą kulturą zarządzania (odpowiedzialność za wyniki, etyka AI) daje najlepsze efekty³⁰ – zapewnia optymalizację procesów finansowych bez utraty zaufania do raportów.

Przegląd przypadków zastosowania AI w zarządzaniu finansami przedsiębiorstwa, przy wsparciu rozwiązań systemu SAP

Do przygotowania syntezy obszarów zastosowania AI w finansach, przedstawionych szczegółowo w akapitach powyżej, zastosowano Systematyczną Analizę Macierzową (SAM), która umożliwiła strukturalne uporządkowanie przypadków użycia, korzyści i wyzwań w przejrzystym formacie tabelarycznym.

28 Sztuczna inteligencja w controllingu..., op. cit.

29 Tamże.

30 Sztuczna inteligencja w finansach..., op. cit.

Tabela 1. Przypadki zastosowania, korzyści i wyzwania AI w zarządzaniu finansami przedsiębiorstwa, przy wsparciu rozwiązań systemu SAP

Przypadek użycia	Korzyści	Wyzwania (co wymaga uwagi)
Planowanie i prognozowanie finansowe	- Bardziej precyzyjne prognozy dzięki analizie AI (uwzględnienie trendów, sezonowości, zależności) - Szybsze budżetowanie – automatyczne generowanie planów skraca cykl planistyczny z tygodni do dni - Lepsze wsparcie decyzji – scenariusze „what – if” i rekomendacje AI pozwalają zarządcy podejmować trafniejsze decyzje strategiczne	- Jakość danych – model AI wymaga wiarygodnych, zintegrowanych danych, inaczej prognozy mogą być błędne - Zaufanie do AI – konieczne wyjaśnialne modele; decydenci muszą rozumieć podstawy prognoz - Kompetencje – potrzeba specjalistów data science w finansach; początkowo krzywa nauki przy wdrożeniu
Zamknięcia finansowe i księgowość (R2R)	- Skrócony czas zamknięcia – AI automatyzuje uzgadnianie kont (np. GR/IR) i księgowania korekt - Mniej błędów – wyeliminowanie ręcznego wprowadzania danych redukuje pomyłki - Większa efektywność – księgowi skupiają się na analizie, a nie na wprowadzaniu danych; łatwiejsze zapewnienie zgodności	- Integracja z kontrolą – potrzebne procedury zatwierdzania automatycznych księgowoń - Zmiana procesów – reorganizacja pracy działu księgowości, szkolenia personelu - Bezpieczeństwo – boty/AI operujące na danych finansowych muszą być odpowiednio zabezpieczone
Zarządzanie należnościami (AR)	- Poprawa płynności – szybsza windykacja i lepsze priorytety zadań - Automatyczne dopasowanie płatności – AI zwiększa odsetek automatycznie rozliczonych wpłat - Lepsza koncentracja – windykatorzy skupiają się na najważniejszych dłużnikach	- Ustawienie kryteriów AI – wymaga dostosowania modelu scoringowego klientów - Adaptacja modelu – AI potrzebuje czasu na naukę wzorców płatności - Regulacje i wizerunek – działania AI muszą być zgodne z dobrymi praktykami windykacji
Obsługa zobowiązań (AP)	- Automatyzacja pracy manualnej – odczyt faktur przez AI przyspiesza proces kilkukrotnie - Szybsze płatności – możliwość negocjowania rabatów za terminowe płatności - Kontrola zgodności – system wychwytuje duplikaty, błędy kwot, niezgodności	- Dokładność OCR/AI – może wymagać ręcznej weryfikacji przy nietypowych dokumentach - Wyjątki bez PO – trudność z klasyfikacją wydatków nietypowych - Zwrot z inwestycji – przy małej skali może być nieopłacalna

Przypadek użycia	Korzyści	Wyzwania (co wymaga uwagi)
Kontroling i wykrywanie anomalii	• Wczesne ostrzeganie – AI wykrywa nieprawidłowości przed ich eskalacją • Lepsza informacja zarządcza – bieżąca analiza wielu wskaźników • Analiza ryzyka – symulowanie wpływu zmian na wyniki finansowe	• Fałszywe alarmy – wymagana kalibracja systemu • Akceptacja rekomendacji – menedżerowie muszą zaufać wnioskowi AI • Integracja danych zewnętrznych – wymaga dodatkowych integracji spoza SAP

Źródło: opracowanie własne.

Podsumowanie

Sztuczna inteligencja staje się nieodłącznym elementem nowoczesnych finansów, a nowoczesne rozwiązania firmy SAP stanowią doskonałą platformę do wykorzystania potencjału AI. Polskie firmy – od globalnych koncernów (jak Orlen) po średnie przedsiębiorstwa – coraz śmielej inwestują w rozwiązania AI w obszarze finansów, dostrzegając ich wymierne korzyści. Automatyzacja żmudnych procesów, skrócenie czasu raportowania, poprawa dokładności i lepsze prognozy – to wszystko prowadzi do zwiększenia efektywności funkcji finansowej i roli finansów, w znaczeniu osób pracujących w tym dziale i wykorzystujących narzędzia AI, jako strategicznego partnera biznesu. Co istotne, korzyści często przerażają oczekiwania – ponad połowa firm będących liderami w AI ocenia, że uzyskała więcej niż zakładała³¹.

Nie oznacza to jednak, że droga jest wolna od wyzwań. Technologia musi iść w parze z kompetencjami i zaufaniem. Kluczem do sukcesu jest odpowiednie przygotowanie danych, szkolenie zespołów oraz ustanowienie właściwych ram nadzoru nad AI (governance). Firmy w Polsce zwracają uwagę na bariery, takie jak bezpieczeństwo danych (ważne dla 41% firm) oraz braki kadrowe w obszarze AI³², jednak aktywnie działają, by je pokonać, inwestując w rozwój kompetencji i strategię odpowiedzialnego korzystania z AI.

Dotychczasowe doświadczenia pokazują, że najlepsze efekty przynosi synergia ludzi i sztucznej inteligencji. AI przejmuje powtarzalne, obliczeniowe aspekty pracy, dając profesjonalistom finansowym więcej czasu na analizę, interpretację i decyzje. Jak podkreśla Andrzej Gałkowski z KPMG, firmy, które łączą

³¹ Tamże.

³² Tamże.

zaawansowane technologie z odpowiednim nadzorem i kulturą organizacyjną, budują trwałą przewagę konkurencyjną³³.

Polski sektor finansowy stoi u progu nowej ery, gdzie raporty finansowe tworzą się niemal same, prognozy są aktualizowane na bieżąco przez inteligentne algorytmy, a CFO otrzymuje na pulpicie tylko te informacje, które wymagają jego uwagi. Śmiałe projekty, takie jak wdrożenie sztucznej inteligencji w Orlenie, pokazują wizję, w której AI staje się integralnym „współpracownikiem” człowieka – nie zastępuje go, lecz wspiera we wszystkich aspektach zarządzania finansami³⁴.

Podsumowując, AI w rozwiązaniach SAP dla finansów to już nie futurystyczna ciekawostka, lecz rzeczywistość w wielu polskich firmach. Ci, którzy umiejętnie z niej skorzystają – automatyzując planowanie, zamknięcia, rozrachunki czy kontroling – zyskują na efektywności, redukują koszty i lepiej zarządzają ryzykiem. W dobie cyfrowej transformacji finansów pozostawanie przy tradycyjnych metodach może oznaczać pozostanie w tyle. Przedsiębiorstwa powinny więc odważnie, ale i rozważnie, integrować AI ze swoimi systemami SAP, tworząc nowoczesną funkcję finansową, gotową na wyzwania przyszłości. Polski rynek dostarcza już konkretnych przykładów sukcesu na tej drodze – pora, by dołączyły do nich kolejne organizacje, wykorzystując sztuczną inteligencję jako dźwignię rozwoju i doskonałości operacyjnej.

Bibliografia

- Anderson, G. (2019). *SAP S/4HANA Finance: The Reference Guide to What's New*. SAP Press.
- Kaczor, A. (2021). *Technologie chmurowe w systemach ERP: Przypadek SAP S/4HANA*. Wydawnictwo Naukowe PWN.
- Kowalski, P. (2021). *Digital Transformation in Finance: The Role of SAP S/4HANA*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Lettini, R. (2023). *The Future of Finance: Innovations in SAP S/4HANA and Beyond*. Springer.
- Meyer, T. (2024). *Sustainability in Finance: Integrating ESG Factors into SAP Financial Solutions*. Wiley.
- Nowacka, M. (2021). *Automatyzacja procesów księgowych w przedsiębiorstwie za pomocą systemu SAP*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Nowak, A. (2020). *Zastosowanie modułu FI systemu SAP do wspomagania procesów finansowo-księgowych w przedsiębiorstwie*. Wydawnictwo Naukowe PWN.

³³ Tamże.

³⁴ ORLEN – największe wdrożenie AI w Europie Środkowej..., op. cit.

- Padhi, S. N. (2021). *Configuring SAP ERP Financials and Controlling*. SAP Press.
- Pokala, P. (2023). Artificial Intelligence (AI) and Data Science Integration in SAP S/4HANA Finance. *JETIR*, 10(11), 677–682.
- Ray, P. (2023). *Implementing SAP S/4HANA: Realizing Business Value in Finance*. SAP Press.
- Salmon, J. (2020). *SAP S/4HANA Finance: An Introduction*. SAP Press.
- Sharma, C., Sharma, R., & Sharma, K. (2025). The evolution of finance and controlling: SAP and intelligent systems. *World Journal of Advanced Research and Reviews*, 25(01), 1786–1795.
- Schneider, M. (2021). *SAP S/4HANA: Financial Accounting and Controlling*. Galileo Press.

Źródła internetowe

- European Union Agency for Cybersecurity (ENISA), *Securing Machine Learning Algorithms*, ENISA Report, December 2021. <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>
- Firmy i pracownicy w Polsce coraz bardziej przychylni AI. <https://altab.pl/inne/firmy-i-pracownicy-w-polsce-coraz-bardziej-przychylni-ai/>
- Lepsze planowanie i raportowanie w środowisku SAP – Aktualności – KBJ. <https://www.kbj.com.pl/aktualnosci/lepsze-planowanie-i-raportowanie-w-srodowisku-sap.html>
- National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, January 2023. <https://www.nist.gov/itl/ai-risk-management-framework>
- PwC's 28nd Annual Global CEO Survey, *Reinvention on the edge of tomorrow*. PwC Global, 2025. <https://www.pwc.com/gx/en/ceo-survey/2025/28th-ceo-survey.pdf>
- Reviewing the Use Cases of SAP Business AI in Finance. https://learning.sap.com/learning-journeys/discovering-new-ai-capabilities-for-sap-finance/reviewing-the-use-cases-of-sap-business-ai-in-finance_d4221311-941d-4196-bd17-2c712f6d65a2
- SAP SE, *SAP Analytics Cloud: Planning and Analytics*, SAP Help Portal, 2024. https://help.sap.com/docs/SAP_ANALYTICS_CLOUD
- SAP SE, *SAP S/4HANA: Digital Core for the Intelligent Enterprise*, SAP Documentation, 2023. https://help.sap.com/docs/SAP_S4HANA_ON-PREMISE
- Sztuczna inteligencja w controllingu – FlexiEPM. <https://flexiepm.pl/sztuczna-inteligencja-w-controllingu/>
- Sztuczna inteligencja w finansach – KPMG Poland. <https://kpmg.com/pl/pl/home/insights/2025/02/sztuczna-inteligencja-w-finansach.html>

Zarządzanie wdrożeniami sztucznej inteligencji (ZT-AI) w przedsiębiorstwach sektora energetycznego

Streszczenie

Celem artykułu jest zaprezentowanie modelowego ujęcia zarządzania wdrożeniami sztucznej inteligencji (AI) w przedsiębiorstwach sektora energetycznego. Proponowany model łączy zasady bezpieczeństwa Zero Trust (ZT) z systemem zarządzania jakością (QMS), podejściem PDCA oraz zarządzaniem ryzykiem – w taki sposób, aby wdrożenia AI (w tym agentów opartych na LLM) były mierzalne, audytowalne i zgodne z regulacjami. Artykuł porządkuje narzędzia w pięciu warstwach (procesowa, danych, modeli, agentów, operacyjna), łączy je ze zdefiniowanymi rolami (RACI), mapą interesariuszy, KPI oraz metrykami kosztów jakości (CoQ). W opracowaniu zaproponowano sekwencję wdrożenia (0–30, 31–90, 91–180 dni) z bramkami jakości (go/no-go) i mierzalnymi kamieniami milowymi, a dla nadzoru strategicznego zbudowano kartę wyników BSC dla AI w energetyce. Wkład artykułu ma charakter koncepcyjno-metodyczny: przedstawia spójny, sektorowy model zarządzania jakością i ryzykiem AI, osadzony w realiach OT/IT oraz niezawodności sieci (SAIDI/SAIFI, ETR). Ograniczeniem jest brak badań empirycznych – wskazane są przy tym jednak konkretne ścieżki walidacji.

Słowa kluczowe: energetyka, zarządzanie jakością, Zero Trust, AI, KPI, OT/IT, CoQ, PDCA

1 Posel na Sejm X kadencji

Management of ZT-AI Implementations in Energy Companies

Abstract

The aim of this article is to present a model-based approach to managing the deployment of artificial intelligence (AI) in energy-sector enterprises. The proposed model integrates Zero Trust (ZT) security with a Quality Management System (QMS), the PDCA cycle, and risk management so that AI deployments—including LLM-based agents—are measurable, auditable, and compliant with regulations. The paper organizes tools across five layers (process, data, model, agent, and operations) and links them to defined roles (RACI), a stakeholder map, KPIs, and cost-of-quality (CoQ) metrics. It proposes an implementation sequence (0–30, 31–90, 91–180 days) with quality gates (go/no-go) and measurable milestones, and builds a Balanced Scorecard (BSC) for AI in the energy sector to provide strategic oversight. The contribution is conceptual and methodological: it presents a coherent, sector-specific model for managing AI quality and risk, anchored in OT/IT realities and grid-reliability metrics (SAIDI/SAIFI, ETR). A limitation is the absence of empirical studies; specific validation pathways are indicated.

