

**Załącznik nr 3 do wniosku o przeprowadzenie postępowania habilitacyjnego**

**Autoreferat**

przedstawiający opis kariery zawodowej, dorobku oraz osiągnięć w pracy naukowo-badawczej

dr Jacek Maślankowski

Katedra Informatyki Ekonomicznej

Wydział Zarządzania

Uniwersytet Gdański

e-mail: [jacek.maslankowski@ug.edu.pl](mailto:jacek.maslankowski@ug.edu.pl)

## Spis treści

1. Imię i nazwisko.....	3
2. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej. ....	3
3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych.....	4
4. Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.).....	5
4.1. Główne osiągnięcie .....	5
4.2. Pozostałe osiągnięcia mające wpływ na rozwój dziedziny nauk społecznych w dyscyplinie nauk o zarządzaniu i jakości.....	30
5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej. ....	40
5.1. Współpraca w międzynarodowej sieci naukowej European Research Consortium for Information Systems (ERCIS), Niemcy .....	40
5.2. Współpraca z SNSPA, National University of Political Studies and Public Administration, Rumunia .....	42
5.3. Współpraca z TuE – Technical University of Eindhoven, Niderlandy .....	42
5.4. Inne instytucje naukowe .....	43
5.5. Konferencje naukowe.....	44
6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę.....	46
6.1. Osiągnięcia dydaktyczne.....	46
6.2. Osiągnięcia organizacyjne .....	52
6.3. Osiągnięcia popularyzujące naukę.....	56
7. Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej. ....	58

## Autoreferat

### 1. Imię i nazwisko.

Jacek Maślankowski

### 2. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej.

<b>Stopień doktora nauk ekonomicznych</b>	
Nadający stopień	Wydział Zarządzania Uniwersytetu Gdańskiego
Tytuł	Studium uwarunkowań wdrożeń hurtowni danych w organizacjach gospodarczych
Promotor	prof. dr hab. Stanisław Wrycza
Recenzenci	prof. dr hab. Mirosława Krystyna Lasek prof. dr hab. Bernard Franciszek Kubiak
Rok uzyskania	10 grudnia 2009 r.

<b>Studia magisterskie</b>	
Nazwa uczelni	Uniwersytet Gdański, Wydział Zarządzania
Kierunek	Informatyka i ekonometria
Rok uzyskania	czerwiec 2003 r.

### Certyfikaty międzynarodowe uprawniające do prowadzenia certyfikowanych zajęć

- Cisco Certified Networking Academy (CCNA)
- Cisco Security
- PCAP (Python Certified Associate Programming)

Poza tym liczne certyfikaty potwierdzające umiejętności, zdobywane w czasie kariery zawodowej na uczelni, m.in. wydawane przez Microsoft, Google czy SPSS, jak również językowe np. First Certificate in English.

### 3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych.

#### Zatrudnienie w Uniwersytecie Gdańskim

<i>Uniwersytet Gdański, Wydział Zarządzania, Katedra Informatyki Ekonomicznej</i>	
1.10.2003 r. – 31.01.2010 r.	Asystent
1.02.2010 r. – obecnie	Adiunkt

#### Zatrudnienie poza Uniwersytetem Gdańskim

<i>Uniwersytet Dolnośląski we Wrocławiu</i>	
Okres	od października 2023 r. (umowa cywilno-prawna)
Pełniona funkcja	wykładowca
Zakres obowiązków	Prowadzenie zajęć dydaktycznych z przedmiotu Programowanie w języku Python.

<i>Chongqing Technology and Business University, Chiny</i>	
Okres	marzec-czerwiec 2021 r. (umowa cywilno-prawna)
Pełniona funkcja	wykładowca
Zakres obowiązków	Prowadzenie zajęć z przedmiotu Big Data Analytics and Applications.

<i>Uniwersytet Warszawski w Warszawie</i>	
Okres	Okres: czerwiec-lipiec 2019 r. (umowa cywilno-prawna)
Pełniona funkcja	wykładowca
Zakres obowiązków	Zakres obowiązków: prowadzenie zajęć w języku angielskim z przedmiotów Databases oraz Data Visualization w ramach szkoły letniej Uniwersytetu Warszawskiego.

<i>Wyższa Szkoła Zarządzania w Gdańsku</i>	
Okres	październik 2004 r. – marzec 2007 r. (umowa cywilnoprawna)
Pełniona funkcja	wykładowca
Zakres obowiązków	Prowadzenie zajęć z przedmiotów dotyczących: języków programowania, sieci komputerowych, baz danych.

*Urlopy naukowe, macierzyńskie, rodzicielskie*

Nie dotyczy

4. **Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.).**

#### **4.1. Główne osiągnięcie**

##### **Tytuł osiągnięcia naukowego:**

Zastosowanie źródeł danych i metod big data w badaniach naukowych oraz opracowywaniu danych statystycznych

##### **Monografia naukowa:**

*Maślankowski J., (2025) Wykorzystanie metod big data do badania organizacji gospodarczych obecnych w Internecie. Teoria, metodologia, wyniki. Wydawnictwo Uniwersytetu Gdańskiego, 2025, ISBN: 978-83-8206-744-6*

##### **Monografię habilitacyjną recenzowali:**

dr hab. Maria Mach Król, Associate Professor VSB, Technical University of Ostrava

dr hab. Jakub Swacha, prof. Uniwersytetu Szczecińskiego, Uniwersytet Szczeciński

##### **Uzasadnienie wyboru problematyki badawczej**

Dynamiczny rozwój infrastruktury informatycznej, w szczególności w zakresie mocy obliczeniowej, architektur przetwarzania rozproszonego oraz dostępności narzędzi

analitycznych, doprowadził do zasadniczej zmiany możliwości pozyskiwania, przechowywania i analizy informacji cyfrowych. W konsekwencji możliwe stało się systematyczne przetwarzanie danych pochodzących z zasobów sieci Web w skali masowej, przy relatywnie niskich kosztach jednostkowych. Potencjał tego rodzaju danych w zastosowaniach biznesowych został zauważony podczas upowszechniania się i gwałtownego rozwoju Internetu w Polsce, m.in. przez Abramowicza, Kalczyńskiego oraz Węcla (2002). Zjawisko pozyskiwania danych z Internetu stworzyło nowe uwarunkowania dla prowadzenia badań empirycznych opartych na danych zastanych, w tym danych generowanych i udostępnianych przez organizacje za pośrednictwem ich stron internetowych, których potencjał badawczy pozostaje w istotnym stopniu niewykorzystany.

Wraz z dynamicznym rozwojem metod przetwarzania dużych zbiorów danych, dane powstałe zarówno w Internecie, jak również w procesie funkcjonowania firm, m.in. w mediach społecznościowych, serwisach internetowych czy sklepach internetowych, zostały zidentyfikowane jako istotne źródła danych w biznesie. W początkowym etapie rozwoju big data, pytania badawcze koncentrowały się wokół synergicznego efektu tradycyjnych systemów informatycznych oraz big data (Nycz, Półkowski, 2015), czy też konwergencji systemów przetwarzania w chmurze oraz big data (Wielki, 2016).

Przełomem w rozwoju big data było poszukiwanie możliwości wykorzystania tego rodzaju narzędzi. I tak, w literaturze naukowej związanej z big data uściślano teorię dotyczącą m.in. wykorzystania narzędzi i metod big data w zarządzaniu jednostkami służby zdrowia (Jelonek, Chluski, 2018), czy zastosowania logów z serwisów internetowych takich jak sklepy internetowe w celach marketingowych (Korczak, Pondel, 2017). Oceniano również potencjał mediów społecznościowych jako źródła informacji biznesowej (Chomiak-Orsa, Mrozek, 2018). Ocena źródeł danych typu big data w kontekście ich zastosowań w biznesie okazała się być jednym z wielu wyzwań naukowych (Gmiterek, 2019), jak również do oceny jakości tego rodzaju źródeł w kontekście wykorzystywania wyników badań w celach statystycznych (Szreder, 2019). Zaawansowane metody związane z przetwarzaniem danych big data, takie jak uczenie maszynowe, dostarczały możliwości m.in. badania ofert pracy (Lula, Oczkowska, Wiśniewska, Wójcik, 2018) czy predykcji bankructwa (Wyrobek, 2019) oraz niewypłacalności firm (Paliński, 2018). Należy zauważyć też, że artykuły podkreślające dynamiczny rozwój metod uczenia maszynowego, stosowały te algorytmy również do zbiorów danych powstających wewnątrz organizacji (Pospieszny, Czarnacka-Chrobot, Kobyliński 2018). Uzupełnieniem tego różnego rodzaju analiz były również te, które dotyczyły makroekonomicznych uwarunkowań, np. wykorzystania big data w celu monitorowania zagrożeń związanych z bezpieczeństwem narodowym (Stanik, Protasowicki, 2019).

Istotne jest to, że dane, metody oraz narzędzia powiązane z big data wymagają zastosowania modelu biznesowego dostosowanego do wymogów i potrzeb analitycznych firm (Mach-Król, Modrzejewska, 2017; Mach-Król, 2018), jak również zdefiniowania architektury systemowej na potrzeby przetwarzania wspomnianych wcześniej dużych zbiorów danych (Pańkowska, 2019). Oznacza to, że powstanie koncepcji big data, zbiorów danych i metod przetwarzania skutkowało również dynamicznym rozwojem badań naukowych, które traktują big data jako dane empiryczne. Należy podkreślić, że wymienione wcześniej przykłady

dotyczyły artykułów wydawanych w czasie największego rozwoju metod, narzędzi oraz danych big data, który miał miejsce w drugiej dekadzie XXI wieku.

Obecnie prace wokół danych webowych znanych jako big data koncentrują się na bardziej wyrafinowanych technikach ekstrakcji danych, wykorzystując między innymi techniki takie jak NLP (Natural Language Processing) czy LLM (Large Language Models) w celu pozyskania ustrukturyzowanych danych (Liu et al., 2024). Co więcej, stosuje się również modele integracji danych bazujące na ontologii (Ontology-Based Data Integration), które pozwalają na połączenie danych webowych z innymi źródłami danych, dzięki m.in. zastosowaniu semantycznej integracji danych (Masmoudi et al., 2024). Istotnym uzupełnieniem tych metod są badania na danych webowych wspierane grafami. Przykładem są badania interakcji użytkowników Internetu, np. w zakresie publikowanych ofert pracy online (Kwieciński, Górecki, Filipowska, 2022). Nieustannie autorzy artykułów naukowych wykorzystują metody web scrapingu w celu prowadzenia badań empirycznych, a przykładów nowych artykułów naukowych, również polskich autorów, bazujących na tego rodzaju danych jest wiele. Dotyczy to takich aspektów jak analiza rynku cen wybranych grup towarów (Juszczak, 2023), czy też wpływ ograniczania asymetrii informacji za pomocą danych webowych na ogólną efektywność rynku e-commerce (Hadasik, 2024). Przykłady te pokazują, że narzędzia do pobierania danych z Internetu znajdują i nadal będą znajdować swoje miejsce w badaniach empirycznych prowadzonych w środowisku naukowym. Jednocześnie warto zaznaczyć, że zgodnie z badaniem przeprowadzonym przez Gołuchowskiego, Jelonek i Węcla (2025) big data jest jednym z pojęć dotyczących technologii informatycznych stosowanych w informatyce ekonomicznej, które wpisują się w nauki o zarządzaniu i jakości.

Badania przedstawione w monografii stanowią odpowiedź na zidentyfikowaną lukę badawczą o charakterze zarówno metodologicznym, jak i empirycznym, polegającą na braku spójnego, zintegrowanego oraz skalowalnego podejścia umożliwiającego prowadzenie analiz organizacji gospodarczych na dużą skalę z wykorzystaniem danych pozyskiwanych z ich witryn internetowych. Pomimo istnienia bogatej literatury dotyczącej ekstrakcji danych z serwisów WWW, dotychczasowe opracowania koncentrują się głównie na wybranych etapach procesu analitycznego lub na eksperymentach o ograniczonym zakresie. W szczególności brak jest podejść, które integrowałyby pozyskiwanie danych, ich automatyczne przetwarzanie, analizę oraz ocenę jakości w ramach jednolitej, powtarzalnej procedury badawczej możliwej do zastosowania w skali masowej.

Oryginalność zaproponowanego podejścia polega na opracowaniu i empirycznej weryfikacji autorskiej metodologii umożliwiającej kompleksową analizę organizacji na podstawie informacji pochodzących z ich stron internetowych, z wykorzystaniem współczesnych metod analitycznych, w tym technik uczenia maszynowego. Metodologia ta obejmuje pełen cykl badawczy — od automatycznego pozyskiwania danych z zasobów Internetu, poprzez ich standaryzację, integrację i analizę, aż po systematyczną ocenę jakości uzyskiwanych wyników. W odróżnieniu od dotychczasowych badań, zaproponowane rozwiązanie zostało zaprojektowane z myślą o zastosowaniach na dużych, zróżnicowanych zbiorach danych, co umożliwia prowadzenie analiz o charakterze porównawczym i przekrojowym.