Keywords: Zero Trust, AI, KPI, OT/IT, CoQ, PDCA

Lista skrótów

- Agent AI: oprogramowanie sterowane modelem AI (np. LLM), które autonomicznie wykonuje sekwencje działań w systemach IT/OT w granicach przyznanych uprawnień.
- AI: sztuczna inteligencja tworząca systemy, które wykonują zadania zwykle wymagające ludzkiej inteligencji: uczą się na danych, rozpoznają wzorce, wnioskuje, podejmują decyzje i komunikują się językiem naturalnym.
- API: Application Programming Interface (interfejs programistyczny) – ustalony zestaw reguł, formatów i punktów dostępowych, który pozwala systemom wymieniać dane i wywoływać funkcje w sposób zautomatyzowany i kontrolowany.
- Audit coverage: odsetek zmian/obiektów, które mają kompletny i wymagany audyt.
- CAB: Change Advisory Board (Komisja ds. Zmian) – międzydziałowy zespół doradczo-decyzyjny, który ocenia, klasyfikuje i zatwierdza zmiany w systemach i procesach (szczególnie OT/IT), weryfikując ryzyko, wpływ, zgodność, plan testów i plan wycofania.
- CIO: Chief Information Officer (Dyrektor ds. Informatyki).
- CMMS: Computerized Maintenance Management System – skomputeryzowany system zarządzania utrzymaniem ruchu, który planuje i rejestruje prace utrzymaniowe (prewencyjne, predykcyjne i korekcyjne) na majątku technicznym.
- CoQ (Cost of Quality) – koszty jakości: prewencji, oceny i błędów (Juran); w AI obejmuje m.in. koszty red-teamingu, walidacji i incydentów.
- CSAT (Customer Satisfaction): satysfakcja z konkretnej interakcji/procesu (np. kontakt z BOK, zgłoszenie awarii).
- Dashboard KPI: jedna, zwięzła tablica wyników odnosząca się do bieżącego nadzoru nad tym, czy procesy realizują cele.
- Data Steward: operacyjny opiekun jakości i metadanych w danej domenie danych; pracuje „pomiędzy” biznesem a IT, wdraża reguły jakości, prowadzi katalog danych i dba o zgodność (RODO, regulacje), w duchu Zero Trust.
- DER (Distributed Energy Resources): rozproszone źródła/zasoby energii, magazyny energii oraz elastyczne odbiory przyłączone zwykle blisko miejsca zużycia („na brzegu” systemu); obejmują m.in. instalacje prosumenckie (PV na dachach), małe turbiny wiatrowe i biogazownie, stacjonarne magazyny baterii.
- ETR: Estimated Time of Restoration (czas usunięcia awarii).
- FPY (First Pass Yield): odsetek wyrobów/zadań, które przechodzą cały proces za pierwszym razem.
- KPI: kluczowe wskaźniki efektywności powiązane z celami biznesowymi i bramkami jakości.

- LLM: wielkie modele językowe uczone przewidywania kolejnych „tokenów” w tekście; dzięki temu modele potrafią rozumieć i generować język naturalny, streszczać, tłumaczyć, odpowiadać na pytania, a także działać jako agent wywołujący narzędzia (API), wyszukiwarki czy bazy danych.
- MOC: Management of Change (zarządzanie zmianą) – sformalizowany proces identyfikacji, oceny, zatwierdzania, wdrożenia i przeglądu zmian.
- NPS (Net Promoter Score): lojalność do firmy/marki („na ile prawdopodobne, że polecisz nas znajomym?”).
- OMS: Outage Management System (System Zarządzania Przerwami w Dostawie Energii) – system, który zbiera zgłoszenia o przerwach, lokalizuje i grupuje zdarzenia w sieci (predykcja miejsca uszkodzenia), koordynuje brygady terenowe, estymuje ETR (Estimated Time of Restoration) i komunikuje status do klientów oraz regulatora.
- OT/IT: OT to systemy sterowania procesami fizycznymi; IT to systemy do przetwarzania informacji biznesowych (CRM, e-mail); konwergencja OT/IT łączy świat procesu z back-office i rodzi specyficzne ryzyka bezpieczeństwa oraz zgodności.
- PDCA: plan-do-check-act (Zaplanuj–Wykonaj–Sprawdź–Działaj) – iteracyjny cykl zarządzania jakością, który zapewnia ciągłe doskonalenie procesów, produktów i usług oraz służy jako „kręgosłup” QMS (ISO 9001).
- Prompt-injection success: wskaźnik bezpieczeństwa dla agentów/LLM pokazujący, w ilu próbach atak „wstrzyknięcia polecenia” faktycznie doprowadził do niepożądanego działania modelu lub agenta.
- QMS: Quality Management System (system zarządzania jakością) – uporządkowany zestaw zasad, ról, procesów, dokumentów i mierników, który w cyklu PDCA (Plan–Do–Check–Act) zapewnia zgodność wyrobów/usług z wymaganiami klienta i prawa oraz ciągłe doskonalenie organizacji.
- RACI: Responsible–Accountable–Consulted–Informed (macierz ról i odpowiedzialności) – prosta macierz przypisująca do każdego zadania/procesu rolę; celem RACI jest skrócenie ścieżek decyzyjnych, jasność odpowiedzialności i lepsza audytowalność.
- RCM: reliability-centered maintenance (utrzymanie zorientowane na niezawodność) – metodyka doboru strategii utrzymania ruchu zależnie od skutków awarii (konsekwencji), a nie od kalendarza.
- SAIDI/SAIFI, ETR: miary niezawodności sieci (np. czas i częstość przerw; szacowany czas przywrócenia zasilania).
- ZERO TRUST (ZT): strategia bezpieczeństwa zakładająca brak domyślnego zaufania – każdy dostęp i każda transakcja w tej strategii wymagają weryfikacji, autoryzacji i ciągłego monitoringu.

Wprowadzenie

Sektor energetyczny musi nieprzerwanie dostarczać prąd i szybko usuwać awarie. Jednocześnie rośnie udział OZE i rozproszonych źródeł u prosumentów, co utrudnia planowanie i sterowanie siecią. Firmy działają na styku systemów technicznych (OT) i biznesowych (IT), więc ważne jest cyberbezpieczeństwo i jasne zasady dostępu do danych. Klienci oczekują rzetelnych informacji o czasie przywrócenia zasilania (ETR) i sprawnej komunikacji. Menedżerowie potrzebują prostych narzędzi do zarządzania: czytelnych ról i odpowiedzialności (RACI), kontroli zmian (MOC/CAB), regularnych mierników (np. niezawodność sieci, satysfakcja klientów) oraz podejścia Zero Trust i QMS/PDCA, które pomagają zmniejszać liczbę błędów i szybciej podejmować decyzje. Taki porządek pracy realnie obniża koszty awarii i buduje zaufanie klientów oraz regulatora. Sektor energetyczny stoi wobec jednoczesnych, niespotykanych wcześniej presji: integracji OZE i zasobów rozproszonych (DER), dekarbonizacji i rosnących wymogów regulacyjnych, cyberzagrożeń na styku OT/IT, oczekiwań klientów nakierowanych na niezawodność, a także presji kosztowej. AI może być widziana jako narzędzie wpływające na poprawę prognozowania, optymalizację dyspozycji, wsparcie decyzji i obsługę klienta. Jednak bez ram jakości, jasnych odpowiedzialności i polityk bezpieczeństwa wdrożenia AI generują wysokie koszty błędów, luki zgodności oraz ryzyka operacyjne. Artykuł odpowiada na to wyzwanie przez przedstawienie spójnego modelu ZT-AI zarządzanego w logice QMS/PDCA.

Nie ulega wątpliwości, że AI daje szerokie możliwości dla sektora energetycznego na poziomie predyktywnym (przewidującym oraz klasyfikującym) oraz generatywnym (tworzącym nowe treści). Łącząc wszystkie poziomy AI z polityką Zero Trust oraz jakością w przedsiębiorstwach energetycznych, należy odpowiedzieć na kluczowe pytania:

- Jak zintegrować ZT i QMS, aby sterować wartością i ryzykiem projektów AI w energetyce?
- Jakie KPI i metryki CoQ mogą opisywać dojrzałość wdrożeń AI dla niezawodności i zgodności?
- Jak rozdzielić role i odpowiedzialności (RACI) w środowisku OT/IT oraz zmapować interesariuszy?
- Jak ustrukturyzować plan wdrożenia i nadzór strategiczny (BSC), aby był realistyczny dla spółek energetycznych?

Analizując ZA-AI, zakresem prac należy objąć: zarządzanie, jakość, kulturę organizacyjną w projektach AI, praktyki bezpieczeństwa (ZT-AI) oraz wątki sektorowe (energetyka, OT/IT). Uwzględnić także należy kryteria włączenia: prace z metrykami, praktykami wdrożeniowymi i/lub wytycznymi – wykluczając

materiały czysto marketingowe. Wyniki zmapowane w formie tabel: (tabela 1) „Kluczowe procesy QMS i przykładowi właściciele”, (tabela 2) „KPI i koszty jakości – przykładowy zestaw dla AI w energetyce”, (tabela 3) „Interesariusze i sposoby komunikacji”, (tabela 4) „Przykładowa macierz RACI” – mogą mieć zastosowanie w branży energetycznej.

Energetyka jest sektorem infrastruktury krytycznej, gdzie tolerancja na ryzyko jest minimalna. W energetyce charakterystyczne są: rozległa sieć systemów OT powiązana z IT, silny nadzór regulacyjny i wymogi raportowe, istotność wskaźników niezawodności oraz bezpieczeństwa prac, wysokie koszty przerw w dostawach i błędnych decyzji operacyjnych, presja transformacji (np. OZE, DER, elektromobilność) oraz konieczność odporności cybernetycznej. Z tych przesłanek wynika specyficzna potrzeba: AI musi być zarządzana jak element systemu jakości i bezpieczeństwa, a nie jedynie jak inicjatywa IT. Integralną częścią wdrażania rozwiązań IT jest oczywiście analiza głównych czynników wpływających na wybór platformy AI, jednak platforma musi być dostosowana do kompleksowej polityki przedsiębiorstw – większość z nich niemal zawsze dąży do optymalizacji produkcji energii, integracji nowych źródeł energii, utrzymania potrzebnej i krytycznej infrastruktury, trafnej prognozy zapotrzebowania oraz, na koniec, efektywnego zarządzania cenami energii.

Kształtowanie polityki wspieranej przez AI – ujęcie modelowe w branży energetycznej

Model zarządzania ZT-AI w energetyce

Model zarządzania ZT-AI rozumiany jest w artykule jako spójny sposób kierowania wdrożeniami sztucznej inteligencji w spółkach energetycznych, łączący: Zero Trust (bezpieczeństwo i minimalne uprawnienia) z QMS/PDCA (jakość i ciągłe doskonalenie) oraz zarządzanie ryzykiem i zgodnością. Na etapie analizy efektywności wdrażania modelu zarządzania ZT-AI należy uwzględnić kilka warstw oddziaływań. Kluczowe są:

- **warstwa procesowa** (QMS/PDCA) – planowanie celów jakości i ryzyka;
- **warstwa danych** – kuraacja, zgodność i audytowalność;
- **warstwa modeli** – odtwarzalność, atestacja, akceptacja jakości;
- **warstwa agentów** – zasada najmniejszych uprawnień;
- **warstwa operacyjna** – monitoring KPI/CoQ, przeglądy i doskonalenie.

Działania realizowane w rozważanych warstwach muszą być spójne, tzn. wejścia/wyjścia sąsiadujących warstw są objęte bramkami jakości (go/no-go), śladem audytu i jednoznaczną odpowiedzialnością (RACI). W praktyce oznacza to pętlę składającą się z mierników z warstwy operacyjnej (KPI/CoQ) inicjującą działania korygujące w danych, modelach i politykach agentów. W realiach

OT/IT (energetyka) każda warstwa dodatkowo respektuje wymogi bezpieczeństwa i zgodności – od atestacji artefaktów po zasadę najmniejszych uprawnień i kontrolę dostępu do krytycznych systemów.

W badaniach stosowanych (jak zarządzanie jakością i ryzykiem AI w energetyce) model nie jest prostą listą narzędzi, lecz spójną strukturą łączącą założenia, pojęcia, relacje/przyczynowość, warunki brzegowe i mechanizm operacjonalizacji, np. KPI. W tej logice pięciowarstwowy model ZT-AI (procesy – dane – model – agenci – operacje) jest modelem naukowo-praktycznym, bowiem ma strukturalne warstwy i przepływy (wejścia/wyjścia, bramki jakości, odpowiedzialności), opiera się na zdefiniowanych konstruktach, definiuje testowalne zależności i kryteria akceptacji oraz zakłada weryfikację i doskonalenie.

Procesy kluczowe i ich właściciele

Procesy kluczowe i ich właściciele są elementem ładu (governance) warstwy procesowej. Istotne jest zestawienie, które porządkuje kluczowe procesy QMS wspierające cykl życia AI i przypisuje im typowych właścicieli ról odpowiedzialnych za ich realizację. Zestawienie powinno zostać dostosowane do lokalnej struktury RACI, poziomu ryzyka oraz podziału obowiązków między obszarami OT/IT w branży. Konieczne jest sporządzenie czytelnej tabeli, która pozwala na wpisanie przy każdym procesie osób odpowiedzialnych. Konieczna może być krótka lista wymaganych dokumentów, np. karty danych czy raportu z testów. Tabelę należy aktualizować po każdej większej zmianie i przeglądać regularnie, aby każdy z uczestników procesu wiedział, kto za co odpowiada. Kluczowe zmiany będą wpływały na dostosowywanie modelu AI w przedsiębiorstwie, bowiem własne zbiory danych będą zawsze odpowiadały ich przypadkom użycia.