Istotnym elementem wkładu habilitacyjnego jest również zaproponowanie procedur obiektywnej oceny skuteczności zastosowanych metod analitycznych, umożliwiających porównywanie wyników uzyskiwanych w różnych wariantach modelowych oraz na różnych etapach przetwarzania danych. Wprowadzenie mierzalnych wskaźników jakości danych i wyników analiz stanowi istotne rozszerzenie dotychczasowych podejść, w których zagadnienia te były traktowane marginalnie lub miały charakter opisowy. Rozwiązanie to pozwala na replikowalność badań oraz tworzy podstawy do systematycznego doskonalenia stosowanych metod.

Dodatkowym aspektem nowości zaprezentowanej metodologii jest jej potencjał aplikacyjny w obszarach wymagających analiz prowadzonych w skali masowej, takich jak badania struktury i dynamiki rynków, analizy sektorowe, monitoring działalności organizacji czy rozwój nowoczesnych systemów statystyki publicznej opartych na alternatywnych źródłach danych. Zaproponowane podejście wpisuje się tym samym w aktualne kierunki rozwoju badań empirycznych wykorzystujących metody big data oraz stanowi istotny wkład w rozwój metodologii analiz danych internetowych w naukach ekonomicznych i społecznych.

W świetle aktualnego stanu badań monografia stanowi pierwsze opracowanie prezentujące wyniki analiz o porównywalnej skali, stopniu integracji metod oraz poziomie formalizacji procedur badawczych, realizowanych w oparciu o zaproponowaną autorską metodologię. Osiągnięte rezultaty mają charakter nie tylko poznawczy, lecz również metodologiczny, dostarczając narzędzi i procedur możliwych do adaptacji w dalszych badaniach empirycznych oraz w praktyce instytucjonalnej.

### **Cele i hipotezy**

Celem głównym monografii było zaprezentowanie autorskiej metody badań dotyczącej pozyskiwania i przetwarzania danych pochodzących z Internetu. Cele podrzędne dotyczyły zastosowania tej metody w praktyce w badaniu empirycznym, którego wyniki zaprezentowano również w pracy, jak również dokonanie przeglądu literatury w zakresie teoretycznych aspektów rozwoju metod i źródeł danych big data. W zakresie celu głównego monografii skoncentrowano się na obszarach związanych z wybranymi atrybutami przedsiębiorstw oraz ogólnie wykorzystaniem *big data* w organizacjach działających w Polsce.

Aby zrealizować cel główny, postawiono trzy pytania badawcze:

- PB1: Czy strona internetowa jest wystarczającym zasobem informacji pozwalającym na identyfikację rodzaju działalności organizacji, której dotyczy?
- PB2: Czy oferty pracy online pozwalają na identyfikację trendów w zakresie wykorzystania technologii *big data* w przedsiębiorstwach?
- PB3: Czy istnieje możliwość pozyskania bazy danych adresów URL w celu wiarygodnego badania polskich przedsiębiorstw?

Do każdego z pytań badawczych przyporządkowano następujące hipotezy:

- H1: Treść strony internetowej wybranej organizacji pozwala na określenie rodzaju jej działalności.

- H2: Badanie ofert pracy pozwala zidentyfikować trendy w zawodach powiązanych z *big data*.
- H3: Obecne źródła dostępne online pozwalają na budowę operatu do badań stron internetowych organizacji.

Sformułowane pytania badawcze oraz odpowiadające im hipotezy odnoszą się do fundamentalnych problemów metodologicznych i poznawczych w naukach o zarządzaniu, związanych z wykorzystaniem alternatywnych źródeł danych w badaniach organizacji. Ich istotność wynika z rosnącej roli danych cyfrowych generowanych przez przedsiębiorstwa w środowisku internetowym oraz z potrzeby rozwoju metod umożliwiających systematyczne i skalowalne analizy procesów organizacyjnych i rynkowych.

Pytanie badawcze PB1, dotyczące możliwości wykorzystania treści stron internetowych do identyfikacji rodzaju działalności organizacji, odnosi się bezpośrednio do problemu operacjonalizacji podstawowych cech organizacji w badaniach empirycznych. W naukach o zarządzaniu klasyfikacja podmiotów według profilu działalności stanowi punkt wyjścia do analiz strategii, konkurencyjności czy struktur rynkowych. Weryfikacja hipotezy H1, zakładającej, że treść strony internetowej umożliwia określenie rodzaju działalności organizacji, ma istotne znaczenie metodologiczne, gdyż pozwala ocenić, w jakim stopniu dane internetowe mogą zastępować lub uzupełniać tradycyjne źródła informacji, takie jak ankiety czy rejestry administracyjne. Wyniki w tym zakresie przyczyniają się do rozwoju nowych metod badań organizacji opartych na danych zastanych, ograniczających koszty i bariery dostępu do informacji.

Pytanie badawcze PB2 koncentruje się na możliwości wykorzystania ofert pracy publikowanych online jako źródła wiedzy o trendach technologicznych w przedsiębiorstwach, w szczególności w zakresie wykorzystania technologii *big data*. W naukach o zarządzaniu analiza kompetencji i zapotrzebowania na określone kwalifikacje jest kluczowa dla badań nad rozwojem organizacji, zarządzaniem zasobami ludzkimi oraz adaptacją przedsiębiorstw do zmian technologicznych. Weryfikacja hipotezy H2, zgodnie z którą analiza ofert pracy pozwala na identyfikację trendów w zawodach powiązanych z *big data*, wnosi nową perspektywę metodologiczną do badań nad transformacją cyfrową organizacji. Pozwala ona na prowadzenie analiz dynamicznych i prognostycznych, opartych na danych generowanych w sposób ciągły, co stanowi istotne rozszerzenie dotychczasowych podejść bazujących na badaniach z wykorzystaniem kwestionariuszy, które zwykle odnoszą się do okresu minionego, czyli nieaktualnych w danym momencie.

Pytanie badawcze PB3 dotyczy możliwości pozyskania wiarygodnej bazy adresów URL przedsiębiorstw w celu prowadzenia badań empirycznych na dużą skalę. Problem ten ma fundamentalne znaczenie dla rozwoju badań ilościowych w naukach o zarządzaniu, gdyż wiarygodny operat badawczy stanowi warunek poprawności metodologicznej analiz porównawczych i przekrojowych. Weryfikacja hipotezy H3, zakładającej, że dostępne źródła online umożliwiają budowę operatu do badań stron internetowych organizacji, przyczynia się do rozwiązania jednego z kluczowych ograniczeń badań opartych na danych internetowych, tj. problemu doboru próby i reprezentatywności. Tym samym pytanie to ma charakter

metodyczny, wykraczający poza pojedyncze zastosowanie empiryczne, i stanowi istotny wkład w rozwój infrastruktury badawczej nauk o zarządzaniu.

Łącznie sformułowane pytania badawcze i hipotezy tworzą spójny program badawczy, którego celem jest empiryczna weryfikacja możliwości wykorzystania danych internetowych jako pełnoprawnego źródła informacji w badaniach organizacji. Ich realizacja przyczynia się do rozwoju nauk o zarządzaniu poprzez rozszerzenie repertuaru metod badawczych, umożliwienie prowadzenia analiz w skali masowej oraz wzmocnienie podstaw metodologicznych badań nad funkcjonowaniem i rozwojem współczesnych organizacji w gospodarce cyfrowej.

### **Cele o charakterze poznawczym**

Cele o charakterze poznawczym obejmowały przede wszystkim przeprowadzenie pogłębionego przeglądu literatury przedmiotu. W ramach przygotowań do opracowania monografii dokonano analizy kilkuset publikacji naukowych dotyczących problematyki big data w kontekście ekonomicznym, organizacyjnym oraz metodologicznym. Spośród nich ponad 250 pozycji uznano za kluczowe dla realizacji celów badawczych i ujęto w końcowej bibliografii.

Ze względu na interdyscyplinarny oraz wielowymiarowy charakter zagadnienia big data, przegląd literatury naukowej został uporządkowany według kryterium treściowego. W rozdziale pierwszym przedstawiono genezę i ewolucję pojęcia big data, ze szczególnym uwzględnieniem jego roli w procesach gospodarczych i zarządzaniu informacją. Rozdział drugi poświęcono typologii zbiorów danych big data, w tym klasyfikacji danych według źródeł ich powstawania oraz charakteru informacyjnego, co pozwala na ich systematyzację w kontekście zastosowań ekonomicznych.

Rozdział trzeci koncentruje się na aspektach metodologicznych badań nad big data, obejmujących m.in. stosowane metody analityczne, modele biznesowe, dobór prób badawczych, architektury systemowe i procesowe, a także metody wykorzystywane w analizie danych, takie jak uczenie maszynowe czy techniki przetwarzania danych tekstowych. Analiza tych zagadnień stanowi zasadniczy cel poznawczy monografii.

W rozdziale czwartym dokonano przeglądu literatury dotyczącej problematyki jakości danych, w tym wpływu jakości informacji na procesy decyzyjne w przedsiębiorstwach oraz na efektywność analiz ekonomicznych. Z kolei rozdział piąty obejmuje analizę badań empirycznych, w których dane źródłowe pozyskiwane są w dominującym stopniu z Internetu, co pozwala na ocenę ich przydatności w badaniach gospodarczych oraz w procesach decyzyjnych opartych na danych.

### **Cele o charakterze metodycznym**

Cele o charakterze metodycznym zostały sformułowane przede wszystkim do utworzenia spójnej i kompleksowej metodyki dotyczącej badania stron internetowych w dużej skali.

Celem nadrzędnym o charakterze metodycznym monografii było opracowanie, empiryczne przetestowanie i metodologiczne ugruntowanie autorskiego podejścia badawczego do analizy organizacji gospodarczych na podstawie ich obecności w Internecie z wykorzystaniem metod big data, możliwego do zastosowania w naukach o zarządzaniu i jakości.

W ramach celu nadrzędnego przeprowadzono następujące prace o charakterze metodycznym:

- Opracowanie procedury systematycznego pozyskiwania i analizy danych ze stron internetowych organizacji gospodarczych, dostosowanej do wymogów badań empirycznych w naukach o zarządzaniu i jakości.
- Wypracowanie metody operacjonalizacji obecności w Internecie organizacji gospodarczych, umożliwiającej przejście od teoretycznej kategorii organizacji do empirycznej jednostki analizy opartej na danych internetowych.
- Zaprojektowanie i empiryczna weryfikacja procedury budowy operatu badawczego organizacji gospodarczych na podstawie publicznie dostępnych źródeł online.
- Opracowanie metody przetwarzania i normalizacji danych nieustrukturyzowanych, pozwalającej na ich wykorzystanie w analizach ilościowych i jakościowych.
- Adaptacja i integracja metod text mining oraz uczenia maszynowego do badań treści organizacyjnych w środowisku cyfrowym, z uwzględnieniem specyfiki nauk o zarządzaniu i jakości.
- Opracowanie zestawu mierników strukturalno-funkcjonalnych stron internetowych, umożliwiających empiryczną ocenę jakości i dojrzałości cyfrowej organizacji.
- Wypracowanie procedur walidacji danych i wyników analiz big data, zapewniających rzetelność, trafność i replikowalność badań empirycznych.
- Empiryczne przetestowanie spójności i użyteczności zaproponowanego podejścia metodologicznego w badaniu dużej liczby organizacji gospodarczych funkcjonujących w Internecie.
- Wskazanie możliwości replikacji i dalszego rozwoju zaproponowanych metod w przyszłych badaniach empirycznych w naukach o zarządzaniu i jakości.

### **Cele o charakterze empirycznym**

Celem nadrzędnym o charakterze empirycznym monografii było rozpoznanie cech organizacji gospodarczych funkcjonujących w środowisku cyfrowym na podstawie analizy danych pozyskanych z ich stron internetowych, z wykorzystaniem metod big data.

W ramach celu nadrzędnego przeprowadzono następujące prace:

- Empiryczna identyfikacja rodzaju działalności organizacji gospodarczych na podstawie treści ich stron internetowych.
- Analiza treści strony internetowej organizacji gospodarczych w celu identyfikacji dominujących tematów i wzorców komunikacyjnych.
- Empiryczna analiza atrybutów strukturalnych i funkcjonalnych stron internetowych organizacji gospodarczych.

- Identyfikacja zróżnicowania obecności w Internecie organizacji gospodarczych ze względu na wybrane cechy organizacyjne.
- Empiryczna analiza ofert pracy publikowanych przez organizacje gospodarcze w kontekście zapotrzebowania na kompetencje związane z obszarem big data.
- Identyfikacja trendów w zawodach i kompetencjach powiązanych z big data na podstawie treści ofert pracy.
- Empiryczna ocena możliwości budowy operatu badawczego organizacji gospodarczych na podstawie publicznie dostępnych źródeł online.

### **Cele dotyczące aplikacji rozwiązania**

Celem nadrzędnym o charakterze aplikacyjnym monografii było wskazanie możliwości praktycznego wykorzystania wyników badań empirycznych oraz zaproponowanego podejścia badawczego w analizie i ocenie funkcjonowania organizacji gospodarczych w Internecie.