Tabela 1. Kluczowe procesy QMS i przykładowi właściciele

Proces	Cel	Właściciel (przykład)
Kuracja danych	Jakość i zgodność danych wejściowych	Owner danych (osoba odpowiedzialna)
Walidacja modeli	Akceptacja jakości i bezpieczeństwa	Owner modelu w obszarze zapewnienia jakości
Zarządzanie agentami	Polityki narzędzi i uprawnień	Owner platformy AI
Monitoring i incydenty	Wczesne wykrywanie odchyleń	Centrum Operacji Bezpieczeństwa
Przegląd zmian (MOC)	Kontrolowana adopcja zmian	CAB/CIO

Źródło: opracowanie własne na podstawie: ISO 9001:2015, ISO 55000:2014, NIST AI RMF 1.0 (2023).

KPI i koszty jakości (CoQ)

KPI i CoQ to metryki „przecinające” w energetyce cały model zarządzania ZT-AI. Funkcjonują w warstwie procesowej – tu są „własnością” ładu (governance), w warstwie operacyjnej – są narzędziem doskonalenia oraz w warstwach danych, modeli i agentów – w nich pozwalają dokonywać pomiarów i wskazują, jak należy liczyć koszty.

Zestaw KPI oraz miar kosztów jakości (CoQ) umożliwia obiektywne monitorowanie skuteczności i dojrzałości wdrożeń AI, z rozbiciem na prewencję, ocenę i koszty błędów. Wskaźniki te są powiązane z bramkami jakości (go/no-go) i cyklem PDCA – stanowią kryteria akceptacji zmian, podstawę przeglądów zarządzania oraz punkt wyjścia do działań korygujących. Takie podejście staje się niezbędne w procesie wdrażania AI w energetyce.

Tabela 2. KPI i koszty jakości – przykładowy zestaw dla AI w energetyce

Perspektywa	KPI (przykład)	Definicja/cel (przykład)
Jakość	FPY – First Pass Yield	Odsetek zadań/zgłoszeń bez poprawek; cel $\geq 90\%$
Bezpieczeństwo	Prompt-injection success	Odsetek skutecznych prób; cel $\leq 1\%$
Zgodność	Audit coverage	Zmiany objęte przeglądem; cel $\geq 95\%$
Klient	NPS/CSAT	Satysfakcja klienta; trend wzrostowy
Operacje	SAIDI/SAIFI (energetyka)	Niezawodność sieci; trend spadkowy
Koszty	CoQ (prevention/appraisal/failures)	Udział kosztów zapobiegania i oceny rośnie; przestają dominować koszty błędów

Źródło: opracowanie własne na podstawie: Juran (1999), Deming (1986), ISO 9001:2015, NIST (2023); kontekst sektorowy: Ginevičius et al. (2022), Zamasz (2018).

W sektorze energetycznym błąd jakości nie kończy się jedynie na reklamacji – może skutkować: przerwą w dostawie (spadkiem niezawodności), zagrożeniem bezpieczeństwa pracy i klientów, kosztami systemowymi, ryzykiem regulacyjnym (karami, obowiązkiem informacyjnym) czy eskalacją operacyjną (np. wzrostem nadgodzin lub wdrożeniem łańcucha komunikacji kryzysowej). Stąd uwzględnienie KPI oraz kosztów jakości podczas wdrażania polityki opartej na AI jest niezbędne.

Mapa interesariuszy i komunikacja

W energetyce mapa interesariuszy i komunikacja stają się narzędziem operacyjnym, a nie formalnością – każda decyzja komunikacyjna może wpływać na niezawodność sieci, bezpieczeństwo pracy, zgodność regulacyjną oraz zaufanie klientów. Mapa interesariuszy porządkuje kluczowe role w projektach AI w środowiskach OT/IT oraz ich cele, obawy i potrzeby informacyjne, co umożliwia świadome zarządzanie oczekiwaniami i odpowiedzialnościami. Zasadne jest, aby każdej grupie przypisać adekwatne kanały i częstotliwość komunikacji w celu skrócenia czasu decyzji i zapewnienia transparentności zmian oraz pełnego śladu audytu. W energetyce mapa interesariuszy wspiera działania kryzysowe oraz jasno wskazuje, kogo należy natychmiast powiadomić przy awarii.

Tabela 3. Interesariusze i sposoby komunikacji

Interesariusz	Potrzeby/obawy	Kanały komunikacji
Zarząd	Wartość, ryzyko, zgodność	Dashboard KPI, przeglądy kwartalne
Operacje (OT)	Bezpieczeństwo, niezawodność	CAB/MOC, procedury
IT/AI	Przepustowość, jakość, SLA	Raporty jakości
Regulator	Zgodność i raportowanie	Raporty cykliczne
Klient	Niezawodność i obsługa	Badania CSAT/NPS, komunikaty

Źródło: opracowanie własne na podstawie: ISO 9001:2015 (komunikacja/odpowiedzialności), NIST AI RMF (2023) (governance).

RACI – role i odpowiedzialności

Błędne przypisanie ról (kto robi, kto odpowiada, kto konsultuje, kogo informować) może skutkować wzrostem liczby opóźnionych decyzji, niepotrzebnymi wyjazdami brygad, wzrostem ryzyka i niepotrzebnych incydentów oraz niezgodnością regulacyjną. Stąd konieczne jest wdrożenie i stosowanie macierzy RACI, która jednoznacznie przypisuje: wykonanie (R), odpowiedzialność właścicielską (A), role konsultowane (C) i role informowane (I) do kluczowych procesów, co skraca ścieżki decyzyjne i ułatwia audyt. W energetyce pomaga jasno przypisać odpowiedzialności w krytycznych procesach, co przyspiesza decyzje i ułatwia audyt. W warstwie procesowej modelu zarządzania ZT-AI w energetyce decyzje zapadają na styku OT/IT, często pod presją czasu i w warunkach funkcjonowania infrastruktury krytycznej.

Tabela 4. Przykładowa macierz RACI

Obszar	R	A	C	I
Dane	Data Steward	CIO	Legal/ Compliance	Operacje
Model	Owner modelu	CTO/Chief Data Officer	QA/Sec	Zarząd
Agent	Owner platformy AI	CISO/CIO	SOC/IT	Użytkownicy
Zmiany	CAB/Change Manager	CIO	Operacje/AI/QA	Regulator

Źródło: opracowanie własne na podstawie: ISO 9001:2015 (role/kompetencje), praktyka RACI.

Ryzyka organizacyjne

Wdrożenia AI w energetyce częściej potykają się o ryzyka organizacyjne niż o same modele czy narzędzia. Ryzyka organizacyjne są „przekrojowe”, ale ich ład znajduje się w warstwie procesowej, która jest jedną z pięciu warstw modelu zarządzania ZT-AI w energetyce. Antywzorce „podbijają” koszty jakości (CoQ), rozmywają odpowiedzialności i osłabiają skuteczność kontroli ZT-AI. Każdy z nich powinien mieć przypisanego właściciela oraz stały wpis w rejestrze ryzyk, omawiany cyklicznie w przeglądach PDCA/MOC. Brak powiązania AI z celami jakości i ryzyka (odwrócona przyczynowość: „bo AI jest modne”) oraz ignorowanie kosztów jakości (CoQ) – co uwidacznia się oszczędzaniem na prewencji – skutkuje kosztownymi poprawkami. Prawdziwym wyzwaniem są także niedookreślone role, które wynikają z braku właścicieli danych, modeli lub agentów. Nie bez znaczenia jest także tzw. change fatigue, czyli zmęczenie zmianą w zespole bez planu komunikacji i szkoleń. Sukces wdrożeń AI bezsprzecznie zależy od opanowania ryzyk organizacyjnych — od jasnego podziału ról po sprawną komunikację i zarządzanie zmianą.

Plan wdrożenia

Plan wdrożenia polityki opartej na AI powinien zakładać horyzont czasowy. W celu szybkiego osiągnięcia mierzalnych korzyści, przy kontrolowanym ryzyku, najczęściej zakłada się horyzont 90–180 dni, z jasno określonymi kamieniami milowymi, bramkami jakości (go/no-go) i właścicielami zadań. Przy ustalaniu planu wdrożenia najpierw ustalone powinny zostać cele/KPI i fundamenty działań (np. dane i role), następnie powinien zostać uruchomiony pilotaż oraz cykl PDCA. Okresy: 0–30/31–90/91–180 dni odzwierciedlają typowy rytm PDCA (planowanie, pilotaż, stabilizacja i skalowanie) oraz cykle decyzyjne spółek energetycznych..

Oto jedna z propozycji sekwencji działań z zakresem i kluczowymi produktami każdego etapu:

- 0–30 dni: cele jakości i ryzyka; mapa interesariuszy; miary KPI/CoQ; szybki przegląd danych;
- 31–90 dni: zasady kuracji danych; bramki akceptacji modeli; RACI; polityki agentów; dashboard KPI;
- 91–180 dni: pilotaż w jednym procesie (np. OMS/obsługa zdarzeń); przegląd zgodności; plan skalowania.

Balanced Scorecard (BSC) dla AI w energetyce

Aby włączyć AI do zarządzania, w warstwie procesowej należy również uwzględnić BSC wraz z celami, miarami i inicjatywami w czterech perspektywach. BSC spina KPI jakości, bezpieczeństwa, zgodności i klienta, a wyniki regularnie przegląda Zarząd (PDCA). Dzięki BSC Zarząd i zespoły mogą zobaczyć na jednej stronie: cele, miary oraz potrzebne działania korygujące – inicjatywy. W energetyce BSC dla AI nie jest opcjonalnym dodatkiem, tylko mechanizmem dowodzenia i kontroli – przekłada strategię na mierzalne cele i równoważy wartość i ryzyko. Dzięki BSC AI podnosi niezawodność, bezpieczeństwo i zgodność, zamiast je przypadkowo obniżać.

Tabela 5. Balanced Scorecard (BSC) – przykład dla AI w energetyce

Perspektywa	Cel (przykład)	Miara/KPI (przykład)	Inicjatywy (przykład)
Finanse	Zmniejszyć koszty przerw i strat operacyjnych	CoQ	Predykcja awarii; optymalizacja dyspozycji i prac
Klient	Podnieść satysfakcję i skrócić czas odpowiedzi	średni czas odpowiedzi; trafność ETR	Asystent LLM dla BOK; standardy odpowiedzi i komunikacji kryzysowej
Procesy	Ustabilizować działanie modeli i agentów	FPY, odsetek rollbacków pokazujący, jaki procent wdrożeń/zmian trzeba było wycofać w danym okresie	Bramki jakości (go/no-go); polityki agentów
Uczenie/rozwój	Podnieść kompetencje zespołów OT/IT/AI	ukończone szkolenia/egzamin; liczba ćwiczeń	Program szkoleń AI + OT/IT; ćwiczenia red-team naśladujące działania „napastnika”, żeby celowo złamać zasady i znaleźć luki

Źródło: opracowanie własne na podstawie Kaplan & Norton (1996), Deming (1986) (PDCA), NIST AI RMF (2023).

Zakończenie

Zestawienie ZT-AI z QMS pozwala włączyć AI do systemu zarządzania, a nie traktować ją jako eksperyment technologiczny. Model ZT-AI odpowiada na zgłaszane w literaturze potrzeby łączenia jakości, bezpieczeństwa i zarządzania ryzykiem w złożonych środowiskach technologicznych w branży energetycznej. Wkład praktyczny polega na ustrukturyzowaniu narzędzi w logice warstw i PDCA oraz na dopasowaniu metryk (KPI/CoQ) do celów niezawodności. AI może stać się dźwignią doskonałości operacyjnej w energetyce, jeśli będzie zarządzana jak element QMS i bezpieczeństwa. Połączenie ZT, PDCA i metryk CoQ pozwala równoważyć wartość i ryzyko. Zaproponowany model ZT-AI – warstwy, RACI, KPI/CoQ, BSC oraz plan 0–180 dni – tworzy zarys dla ramy wdrożeniowej dla przedsiębiorstw energetycznych.

Dalsze wyzwania, które stoją przed branżą energetyczną, obejmują m.in. standaryzację metryk dla agentów LLM, integrację kultury jakości z kulturą bezpieczeństwa oraz pozatechniczne bariery (kompetencje, komunikacja czy czas decyzyjny). Kierunkiem dalszych analiz powinny zostać objęte obszary związane ze standaryzacją KPI dla agentów LLM w energetyce, modele kosztów jakości (CoQ) specyficzne dla wdrożeń AI w OT/IT oraz wpływ zarządzania zmianą na ROI z AI. AI może stać się dźwignią doskonałości operacyjnej, jeśli jest zarządzana przez pryzmat jakości, ryzyka i odpowiedzialności. Proponowane ramy do zarządzania wdrożeniami sztucznej inteligencji (AI) w przedsiębiorstwach sektora energetycznego, łączące ZT-AI z QMS i narzędziami zarządczymi (KPI, CoQ, RACI), oraz mądre wykorzystanie AI w procesie budowania strategii w warunkach transformacji cyfrowej mogą tworzyć nowe wytyczne dla dalszego kierunku rozwoju w sektorze energetycznym.

Bibliografia

- Dacko-Pikiewicz, Z., Szczepańska-Woszczyna, K., & Lis, M. (2025). *Horyzonty sztucznej inteligencji a Przemysł 5.0*. Wydawnictwo Naukowe AWSB w Dąbrowie Górniczej.
- Deming, W. E. (1986). *Out of the Crisis*. MIT Press.
- Ginevičius, R., Trishch, R., Bilan, Y., Lis, M., & Pěnčík, J. (2022). Assessment of the Economic Efficiency of Energy Development in the Industrial Sector of the EU Area Countries. *Energies*, 15(9), 3322. <https://doi.org/10.3390/en15093322>
- International Organization for Standardization ISO 9001. (2015). *Quality management systems – Requirements*.
- International Organization for Standardization ISO 55000. (2024). *Asset management – Vocabulary, overview and principles*.