W ramach celu nadrzędnego sformułowano następujące cele szczegółowe:

- Wskazanie możliwości wykorzystania wyników analizy treści stron internetowych do wspomaganie identyfikacji profilu działalności organizacji gospodarczych.
- Określenie potencjalnych zastosowań analizy obecności w Internecie organizacji w procesach zarządzania strategicznego i komunikacji z interesariuszami.
- Zaproponowanie sposobu wykorzystania analizy atrybutów stron internetowych jako narzędzia wspomagającego ocenę jakości i dojrzałości cyfrowej organizacji.
- Wskazanie możliwości zastosowania analizy ofert pracy do monitorowania zapotrzebowania na kompetencje i zawody związane z obszarem big data.
- Określenie potencjalnych zastosowań zaproponowanego operatu badawczego w dalszych badaniach empirycznych oraz analizach porównawczych organizacji gospodarczych.
- Zarysowanie możliwości wykorzystania zaproponowanego podejścia badawczego przez instytucje publiczne, organizacje branżowe oraz badaczy do analizy struktur i dynamiki gospodarki cyfrowej.

### **Źródła danych**

Badanie empiryczne przedstawione w monografii zostało oparte na połączeniu danych pierwotnych i wtórnych, pozyskiwanych z wielu komplementarnych źródeł, co umożliwiło łączenie danych oraz zwiększenie rzetelności i wiarygodności uzyskanych wyników.

#### **1. Dane pierwotne**

Dane pierwotne stanowią zasadniczą podstawę empiryczną badania i zostały pozyskane bezpośrednio przez autora w ramach zaprojektowanej procedury badawczej. Obejmują one w szczególności:

- treści tekstowe stron internetowych organizacji gospodarczych (np. opisy działalności, zakładki informacyjne),

- treści ofert pracy publikowanych na stronach internetowych organizacji,
- metadane oraz elementy strukturalne i funkcjonalne stron internetowych.

Dane pierwotne zostały zgromadzone w sposób systematyczny z wykorzystaniem metod automatycznego pozyskiwania danych i poddane dalszemu przetwarzaniu oraz analizie ilościowej i jakościowej.

## 2. Dane wtórne

Dane wtórne zostały wykorzystane w celu budowy operatu badawczego, rozszerzenia zakresu analiz oraz walidacji analiz empirycznych. Obejmują one:

- dane pochodzące z publicznych rejestrów organizacji gospodarczych (np. rejestry statystyczne i administracyjne),
- informacje z baz instytucjonalnych i branżowych,
- dane dostępne w publicznych zasobach internetowych (np. katalogi firm, serwisy informacyjne),
- dane pochodzące z literatury naukowej i raportów branżowych.

Dane wtórne umożliwiły identyfikację populacji badawczej, weryfikację poprawności klasyfikacji oraz interpretację uzyskanych wyników empirycznych.

## 3. Szereg i komplementarność źródeł danych

W badaniu wykorzystano szereg zróżnicowanych źródeł danych, zarówno o charakterze treściowym, jak i strukturalnym, co pozwoliło na:

- budowę empirycznego operatu badawczego organizacji gospodarczych obecnych w Internecie,
- porównanie danych pochodzących z różnych źródeł,
- przeprowadzenie łączenia danych i metod badawczych.

Zastosowanie wielu źródeł danych umożliwiło kompleksowe ujęcie obecności w Internecie organizacji gospodarczych oraz zwiększyło trafność wniosków formułowanych w monografii.

## **Weryfikacja hipotez**

H1: Treść strony internetowej wybranej organizacji pozwala na określenie rodzaju jej działalności.

Weryfikacja hipotezy została przeprowadzona na podstawie empirycznej analizy treści stron internetowych organizacji gospodarczych. Treści pozyskane w drodze automatycznego pozyskiwania danych zostały poddane przetwarzaniu, normalizacji oraz ilościowej analizie tekstu z wykorzystaniem metod text mining. Zastosowano metody klasyfikacji treści, umożliwiające przypisanie analizowanych stron do kategorii działalności organizacyjnej. Trafność klasyfikacji została poddana walidacji poprzez porównanie wyników analizy automatycznej z informacjami pochodzącymi z publicznych rejestrów oraz oceną ekspercką. Uzyskane rezultaty pozwoliły na ocenę, czy treść strony internetowej stanowi wystarczające źródło informacji do identyfikacji rodzaju działalności organizacji.

H2: Badanie ofert pracy pozwala zidentyfikować trendy w zawodach powiązanych z big data.

Hipoteza została zweryfikowana na podstawie analizy treści ofert pracy publikowanych na stronach internetowych organizacji gospodarczych. Oferty pracy zostały wyodrębnione z treści stron internetowych i poddane analizie ilościowej z wykorzystaniem metod text mining oraz technik grupowania i modelowania tematów. Analiza częstości występowania pojęć oraz identyfikacja wzorców tematycznych umożliwiły określenie dominujących kompetencji i stanowisk związanych z obszarem big data. Wyniki analiz pozwoliły na identyfikację trendów w strukturze zapotrzebowania na zawody i kompetencje w badanym obszarze.

H3: Obecne źródła dostępne online pozwalają na budowę operatu do badań stron internetowych organizacji.

Weryfikacja hipotezy została przeprowadzona poprzez empiryczne sprawdzenie możliwości identyfikacji i operacjonalizacji populacji badawczej organizacji gospodarczych na podstawie publicznie dostępnych źródeł online. Wykorzystano dane pochodzące z publicznych rejestrów, baz instytucjonalnych oraz zasobów internetowych w celu przypisania jednoznacznych adresów URL do jednostek organizacyjnych. Ocena kompletności, spójności i użyteczności uzyskanego operatu badawczego pozwoliła na weryfikację, czy dostępne źródła online umożliwiają budowę wiarygodnej ramy do badań empirycznych stron internetowych organizacji.

### **Zastosowane metody i narzędzia badawcze**

#### **1. Systematyczny i krytyczny przegląd literatury**

Dokonano pogłębionej syntezy dorobku naukowego dotyczącego badań organizacji gospodarczych w środowisku cyfrowym, integrując rozproszone podejścia teoretyczne z zakresu zarządzania, jakości oraz analizy danych big data. Osiągnięciem habilitacyjnym jest wypracowanie autorskiego ujęcia obecności cyfrowej organizacji jako kategorii analitycznej, osadzonej w naukach o zarządzaniu i jakości, oraz identyfikacja luki metodologicznej polegającej na niedostatecznym wykorzystaniu danych internetowych jako źródła badań empirycznych organizacji, co wynika przede wszystkim z braku spójnego metodologicznego podejścia do wykorzystania danych internetowych w badaniach naukowych w dyscyplinie nauk o zarządzaniu i jakości.

#### **2. Identyfikacja i operacjonalizacja populacji badawczej organizacji gospodarczych**

Opracowano oryginalną procedurę operacjonalizacji populacji badawczej organizacji gospodarczych na podstawie ich obecności w Internecie, umożliwiającą przejście od abstrakcyjnej kategorii organizacji do empirycznie obserwowalnej jednostki analizy. Osiągnięciem jest metodologiczne ugruntowanie stron internetowych jako reprezentatywnego nośnika cech organizacyjnych, co stanowi nowatorski wkład w badania ilościowe organizacji w warunkach transformacji cyfrowej.

#### **3. Automatyczne pozyskiwanie danych cyfrowych ze stron internetowych**

Zaprojektowano i zastosowano autorskie procedury automatycznego pozyskiwania danych empirycznych z wykorzystaniem metod znanych powszechnie jako web scraping. Osiągnięciem habilitacyjnym jest adaptacja technik charakterystycznych dla nauk

informatycznych do badań organizacyjnych, z zachowaniem rygorów metodologicznych właściwych dla nauk społecznych.

#### 4. Przetwarzanie, normalizacja i analiza treści cyfrowych

Opracowano spójny model przetwarzania i analizy treści cyfrowych organizacji gospodarczych, umożliwiający systematyczne badanie komunikacji organizacyjnej na dużą skalę. Osiągnięciem jest połączenie ilościowej analizy zawartości z jakościową interpretacją treści, co pozwoliło na uchwycenie zarówno strukturalnych, jak i znaczeniowych aspektów komunikacji organizacji w Internecie.

#### 5. Analiza strukturalno-funkcjonalnych atrybutów stron internetowych

Dokonano operacjonalizacji cech strukturalnych i funkcyjnych stron internetowych jako mierników dojrzałości organizacyjnej i jakościowej, co stanowi istotny wkład w rozwój badań nad zarządzaniem jakością w środowisku internetowym. Osiągnięciem habilitacyjnym jest empiryczne powiązanie cech technicznych stron internetowych z kategoriami zarządczymi, dotychczas rzadko analizowanymi w sposób systematyczny.

#### 6. Ilościowa analiza tekstu z wykorzystaniem metod text mining i uczenia maszynowego

Zastosowano i zaadoptowano zaawansowane metody text mining oraz uczenia maszynowego (nadzorowanego i nienadzorowanego) do analizy treści organizacyjnych, wykazując ich użyteczność w badaniach nad strategiami, orientacjami i praktykami zarządczymi. Osiągnięciem jest wprowadzenie skalowalnych metod analizy danych nieustrukturyzowanych do głównego nurtu badań empirycznych w naukach o zarządzaniu i jakości.

#### 7. Walidacja danych, procedur i wyników badań

Zaprojektowano i zastosowano wieloetapowy proces walidacji danych i wyników, obejmujący kontrolę jakości danych oraz testowanie stabilności modeli analitycznych. Osiągnięciem habilitacyjnym jest podniesienie standardów rzetelności badań big data w naukach o zarządzaniu i jakości, w obszarze często krytykowanym za nadmierny determinizm techniczny.

#### 8. Empiryczne badanie organizacji gospodarczych na podstawie ich obecności cyfrowej

Całość procedur badawczych składa się na autorską koncepcję empirycznego badania organizacji gospodarczych na podstawie ich obecności cyfrowej, stanowiącą samodzielne osiągnięcie naukowe. Osiągnięciem habilitacyjnym jest wykazanie, że strony internetowe mogą być traktowane jako pełnoprawne źródło danych empirycznych w naukach o zarządzaniu i jakości, umożliwiające formułowanie oryginalnych wniosków teoretycznych i aplikacyjnych.

W ramach w/w metod badawczych wykorzystano cały szereg narzędzi, zarówno informatycznych, statystycznych, jak i baz danych, zaprezentowanych w tabeli poniżej. Z perspektywy technicznej, głównie wykorzystywano środowisko data science z wiodącym językiem Python, jak również bazy relacyjne SQL oraz nierelacyjne NoSQL.

**Tabela 1. Zastosowane w monografii metody i narzędzia badawcze**

<b>Lp.</b>	<b>Zastosowana metoda badawcza</b>	<b>Narzędzia badawcze</b>
1.	Systematyczny i krytyczny przegląd literatury	Bazy publikacji naukowych, autorskie arkusze analizy literatury
2.	Identyfikacja i operacjonalizacja populacji badawczej	Publiczne rejestry organizacji gospodarczych, bazy instytucjonalne, procedury standaryzacji i deduplikacji danych
3.	Automatyczne pozyskiwanie danych z Internetu	Narzędzia automatycznego pozyskiwania danych ze stron internetowych, autorskie procedury pozyskiwania danych
4.	Przetwarzanie i analiza treści stron internetowych	Narzędzia przetwarzania danych tekstowych, procedury normalizacji językowej, metody analizy treści
5.	Analiza atrybutów stron internetowych	Narzędzia analizy struktury i funkcjonalności stron, autorskie arkusze oceny atrybutów
6.	Text mining i uczenie maszynowe	Środowiska analizy danych, algorytmy analizy tekstu, modele klasyfikacyjne i grupujące
7.	Walidacja danych i wyników analiz	Procedury walidacji krzyżowej, miary jakości modeli, łączenie różnych metod
8.	Empiryczne badanie organizacji gospodarczych	Zintegrowany zestaw narzędzi analitycznych i baz danych

### **Struktura monografii**

Opracowanie stanowi kompleksowe studium poświęcone tematyce big data, od podstaw teoretycznych, przez aspekty metodologiczne i jakościowe, aż po praktyczne zastosowanie w badaniach naukowych i analizie stron internetowych organizacji w Polsce.

Wprowadzenie przedstawia cel pracy, jej znaczenie i kontekst badawczy, przygotowując grunt pod dalsze rozdziały. W rozdziale pierwszym omówiono pojęcie big data, jego rozwój oraz zastosowania w praktyce gospodarczej i naukowej. Zawiera on także krytykę podejść do przetwarzania dużych zbiorów danych oraz wskazuje na strony internetowe jako jedno z ważnych źródeł danych biznesowych.

Rozdział drugi klasyfikuje dane big data według różnych kryteriów, tj. źródła pochodzenia, rodzaju oraz stopnia ustrukturyzowania. Celem tej części jest uporządkowanie pojęć i stworzenie podstaw dla dalszej analizy metod przetwarzania danych.

W rozdziale trzecim skupiono się na metodologii badań big data, obejmującej zarówno sposoby analizy danych, jak i wdrażanie rozwiązań opartych na danych. Omówiono tu metody statystyczne, uczenie maszynowe, web mining (wraz z technikami takimi jak TF-IDF czy modele przestrzeni wektorowych), a także kwestie próbkowania, łączenia zbiorów danych i zarządzania ich zmianami. Rozdział ten stanowi trzon teoretyczno-metodologiczny pracy.

Rozdział czwarty dotyczy zagadnień jakości danych big data. Opisano tam stosowane podejścia i mierniki jakości, wymiary oceny danych oraz zaproponowano syntetyczne metody oceny dużych zbiorów danych.

W piątym rozdziale przedstawiono zastosowania big data w badaniach naukowych opartych na danych z Internetu. Zawiera on przegląd ośrodków naukowych zajmujących się big data, przykłady badań w różnych dziedzinach (od edukacji i rynku pracy po finanse i środowisko) oraz omówienie wyzwań naukowych wynikających z pracy z danymi masowymi.

Rozdział szósty opisuje autorskie badanie stron internetowych polskich organizacji, od metodologii pozyskiwania i przetwarzania danych, przez dobór próby, aż po walidację i ocenę jakości uzyskanych informacji. Stanowi on praktyczne zastosowanie omówionych wcześniej koncepcji i metod big data.

W rozdziale siódmym zaprezentowano rezultaty badania, analizując m.in. cechy stron internetowych, ich język, obecność w mediach społecznościowych, odniesienia do technologii big data oraz oferty pracy związane z tą tematyką. Uwzględniono również wykorzystanie metod uczenia maszynowego i klasteryzacji w analizie danych. Rozdział kończy się dyskusją o wyzwaniach prawnych i metodologicznych w badaniach opartych na danych internetowych oraz podsumowaniem wyników.

Całość zamyka zakończenie, w którym autor syntetyzuje wnioski z pracy, podkreśla znaczenie big data w badaniach naukowych i praktyce gospodarczej oraz wskazuje możliwe kierunki dalszych badań.

### **Szczegółowy spis treści monografii**

Wprowadzenie
Rozdział 1. Pojęcie i zastosowania big data
1.1. Przegląd literatury naukowej
1.1.1. Koncepcja i rozwój big data
1.1.2. Podstawowe definicje
1.1.3. Zastosowanie big data w praktyce
1.2. Rozwój systemów informatycznych wspierających przetwarzanie dużych zbiorów danych
1.3. Krytyka podejść związanych ze stosowaniem big data
1.4. Strony internetowe jako źródło danych biznesowych
Rozdział 2. Typologia danych big data

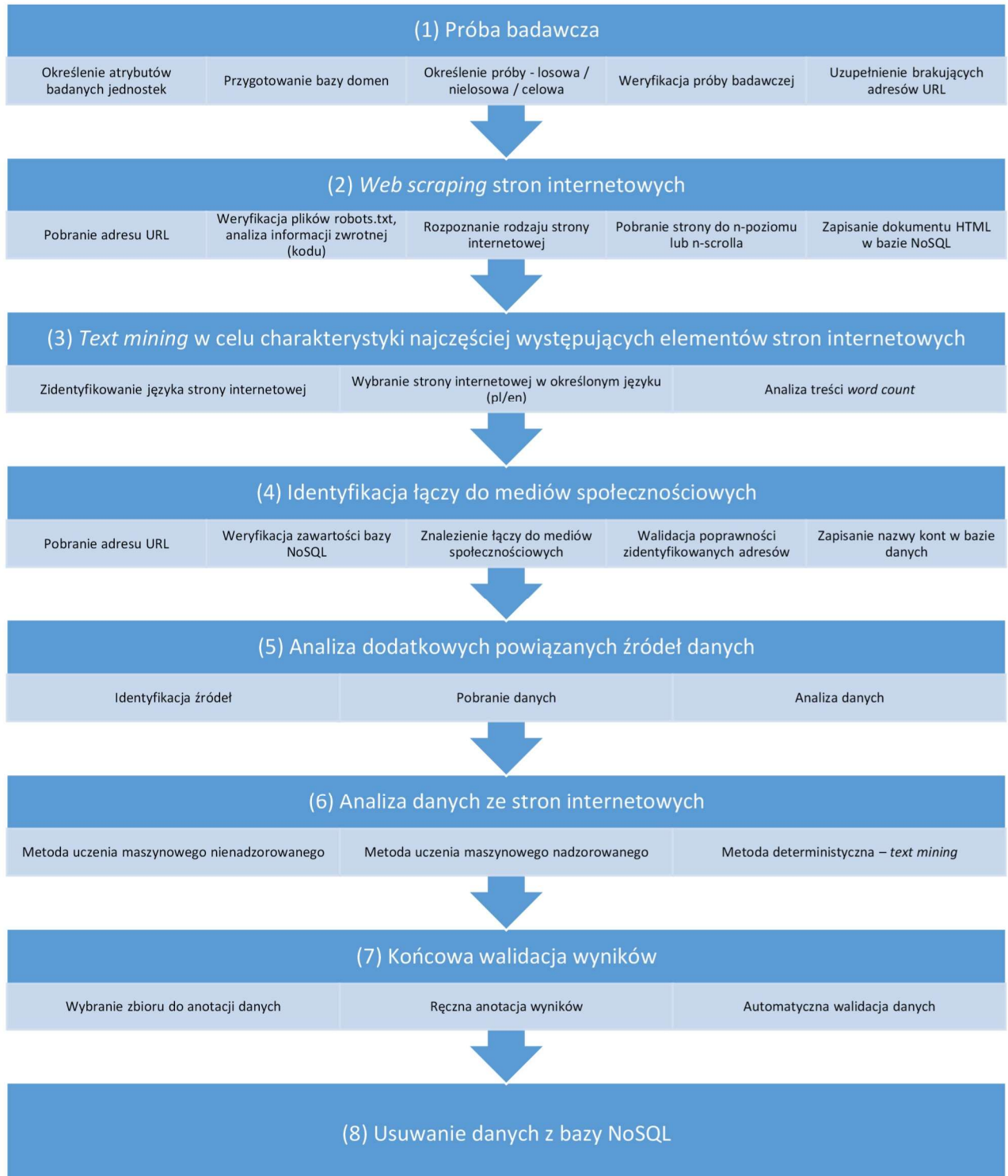
2.1. Podział danych
2.2. Rodzaje danych ze względu na źródło powstawania
2.3. Stopień ustrukturyzowania danych
Rozdział 3. Metodologiczne aspekty badań big data
3.1. Wybrane metody związane z analizą dużych zbiorów danych
3.2. Metody wdrożeń rozwiązań wykorzystujących źródła big data
3.3. Model biznesowy sterowany danymi
3.4. Próba badawcza
3.4.1. Ogólne uwagi dotyczące próbkowania danych
3.4.2. Próby nielosowe
3.5. Architektura systemowa i procesowa
3.6. Metody statystyczne
3.7. Metody uczenia maszynowego
3.8. Metody web mining
3.8.1. Rodzaje i zastosowanie
3.8.2. Model Bag-of-words
3.8.3. Przykłady Vector Space: Document Term Matrix oraz Term Document Matrix
3.8.4. Metoda TF-IDF
3.9. Łączenie zbiorów danych
3.10. Zarządzanie zmianami w źródłach danych
3.11. Badania tradycyjne a big data
Rozdział 4. Zagadnienia jakości danych typu big data
4.1. Przegląd stosowanych podejść do oceny jakości danych
4.2. Wymiary jakości danych
4.3. Syntetyczne ujęcie mierników jakości danych
4.5. Sugerowane podejście do oceny jakości danych typu big data
Rozdział 5. Big data w badaniach naukowych bazujących na danych z Internetu
5.1. Big data a dziedziny nauki
5.2. Big data w środowisku naukowym
5.2.1. Ośrodki naukowe big data
5.2.2. Przykłady zastosowań big data według dziedzin nauki
5.3. Przegląd badań korzystających z metod i źródeł big data
5.3.1. Badania edukacyjne
5.3.2. Badania rynku pracy
5.3.3. Badania serwisów społecznościowych
5.3.4. Badania językowe
5.3.5. Badania trendów
5.3.6. Badania społeczne
5.3.7. Badania finansowe
5.3.8. Badania środowiska
5.3.9. Badania przedsiębiorstw
5.4. Wyzwania w obszarze nauki związane z big data

<b>Rozdział 6. Badanie stron internetowych organizacji działających w Polsce</b>
6.1. Metodyka badania 6.1.1. Wprowadzenie 6.1.2. Wyszukiwanie adresów stron internetowych 6.1.3. Pobieranie zawartości stron internetowych 6.1.4. Identyfikacja kont w mediach społecznościowych 6.1.5. Pobieranie danych dotyczących ofert pracy 6.1.6. Przetwarzanie i analiza danych ze stron internetowych 6.1.7. Sugerowana metodyka prowadzenia badań na podstawie stron internetowych 6.2. Opis badania 6.2.1. Próba badania 6.2.2. Hipotezy badawcze 6.2.3. Charakterystyka czasowa i przestrzenna 6.3. Walidacja danych 6.4. Ocena jakości danych
<b>Rozdział 7. Rezultaty badania</b>
7.1. Podział wniosków badania 7.2. Zainteresowanie terminem big data na badanym obszarze 7.3. Wielkość jednostek obecnych w Internecie oraz rodzaj ich działalności 7.4. Atrybuty stron internetowych firm 7.4.1. Obecność pliku robots.txt 7.4.2. Język stron internetowych 7.4.3. Analiza treści strony na podstawie słów kluczowych 7.4.4. Analiza łączy do mediów społecznościowych 7.4.5. Analiza ofert pracy związanych z big data 7.4.6. Podział stron internetowych na klastry z wykorzystaniem uczenia nienadzorowanego 7.4.7. Rozpoznawanie innowacyjnej działalności z wykorzystaniem uczenia maszynowego nadzorowanego 7.4.8. Odniesienie do technologii big data na stronach internetowych 7.5. Metody i wyniki walidacji danych 7.5.1. Adresy URL 7.5.2. Obecność pliku robots.txt 7.5.3. Język stron internetowych 7.5.4. Analiza treści strony na podstawie słów kluczowych 7.5.5. Analiza łączy do mediów społecznościowych 7.5.6. Analiza ofert pracy związanych z big data 7.5.7. Podział stron internetowych na klastry z wykorzystaniem uczenia nienadzorowanego 7.5.8. Rozpoznawanie innowacyjnej działalności z wykorzystaniem uczenia maszynowego nadzorowanego 7.5.9. Odniesienie do technologii big data na stronach internetowych 7.6. Dyskusja dotycząca wyzwań związanych z prowadzeniem badań na podstawie źródeł big data w postaci stron internetowych 7.6.1. Próba badawcza 7.6.2. Prawne aspekty pozyskiwania danych z Internetu 7.7. Podsumowanie wyników badań
Zakończenie

## Główne wnioski wynikające z badania przedsiębiorstw na podstawie stron internetowych

W pracy zaproponowano następującą metodykę badania związaną z pozyskiwaniem informacji ze stron internetowych przedsiębiorstw, zgodnie z zaprezentowanym poniżej rysunkiem.

**Rysunek 1. Proponowana metodyka prowadzenia badań empirycznych z wykorzystaniem danych typu big data jakimi są strony internetowe**



Źródło: (Maślankowski 2025).

Opracowana metodyka stanowi autorskie, zintegrowane podejście badawcze do analizy cyfrowej obecności organizacji, traktowanej jako istotny element zarządzania strategicznego, komunikacji organizacyjnej oraz relacji z interesariuszami. Metodyka odpowiada na rosnącą potrzebę stosowania skalowalnych i replikowalnych narzędzi empirycznych w naukach o zarządzaniu i jakości, umożliwiających analizę rzeczywistych praktyk organizacyjnych ujawnianych w środowisku cyfrowym.

### 1. Projektowanie próby badawczej organizacji

Pierwszy etap metodyki obejmuje konstrukcję próby badawczej organizacji, z uwzględnieniem celów badania oraz założeń teoretycznych właściwych dla nauk o zarządzaniu. Na tym etapie definiowane są atrybuty badanych jednostek, takie jak sektor działalności, wielkość organizacji, forma własności, zasięg działania czy kontekst instytucjonalny. Atrybuty te umożliwiają późniejszą analizę porównawczą i segmentacyjną.

Następnie tworzona jest baza domen internetowych organizacji, stanowiąca empiryczny odpowiednik próby badawczej. Dobór próby może mieć charakter losowy, nielosowy lub celowy, w zależności od problemu badawczego (np. analiza najlepszych praktyk, badanie liderów rynkowych, studia porównawcze sektorów). Każdorazowo przeprowadzana jest weryfikacja próby pod kątem dostępności i aktualności stron internetowych oraz uzupełnienie brakujących adresów URL, co zapewnia spójność i kompletność zbioru danych.

### 2. Pozyskiwanie danych ze stron internetowych jako źródła wiedzy o praktykach zarządczych

Drugi etap polega na automatycznym pozyskiwaniu danych ze stron internetowych organizacji, traktowanych jako oficjalny kanał komunikacji strategicznej z interesariuszami. Proces rozpoczyna się od weryfikacji dostępności strony oraz zgodności działań badawczych z zasadami etycznymi i technicznymi (analiza plików robots.txt czy kodów odpowiedzi serwera).

Kolejnym krokiem jest identyfikacja rodzaju strony internetowej oraz pobranie jej zawartości do określonego poziomu strukturalnego lub liczby przewinięć. Pozyskane dokumenty HTML zapisywane są w bazie danych typu NoSQL, co umożliwia elastyczne zarządzanie danymi nieustrukturyzowanymi i ich dalszą analizę w kontekście procesów zarządczych i komunikacyjnych.

### 3. Analiza treści stron internetowych w kontekście komunikacji organizacyjnej

Na trzecim etapie stosowane są metody eksploracji tekstu w celu identyfikacji dominujących treści i przekazów komunikacyjnych organizacji. Automatyczna identyfikacja języka strony umożliwia selekcję materiału badawczego w określonym języku, co jest istotne w badaniach międzynarodowych i porównawczych.