- Juran, J. M. (2023). *Juran's Quality Handbook*, McGraw-Hill.
- Kaplan, R. S., Norton, D. P. (1996). *The Balanced Scorecard*. Harvard Business School Press.
- Mucha-Kuś, K., Sołtysik, M., Zamasz, K., & Szczepańska-Woszczyna, K. (2021). Coopetitive Nature of Energy Communities – The Energy Transition Context. *Energies*, 14(4), 931. <https://doi.org/10.3390/en14040931>
- NIST AI 100-1. (2023). *AI Risk Management Framework 1.0*, National Institute of Standards and Technology. <https://www.nist.gov/itl/ai-risk-management-framework>
- Zamasz, K. (2018). Energy company in a competitive energy market. *Polityka Energetyczna – Energy Policy Journal*, 21(2), 35–48. <https://doi.org/10.24425/122768>

Industry 5.0 i sztuczna inteligencja jako źródło wyzwań dla polityki rozwoju regionalnego województwa śląskiego

Streszczenie

W artykule omówiono kluczowe dylematy, które powinny być uwzględnione w procesie sporządzania strategii regionalnej w dynamicznie zmieniającym się świecie. Podkreślono, że współczesne procesy mające wpływ na kształtowanie polityki rozwoju cechuje duża dynamika i zmienność, obejmujące wymiar technologiczny, klimatyczny i społeczny. Determinuje to konieczność elastycznego podejścia do strategii, umożliwiającego adaptację do zmieniających się warunków otoczenia. Artykuł analizuje znaczenie koncepcji Industury 5.0 i sztucznej inteligencji (ang. artificial intelligence, AI) dla procesów rozwoju województwa śląskiego, które niosą ze sobą szereg wyzwań, takich jak zmiany na rynku pracy, braki w infrastrukturze cyfrowej i energetycznej oraz wysokie koszty transformacji technologicznej i procesowej przedsiębiorstw. Przedstawiono także rekomendacje dla samorządu w procesie zmiany strategii rozwoju województwa śląskiego, uwzględniające potrzebę ujęcia dynamiki zmian technologicznych, a także rekomendacje do procesu budowy i wdrażania tego dokumentu.

Słowa kluczowe: Industury 5.0, sztuczna inteligencja, strategia, polityka regionalna

1 Akademia WSB w Dąbrowie Górniczej

Industry 5.0 and Artificial Intelligence as a Source of Challenges for the Regional Development Policy of the Silesian Voivodeship

Abstract

This article discusses key dilemmas that should be considered in the process of developing a regional strategy in a dynamically changing world. It emphasizes that contemporary processes influencing development policy are characterized by high dynamics and volatility, encompassing technological, climatic, and social dimensions. This determines the need for a flexible approach to strategy that allows for adaptation to changing environmental conditions. The article analyzes the importance of Industry 5.0 and Artificial Intelligence (AI) concepts for the development processes of the Silesian Voivodeship, which pose a number of challenges, such as changes in the labor market, deficiencies in digital and energy infrastructure, and the high costs of technological and process transformation for enterprises. The article also presents recommendations for local governments in the process of amending the Silesian Voivodeship's development strategy, taking into account the need to address the dynamics of technological change, as well as recommendations for the process of developing and implementing this document.

Keywords: Industry 5.0, artificial intelligence, strategy, regional policy

Wstęp

Współczesne procesy mające wpływ na kształtowanie polityki rozwoju, zarówno na poziomie makro, mezo, jak i mikro, cechuje duża dynamika i zmienność. Do tych kluczowych obszarów zmian zaliczyć należy wymiar technologiczny, klimatyczny i społeczny. Z tego też względu klasyczna definicja strategii mówiąca, że strategia to określenie podstawowych, długofalowych celów organizacji, przyjęcie kierunków działania oraz alokacja zasobów niezbędnych do osiągnięcia tych celów (Chandler, 1962), wymaga zastosowania podejścia zaproponowanego przez M. Portera, który wskazuje, że strategia to tworzenie unikalnej i wartościowej pozycji poprzez odmienny zestaw działań, której celem jest osiągnięcie trwałej przewagi konkurencyjnej (Porter, 1996). Jednocześnie, biorąc pod uwagę znaczną dynamikę zmian, należy przyjąć podejście H. Mintzberga, traktującego strategię jako pewien wzorzec działań, który może być zarówno zaplanowany, jak i emergentny – wyłaniający się w odpowiedzi na zmieniające się warunki otoczenia, i powinna być ona elastyczna, umożliwiając organizacji adaptację do zachodzących zmian (Mintzberg, Ahlstrandand, Lampel, 2005). Ta dynamika, złożoność, nieprzewidywalność i zmienność otoczenia procesu wdrożenia strategii wiążą się nie tylko z realizacją opracowanego scenariusza, ale także z koniecznością strategicznej, kreatywnej elastyczności (Szczepańska-Woszczyna, 2018).

Jednocześnie należy zauważyć, że zarządzanie strategiczne w sektorze publicznym coraz częściej przyjmuje ramy i narzędzia ze strategii sektora prywatnego – szczególnie w obszarach takich jak pomiar wydajności, alokacja zasobów i planowanie strategiczne – w celu poprawy efektywności, elastyczności oraz tworzenia wartości publicznej (Ferlie, Ongaro, 2015).

Dostrzegając zatem potrzebę traktowania strategii jako planu umożliwiającego adaptowanie się do zachodzących zmian, przy jednoczesnej potrzebie myślenia o długookresowych skutkach naszych obecnych decyzji oraz celach społecznych, gospodarczych i środowiskowych w dłuższej perspektywie czasu, do kluczowych mechanizmów, jakie powinny stanowić fundament zarządzania strategicznego zarówno w sektorze biznesowym, jak i szczególnie w sektorze publicznym, zaliczyć należy koncepcję Industry 5.0.

Ramy teoretyczne w zakresie Industry 5.0 i sztucznej inteligencji – kontekst strategii rozwoju regionalnego

Industry 5.0 stawia człowieka w centrum procesu produkcyjnego, promując zrównoważony rozwój i odporność, uzupełniając cele efektywności i automatyzacji Industry 4.0 (European Commission, 2021). W ramach tej koncepcji

technologie takie jak: sztuczna inteligencja, robotyka czy cyfryzacja mają wspierać nie tylko wydajność, lecz także lokalne wartości społeczne i środowiskowe (Pereira, dos Santos, 2023). Należy zauważyć, że to właśnie współczesne technologie stanowią o tempie rozwoju przedsiębiorstw (Ginevičius, Nazarko, Gedvilaitė, Dacko-Pikiewicz, 2021). Jednocześnie, jak wskazuje M. Lis, rozwój technologii cyfrowych powinien być traktowany jako narzędzie, a nie cel transformacji (Lis, Brike, 2023). Z tego też względu koncepcja Industy 5.0 powinna wywierać istotny wpływ na zagadnienia kształtowania polityk rozwojowych w regionach i tworzonych w administracji publicznej strategiach rozwoju. W kontekście regionów europejskich Industy 5.0 promuje model rozwoju inkluzywnego i, jak wskazuje R. Verma, regiony, które inwestują w rozwój kompetencji cyfrowych oraz integrację człowieka i technologii, zwiększają swoją odporność na kryzysy (np. energetyczne czy zdrowotne), a także przyciągają inwestycje wspierające lokalny ekosystem innowacji (Verma, 2024). Takie podejście podkreślono również w raporcie „Industy 5.0 – towards more sustainable, resilient and humancentric industry”, gdzie zarysowano wizję Industy 5.0 jako kolejnego kroku w rozwoju przemysłu w Europie, który uzupełnia i poszerza koncepcję Industy 4.0. Zgodnie z zapisami raportu koncepcja opiera się na trzech filarach (European Commission, 2021):

1. Human-centric

- człowiek jest w centrum procesów przemysłowych,
- technologie mają wspierać, a nie zastępować ludzi,
- istotne są godne warunki pracy, rozwój kompetencji i partycypacja społeczna,

2. Sustainable

- ograniczenie wpływu przemysłu na środowisko,
- wspieranie gospodarki cyrkularnej i zielonej transformacji,
- integracja celów klimatycznych UE z działalnością przemysłową,

3. Resilient

- budowa odpornych łańcuchów dostaw i systemów produkcji,
- umiejętność szybkiego reagowania na kryzysy,
- cyfrowa i operacyjna adaptacyjność.

Różnice między Industy 4.0 a Industy 5.0 dotyczą zatem głównie celów, roli człowieka oraz wartości środowiskowych.

Tabela 1. Kluczowe różnice pomiędzy koncepcją Industry 4.0 i Industry 5.0

Aspekt	Industry 4.0	Industry 5.0
Wartości i cele	Automatyzacja i cyfryzacja procesów przemysłowych ukierunkowana na optymalizację i wysoką efektywność	Współpraca człowieka z maszyną i technologią z uwzględnieniem celów zrównoważonego rozwoju
Technologie kluczowe	IoT (Internet rzeczy), big data, chmura, robotyka	Sztuczna inteligencja, coboty (roboty współpracujące), personalizacja produkcji
Relacja człowiek–maszyna, technologia, sztuczna inteligencja	Subsydiarna – zastępują człowieka	Komplementarna – wspierają i współpracują z człowiekiem
Środowisko	Optymalne wykorzystanie zasobów przez automatyzację i optymalizację technologii i procesów	Realizacja zrównoważonego modelu gospodarki z minimalizacją wpływu na klimat

Źródło: opracowanie własne.

Należy jednak zauważyć, że wdrażanie koncepcji Industry 5.0 niesie ze sobą szereg wyzwań i ryzyk. Jednym z istotnych wyzwań są zmiany na rynku pracy oraz niedostosowanie kompetencyjne pracowników na nim (European Commission, 2021), w tym brak odpowiednio przygotowanej kadry do obsługi zaawansowanych technologii, takich jak sztuczna inteligencja, robotyka kolaboratywna czy systemy cyberfizyczne (European Commission, 2021; Verma, 2024). Jak wskazuje raport World Economic Forum (WEF) 41% firm planuje redukcję zatrudnienia do 2030 roku, ale jednocześnie aż 77% prowadzi działania aktywnego przekwalifikowania i podnoszenia kompetencji (reskilling/upskilling). WEF prognozuje, że do 2030 roku udział zadań powierzonych wyłącznie ludziom spadnie, a równy podział nastąpi pomiędzy zadaniami w pełni zautomatyzowanymi i tymi realizowanymi przez człowieka we współpracy z AI. Z kolei AI i technologie informacyjne radykalnie zmienią model biznesowy organizacji (World Economic Forum, 2025). Szacuje się, że do 2030 roku aż 39% obecnych umiejętności stanie się przestarzałe, a najszybciej rozwinie się zapotrzebowanie na zawody związane z big data, AI, oprogramowaniem, zieloną energią oraz opieką zdrowotną i edukacją. Natomiast najbardziej zagrożone automatyzacją są stanowiska takie jak kasjerzy, asystenci administracyjni, urzędnicy pocztowi, sekretarki i graficy (World Economic Forum, 2025).

Szybkie wdrażanie automatyzacji i nowoczesnych technologii, jak robotyka, powoduje, że sektory gospodarki i gospodarki regionalne są w stanie zwiększyć

wydajność, co pozwala na produkcję większego wolumenu produktów przy mniejszym zatrudnieniu.

W obszarze podnoszenia wydajności oraz ograniczenia i zmiany zakresu pracy i aktywności człowieka ogromne znaczenie będzie miało upowszechnienie sztucznej inteligencji. Do 2050 roku AI zautomatyzuje większość rutynowych zadań, ale stworzy miliony miejsc pracy w sektorach kreatywnych, zarządzaniu AI i zdrowiu (PwC Europe 2025; OECD, 2024). Szacuje się, że do 2030 roku zniknie około 20 mln miejsc pracy w UE z powodu automatyzacji i AI, ale powstanie 30 mln nowych w branżach technologicznych i kreatywnych, co będzie istotnie wpływać na pogłębianie się nierówności regionalnych (McKinsey Global Institute, 2024; WEF, 2025).

Kolejnym wyzwaniem związanym z wdrożeniem koncepcji Industy 5.0 są braki w infrastrukturze cyfrowej i energetycznej. Wykorzystanie technologii chmurowych, Internetu rzeczy (IoT), cyfrowych bliźniaków czy analityki czasu rzeczywistego wymaga wysokiej jakości infrastruktury, której brak zwiększa ryzyko cyfrowego wykluczenia (Pereira, dos Santos, 2023). Według prognoz, w związku z rozwojem centrów danych i AI, do 2030 roku zużycie energii podwoi się, a w przypadku centrów dedykowanych AI wzrośnie czterokrotnie (IEA, 2024). Jednocześnie AI generuje skokowe zapotrzebowanie, np. centra AI mogą wywołać chwilowe wzrosty obciążenia sieci nawet dziesięciokrotnie, co grozi destabilizacją systemów energetycznych (Hook, 2025). Z drugiej strony, jak wskazuje M. Lis, sztuczna inteligencja wspomaga analizę danych w czasie rzeczywistym, optymalizując zużycie energii w zakładach przemysłowych, czym wspiera analizę procesów produkcyjnych, pozwalając na zmniejszenie zużycia surowców i redukcję odpadów (Lis, Kądziałowski, 2025).

Wdrożenie nowoczesnych rozwiązań niewątpliwie powodować będzie wysokie koszty transformacji technologicznej i procesowej przedsiębiorstw (Pereira, dos Santos, 2023; ManpowerGroup, 2023). Wyzwanie to dostrzega również Komisja Europejska – wskazuje, że wdrożenie niezbędnych procesów transformacyjnych może być utrudnione z powodu barier kosztowych lub kulturowych (European Commission, 2021).

Kształtowanie procesów transformacji w kierunku wdrożenia koncepcji Industy 5.0, w tym szerokiego wykorzystania AI, będzie również warunkowane kulturą organizacyjną oraz nastawieniem pracowników (World Economic Forum, 2023). Ma to znaczenie zwłaszcza w sektorze publicznym, w którym złożoność integracji celów zrównoważonego rozwoju oraz trudności instytucjonalne i niedopasowanie polityk regionalnych wymagają szczególnej koordynacji działań i rozwoju współpracy międzysektorowej, w tym z podmiotami naukowymi, organizacjami pozarządowymi i przedsiębiorstwami. Jednak w wielu regionach

dostrzega się brak struktur i kompetencji do prowadzenia takiej współpracy (Springer, 2025; Maddikunta et al., 2022).