Analiza częstości występowania słów i fraz pozwala na identyfikację kluczowych obszarów komunikacji (np. orientacja na klienta, innowacyjność, zrównoważony rozwój, odpowiedzialność społeczna), stanowiących empiryczne odzwierciedlenie priorytetów strategicznych organizacji.

#### 4. Identyfikacja relacji organizacji z interesariuszami w mediach społecznościowych

Czwarty etap obejmuje analizę powiązań pomiędzy stronami internetowymi organizacji a ich obecnością w mediach społecznościowych. Wyszukiwanie i walidacja łączy do profili organizacji gospodarczych w mediach społecznościowych, umożliwia identyfikację kanałów dialogu z interesariuszami oraz ocenę stopnia integracji komunikacji cyfrowej.

Zapisanie nazw kont i adresów profili w bazie danych pozwala na dalsze analizy dotyczące spójności komunikacji międzykanałowej, intensywności obecności cyfrowej oraz strategii zarządzania relacjami z interesariuszami.

#### 5. Integracja dodatkowych źródeł danych w analizie zarządczej

W piątym etapie metodyki następuje identyfikacja i analiza dodatkowych źródeł danych powiązanych z badanymi organizacjami, takich jak dane rejestrowe, raporty niefinansowe, dane branżowe czy otwarte bazy statystyczne. W badaniu empirycznym opisywanym w monografii dodatkowym źródłem danych były strony internetowe z ofertami pracy, co pozwoliło zbadać obecne trendy w zakresie big data w polskich organizacjach. Integracja dodatkowych danych umożliwia pogłębioną analizę zależności pomiędzy deklarowanym na stronie internetowej charakterem organizacji a jej kontekstem instytucjonalnym i rynkowym.

#### 6. Analiza danych z wykorzystaniem metod analitycznych stosowanych w naukach o zarządzaniu

Szósty etap obejmuje zastosowanie metod analizy danych właściwych dla współczesnych badań w naukach o zarządzaniu. Metody uczenia maszynowego nienadzorowanego wykorzystywane są do identyfikacji typologii organizacji oraz wzorców komunikacyjnych, natomiast metody nadzorowane umożliwiają klasyfikację treści zgodnie z przyjętymi konstruktami teoretycznymi (np. orientacja strategiczna, dojrzałość komunikacyjna).

Uzupełnieniem są deterministyczne metody eksploracji tekstu, oparte na regułach i słownikach pojęć ze słowami kluczowymi, które zwiększają interpretowalność wyników i ich użyteczność teoretyczną.

#### 7. Walidacja wyników i zapewnienie rzetelności badań

Siódmy etap metodyki poświęcony jest walidacji wyników. Obejmuje on wybór podzbioru danych do ręcznej anotacji eksperckiej, co pozwala na ocenę trafności i rzetelności wyników uzyskanych metodami automatycznymi. Następnie przeprowadzana jest automatyczna walidacja danych z wykorzystaniem miar jakości, co zapewnia zgodność metodyki z rygorami metodologicznymi nauk o zarządzaniu.

#### 8. Zarządzanie danymi i odpowiedzialność badawcza

Ostatni etap dotyczy zarządzania cyklem życia danych badawczych, w tym usuwania danych z bazy NoSQL po zakończeniu analiz. Etap ten realizowany jest zgodnie z zasadami odpowiedzialności badawczej, ochrony danych oraz transparentności procesu badawczego, co ma szczególne znaczenie w badaniach empirycznych dotyczących organizacji.

Proponowana metodyka została przetestowana i zwalidowana poprzez badanie prowadzone cyklicznie na potrzeby niniejszej monografii. Główne badanie zostało przeprowadzone trzykrotnie, rok po roku i obejmowało od 366 do 390 tys. stron internetowych. Do badania zostały włączone wszystkie jednostki działające w Polsce, dla których zidentyfikowano adres strony internetowej. W praktyce oznacza to, że liczba jednostek została znacząco ograniczona i wykluczono z niej większość małych przedsiębiorstw. Jednak to najczęściej duże jednostki posiadają zasoby i warunki, aby wspierać się technologiami big data.

Pomocnicze badanie na podstawie dodatkowych źródeł danych obejmowało analizę ofert pracy znajdujących się w portalach publikujących oferty pracy oraz analizę dostępności stron internetowych na podstawie pliku wskazującego na możliwość pozyskania danych przy pomocy robota skrapującego.

## **Wyniki badania**

Głównym rezultatem przedstawionego osiągnięcia naukowego jest opracowanie i empiryczna weryfikacja autorskiej metodyki prowadzenia badań z wykorzystaniem stron internetowych organizacji. Metodyka ta obejmuje pełny cykl badawczy, tj. od pozyskania danych ze źródeł internetowych, poprzez ich przetwarzanie, analizę i walidację, aż po udostępnianie wyników. Szczegółowy opis tej metodyki został przedstawiony we wcześniejszej części autoreferatu.

Istotnym rezultatem badań jest również uporządkowanie pojęciowe i typologiczne zbiorów danych typu big data. W monografii przeprowadzono analizę porównawczą różnych sposobów klasyfikacji takich zbiorów, funkcjonujących w literaturze przedmiotu. Szczególną uwagę poświęcono klasyfikacji umożliwiającej rozróżnienie danych tworzonych przez człowieka, danych generowanych w procesach biznesowych oraz danych powstających w sposób automatyczny (maszynowy). Typologia ta stanowi punkt odniesienia dla dalszych analiz empirycznych oraz pozwala na lepsze osadzenie badań internetowych w kontekście nauk o zarządzaniu.

Kolejnym rezultatem monografii jest syntetyczna analiza literatury dotyczącej dostępności metod analizy danych typu big data oraz ich zastosowań w praktyce organizacyjnej. W pracy zaprezentowano kilkanaście metod analitycznych wspierających przetwarzanie i analizę dużych zbiorów danych, ze szczególnym uwzględnieniem ich przydatności w badaniach nad funkcjonowaniem organizacji.

Do istotnych wyników należy również analiza porównawcza literatury w zakresie metod oceny jakości zbiorów danych typu big data. Uzupełnieniem tej analizy jest propozycja procedur walidacji danych pochodzących z Internetu, które zostały następnie wykorzystane i zweryfikowane empirycznie w dalszej części monografii. Zaproponowane podejście umożliwi bardziej obiektywną ocenę wiarygodności danych internetowych oraz skuteczności stosowanych metod analitycznych.

Ważnym elementem monografii jest także przegląd zastosowań źródeł danych typu big data w badaniach naukowych, który pozwala na osadzenie zaproponowanej metodyki w szerszym kontekście rozwoju współczesnych metod empirycznych w naukach o zarządzaniu i jakości.

Centralną część monografii stanowi jednak badanie empiryczne, przeprowadzone z wykorzystaniem zaproponowanej metodyki. Badanie to umożliwiło sformułowanie wniosków dotyczących charakterystyki przedsiębiorstw, instytucji oraz organizacji posiadających własne strony internetowe, funkcjonujących na terenie Polski. Uzyskane wyniki zostały uporządkowane w kilka zasadniczych obszarów analitycznych, obejmujących:

- ogólną charakterystykę organizacji gospodarczych posiadających strony internetowe na podstawie danych pochodzących z rejestrów i baz administracyjnych,
- analizę zawartości stron internetowych organizacji gospodarczych, w tym odniesień do technologii big data,
- analizę ogłoszeń o pracę publikowanych w formie elektronicznej,
- ocenę obecności organizacji gospodarczych w mediach społecznościowych oraz możliwości ich charakterystyki na podstawie treści publikowanych na stronach internetowych,
- grupowanie stron internetowych organizacji według dominujących treści.

Przeprowadzone w monografii analizy pozwoliły stwierdzić, że zainteresowanie terminem big data utrzymywało się na względnie stabilnym poziomie, przy jednoczesnym wyraźnym trendzie wzrostowym w latach 2012-2016. Wnioski te zostały sformułowane na podstawie łączenia danych pochodzących z różnych źródeł, w tym analizy zapytań w wyszukiwarkach internetowych, treści publikowanych w mediach społecznościowych, zawartości stron internetowych oraz ofert pracy zamieszczanych w Internecie. W przeciwieństwie do narzędzi takich jak Google Trends, które jedynie pośrednio odzwierciedlają zainteresowanie danym terminem, zaprezentowane badanie umożliwia bardziej rzetelną ocenę faktycznego wykorzystania technologii big data w organizacjach.

W ramach badania właściwego, tj. na stronach internetowych organizacji gospodarczych działających w Polsce, zidentyfikowano 503,7 tys. stron internetowych powiązanych z ponad 2,05 mln przedsiębiorstw. Ze względu na niekompletność informacji w dostępnych bazach danych, w szczególności dotyczących liczby zatrudnionych, nie było możliwe pełne określenie dostępności stron internetowych dla całego operatu badawczego według kryterium wielkości zatrudnienia. Analiza obejmowała jednak zbiór około 1,5 mln przedsiębiorstw, dla których takie dane były dostępne.

Uzyskane wyniki wskazują na istotne zróżnicowanie dostępności stron internetowych w zależności od wielkości przedsiębiorstwa. Strony WWW posiadało 95,7% dużych przedsiębiorstw, 88,7% przedsiębiorstw średnich, 61,6% małych firm oraz jedynie 21,3% mikroprzedsiębiorstw. Wyniki te pozostają zbieżne z danymi statystyki publicznej, co potwierdza trafność przyjętego operatu badawczego.

Analiza treści dostępnych na stronach internetowych umożliwiła wstępną klasyfikację działalności przedsiębiorstw na podstawie występowania określonych słów kluczowych. Zastosowane mechanizmy czyszczenia i walidacji danych pozwalają na efektywne identyfikowanie określonych typów działalności, takich jak handel elektroniczny czy sprzedaż internetowa.

Istotnym elementem badania była również analiza wykorzystania mediów społecznościowych przez przedsiębiorstwa. Informacje te zostały pozyskane na podstawie analizy kodu źródłowego stron internetowych, co umożliwiło identyfikację i ilościową charakterystykę profili organizacji w mediach społecznościowych.

W zakresie analizy ofert pracy związanych z technologiami big data przeprowadzono badanie obejmujące dane zbierane w granulacji dziennej przez okres pięciu lat. Łącznie pozyskano blisko 42,8 mln ofert pracy, które po usunięciu duplikatów ograniczono do 8,8 mln unikatowych ogłoszeń. Analiza opierała się na tytułach ofert pracy, co pozwoliło na ograniczenie wpływu nadmiernie rozbudowanych i często deklaratywnych opisów wymagań.

Kolejnym etapem badania była identyfikacja grup stron internetowych organizacji z wykorzystaniem metod uczenia maszynowego nienadzorowanego. Zastosowanie algorytmu k-średnich umożliwiło wyodrębnienie klastrów stron internetowych na podstawie treści stron głównych, przy czym kryterium podziału stanowiły słowa kluczowe najlepiej charakteryzujące poszczególne grupy.

W dalszej części badania określono skalę wykorzystania innowacyjnych rozwiązań w przedsiębiorstwach. Analiza ta została przeprowadzona zarówno z wykorzystaniem metod uczenia maszynowego nadzorowanego, opartych na zbiorach treningowych, jak i metod deterministycznych, bazujących na zestawach słów kluczowych (tzw. golden records). Zastosowane podejścia umożliwiły identyfikację przedsiębiorstw wykorzystujących technologie big data, sztuczną inteligencję oraz inne innowacyjne rozwiązania.

## **Wkład autora w rozwój dziedziny nauk społecznych w dyscyplinie nauk o zarządzaniu i jakości**

### **1. Wymiar teoretyczny**

Wkład teoretyczny monografii polega na rozszerzeniu sposobu konceptualizacji organizacji oraz jej otoczenia informacyjnego w naukach o zarządzaniu. Praca wprowadza perspektywę, w której strona internetowa organizacji traktowana jest jako trwały, empirycznie obserwowalny element zarządzania, odzwierciedlający strategię, priorytety komunikacyjne oraz relacje z interesariuszami.

W tym ujęciu:

- komunikacja cyfrowa organizacji staje się źródłem danych umożliwiających analizę praktyk zarządczych poza deklaratywnymi formami znanymi z badań ankietowych czy dokumentów formalnych,
- możliwe jest prowadzenie analiz porównawczych i przekrojowych dużych populacji organizacji, co poszerza dotychczasowe ramy badań empirycznych w zarządzaniu,
- praca przyczynia się do rozwoju teorii zarządzania poprzez empiryczne osadzenie takich elementów jak dojrzałość organizacyjna czy zarządzanie relacjami z interesariuszami, tj. odbiorcami informacji dostępnej na stronie internetowej.

Tym samym monografia wzmacnia nurt badań opartych na danych obserwacyjnych w naukach o zarządzaniu, oferując alternatywę lub uzupełnienie dla klasycznych metod deklaratywnych.

## **2. Wymiar metodyczny**

Najistotniejszy wkład pracy ujawnia się w wymiarze metodycznym. Na potrzeby monografii opracowano autorską, zintegrowaną metodologię analizy organizacji z wykorzystaniem danych pochodzących z witryn internetowych, obejmującą pełen cykl badawczy: od pozyskiwania danych, przez ich standaryzację i integrację, aż po analizę oraz ocenę jakości wyników. Należy podkreślić, że metodyka ta jest rezultatem ponad 10-letniego doświadczenia autora niniejszego wniosku w badaniach naukowych i projektach bazujących na metodach i źródłach big data.

Wkład metodyczny obejmuje w szczególności:

- Opracowanie spójnego i replikowalnego schematu badawczego, umożliwiającego prowadzenie analiz w skali masowej, co stanowi istotne rozwinięcie dotychczasowych, często fragmentarycznych podejść stosowanych w badaniach stron internetowych organizacji.
- Rozszerzenie metod ekstrakcji danych z Internetu o perspektywę analiz porównawczych i przekrojowych, pozwalających na systematyczne badanie dużych i zróżnicowanych populacji organizacji.
- Integrację procesów pozyskiwania danych z zaawansowanymi technikami analitycznymi, w tym metodami uczenia maszynowego nadzorowanego i nienadzorowanego, co umożliwi automatyczną identyfikację oraz klasyfikację atrybutów organizacji na podstawie treści publikowanych w Internecie.
- Wprowadzenie procedur obiektywnej oceny jakości danych i skuteczności metod analitycznych, opartych na mierzalnych wskaźnikach, co zwiększa transparentność procesu badawczego oraz umożliwia porównywalność i replikację wyników.

Metodyka ta stanowi istotne uzupełnienie warsztatu badawczego nauk o zarządzaniu, odpowiadając na rosnące znaczenie danych typu big data oraz potrzebę ich rzetelnej analizy.

## **3. Wymiar empiryczny**

Wkład empiryczny monografii polega na szeroko zakrojonej weryfikacji zaproponowanej metodologii w badaniach obejmujących zbiory danych o skali i zróżnicowaniu istotnie przekraczających zakres analiz prezentowanych w dotychczasowej literaturze przedmiotu.

Empiryczna część pracy:

- potwierdza użyteczność i skalowalność zaproponowanych rozwiązań metodologicznych,
- umożliwia identyfikację wzorców i zróżnicowania praktyk komunikacyjnych organizacji związanych z treścią udostępnianą na stronie internetowej, w różnych kontekstach sektorowych i instytucjonalnych,
- dostarcza wyników, które nie byłyby możliwe do uzyskania przy użyciu tradycyjnych metod empirycznych.

Dodatkowo, praca wskazuje na szerokie możliwości aplikacyjne zaproponowanej metodologii, zarówno w badaniach akademickich, jak i w praktyce instytucjonalnej, w tym w analizach rynkowych, badaniach sektorowych oraz rozwoju nowoczesnych systemów statystyki publicznej opartych na alternatywnych źródłach danych.

Całościowo przedstawiony dorobek habilitacyjny wnosi trwały wkład w rozwój nauk o zarządzaniu i jakości, łącząc nową perspektywę teoretyczną, oryginalne rozwiązania metodyczne oraz ich empiryczną weryfikację w skali masowej. Wkład ten wzmacnia pozycję badań opartych na danych internetowych jako pełnoprawnego nurtu empirycznego w dyscyplinie.

#### **4. Uszczegółowienie wkładu w rozwój dziedziny nauk społecznych w dyscyplinie nauk o zarządzaniu i jakości**

Opracowana autorska metodyka badań stron internetowych organizacji wnosi istotny i oryginalny wkład do dyscypliny nauki o zarządzaniu i jakości poprzez wprowadzenie spójnego, wieloetapowego podejścia empirycznego, umożliwiającego systematyczną analizę cyfrowej komunikacji organizacji jako przejawu ich praktyk zarządczych i strategicznych. Metodyka odpowiada na lukę badawczą polegającą na braku wystandaryzowanych, skalowalnych i replikowalnych procedur analizy stron internetowych w badaniach zarządzania, które dotychczas były najczęściej analizowane w sposób fragmentaryczny, jakościowy lub ograniczony do studiów przypadków.

##### **1. Nowe ujęcie strony internetowej jako empirycznego artefaktu zarządzania**

Pierwszym elementem oryginalności metodyki jest konceptualizacja strony internetowej organizacji jako empirycznego artefaktu zarządzania, odzwierciedlającego zarówno deklarowane strategie, jak i rzeczywiste praktyki komunikacyjne oraz relacyjne organizacji. W przeciwieństwie do dominujących w literaturze podejść traktujących strony internetowe głównie jako narzędzie marketingowe, zaproponowana metodyka umożliwia ich analizę jako źródła danych o orientacji strategicznej, dojrzałości zarządczej oraz sposobie zarządzania relacjami z interesariuszami.

##### **2. Integracja metod informatycznych z metodologią badań zarządzania**

Istotnym wkładem metodycznym jest integracja zaawansowanych technik informatycznych (web scraping, eksploracja tekstu, uczenie maszynowe) z uwarunkowaniami metodologicznymi właściwymi dla nauk o zarządzaniu i jakości. Metodyka nie ogranicza się do zastosowania narzędzi analitycznych, lecz osadza je w logicznej strukturze badawczej, obejmującej projektowanie próby, walidację danych oraz łączenie różnych źródeł danych.

Dzięki temu możliwe jest prowadzenie badań na dużych zbiorach danych empirycznych, przy jednoczesnym zachowaniu kontroli nad jakością i interpretowalnością wyników, co stanowi istotny krok w kierunku rozwoju badań ilościowych opartych na danych cyfrowych w zarządzaniu.

##### **3. Autorska procedura konstrukcji i weryfikacji próby badawczej organizacji**

Nowość metodyczna przejawia się również w opracowaniu autorskiej procedury konstrukcji próby badawczej organizacji na podstawie ich oficjalnych domen internetowych. Procedura ta umożliwia elastyczne dostosowanie sposobu doboru próby (losowy, celowy, nielosowy) do problemu badawczego, przy jednoczesnym zapewnieniu porównywalności wyników między badaniami.

Wprowadzenie etapu systematycznej weryfikacji i uzupełniania adresów URL (Uniform Resource Locator) zwiększa rzetelność badań empirycznych, eliminując częsty w badaniach bazujących na danych internetowych, problem niekompletności i niereprezentatywności danych.

#### 4. Standaryzacja analizy treści stron internetowych organizacji

Kolejnym elementem wkładu metodycznego jest standaryzacja analizy treści stron internetowych poprzez połączenie deterministycznych metod eksploracji tekstu z metodami uczenia maszynowego nadzorowanego i nienadzorowanego. Takie podejście umożliwia jednocześnie wykrywanie wzorców komunikacyjnych oraz ich interpretację w odniesieniu do kategorii teoretycznych funkcjonujących w naukach o zarządzaniu (np. orientacja strategiczna, odpowiedzialność społeczna, innowacyjność).

W rezultacie metodyka pozwala na przejście od opisowej analizy treści do systematycznego wnioskowania empirycznego, co zwiększa jej wartość poznawczą i aplikacyjną.

#### 5. Metodyczna analiza relacji organizacji z interesariuszami w środowisku transformacji cyfrowej

Oryginalnym elementem metodyki jest również systematyczna analiza powiązań pomiędzy stronami internetowymi organizacji a ich obecnością w mediach społecznościowych. Ujęcie to umożliwia empiryczną analizę strategii zarządzania relacjami z interesariuszami w warunkach transformacji cyfrowej, wykraczając poza deklaratywne ujęcia znane z raportów czy ankiet.

Metodyka tworzy podstawy do badań nad spójnością komunikacji międzykanałowej oraz stopniem integracji narzędzi cyfrowych w zarządzaniu organizacją.

#### 6. Procedura walidacji wyników dostosowana do badań cyfrowych w zarządzaniu

Istotnym wkładem metodycznym jest także opracowanie procedury walidacji wyników, łączącej ręczną anotację ekspercką z automatyczną walidacją danych. Takie podejście zwiększa wiarygodność wyników uzyskiwanych z wykorzystaniem metod obliczeniowych i odpowiada na krytykę dotyczącą „czarnej skrzynki” algorytmów w badaniach społecznych i zarządczych.

#### 7. Znaczenie metodyki dla rozwoju dyscypliny

Całościowo autorska metodyka przyczynia się do rozwoju nauk o zarządzaniu i jakości poprzez:

- poszerzenie zasobu metod empirycznych o skalowalne analizy danych internetowych,
- umożliwienie prowadzenia badań porównawczych w długich szeregach czasowych,
- zwiększenie replikowalności i transparentności badań empirycznych,
- utworzenie pomostu metodologicznego między badaniami jakościowymi a ilościowymi w zarządzaniu.

Wkład ten ma charakter zarówno teoretyczny, jak i aplikacyjny, dostarczając narzędzi do analizy rzeczywistych praktyk zarządczych ujawnianych w środowisku cyfrowym.

## Bibliografia

- Abramowicz, W., Kalczyński, P., Węcel, K. (2002). *Filtering the Web to Feed Data Warehouses*. Springer.
- Chomiak-Orsa, I., Mrozek, B. (2017). Główne perspektywy wykorzystania Big Data w mediach społecznościowych. *Informatyka Ekonomiczna / Uniwersytet Ekonomiczny we Wrocławiu*, 3(45), 44-54. <https://doi.org/10.15611/ie.2017.3.04>.
- Gmiterek, G. (2019). Challenges Related to Identifying Sources and Document Collection for Big Data Analyses, *e-mentor*, 3(80), 4-9.
- Gołuchowski, J., Jelonek, D., Węcel, K. (2025) Miejsce informatyki ekonomicznej w przestrzeni badawczej nauk o zarządzaniu i jakości, *Przegląd Organizacji / Organization Review / Nr 1(1004)*, s. 21-32, DOI: 10.33141/po.2025.01.02
- Hadasik, B. (2024). Reduction of information asymmetry in e-commerce: the web scraping approach. *Wyd. UE Katowice*, DOI: 10.22367/9788378758907.
- Jelonek, D., Chluski, A. (2018). The use of Big Data Resources in Patient Experience Management. *Zeszyty Naukowe. Organizacja i Zarządzanie / Politechnika Śląska*, 117, 199-212. 10.29119/1641-3466.2018.117.13.
- Juszczak, A. (2023). The use of web-scraped data to analyse the dynamics of clothing and footwear prices. *Wiadomości Statystyczne. The Polish Statistician*, 68(9), 15–33. <https://doi.org/10.59139/ws.2023.09.2>.
- Korczak, J., Pondel, M. (2017). Metodyczne podejście do analizy i eksploracji danych marketingowych. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 342, 52-71.
- Kwieciński, R., Górecki, T., & Filipowska, A. (2022). Learning edge importance in bipartite graph-based recommendations. In *Proceedings of the 17th Conference on Computer Science and Intelligence Systems (FedCSIS)* (pp. 227–233). <https://doi.org/10.15439/2022F191>
- Lula, P., Oczkowska, R., Wiśniewska, S., & Wójcik, K. (2018). Ontology-based system for automatic analysis of job offers. *Information Technology for Practice*, 2018, 205-212
- Liu, P., Gao, W., Dong, W., Ai, L., Gong, Z., Huang, S., Li, Z., Hoque, E., Hirschberg, J., & Zhang, Y. (2024). A survey on open information extraction from rule-based model to large language model. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 9586–9608). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.findings-emnlp.560>
- Mach-Król, M. (2018). Recepcja Temporalnego Modelu Dojrzałości do Big Data wśród menedżerów – badania wstępne. *Przegląd Organizacji*, 4(939), 27–34. <https://doi.org/10.33141/po.2018.04.04>.
- Mach-Król, M., Modrzejewska, D. (2017). Potrzeby Analityczne Polskich Firm a Big Data. *Business Informatics / Informatyka Ekonomiczna*, 2(44), 82-93. <https://doi.org/10.15611/ie.2017.2.07>.
- Masmoudi, M., Ben Abdallah Ben Lamine, S., Karray, M. H., Archimede, B., & Baazaoui Zghal, H. (2024). Semantic data integration and querying: A survey and challenges. *ACM Computing Surveys*, 56(8), Article 209. <https://doi.org/10.1145/3653317>

- Nycz, M., Pólkowski, Z. (2015). Big Data a Business Intelligence. *Prace Naukowe / Uniwersytet Ekonomiczny w Katowicach, Systemy wspomagania organizacji SWO* 2015, 196-205.
- Paliński, A. (2018). Metody uczenia maszynowego w prognozowaniu niewypłacalności, *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 358, 173-181.
- Pańkowska, M. (2019). Architektura systemowa dla modelowania i zarządzania danymi masowymi [w:] M. Pańkowska (red.), *Procesy modelowania danych i podejmowania decyzji w architekturze systemowej przedsiębiorstwa*, 100-118. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2018). An effective approach for software project effort and duration estimation with machine learning algorithms. *Journal of Systems and Software*, 137, 184–196. <https://doi.org/10.1016/j.jss.2017.11.066>
- Protasowicki, T., Stanik, J., 2016. Big data w analizie zagrożeń bezpieczeństwa narodowego. *Ekonomiczne Problemy Usług*, 123, 275-286. <https://doi.org/10.18276/epu.2016.123-26>.
- Szreder, M. (2019). Istotność statystyczna w czasach big data. *Wiadomości Statystyczne, The Polish Statistician*, 64(11), 42-57.
- Wielki, J. (2016). Analiza szans, możliwości oraz wyzwań związanych z wykorzystaniem Cloud Computingu i Big Data jako technologii konwergentnych. *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, 50(2), 163-173.
- Wyrobek, J. (2019). Predicting Bankruptcy at Polish Companies: A Comparison of Selected Machine Learning and Deep Learning Algorithms. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie / Cracow Review of Economics and Management*, 6(978), 41-60. <https://doi.org/10.15678/ZNUEK.2018.0978.0603>.