Ta zarysowana powyżej niewielka część wyzwań związanych z rozwojem technologicznym, obok zmian klimatycznych, społecznych czy politycznych, powinna stanowić fundament myślenia o polityce rozwoju regionów w Polsce, której emanacją jest opracowanie i wdrażanie regionalnej strategii rozwoju.

Polskie regiony – województwa – zgodnie z przepisami prawa mają obowiązek opracowania i wdrażania strategii rozwoju. Zgodnie z zapisami art. 11 ust. 1c Ustawy o samorządzie województwa strategia zawiera:

- wnioski z diagnozy,
- cele strategiczne w wymiarze społecznym, gospodarczym, przestrzennym i klimatyczno-środowiskowym,
- kierunki działań podejmowanych dla osiągnięcia celów strategicznych,
- oczekiwane rezultaty planowanych działań, w tym w wymiarze przestrzennym, oraz wskaźniki ich osiągnięcia,
- model struktury funkcjonalno-przestrzennej województwa,
- ustalenia i rekomendacje w zakresie kształtowania i prowadzenia polityki przestrzennej w województwie,
- obszary strategicznej interwencji, o których mowa w art. 12a ust. 1 pkt 5 ustawy z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju, wraz z zakresem planowanych działań,
- obszary strategicznej interwencji kluczowe dla województwa, jeżeli takie zidentyfikowano, wraz z zakresem planowanych działań,
- system realizacji strategii, w tym wytyczne do sporządzania dokumentów wykonawczych,
- ramy finansowe i źródła finansowania.

Powyższa ustawa określa ramy procesu opracowania strategii regionalnej i, niestety, w istotny sposób go formalizuje, powodując, że zarówno przygotowanie dokumentu strategii, jaki i wprowadzenie do niej zmian – np. w wyniku zmiany warunków realizacyjnych – jest skomplikowane i długotrwałe proceduralnie. Ustawa ta wskazuje, że strategia rozwoju województwa podlega aktualizacji, jeżeli wymaga tego sytuacja społeczna, gospodarcza lub przestrzenna województwa albo gdy jest to konieczne dla zachowania jej spójności ze średniookresową strategią rozwoju kraju lub krajową strategią rozwoju regionalnego; przy czym taka aktualizacja wymaga procedowania strategii, jak przy jej pierwotnym opracowywaniu (art. 11 Ustawy o samorządzie województwa).

Jednocześnie w systemie prawnym istnieje szereg zapisów rozszerzających zakres strategii regionalnych, w tym np.:

- Ustawa z dnia 6 grudnia 2006 r. – o zasadach prowadzenia polityki rozwoju wprowadza definicję strategii, określa zagadnienie spójności strategicznej

oraz rozszerza zakres procesu konsultacyjnego w toku przygotowania projektu strategii,

- Ustawa z dnia 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko wprowadza proces tzw. strategicznej oceny oddziaływania na środowisko, w tym udział społeczeństwa w procedurze przygotowania projektu strategii oraz opiniowanie przez określone organy – Regionalny Dyrektor Ochrony Środowiska oraz Państwowy Wojewódzki Inspektor Sanitarny,
- Ustawa z dnia 20 lipca 2017 r. Prawo wodne – wskazuje obowiązek uwzględnienia m.in. obszarów szczególnego zagrożenia powodzią oraz ustalenia dokumentów planistycznych: planów zagospodarowania wodami na obszarach dorzeczy; planów zarządzania ryzykiem powodziowym; planów przeciwdziałania skutkom suszy.

Tak szczegółowe i rozporoszone określenie zakresu i trybu opracowania strategii rozwoju województwa oraz formalizacja procesu jej sporządzenia (opatrzona procedurami administracyjnymi) powodują, że dokument ten traci swoją użyteczność jako rzeczywiste i skuteczne narzędzie zarządzania rozwojem w szybko zmieniających się uwarunkowaniach, w jakich współczesne regiony funkcjonują i konkurują. W warunkach dynamicznych przemian – takich jak cyfryzacja, rozwój sztucznej inteligencji, transformacja energetyczna, zmiany demograficzne czy globalne kryzysy – skuteczna strategia musi być elastyczna i adaptacyjna. Nie może funkcjonować jako sztywny plan, lecz jako żywy mechanizm zarządzania zmianą, umożliwiający szybkie reagowanie na pojawiające się wyzwania i szanse, a także zmieniający się potencjał i sprawność podmiotu wdrażającego. Elastyczne kształtowanie strategii oznacza także konieczność jej stałej weryfikacji – zarówno poprzez monitorowanie i ewaluowanie efektów realizowanych działań, jak i poprzez budowanie nowych relacji w kształtowaniu i wdrażaniu polityki regionalnej. Niezbędna jest zatem zmiana podejścia zarówno do zakresu strategii, wymagającego uwzględnienia dynamiki przeobrażeń, jak i miejsca oraz funkcji tego narzędzia w zarządzaniu publicznym.

Rekomendacje dla samorządu w procesie zmiany strategii

Opierając się na wieloletniej obserwacji uczestniczącej w zakresie planowania procesów rozwoju województwa śląskiego, wyrażonych w strategii rozwoju, poniżej sformułowano kluczowe rekomendacje do zmiany w aktualnej strategii rozwoju województwa, uwzględniające potrzebę ujęcia w niej dynamiki zmian technologicznych oraz rekomendacje do procesu budowy i wdrażania tego dokumentu.

Obecnie wdrażana jest piąta edycja strategii w województwie śląskim, tj. Strategia Rozwoju Województwa Śląskiego „Śląskie 2030. Zielone Śląskie”, która stanowi kontynuację wcześniejszych dokumentów strategicznych z lat: 2000, 2005, 2010 oraz 2013. W aktualnej strategii województwa śląskiego skoncentrowano się na:

- procesach transformacji i budowie nowoczesnej gospodarki,
- rozwoju kapitału ludzkiego i społecznego,
- atrakcyjności przestrzeni i poprawie mobilności,
- rozwoju współpracy i partnerstw.

Analiza dokumentu wskazuje, że strategia w głównej mierze odnosi się do koncepcji Industry 4.0 poprzez promocję digitalizacji i automatyzacji, a także łączenie ich z transformacją energetyczną, natomiast w bardzo ograniczonym zakresie uwzględnia koncepcję Industry 5.0, która w pełniejszy sposób respektuje zasady zrównoważonego rozwoju, zmianę roli technologii i uwzględnia dynamiczny wpływ sztucznej inteligencji na procesy rozwoju.

W zakresie ww. kluczowych obszarów rozwoju województwa śląskiego należy zwrócić uwagę i podejmować działania związane z wdrażaniem koncepcji Industry 5.0, w tym z wykorzystaniem AI. Poniżej przedstawiono zakres rekomendowanych kierunków działań do aktualnego dokumentu strategii rozwoju województwa śląskiego.

Tabela 2. Rekomendowane kierunki działań do Strategii Województwa Śląskiego z obszaru Industry 5.0

Cel strategiczny rozwoju województwa	Rekomendowane kierunki działań implementujące koncepcję Industry 5.0
Województwo śląskie regionem odpowiedzialnej transformacji gospodarczej	• Wspieranie przedsiębiorstw we wdrażaniu robotyki współpracującej (cobots) i AI – finansowanie, doradztwo, • Promowanie personalizacji produkcji (mass customization) – wsparcie dla firm wdrażających elastyczne linie produkcyjne z użyciem AI, • Tworzenie centrów kompetencji Industry 5.0 – regionalne huby innowacji, które łączą uczelnie, startupy i przemysł, • Transformacja przemysłu z pomocą AI – optymalizacja procesów produkcyjnych pod kątem zużycia surowców i energii, • Promowanie w przedsiębiorstwach aktywnego reskillingu i upskillingu, • Wdrażanie standardów etycznych i odpowiedzialnego wykorzystania AI,

Cel strategiczny rozwoju województwa	Rekomendowane kierunki działań implementujące koncepcję Industy 5.0
	• Podnoszenie kwalifikacji pracowników poprzez szkolenia z zakresu AI, robotyki współpracującej, etyki AI oraz cyberbezpieczeństwa, • Promowanie „human-centric design” w przemyśle, w tym budowanie partycypacji pracowników i wzmocnienie bezpieczeństwa pracy przy współpracy z maszynami.
Województwo śląskie regionem przyjaznym dla mieszkańca	• Uowszechnianie cyfrowej inkluzji z wykorzystaniem szkoleń tworzonych i realizowanych przy wykorzystaniu AI i technologii dla grup wykluczonych (seniorzy, osoby z niepełnosprawnościami, obszary peryferyjne), • Wspieranie zmian w systemie edukacji, w tym położenie nacisku na interdyscyplinarność poprzez tworzenie i realizację programów łączących nauki ścisłe, społeczne i humanistyczne, • Włączenie do programów nauczania kompetencji miękkich i technologicznych obejmujących korzyści i koszty współpracy człowiek – maszyna, mechanizmy analizy danych oraz zagadnień etycznych związanych z technologią AI, • Wspieranie i promowanie współpracy uczelni ponad podziałami branżowymi, np. technicznych z uczelniami kształcącymi w obszarach humanistycznych i artystycznych czy naukach medycznych, • rozwój kompetencji nauczycieli w zakresie nowoczesnych technologii wykorzystujących narzędzia AI, • Wdrożenie AI jako asystenta dydaktycznego w procesach analizy wyników uczniów czy opracowanie spersonalizowanych narzędzi, materiałów i pomocy edukacyjnych, • Personalizacja opieki zdrowotnej z wykorzystaniem AI, w tym wdrażanie rozwiązań AI w diagnostyce i profilaktyce (np. analiza obrazów medycznych, przewidywanie ryzyka chorób), • Wspomaganie opieki senioralnej i osób zależnych z wykorzystaniem technologii monitorujących i diagnozujących (np. czujniki upadku, analiza snu, detekcja depresji), • Wspieranie modelu uczenia przez doświadczenie, w tym poprzez rozwój laboratoriów Przemysłu 5.0 i kształcenia dualnego, • Rozwijanie etycznego podejścia do AI – edukacja obywateli nt. wpływu AI na prywatność, pracę, bezpieczeństwo, • Wspieranie „human-centric AI” w usługach publicznych – AI w opiece zdrowotnej, edukacji, polityce społecznej, • Budowanie społeczeństwa prosumenckiego – zachęty do współtworzenia produktów i usług z wykorzystaniem AI.

Cel strategiczny rozwoju województwa	Rekomendowane kierunki działań implementujące koncepcję Industry 5.0
Województwo śląskie regionem wysokiej jakości środowiska i przestrzeni	• Wykorzystanie AI do monitoringu i zarządzania zasobami naturalnymi – np. inteligentne systemy zarządzania wodą, energią, emisjami, • Mapowanie zagrożeń klimatycznych z użyciem uczenia maszynowego i AI – planowanie adaptacji do zmian klimatu na poziomie lokalnym, • Rozwój inteligentnych systemów transportowych (ITS) – AI w zarządzaniu ruchem, bezpieczeństwem, planowaniu sieci komunikacyjnych, • Wzmocnienie odporności infrastruktury krytycznej – wykrywanie zagrożeń, cyberbezpieczeństwo z wykorzystaniem AI, • Modernizacja i rozwój systemu energetycznego regionu (produkcja, przesył, magazynowanie) z wykorzystaniem narzędzi AI w sterowaniu produkcji energii ze źródeł odnawialnych i konwencjonalnych, w tym wdrażanie inteligentnych sieci energetycznych, które automatycznie zarządzają przepływem energii, reagując aktywnie na jej produkcję, zużycie i zakłócenia.
Województwo śląskie regionem sprawnie zarządzanym	• Promowane koncepcji smart city i smart village – integracja AI w zarządzaniu miastami i wsiami (energia, odpady, mobilność, przestrzeń publiczna), • Wdrożenie cyfrowych narzędzi partycypacji społecznej, w tym tworzenie platform opartych na AI do konsultacji społecznych i budżetów obywatelskich, • Wykorzystanie AI w analityce predykcyjnej ukierunkowanej na przewidywanie przyszłych zdarzeń lub trendów, • Wykorzystywanie AI do tworzenia dynamicznych raportów z realizacji strategii oraz jej dostosowywania do zidentyfikowanych ryzyk w zarządzaniu jednostkami terytorialnymi.

Źródło: opracowanie własne.

Biorąc pod uwagę doświadczenia i olbrzymi dorobek samorządu województwa śląskiego w zakresie partycypacyjnego formułowania strategii rozwoju regionu, w dobie dynamicznych zmian oraz przy usztywnieniu procesu tworzenia strategii – wskazać można dwie grupy dylematów, jakie należy rozstrzygnąć, przystępując do kolejnych edycji opracowania Strategii, tj.:

- dylematy wyboru ścieżki rozwojowej,
- dylematy procesy budowy strategii.