#### **4.2. Pozostałe osiągnięcia mające wpływ na rozwój dziedziny nauk społecznych w dyscyplinie nauk o zarządzaniu i jakości**

Jako inne osiągnięcia naukowe wybrano publikacje powiązane z zastosowaniem metod big data, ściśle związane tematycznie z monografią stanowiącą główne osiągnięcie naukowe. Osiągnięcia te zostały uporządkowane w trzy główne obszary tematyczne:

1. Jakość danych big data.
2. Pozyskiwanie informacji z Internetu.
3. Analiza danych z mediów społecznościowych.

W pierwszym obszarze, dotyczącym jakości danych big data, wykazano publikacje z lat 2014–2015. Stanowią one jedno z pierwszych opracowań w literaturze krajowej podejmujących problematykę jakości danych masowych. Prace te stworzyły podstawy dla kolejnych badań związanych zarówno z przetwarzaniem danych pochodzących z Internetu, jak i analizą danych generowanych w mediach społecznościowych.

W drugim obszarze tematycznym ujęto publikacje z lat 2019–2023, w których rozwinięto zagadnienia stanowiące istotny wpływ na metodykę proponowaną w monografii.

Dotyczą one przede wszystkim wyzwań metodologicznych i praktycznych związanych z przetwarzaniem oraz wykorzystaniem informacji pozyskiwanej z Internetu, a także ograniczeń i problemów towarzyszących tego typu analizom.

Trzeci obszar obejmuje publikacje poświęcone bezpośrednio pozyskiwaniu i przetwarzaniu danych z Internetu, w szczególności z mediów społecznościowych. Przedstawiono w nich wyniki badań empirycznych oraz procesy przetwarzania i integracji tego rodzaju danych, co pozostaje w ścisłym związku z problematyką omówioną w monografii.

Należy podkreślić, że autor wniosku zaprezentował jedynie wybrane publikacje dotyczące big data, które w sposób najbardziej reprezentatywny odzwierciedlają kontekst tematyki ujęty w monografii. Dorobek autora w obszarze big data obejmuje ponad 50 publikacji naukowych i eksperckich, w tym raportów przygotowanych dla UNECE, Komisji Europejskiej oraz OECD, z czego ponad 20 stanowią artykuły naukowe dotyczące ściśle badań z wykorzystaniem zbiorów danych big data.

### **Jakość danych big data**

Badania jakości danych są niezwykle istotne z perspektywy wdrażania rozwiązań big data, które uważane są za słabej jakości źródła danych. Z drugiej strony mierniki dotyczące jakości danych big data są często trudne do zdefiniowania. Stąd w początkowych etapach prac związanych z big data, celem było zaproponowanie kilku podejść związanych z oceną jakości danych. Przede wszystkim starano się zbudować szablon, na podstawie którego byłaby możliwa ocena wybranych źródeł danych big data. Wkład w rozwój dyscypliny w tym przypadku stanowi próba wprowadzenia precyzyjnych mierników, z których możliwy stałby się wybór zestawu do badania wybranego źródła, np. stron internetowych z ofertami pracy. Trzeba zwrócić uwagę, że większość artykułów opisująca jakość danych big data koncentruje się na ogólnych wytycznych, takich jak reprezentatywność, czy precyzja, nie podając szczegółowo w jaki sposób można dokonać pomiaru na różnych źródłach danych big data. Stąd próba doprecyzowania wskaźników z mojej strony przyczyniła się do wzrostu świadomości w tym obszarze. Poniższe artykuły mają obecnie ponad 40 cytowań w bazie Google Scholar.

<b>Tytuł publikacji</b>
Maślankowski, J. (2014) Data quality issues concerning statistical data gathering supported by big data technology, International Conference: Beyond Databases, Architectures and Structures, 92-101
Maślankowski, J., (2015) Analiza jakości danych pozyskiwanych ze stron internetowych z wykorzystaniem rozwiązań big data, Roczniki Kolegium Analiz Ekonomicznych 38, str. 167-177, Wydawnictwo SGH.
Maślankowski J. (2015), Big Data quality issues regarding unstructured data analysis. A case study, In: Information systems development and applications / Wrycza Stanisław (eds.), 2015, Sopot, Faculty of Management University of Gdańsk, pp.79-89, ISBN 978-83-64669-06-4

W ramach tego nurtu badawczego postawiono następującą hipotezę: rozwiązania big data pozwalają na uzyskanie danych dobrej jakości pod warunkiem prawidłowego przygotowania algorytmu analizy składni stron internetowych, które mają zasilać bazę danych wynikowych.

Bardzo często problem jakości danych typu big data ujmuje się w sposób bardzo ogólny, określając przykładowo, że ocenie ma zostać poddana reprezentatywność danych, spójność, czy linkowalność. Niewiele jest opracowań, które dostarczają szczegółowych mierników, którymi można się posłużyć oceniając jakość danych. W poniższej tabeli zaprezentowano przykład ujętych w publikacji (Maślankowski, 2015) mierników, służących do oceny jakości danych big data.

**Tabela 2. Proponowane mierniki oceny jakości danych webowych jako źródła big data**

Lp.	Nazwa miernika i sposób reprezentacji
1.	Jednoznaczność danych: Jednoznaczność odnosi się do stopnia, w jakim obserwacje (rekordy) mogą być interpretowane w sposób jednoznaczny. Metryka ta pozwala określić, jaki odsetek rekordów jest niejednoznaczny i może prowadzić do wielorakiej interpretacji wyników. Niska jednoznaczność może skutkować błędną klasyfikacją lub interpretacją danych, co ogranicza rzetelność analiz empirycznych.
2.	Obiektywizm przetwarzania: Obiektywizm danych mierzy, w jakim stopniu rekordy mogą być niewłaściwie odwzorowane w procesie przetwarzania. Metryka ta identyfikuje potencjalne źródła błędów algorytmicznego lub subiektywnego, które mogą wpływać na spójność i bezstronność wyników.
3.	Zawieranie znacznika czasowego: Metryka ta określa, jaki odsetek obserwacji może być jednoznacznie przyporządkowany do określonego znacznika czasowego. Pozwala to na analizę danych w ujęciu chronologicznym i umożliwia identyfikację trendów oraz dynamiki zmian w badanym zjawisku.
4.	Granularność lub stopień szczegółowości danych: Granularność odnosi się do możliwości przyporządkowania obserwacji do różnych poziomów hierarchii, np. administracyjnej (województwo, powiat, gmina) lub funkcjonalnej (dział, projekt, jednostka organizacyjna). Wyższa granularność zwiększa precyzję analiz, umożliwiając wnioskowanie na różnych poziomach szczegółowości.
5.	Występowanie duplikatów danych: Duplikaty mogą zaburzać analizy statystyczne i prowadzić do przeszacowania częstości występowania zjawisk. Metryka ta określa liczbę powtarzających się rekordów zidentyfikowanych w procesie przetwarzania oraz umożliwia ocenę ryzyka związanej z nadmiernym powielaniem danych.
6.	Kompletność danych: Kompletność mierzy, jaki odsetek obserwacji lub rekordów typu klucz-wartość został odrzucony w procesie przetwarzania z powodu braków danych. Niska kompletność może ograniczać reprezentatywność danych i wpływać na jakość wniosków analitycznych.

Lp.	Nazwa miernika i sposób reprezentacji
7.	Dostępność danych: Metryka dostępności określa, jaki odsetek obserwacji został pozyskany ze źródła bez występowania błędów lub ostrzeżeń generowanych przez algorytmy przetwarzające dane. Wysoka dostępność jest warunkiem uzyskania rzetelnych i powtarzalnych wyników badawczych.
8.	Zgodność z przyjętym wzorcem: Zgodność z wzorcem ocenia, jaki odsetek obserwacji może zostać odzwierciedlony jako klucz-wartość zgodnie z ustalonym schematem lub wyrażeniem regularnym. Brak zgodności może wskazywać na konieczność wprowadzenia dodatkowego przetwarzania lub normalizacji danych.
9.	Klasyfikacja i kategoryzacja: Metryka ta określa, jaki odsetek obserwacji nie może zostać przypisany do przyjętych kategorii lub klasyfikacji. Niezdolność do prawidłowej kategoryzacji ogranicza możliwość agregacji danych i interpretacji wyników w kontekście ustalonej ontologii badawczej.
10.	Precyzja dopasowania: Precyzja odnosi się do utraty informacji w procesie dopasowywania obserwacji do wzorca lub schematu. Niska precyzja może skutkować częściową utratą treści rekordów, co ogranicza szczegółowość i wierność danych w analizie.
11.	Interpretowalność danych: Interpretowalność mierzy, ile rekordów może zostać błędnie zinterpretowanych w wyniku niejednoznaczności lub niewłaściwego mapowania. Niska interpretowalność utrudnia prawidłowe wnioskowanie oraz ogranicza możliwość stosowania danych w badaniach naukowych.
12.	Integralność danych: Integralność określa, jaki odsetek obserwacji ulega modyfikacjom w procesie przetwarzania. Zachowanie integralności danych jest kluczowe dla zapewnienia spójności historycznej i wiarygodności analiz.
13.	Spójność danych: Spójność odnosi się do liczby obserwacji, które zostały przekształcone lub poddane konwersji w trakcie procesu przetwarzania. Zachowanie spójności semantycznej i strukturalnej jest niezbędne do zapewnienia jakości analiz i porównywalności wyników.

Aby zweryfikować powyższe mierniki, posłużono się trzema przypadkami dotyczącymi przetwarzania danych big data pozyskiwanych ze stron internetowych, zgodnie z tabelą poniżej.

**Tabela 3. Przeprowadzone w artykule studia przypadków dla oceny jakości danych**

Lp.	Opis	Charakterystyka
1	analiza wszystkich słów kluczowych	duże raporty analityczne, przypadkowość wyników
2	analiza ściśle określonych słów kluczowych	analitika dobrana do potrzeb, możliwość pominięcia istotnych informacji
3	analiza powiązań pomiędzy słowami kluczowymi i innymi atrybutami	szerszy obraz analiz, możliwość znalezienia nieznanymi dotychczas powiązań

Przeprowadzona analiza potwierdza przyjętą wcześniej hipotezę, że istnieje możliwość uzyskania wyników dobrej jakości na podstawie informacji nieustrukturyzowanej,

przetwarzanej z zastosowaniem rozwiązań big data. Należy jednak dobrać źródła danych w sposób umożliwiający pomiar jakości danych. Ważne jest uzyskanie satysfakcjonujących wartości mierników zaprezentowanych w niniejszym opracowaniu w tabelach 3 i 4. Poza wymienionymi miernikami, istotnym zagadnieniem związanym z wyborem źródła danych stanowi jego reprezentacyjność i kompletność. Tylko w taki sposób istnieje możliwość uzyskania wiarygodnych wyników. Wpływ na wiarygodność danych będzie miał w tym przypadku przede wszystkim operat badania (Maślankowski, 2015).

Artykuły te pozwoliły na późniejsze zaangażowanie autora wniosku, we wdrożenie rozwiązań oceny jakości danych w licznych projektach europejskich, o których mowa w dalszej części wniosku. Przykładem jest współautorstwo raportów Komisji Europejskiej, takich jak:

<b>Tytuł raportu</b>
Maślankowski, J., Salgado, D., Quaresma, S., Ascari G., Brancato, G., Consiglio, L., Righi, P, Tuoto, T., Daas, P., Six, M., Kowarik, A., (2019) Deliverable K1:First Draft of the Quality Guidelines for the Acquisition and Usage of Big Data, Work package K, Deliverable K.1. Eurostat, European Commission.
Consten, A., Chavdarov, V., Daas, P., Horvat, V., Maślankowski, J., Quaresma, S., Six, M., Tuoto, T. (2018) Report describing the quality aspects of big data for official statistics. ESSnet Big Data, Deliverable 8.2. Eurostat, European Commission.
Consten, A., Chavdarov, V., Daas, P., Horvat, V., Maślankowski, J., Quaresma, S., Six, M., Tuoto, T. (2018) Report describing the methodology of using Big Data for official statistics and the most important questions for future studies. ESSnet Big Data I, Work package 8, Deliverable 8.4. Eurostat, European Commission.
Stateva, G., ten Bosch, O., Maślankowski, J., Righi, A., Scannapieco, M., Greenaway, M., Swier, N., Jansson, I., Wu, D. (2017) Legal aspects related to Web scraping of Enterprise Web Sites. ESSnet Big Data I, Work package 2, Deliverable 2.1. Eurostat, European Commission.
Stateva, G., ten Bosch, O., Maślankowski, J., Barcaroli, G., Scannapieco, M., Greenaway, M., Jansson, I., Wu, D. (2017) Methodological and IT Issues and Solutions. ESSnet Big Data I, Work package 2, Deliverable 2.4. Eurostat, European Commission.

Potwierdza to praktyczny wymiar prowadzonych działań i możliwość jego zastosowania w wielu różnych przypadkach pozyskiwania danych z Internetu.