Tabela 3. Kluczowe dylematy wyboru ścieżki rozwoju województwa

Dylemat	Zakres dylematu
Konkurencyjność vs zasady i wartości	Czy sukces strategii powinien być oceniany przez pryzmat wzrostu gospodarczego, czy przez pryzmat rozwoju uwzględniającego jakość zachodzącej zmiany? Skupienie się wyłącznie na efekcie ekonomicznym może prowadzić do naruszenia zasad i wartości istotnych z punktu widzenia odpowiedzialności społecznej i równoważenia rozwoju. Natomiast nadmierne oparcie się na wartościach, bez realistycznych mechanizmów konkurencyjnych, może zahamować dynamikę rozwoju.
Efektywność vs odpowiedzialność	Na ile nasze wybory i decyzje ukierunkowane są na maksymalizację efektów przy minimalnych kosztach i dążeniu do „rentowności” działań, a w jakim zakresie są motywowane odpowiedzialnością, rozumianą jako myślenie systemowe, które uwzględnia skutki społeczne, środowiskowe i gospodarcze strategicznych wyborów.
Ja vs my	Czy region w swoich działaniach promuje indywidualizm i konkurencję jednostek, czy też współpracę, solidarność i działania zbiorowe. Model „ja” sprzyja innowacjom, pojedynczym sukcesom, a model „my” buduje trwałość, zaufanie i kapitał społeczny.
Dziś vs jutro	Z jaką perspektywą czasową wyznaczamy cele i oceniamy skutki podjętych decyzji? „Dziś” to inwestycje i działania przynoszące szybkie rezultaty – często korzystne politycznie. „Jutro” to inwestycje w przyszłość: bezpieczeństwo, edukację, zdrowie, innowacje, klimat, które wymagają czasu, ale dają trwałe efekty i fundamenty do dalszego rozwoju.
Otwartość vs izolacja	Na ile region i jego społeczność są gotowi i dostrzegają potrzebę otwarcia się na zewnętrzne idee, technologie i partnerstwa – bez obawy o utratę swojej tożsamości i ważnych dla nich wartości? W jakich obszarach realizator strategii musi podjąć trwałą współpracę?
Przywództwo transakcyjne vs przywództwo transformacyjne	Na jakim modelu budować rolę samorządu województwa w procesie formułowania i wdrażania strategii? Czy budować model przywództwa transakcyjnego – opartego na kontroli, mechanizmach administracyjnych, czy transformacyjne – oparte na liderstwie wizjonerskim, zdolności do mobilizacji i tworzenia kultury zmiany?

Źródło: opracowanie własne.

Tabela 4. Kluczowe dylematy w procesie formułowania strategii rozwoju województwa

Dylemat	Zakres dylematu
Strategia jako projekt vs strategia jako proces	Czy strategia powinna być jednorazowym, zamkniętym dokumentem (projektem), czy raczej dynamicznym, elastycznym mechanizmem reagującym na zmiany (procesem)? Na ile scharakteryzować mechanizmy wdrożeniowe wiążące podmiot odpowiedzialny za realizację?
Strategicznosc vs operacyjność	Jaka w dokumencie powinna być relacja celów i narzędzi? Czy ważniejsze jest określenie, co chcemy osiągnąć (cele strategiczne), czy jak to osiągniemy (sposoby i środki)?
Aspiracje vs wykonalność	Czy koncentrować się na ambitnych zamiarach, czy realistycznie oceniać, co jest możliwe do wykonania? Jak budować relacje z partnerami dla wspólnego działania na rzecz celów wykraczających poza kompetencje organu tworzącego strategię?
Krótki vs długi horyzont czasowy	W jakiej perspektywie czasowej wyznaczać cele, a w jakiej budować narzędzia wdrożeniowe i jak zadbać o ich adaptacyjność do zmiany?
Reaktywność vs prewencja	Na ile strategia stanowić powinna odpowiedź na bieżące problemy – oparcie wyborów strategicznych na diagnozie już zidentyfikowanych problemów, a na ile na diagnozie prospektywnej, ukierunkowanej na prognozowanie przyszłych uwarunkowań i wyzwań?

Źródło: opracowanie własne.

Biorąc pod uwagę doświadczenia samorządu województwa w obszarze zarządzania strategicznego, należy podkreślić, że obszar formułowania strategii oraz tworzenia analiz i diagnoz jest w fazie dojrzałości procesu. Wyzwaniem natomiast jest optymalizacja i podniesienie efektywności procesu rzeczywistego wdrażania strategii. Strategia rozwoju województwa jest zoperacjonalizowana przez szereg strategii i programów sektorowych oraz programów operacyjnych, w tym z wykorzystaniem narzędzi finansowych UE. Istotnym deficytem natomiast pozostaje powiązanie strategii z bieżącymi operacjami i działaniami samorządu, a także aktywne dostosowywanie działań strategicznych do zmieniających się uwarunkowań poprzez zintegrowany system zarządzania ryzykiem strategicznym.

Biorąc pod uwagę rolę strategii, ramy formalnoprawne oraz zakres procedur, a jednocześnie dynamikę uwarunkowań wpływających na kierunki rozwoju i warunki realizacyjne, kluczowe wyzwania, przed jakim stoi samorząd województwa, są następujące:

- wykreowanie i wdrożenie modelu przywództwa regionalnego, obejmującego dostosowanie struktury administracji regionalnej oraz – w porozumieniu

z administracją samorządów lokalnych – do potrzeb realizacji strategii, do potrzeb wdrażania strategii – w tym optymalizację systemów kontroli, audytu i monitorowania;

- zwiększenie roli i funkcji samorządu regionalnego jako rzeczywistego kreatora i realizatora polityki rozwoju – w tym korzystanie ze strategii w zakresie planowania i realizacji budżetu regionu, opiniowania działań innych podmiotów, realizacji polityki dotacji itd.;
- zwiększenie roli samorządu jako rzeczywistego partnera podmiotów sektora samorządowego i pozasamorządowego w kreowaniu i wdrażaniu strategicznych projektów;
- wypracowanie i wdrożenie aktywnego systemu zarządzania ryzykiem strategicznym – jako elementu komplementarnego do ryzyk na poziomie organizacyjnym określonym w systemie kontroli zarządczej administracji samorządu województwa;
- wykorzystanie narzędzi AI w diagnozowaniu, analizowaniu, prognozowaniu uwarunkowań wewnętrznych i zewnętrznych oraz ocenie efektywności i skuteczności podejmowanych działań.

Podsumowanie

Województwo śląskie stoi przed określeniem skorygowanej ścieżki rozwoju. W toku tego procesu warto postawić sobie kluczowe pytania: w jakim kierunku i jakie wartości powinny determinować trajektorię rozwojową tego dużego i złożonego województwa, a jednocześnie w jaki sposób zbudować mechanizmy wdrożeniowe tak, by były one efektywne i skuteczne, a sama strategia stanowiła rzeczywiste narzędzie zarządzania rozwojem regionalnym.

Bez wątpienia koncepcja Industy 5.0, w której zakłada się komplementarność wykorzystania potencjału człowieka, maszyn i technologii oraz rozwoju i upowszechnienia AI, jest źródłem wielu wyzwań związanych z kształtowaniem potencjału przedsiębiorstw, budowaniem systemu kształtowania kompetencji mieszkańców, zarządzania środowiskiem oraz budowania relacji międzysektorowych. Wyzwaniem będzie także wykorzystanie nowych narzędzi i instrumentów w procesach zarządczych i prognostycznych, których zastosowanie w istotnym stopniu może determinować działania samorządu województwa.

Bibliografia

ACEA. (2023). Share of direct automotive employment in the EU by country. https://www.acea.auto/figure/share-of-direct-automotive-employment-in-the-eu-by-country/?utm_source=chatgpt.com [dostęp: 06.07.2025]

- Chandler, A. D. (1962). *Strategy and Structure: Chapters in the History of the American Industrial Enterprise*. MA: MIT Press.
- European Commission (2021). Industry 5.0 – towards more sustainable, resilient and humancentric industry. European Commission, 7 January [dostęp: 06.07.2025]
- Eurofound. (2004). Employment in the European automotive industry: Trends and prospects. https://www.eurofound.europa.eu/sites/default/files/ef_files/emcc/publications/2004/ef0427en.pdf?utm_source=chatgpt.com [dostęp: 06.07.2025]
- Ferlie, E., & Ongaro, E. (2015). *Strategic Management in Public Services Organizations: Concepts, Schools and Contemporary Issues*. Routledge.
- Ginevičius, R., Nazarko, J., Gedvilaitė, D., & Dacko-Pikiewicz Z. (2021). Quantifying the economic development dynamics of a country based on the lorenz curve. *Ekonomie a Management*, 24(1), 55–65.
- Hook, L. (2025). Hitachi Energy says SI power spikes threaten to destabilise global supply, Financial Times. <https://www.ft.com/content/2789d048-7791-4f9a-8dbe-229c45f8083b>
- IEA. (2024). Data Centres and Data Transmission Networks, International Energy Agency. <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>
- Lis, M. (2023). Transformacja cyfrowa jako nowe wyzwanie dla organizacji i jej otoczenia. W M. Lis & M. Bryke (red.), *Transformacja KAIZEN w perspektywie nowych technologii i cyfryzacji* (ss. 7–47). Wydawnictwo Naukowe AWSB w Dąbrowie Górniczej.
- Lis, M., & Kądziałowski, G. (2025). *Horyzonty sztucznej inteligencji a Przemysł 5.0* (s. 35). Wydawnictwo Naukowe AWSB w Dąbrowie Górniczej.
- Maddikunta, P. K. R., Pham, Q. V., Prabadevi, B. et al. (2022). Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26, 100257.
- ManpowerGroup (2023). *Nastroje polskiego rynku pracy – perspektywa kandydatów* [raport]. ManpowerGroup Polska.
- McKinsey Global Institute. (2024). *The future of work in Europe: Automation, SI and employment trends*. McKinsey & Company.
- Mintzberg, H., Ahlstrand, B., & Lampel, J. (2005). *Strategy Safari: A Guided Tour Through the Wilds of Strategic Management*. Pearson Education.
- Organisation for Economic Co-operation and Development (OECD). (2023). *OECD employment outlook 2023: Artificial intelligence and the labour market*. OECD Publishing. <https://doi.org/10.1787/08785bba-en>
- Pereira, A. C., & dos Santos, M. F. (2023). Neoindustrialization: Reflections on a new paradigmatic approach for the Industry 5.0. *Designs*, 7(3), 43–57.
- Porter, M. E. (1996). What is Strategy? *Harvard Business Review*, 74(6), 61–78.

- PricewaterhouseCoopers. (2025). The 2025 Global AI Jobs Barometer: AI and the future of jobs [Raport]. PwC. <https://www.pwc.com/gx/en/services/ai/ai-jobs-barometer.html>
- Rejeb, A., Rejeb, K., Zrelli, I., & Süle, E. (2025). Industry 5.0 as seen through its academic literature: An investigation using co-word analysis. *Discover Sustainability*. <https://doi.org/10.1007/s43621-025-01166-0>
- Szczepańska-Woszczyna, K. (2018). Strategy, Corporate Culture, Structure and Operational Processes as the Context for the Innovativeness of an Organization. *Foundations of Management*, 10, Article 1. <https://doi.org/10.2478/fman-2018-0004>
- Verma, D. (2024). Industry 5.0: A Human-Centric and Sustainable Approach to Industrial Development. *International Journal of Social Relevance & Concern*, 12(5), 17–21.
- World Economic Forum. (2025). The Future of Jobs Report 2025, Geneva: WEF. <https://www.weforum.org/reports/the-future-of-jobs>

Katowice w cyfrowej transformacji administracji samorządowej – rola Centrów Usług Wspólnych w rozwoju miasta przyszłości

Streszczenie

Artykuł analizuje proces cyfrowej transformacji administracji samorządowej w Katowicach, ze szczególnym uwzględnieniem roli Centrum Usług Wspólnych (CUW) jako kluczowego podmiotu wspierającego rozwój nowoczesnej administracji miejskiej. Na podstawie studium przypadku przedstawiono działania podejmowane w obszarze cyfryzacji procesów administracyjnych, integracji systemów informatycznych oraz wdrażania rozwiązań z zakresu sztucznej inteligencji i automatyzacji. W opracowaniu zaprezentowano również wyzwania związane z bezpieczeństwem danych oraz rozwojem kompetencji cyfrowych pracowników administracji publicznej. Katowice ukazane zostały jako przykład miasta konsekwentnie realizującego model smart administration, w którym technologia wspiera efektywne, transparentne i obywatelsko zorientowane zarządzanie.

Słowa kluczowe: cyfryzacja administracji, Centra Usług Wspólnych, Katowice, transformacja cyfrowa, smart city, sztuczna inteligencja, administracja samorządowa

¹ Wiceprezydent Katowic

Katowice in the Digital Transformation of Local Government Administration – the Role of Shared Services Centres in the Development of the City of the Future

Abstract

The article analyses the process of digital transformation of local government administration in Katowice, with particular emphasis on the role of the Shared Services Centre (SSC) as a key entity supporting the development of modern municipal administration. Based on a case study, the paper presents initiatives undertaken in the areas of digitalisation of administrative processes, integration of IT systems, and the implementation of solutions based on artificial intelligence and automation. The study also discusses challenges related to data security and the development of digital competences among public administration employees. Katowice is presented as an example of a city consistently implementing the "smart administration" model, in which technology supports efficient, transparent, and citizen-oriented governance.

Keywords: administrative digitalisation, Shared Services Centres, Katowice, digital transformation, smart city, artificial intelligence, local government administration

Wprowadzenie

Cyfryzację administracji publicznej zaliczyć można do najważniejszych kierunków rozwoju współczesnego samorządu. Zmiana, która jeszcze dekadę temu była jedynie elementem nowoczesnego zarządzania, dziś jest nieodłącznym warunkiem funkcjonowania jednostek administracji publicznej. W tym procesie szczególnie rolę odgrywają Centra Usług Wspólnych – wyspecjalizowane jednostki organizacyjne, które przejmują obsługę finansowo-księgową, kadrową i administracyjną jednostek samorządowych (OECD, 2023).

Katowice są jednym z miast, które konsekwentnie realizują ideę cyfrowej transformacji, wykorzystując potencjał technologii informatycznych do budowania efektywnej, zintegrowanej i przyjaznej mieszkańcom administracji. Centrum Usług Wspólnych w Katowicach stało się jednym z filarów tego procesu – miejscem, gdzie nowoczesne technologie spotykają się z doświadczeniem urzędników i potrzebami jednostek miejskich. Transformacja cyfrowa Katowic jest procesem przemyślanym i konsekwentnym. Jej podstawą jest dążenie do stworzenia miasta, które nie tylko efektywnie zarządza zasobami, lecz także ułatwia mieszkańcom codzienne kontakty z administracją. Miasto od lat inwestuje w rozwój systemów informatycznych, cyfryzację procesów i budowanie kompetencji cyfrowych pracowników.

Centra Usług Wspólnych zostały powołane, aby usprawnić zarządzanie administracją lokalną poprzez łączenie wspólnych zadań – głównie w obszarach księgowości, płac, kadr, zamówień publicznych i obsługi prawnej. Ich cyfryzacja nie polega wyłącznie na informatyzacji, lecz na głębokiej reorganizacji procesów, która umożliwia działanie w modelu smart administration (Ministerstwo Funduszy i Polityki Regionalnej, 2023).

CUW stały się naturalnym łącznikiem między samorządami a technologią. Integrują systemy, gromadzą dane i wprowadzają nowe modele zarządzania informacją, tworząc podstawy nowoczesnego e-rządu na poziomie lokalnym. Z raportu dr. Pawła Modrzyńskiego wynika, że mimo wzrostu świadomości potrzeby cyfryzacji, aż 48,5% CUW świadczy usługi raczej tradycyjnie, a 20,4% w sposób zdecydowanie tradycyjny. Jedynie 29,1% jednostek działa w pełni cyfrowo (Modrzyński, 2024).

Przyczyną tej dysproporcji są głównie: brak integracji systemów, ograniczone finansowanie oraz niedostateczne przygotowanie kadr. Tylko około 20% CUW wdrożyło pełny elektroniczny obieg dokumentów, a 33% korzysta z jego częściowych funkcji. W 69% jednostek planowane jest jednak zwiększenie funkcjonalności systemów w najbliższych trzech latach – głównie poprzez automatyzację procesów i wdrażanie zintegrowanych rozwiązań ERP (Modrzyński, 2024).

Centrum Usług Wspólnych w Katowicach w systemie zarządzania miastem

Centrum Usług Wspólnych w Katowicach jest jednym z najważniejszych elementów strategii technologicznego rozwoju miasta i usług publicznych. Jednostka ta obsługuje liczne instytucje miejskie – na chwilę obecną wszystkie placówki oświatowe, docelowo jednak może objąć także jednostki pomocy społecznej i instytucje kultury. Dzięki temu możliwe jest centralne zarządzanie finansami, kadrami i dokumentacją w sposób bardziej przejrzysty, spójny i efektywny. Wprowadzenie nowoczesnych systemów klasy ERP umożliwia integrację wielu procesów administracyjnych w jednym środowisku cyfrowym. CUW Katowice z sukcesem wdrożyło elektroniczny obieg dokumentów, dzięki któremu ograniczono papierową dokumentację, skrócono czas realizacji spraw i zwiększono transparentność działań administracji. W praktyce oznacza to m.in. pełną cyfrową kontrolę nad przepływem dokumentów między CUW a jednostkami obsługiwanymi, archiwizację elektroniczną oraz możliwość bieżącego monitorowania statusu każdej sprawy; praktyka wdrożeniowa odzwierciedla główne założenia odnoszące się do powoływania CUW w perspektywie prawnej, organizacyjnej i menedżerskiej (Modrzyński, Gawłowski, Modrzyńska, 2018).

Rozwój CUW Katowice wpisuje się w szerszy kierunek zmian w mieście – strategię Katowice 2030, która zakłada, że nowoczesne technologie będą fundamentem zarządzania miastem. W ramach tej wizji Katowice dążą do budowy modelu „administracji nowej generacji” – opartej na danych, interoperacyjności systemów i automatyzacji procesów.

Współczesna administracja nie może już opierać się na klasycznych, ręcznych procedurach. Sektor nowoczesnych usług wymaga rozwiązań adekwatnych do lokalnego kontekstu ekonomicznego i społecznego, z uwzględnieniem efektywności, wpływu na zatrudnienie oraz miejsca CUW w lokalnych i regionalnych strategiach rozwoju społeczno-gospodarczego (Motowidlak, T., Motowidlak U., 2023). Dzięki cyfryzacji CUW Katowice przejmuje coraz więcej zadań, które dawniej wymagały dużych nakładów pracy i czasu. Systemy informatyczne umożliwiają szybkie raportowanie, automatyczne generowanie zestawień budżetowych oraz elektroniczne zatwierdzanie dokumentów. Takie rozwiązania pozwalają urzędnikom skupić się na analizie danych i podejmowaniu decyzji strategicznych, a nie na samym procesie przetwarzania informacji.

Cyfryzacja CUW Katowice jest integralną częścią strategii Smart City Katowice. To podejście, w którym technologia staje się narzędziem do lepszego zarządzania zasobami, energią, transportem i informacją (Ministerstwo Funduszy i Polityki Regionalnej, 2023). Administracja cyfrowa to także administracja otwarta na mieszkańców. Dzięki wprowadzeniu rozwiązań takich jak platforma e-Urząd

Katowice mieszkańcy mogą załatwiać wiele spraw zdalnie – od wniosków o świadczenia po sprawy podatkowe i administracyjne. CUW Katowice pełni rolę zaplecza organizacyjnego i technologicznego, które zapewnia sprawne funkcjonowanie tych systemów. Praca jednostki opiera się na integracji wielu źródeł danych, które służą zarówno do bieżącego zarządzania, jak i do analizy długofalowych trendów. Dane te wykorzystywane są m.in. do planowania budżetowego, monitorowania wydatków, a także do analiz efektywności działań administracyjnych (Skrzyński, 2021).

Katowice zamierzają w przyszłości wykorzystywać narzędzia analityczne oparte na sztucznej inteligencji i automatyzacji. Algorytmy AI mogłyby wspierać klasyfikację dokumentów, rozpoznawanie treści, a także planowanie finansowe. Dzięki temu miasto poszerzy możliwości podejmowania decyzji opartych na rzeczywistych danych, a nie tylko na szacunkach czy intuicji.

Jednym z kluczowych kierunków rozwoju CUW Katowice jest wprowadzanie narzędzi sztucznej inteligencji i automatyzacji procesów administracyjnych. Wykorzystanie algorytmów uczenia maszynowego umożliwia automatyczne przetwarzanie dokumentów, identyfikowanie błędów, a nawet prognozowanie wydatków w określonych jednostkach miejskich (Dacko-Pikiewicz, Szczepańska-Woszczyna, Lis, 2025).

W Katowicach prowadzone są również wstępne prace nad wykorzystaniem sztucznej inteligencji do analizy danych budżetowych i kadrowych. W dłuższej perspektywie takie rozwiązania mają wspierać zarządzanie finansami miasta oraz optymalizację procesów w jednostkach obsługiwanych przez CUW. Wdrożenie systemów AI pozwoli nie tylko usprawnić obsługę administracyjną, ale również poprawić jakość podejmowanych decyzji i efektywność finansową.

Rozwój technologiczny wymaga równoległego wzmacniania systemów ochrony danych. CUW Katowice kładzie duży nacisk na kwestie cyberbezpieczeństwa, w tym ochronę danych osobowych, polityki dostępu oraz bezpieczne przechowywanie informacji. Zastosowanie rozwiązań chmurowych umożliwia elastyczne zarządzanie danymi, ale również wymaga stosowania najwyższych standardów bezpieczeństwa (Ministerstwo Cyfryzacji, 2022). Katowice współpracują w tym zakresie z ekspertami ds. bezpieczeństwa oraz jednostkami odpowiedzialnymi za ochronę infrastruktury cyfrowej w Górnośląsko-Zagłębiowskiej Metropolii. Audyty, szkolenia oraz wdrażanie polityk bezpieczeństwa pozwalają utrzymać wysoki poziom zaufania do systemów miejskich. Warto dodać, że Katowice należą do miast, które w ostatnich latach pracują nad kompleksową strategią cyberbezpieczeństwa, obejmującą zarówno kwestie technologiczne, jak i organizacyjne. Obejmuje ona systemy monitorowania zagrożeń, zarządzania incydentami, a także programy szkoleniowe dla pracowników administracji.

Cyfryzacja administracji publicznej to nie tylko technologia, ale przede wszystkim ludzie. CUW Katowice, podobnie jak cały Urząd Miasta, inwestuje w rozwój

kompetencji cyfrowych pracowników. Szkolenia, warsztaty i programy edukacyjne pozwalają urzędnikom na efektywne korzystanie z nowych narzędzi informatycznych oraz zrozumienie potencjału, jaki niesie automatyzacja procesów. W tym zakresie miasto współpracuje zarówno z Konsorcjum Akademickim, skupiającym siedem publicznych uczelni, oraz z Akademią WSB. Uczelnie te wspierają rozwój kompetencji cyfrowych, szkolą specjalistów ds. bezpieczeństwa danych i analityki, a także uczestniczą w projektach badawczych związanych z zastosowaniem AI w administracji. Ta rozwijająca się współpraca stanowi przykład synergii między samorządem a środowiskiem akademickim, gdzie nauka wspiera praktyczne rozwiązania w zarządzaniu miastem.

Dzięki tym inicjatywom Katowice budują kulturę cyfrową administracji – otwartą na zmiany, innowacyjną i gotową na wyzwania przyszłości. Transformacja cyfrowa w Katowicach nie ogranicza się wyłącznie do usprawnienia administracji. Jest częścią szerszej wizji rozwoju miasta jako ekosystemu nowoczesnych technologii i inteligentnych rozwiązań miejskich. Koncepcja Smart City Katowice łączy w sobie innowacje technologiczne, zarządzanie oparte na danych oraz aktywną współpracę z mieszkańcami.

CUW Katowice pełni w tej strukturze rolę zaplecza organizacyjnego, które dostarcza dane i wspiera koordynację działań administracyjnych. Dzięki temu miasto może skutecznie zarządzać informacjami o wydatkach, zasobach ludzkich, infrastrukturze i usługach publicznych. Integracja danych pomiędzy jednostkami sprawia, że decyzje podejmowane są szybciej i w sposób bardziej precyzyjny.

Katowice systematycznie rozwijają platformy cyfrowe umożliwiające zarządzanie przestrzenią miejską, transportem, gospodarką odpadami czy energią. Nowoczesne technologie wykorzystywane są również w obszarze komunikacji z mieszkańcami – aplikacje miejskie, systemy zgłaszania usterek, portale dla przedsiębiorców i mieszkańców pozwalają na dwustronną wymianę informacji między obywatelami a urzędem.

Jednym z najważniejszych wyzwań cyfrowej administracji jest integracja systemów informatycznych. W wielu samorządach wciąż funkcjonują rozwiązania działające w oderwaniu od siebie, co powoduje powielanie danych i brak spójności informacji. Katowice dostrzegły ten problem już na wczesnym etapie i podjęły działania na rzecz stworzenia zintegrowanego środowiska cyfrowego.

CUW Katowice prowadzi proces harmonizacji systemów finansowo-księgowych, kadrowych, sprawozdawczych i administracyjnych. Wspólny system pozwala na kontrolę budżetu całego miasta w czasie rzeczywistym, a także na analizę kosztów w poszczególnych jednostkach. Integracja umożliwia także lepsze planowanie wydatków, co jest szczególnie istotne w okresie rosnących kosztów funkcjonowania administracji publicznej (Modrzyński, 2024). Dzięki wprowadzonym rozwiązaniom informatycznym możliwe jest centralne raportowanie danych finansowych

i kadrowych z kilkudziesięciu jednostek podległych, co znacząco usprawnia proces planowania budżetu miejskiego. W praktyce oznacza to większą przejrzystość finansów publicznych, łatwiejsze przygotowanie sprawozdań i szybsze reagowanie na zmiany w otoczeniu gospodarczym. Katowice są jednym z nielicznych miast w Polsce, które konsekwentnie rozwijają własne, interoperacyjne środowisko informatyczne, łączące systemy miejskie z krajowymi rozwiązaniami, takimi jak ePUAP czy Centralna Ewidencja i Informacja o Działalności Gospodarczej.

Kolejnym etapem rozwoju CUW Katowice jest wykorzystanie sztucznej inteligencji w zarządzaniu informacjami i procesami administracyjnymi. W ostatnich latach miasto podjęło działania pilotażowe w zakresie stosowania algorytmów AI do klasyfikowania dokumentów, wspomagania obiegu pism, a także do analizy danych budżetowych. AI może być wykorzystywana do prognozowania wydatków, identyfikacji obszarów wymagających optymalizacji oraz wspierania procesów decyzyjnych w zakresie planowania budżetowego. Dzięki zastosowaniu sztucznej inteligencji urzędnicy zyskują dostęp do narzędzi analitycznych, które pozwalają im lepiej rozumieć mechanizmy finansowe miasta i efektywniej planować działania.

Automatyzacja w CUW Katowice obejmuje również procesy kadrowe i księgowe. Systemy RPA (Robotic Process Automation) umożliwiają realizację powtarzalnych zadań – takich jak księgowanie faktur, przygotowywanie raportów czy rozliczanie delegacji – w sposób szybki i bezbłędny. W efekcie ograniczono czas potrzebny na realizację czynności administracyjnych nawet o kilkadziesiąt procent. Warto podkreślić, że rozwiązania te nie zastępują człowieka, ale wspierają go w codziennej pracy. Automatyzacja procesów zwiększa efektywność, ogranicza ryzyko błędów i pozwala na racjonalne zarządzanie zasobami ludzkimi.

Jednym z priorytetów Katowic jest rozwój analityki danych, która ma wspierać zarządzanie miastem w oparciu o fakty, a nie intuicję. CUW Katowice, jako jednostka centralna, gromadzi ogromne ilości informacji dotyczących funkcjonowania instytucji miejskich. Dane te, odpowiednio analizowane, pozwalają na planowanie inwestycji, monitorowanie efektywności wydatków i przewidywanie przyszłych potrzeb finansowych miasta. System raportowania CUW Katowice umożliwia przygotowywanie zbiorczych zestawień dla całego miasta oraz udostępnianie ich instytucjom nadzorującym, np. Wydziałowi Edukacji Urzędu Miasta Katowice. Umożliwia to lepsze zrozumienie, w jaki sposób wydatkowane są środki publiczne, oraz pozwala na zwiększenie przejrzystości finansów miejskich. Dzięki takim rozwiązaniom Katowice stają się przykładem miasta, które potrafi skutecznie wykorzystywać dane w procesie zarządzania. Cyfrowe raporty, analizy i prognozy tworzą podstawę do podejmowania decyzji strategicznych, a automatyzacja zadań pozwala pracownikom skupić się na działaniach wymagających wiedzy eksperckiej.

Każda transformacja cyfrowa niesie nie tylko korzyści, ale również nowe wyzwania. Wraz z rosnącą liczbą systemów informatycznych, danych i usług online pojawia się potrzeba zapewnienia wysokiego poziomu bezpieczeństwa cyfrowego. Katowice od lat stawiają na konsekwentne budowanie systemu cyberbezpieczeństwa, który obejmuje zarówno aspekty technologiczne, jak i organizacyjne. CUW Katowice, jako jednostka gromadząca i przetwarzająca dane wielu instytucji miejskich, jest szczególnie wrażliwym ogniwem w łańcuchu bezpieczeństwa cyfrowego miasta. Dlatego w jego strukturach wdrożono szereg rozwiązań mających na celu ochronę danych osobowych, integralność systemów oraz kontrolę dostępu. Systemy wykorzystywane w CUW Katowice spełniają wysokie standardy bezpieczeństwa, a ich działanie jest regularnie monitorowane i audytowane. Katowice są również jednym z miast, które wprowadziły politykę cyberbezpieczeństwa opartą na rekomendacjach Ministerstwa Cyfryzacji i Krajowego Systemu Cyberbezpieczeństwa (Ministerstwo Cyfryzacji, 2022). W ramach tych działań prowadzone są regularne szkolenia pracowników, symulacje zagrożeń oraz ćwiczenia reagowania na incydenty. W urzędzie wdrożono wielopoziomowy system uprawnień, a dostęp do danych wrażliwych jest chroniony poprzez autoryzację i szyfrowanie.

Ważnym elementem bezpieczeństwa cyfrowego jest również świadomość pracowników. Katowice systematycznie inwestują w edukację urzędników, którzy uczą się rozpoznawania zagrożeń, stosowania zasad bezpiecznego korzystania z sieci oraz właściwego reagowania na potencjalne incydenty. Dzięki temu bezpieczeństwo staje się nie tylko kwestią technologiczną, ale również elementem kultury organizacyjnej.

Współpraca z Górnośląsko-Zagłębiowską Metropolią pozwala na wymianę doświadczeń w zakresie ochrony systemów informatycznych, wspólne opracowywanie procedur bezpieczeństwa oraz koordynację działań w przypadku incydentów o charakterze regionalnym. W ten sposób budowany jest spójny system cyberbezpieczeństwa metropolitalnego, który chroni zarówno dane administracji, jak i informacje obywateli.

Cyfryzacja administracji nie byłaby możliwa bez ludzi – ich wiedzy, zaangażowania i gotowości do uczenia się nowych technologii. CUW Katowice stanowi przykład jednostki, która równolegle z wdrażaniem systemów informatycznych, prowadzi proces rozwoju kompetencji cyfrowych wśród swoich pracowników. Miasto konsekwentnie inwestuje w szkolenia, warsztaty i programy podnoszenia kwalifikacji, które obejmują obsługę systemów ERP i EOD, jak i zagadnienia związane z ochroną danych, automatyzacją czy wykorzystaniem sztucznej inteligencji. Pracownicy CUW są przygotowywani do pracy w środowisku, gdzie cyfrowe narzędzia są codziennością, a procesy decyzyjne opierają się na danych.

Ważnym elementem tego procesu jest współpraca z uczelniami wyższymi, przede wszystkim z Uniwersytetem Śląskim, Uniwersytetem Ekonomicznym,

Politechniką Śląską i Akademią WSB. Uczelnie te nie tylko kształcą przyszłych specjalistów IT i analityków danych, ale również współpracują z miastem przy projektach badawczo-rozwojowych. Wspólnie opracowywane są rozwiązania w zakresie analizy danych miejskich, wykorzystania sztucznej inteligencji w zarządzaniu finansami publicznymi oraz rozwoju kompetencji cyfrowych w administracji.

W Katowicach funkcjonuje także system wewnętrznej wymiany doświadczeń pomiędzy jednostkami. Pracownicy CUW i innych instytucji miejskich dzielą się wiedzą w ramach warsztatów i spotkań tematycznych, które mają na celu propagowanie dobrych praktyk w obszarze cyfryzacji. Dzięki temu administracja staje się organizacją uczącą się – gotową do adaptacji nowych technologii i reagowania na zmieniające się potrzeby mieszkańców (Modrzyński, 2024).

Centrum Usług Wspólnych w Katowicach – perspektywa rozwoju

Jednym z największych osiągnięć Katowic jest zmiana podejścia do zarządzania urzędem. Transformacja cyfrowa to nie tylko nowoczesne oprogramowanie, ale przede wszystkim zmiana kultury pracy. W CUW Katowice i jednostkach podległych rozwija się kultura oparta na współpracy, odpowiedzialności i otwartości na innowacje. Wdrożenie elektronicznego obiegu dokumentów i systemów ERP wymagało od pracowników zmiany nawyków – przejścia od pracy opartej na papierowych dokumentach do coraz pełniejszej cyfrowej obsługi spraw. Początkowo proces ten wiązał się z wyzwaniami, jednak dzięki konsekwentnym działaniom edukacyjnym i wsparciu organizacyjnemu zaczyna funkcjonować coraz lepiej. Obecnie CUW Katowice to jednostka o wysokim stopniu dojrzałości cyfrowej. Większość procesów realizowana jest w sposób zautomatyzowany, a dane są przechowywane i przetwarzane w środowisku bezpiecznym i dostępnym dla uprawnionych użytkowników. Administracja miejska funkcjonuje coraz bardziej w modelu sieciowym – jednostki współpracują ze sobą, wymieniają dane i korzystają z tych samych platform technologicznych. Takie podejście sprzyja efektywności i transparentności. Każdy dokument ma swoje miejsce w systemie, każdy etap jego obiegu jest widoczny, a procesy można monitorować i doskonalić. Katowice dzięki temu nie tylko usprawniły pracę urzędu, ale również zwiększyły zaufanie mieszkańców do instytucji publicznych.

Perspektywa roku 2030 wyznacza dla Katowic nowy etap rozwoju administracji. W miarę jak technologie cyfrowe stają się coraz bardziej zaawansowane, miasto dąży do stworzenia spójnego ekosystemu usług publicznych, opartego na danych, sztucznej inteligencji i interoperacyjności systemów. Centrum Usług Wspólnych w Katowicach ma odegrać w tym procesie kluczową rolę jako

centralny ośrodek zarządzania informacjami, automatyzacji procesów i rozwoju kompetencji cyfrowych. Już dziś CUW pełni funkcję centrum integracji danych finansowych, kadrowych i organizacyjnych, a w perspektywie najbliższych lat jego zadania będą się rozszerzać o nowe obszary, takie jak analiza predykcyjna, raportowanie w czasie rzeczywistym czy zarządzanie procesami międzyinstytucjonalnymi.

W planach znajduje się także rozwój inteligentnych narzędzi analitycznych, które umożliwią lepsze prognozowanie budżetu miasta, monitorowanie efektywności inwestycji publicznych i wspieranie decyzji strategicznych. Systemy oparte na big data pozwolą na analizę tysięcy danych w krótkim czasie, co przełoży się na większą dokładność planowania i oszczędności finansowe.

Równie ważnym elementem będzie dalsze wdrażanie rozwiązań sztucznej inteligencji (AI), zarówno w procesach administracyjnych, jak i w obsłudze mieszkańców. Planowane jest rozwinięcie systemów wspierających elektroniczną komunikację obywatelską – chatbotów i asystentów głosowych, które pomogą mieszkańcom w załatwianiu spraw urzędowych bez konieczności wizyty w urzędzie.

Katowice pełnią w Górnośląsko-Zagłębiowskiej Metropolii funkcję lidera cyfryzacji. Współpraca z sąsiednimi samorządami w ramach GZM pozwala tworzyć rozwiązania o zasięgu ponadlokalnym – od wspólnych platform informatycznych po systemy raportowania i analizy danych.

CUW Katowice, dzięki swojemu doświadczeniu i rozwiniętej infrastrukturze, jest w stanie wspierać inne jednostki metropolii w procesie cyfrowej transformacji. Wspólne działania obejmują m.in. standaryzację procesów, tworzenie jednolitych baz danych i szkolenia z zakresu kompetencji cyfrowych.

Interoperacyjność, czyli zdolność systemów do współpracy, staje się fundamentem nowoczesnej administracji. Katowice są jednym z miast, które już dziś wdrażają rozwiązania umożliwiające płynny przepływ informacji pomiędzy różnymi systemami, zarówno na poziomie lokalnym, jak i krajowym. Dzięki temu możliwa jest np. integracja danych o mieszkańcach, przedsiębiorcach czy inwestycjach z rejestrami centralnymi.

W dłuższej perspektywie Katowice rozważają stworzenie wspólnego metropolitalnego centrum danych, które pozwoli na efektywne zarządzanie informacjami z wielu gmin, a także wzmocni bezpieczeństwo i odporność cyfrową regionu.

Współczesne wyzwania, takie jak cyberzagrożenia, zmiany demograficzne czy rosnące oczekiwania mieszkańców, wymagają od samorządów elastyczności i innowacyjności. Katowice konsekwentnie budują model administracji odpornej cyfrowo – zdolnej do reagowania na kryzysy, zapewniającej ciągłość działania i bezpieczeństwo danych.

Rozwój infrastruktury chmurowej, cyfrowych repozytoriów dokumentów i elektronicznego obiegu spraw pozwala na funkcjonowanie urzędu w trybie zdalnym, co szczególnie zyskało na znaczeniu w czasie pandemii. Dzięki temu administracja miejska może działać nieprzerwanie, nawet w warunkach ograniczonego dostępu do budynków urzędowych.

Otwartość na innowacje jest jednym z wyróżników Katowic. Miasto aktywnie współpracuje z sektorem IT, startupami i uczelniami, poszukując nowych rozwiązań usprawniających funkcjonowanie urzędu i poprawiających jakość życia mieszkańców. Wdrażane są programy pilotażowe w zakresie inteligentnych systemów miejskich, zrównoważonego transportu czy cyfrowej partycypacji społecznej.

Centrum Usług Wspólnych w Katowicach stało się symbolem profesjonalizacji i nowoczesnego podejścia do zarządzania administracją miejską. Dzięki centralizacji usług wspólnych udało się stworzyć strukturę, która nie tylko obsługuje jednostki miejskie, ale również kreuje nowe standardy jakości w sektorze publicznym.

CUW Katowice to nie tylko miejsce realizacji zadań technicznych – to instytucja, która inspirowa zmiany organizacyjne, wspiera wdrażanie innowacji i stanowi centrum kompetencji cyfrowych miasta. W przyszłości może stać się metropolitalnym hubem wiedzy i innowacji administracyjnych, łączącym potencjał samorządów, uczelni i przedsiębiorstw technologicznych.

Miasto Katowice pokazuje, że cyfrowa transformacja nie jest celem, ale procesem – dynamicznym, wymagającym i długofalowym. Dzięki odpowiedniemu przygotowaniu, konsekwentnej polityce i zaangażowaniu pracowników miasto z powodzeniem buduje wizerunek administracji nowoczesnej, przejrzystej i przyjaznej mieszkańcom.

Podsumowanie

Katowice należą dziś do grona najbardziej zaawansowanych cyfrowo samorządów w Polsce. Dzięki przemyślanej strategii i konsekwentnemu wdrażaniu nowoczesnych rozwiązań miasto stało się liderem cyfrowej administracji. Centrum Usług Wspólnych w Katowicach odgrywa w tym procesie rolę strategiczną – łączy technologię, kompetencje i wizję, tworząc solidny fundament dla dalszego rozwoju miasta. Cyfryzacja w Katowicach to nie tylko systemy informatyczne, ale także zmiana sposobu myślenia o administracji. To otwartość na dane, gotowość do automatyzacji, świadomość roli cyberbezpieczeństwa i inwestowanie w ludzi.

Przyszłość samorządu to współpraca, integracja i inteligentne wykorzystanie technologii. Katowice udowadniają, że cyfrowa administracja może być skuteczna, przejrzysta i bliska mieszkańcom – stanowiąc wzór dla innych miast w Polsce i w Europie.

Bibliografia

- Dacko-Pikiewicz, Z., Szczepańska-Woszczyna, K., & Lis, M. (red.). (2025). *Horyzonty sztucznej inteligencji a Przemysł 5.0*. Wydawnictwo Naukowe Akademii WSB w Dąbrowie Górniczej.
- Komisja Europejska (2024). *Europejska Dekada Cyfrowa 2030*.
- Modrzyński, P. (2024). *Digitalizacja Samorządowych Centrów Usług Wspólnych w Polsce*. Wolters Kluwer Polska.
- Modrzyński, P., Gawłowski, R., & Modrzyńska, J. (red.). (2018). *Samorządowe centra usług wspólnych. Założenia i praktyka*. C.H. Beck.
- Motowidlak, T., & Motowidlak, U. (2023). *Rola i efekty funkcjonowania centrów usług wspólnych w Polsce*. Seria „Ekonomia”. Wydawnictwo Uniwersytetu Łódzkiego.
- Ministerstwo Funduszy i Polityki Regionalnej (2023). *Program Fundusze Europejskie na Rozwój Cyfrowy 2021–2027*.
- Ministerstwo Cyfryzacji. (2022). *Strategia Cyberbezpieczeństwa RP 2022–2026*.
- OECD. (2023). *Digital Government Review of Poland*.
- Skrzyński, D. (2021). *Centra usług wspólnych, czyli zmiany w zadaniach z zakresu finansów publicznych na przełomie roku*. Wydawnictwo Wiedza i Praktyka / OnePress.

Akademia WSB
WSB University

Wydawnictwo Naukowe Akademii WSB
ul. Ciepłaka 1c, 41-300 Dąbrowa Górnicza
www.wsb.edu.pl