### **Pozyskiwanie informacji z Internetu**

Podstawowym zagadnieniem podjętym w monografii, stanowiącej główny wkład autora niniejszego wniosku w rozwój dyscypliny nauk o zarządzaniu i jakości, jest problematyka pozyskiwania danych z Internetu. Autor od wielu lat prowadzi badania w tym obszarze, czego odzwierciedleniem jest szereg publikacji dotyczących przede wszystkim empirycznych analiz big data. Wśród nich szczególne znaczenie mają dwie publikacje, które autor wskazuje jako kluczowe z punktu widzenia dorobku naukowego. Koncentrują się one na metodologicznych aspektach pozyskiwania danych internetowych oraz na możliwościach ich praktycznego wykorzystania, między innymi w analizach statystycznych.

### Tytuł publikacji

Maślankowski, J., (2019) Pozyskiwanie i analiza danych na temat ofert pracy z wykorzystaniem big data, Wiadomości Statystyczne. The Polish Statistician 64 (9), 60–74.

Daas, P., Maślankowski, J. (2023). Current challenges and possible big data solutions for the use of web data as a source for official statistics. Wiadomości Statystyczne. The Polish Statistician, 68(12), 49–64. <https://doi.org/10.59139/ws.2023.12.3>.

Celem pierwszego artykułu było przedstawienie możliwości oraz korzyści wynikających z zastosowania narzędzi automatycznego pozyskiwania danych dotyczących ofert pracy zamieszczanych na stronach internetowych, zaliczanych do źródeł typu big data, na potrzeby statystyki publicznej. W publikacji omówiono również wyzwania związane z wykorzystaniem tego rodzaju danych w procesie statystycznym.

Dla zobrazowania przedstawianych zagadnień zaprezentowano wyniki eksperymentalnego studium przypadku obejmującego pozyskiwanie i przetwarzanie danych internetowych z wykorzystaniem metod web scraping oraz text mining. Celem analizy było pozyskanie informacji dotyczących ofert pracy publikowanych w Internecie, z przeznaczeniem do zastosowania w badaniach statystyki rynku pracy. Badaniem objęto lata 2017 i 2018.

Przedstawione prace wpisują się w szerszy nurt działań prowadzonych ukierunkowanych na wykorzystanie nowych źródeł danych typu big data w celu rozszerzenia zakresu merytorycznego oficjalnych statystyk. Uzyskane wyniki potwierdziły, że techniki web scrapingu mogą stanowić wartościowe narzędzie w statystyce publicznej, umożliwiając pozyskiwanie danych z alternatywnych źródeł, pod warunkiem zapewnienia ich spójności z obowiązującymi badaniami statystycznymi. W artykule wskazano liczne problemy związane m.in. z deduplikacją danych, odwzorowywaniem różnych klasyfikacji takich jak Klasyfikacja Zawodów i Specjalności, jak również samą dostępnością portali z ofertami pracy w zakresie pobrania wiarygodnych informacji.

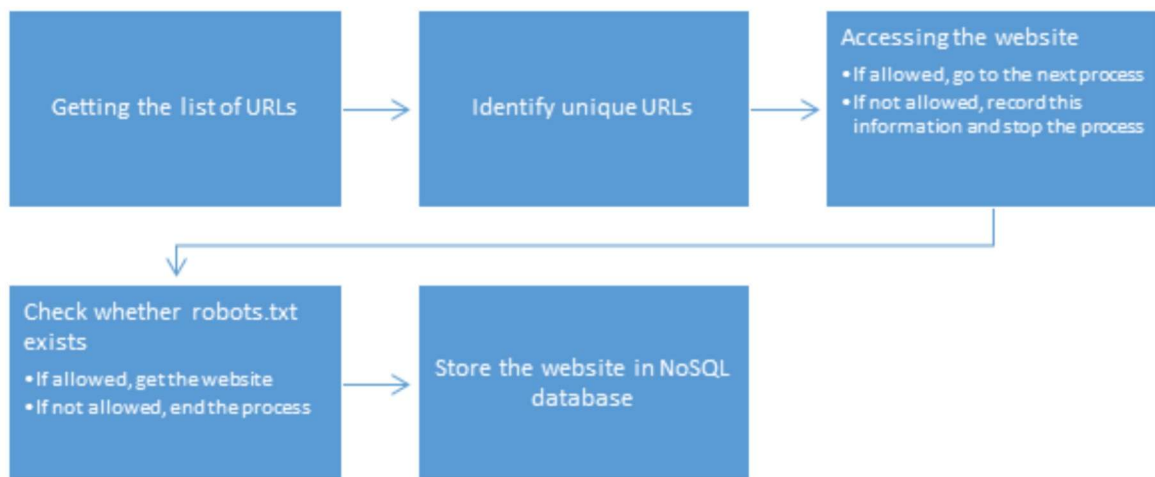
Drugi artykuł, „Current challenges and possible big data solutions for the use of web data as a source for official statistics”, indeksowany w bazie Web of Science, pokazuje bardziej metodologiczną stronę pozyskiwania danych na dużą skalę. Celem artykułu było dokonanie syntetycznego przeglądu współczesnych problemów związanych z dostępem do informacji publikowanych na stronach internetowych, ich ekstrakcją oraz możliwościami praktycznego wykorzystania w analizach opartych na danych big data. W publikacji omówiono zarówno bariery techniczne, jak i ograniczenia metodologiczne, prawne oraz organizacyjne, które mogą wpływać na jakość i użyteczność danych pozyskiwanych z Internetu. Dodatkowo wskazano potencjalne kierunki rozwiązań pozwalających na ograniczenie wpływu tych czynników na proces badawczy.

W ramach artykułu przeprowadzono również studium przypadku dotyczące masowego web scrapingu. Badanie to zrealizowano w 2022 r. na próbie ponad pół miliona stron internetowych, co pozwoliło na praktyczną ocenę stabilności i wiarygodności pozyskiwanych danych. Wyniki analizy wykazały, że nawet przy zastosowaniu zaawansowanych technik

pozyskiwania informacji nie jest możliwe uzyskanie w pełni reprezentatywnego i stabilnego zbioru danych z tak dużej próby. W szczególności stwierdzono, że w trakcie pozyskiwania danych niedostępnych pozostaje od 15% do 20% stron, co jest zgodne z obserwacjami raportowanymi w innych badaniach dotyczących masowego pobierania informacji ze stron internetowych. Dodatkowo podkreślono, że całkowita liczba aktywnych stron internetowych w poszczególnych krajach nie może zostać jednoznacznie ustalona z uwagi na dynamiczny charakter Internetu oraz ciągłą rotację treści.

Zastosowany proces pobierania danych z Internetu obejmował proces, zgodnie z rysunkiem poniżej.

**Rysunek 2. Proces pobierania stron internetowych wykorzystany w publikacji (Daas, Maślankowski, 2023)**



Źródło: (Daas, Maślankowski, 2023).

Zgodnie z powyższym rysunkiem, analizie poddano dostępność strony bazując na kodzie odpowiedzi z serwera, jak również na podstawie pliku robots.txt. Artykuł stanowi podsumowanie wieloletnich doświadczeń badawczych autorów w zakresie pracy z danymi internetowymi, w tym realizacji projektów naukowych i eksperckich dotyczących wykorzystania big data w statystyce publicznej i analizach rynku.

Najważniejszym wkładem publikacji jest identyfikacja i usystematyzowanie jedenastu kluczowych problemów związanych z pobieraniem danych ze stron internetowych, obejmujących m.in. niestabilność źródeł, ograniczenia techniczne, bariery prawne, heterogeniczność struktur stron, trudności w walidacji danych oraz brak pełnej wiedzy o populacji źródeł. Problemy te zostały sklasyfikowane w odpowiednie kategorie, a następnie dla każdej z nich zaproponowano metody ograniczania ich wpływu na proces pozyskiwania i przetwarzania danych, co stanowi istotny wkład metodologiczny w rozwój badań nad wykorzystaniem danych internetowych w naukach o zarządzaniu, statystyce oraz pokrewnych dyscyplinach.

## Badanie mediów społecznościowych

Badanie danych pozyskiwanych ze stron internetowych bardzo często rozszerzane jest na wiedzę z pozyskiwania danych z mediów społecznościowych. W tym przypadku media społecznościowe to źródła internetowe big data charakteryzujące się szybkim napływem dużych porcji informacji. W ramach tego nurtu napisałem kilka publikacji, z czego za najważniejsze uznaję dwie opublikowane w 2020 roku.

Tytuł publikacji
Wrycza, S., & Maślankowski, J. (2020). Social Media Users' Opinions on Remote Work during the COVID-19 Pandemic. Thematic and Sentiment Analysis. <i>Information Systems Management</i> , 37(4), 288–297. DOI 10.1080/10580530.2020.1820631
Maślankowski, J., Brzezicki, Ł. (2020) Wykorzystanie mediów społecznościowych w szkolnictwie wyższym, <i>Wiadomości Statystyczne. The Polish Statistician</i> 65 (2), 30-42. DOI 10.5604/01.3001.0014.0455

Pierwsza publikacja została napisana w roku pandemicznym i dotyczyła aspektów związanych z postrzeganiem pracy zdalnej, co znalazło odzwierciedlenie w liczbie cytowań tej publikacji w artykułach zamieszczanych w międzynarodowych bazach danych (w dniu pisania niniejszego autoreferatu – Google Scholar 113, SCOPUS 56, Web of Science 38).

Celem publikacji było pokazanie wyników badań poprzez zastosowanie metod związanych z pozyskaniem, przetwarzaniem oraz analizą danych w mediach społecznościowych na temat zdalnej pracy w pandemii. Wykorzystano metody takie jak uczenie maszynowe nadzorowane i nienadzorowane.

Przeprowadzone badania dotyczące pracy zdalnej w kontekście aktywności użytkowników mediów społecznościowych w okresie pandemii COVID-19 pozwoliły na sformułowanie szeregu istotnych wniosków o charakterze empirycznym i prognostycznym. W badaniach zastosowano analizę deskryptywną, stanowiącą jedną z wiodących metod analityki mediów społecznościowych, co umożliwiło precyzyjne określenie dynamiki zmian w zakresie dyskursu na temat pracy zdalnej. Ustalono, że w szczytowym okresie pandemii, tj. w marcu 2020 roku, w momencie wprowadzenia powszechnego lockdownu, liczba tweetów odnoszących się do pracy zdalnej była piętnastokrotnie wyższa niż w analogicznym okresie roku 2019. W kolejnych miesiącach, od kwietnia do lipca 2020 roku, odnotowano wahania liczby publikacji, jednak średni poziom aktywności utrzymywał się na istotnie wyższym poziomie niż przed pandemią.

Równolegle zastosowano analizę text mining, która pozwoliła na identyfikację i systematyczne monitorowanie nowych wątków tematycznych związanych z pracą zdalną, diskutowanych w mediach społecznościowych. Wyniki analizy wykazały, że w okresie pandemii media społecznościowe były intensywnie wykorzystywane przez przedsiębiorstwa sektora IT w celu promocji narzędzi i rozwiązań wspierających pracę zdalną, co stanowiło 34,8% wszystkich tweetów dotyczących tej problematyki. Drugim pod względem popularności

zagadnieniem były dyskusje poświęcone kompetencjom niezbędnym do efektywnej pracy zdalnej (26,6%), natomiast trzecim – tematyka wsparcia w zakresie pracy zdalnej, w tym organizacja webinarów i szkoleń online (12,6%).

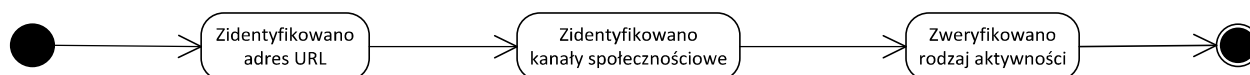
Zastosowanie analizy sentymentu – jednej z kluczowych metod analityki mediów społecznościowych – wykazało, że dominującym wydźwiękiem publikacji dotyczących pracy zdalnej był ton pozytywny, stanowiący 62,23% ogółu wpisów, podczas gdy komentarze o charakterze negatywnym nie przekraczały 12,96%. Wyniki te wskazują na generalnie przychylne postrzeganie pracy zdalnej przez użytkowników mediów społecznościowych w okresie objętym badaniem.

Ponadto, analiza typu wordcloud pozwoliła na identyfikację nowych trendów językowych i obszarów znaczeniowych w komunikacji dotyczącej pracy zdalnej, takich jak „techjobs” (stanowiska techniczne), „remoteteam”, „workanywhere” czy „workfromhome”. Uzyskane wyniki wskazują, że pandemia COVID-19 w sposób trwały wpłynęła na kształtowanie nowych praktyk komunikacyjnych oraz koncepcji pracy zdalnej w przestrzeni cyfrowej.

Rezultaty badań pozwoliły również sformułować wnioski prognostyczne odnoszące się do okresu postpandemicznego. Stwierdzono, że większość rozwiązań technologicznych, metod i narzędzi wykorzystywanych w okresie pandemii utrzyma swoją znaczącą rolę w codziennym funkcjonowaniu organizacji także po ustaniu bezpośredniego zagrożenia epidemiologicznego. Wskazuje to na trwałą transformację modeli organizacyjnych i komunikacyjnych, w których praca zdalna została włączona jako integralny element funkcjonowania instytucji i przedsiębiorstw.

Celem drugiego artykułu, tj. „Wykorzystanie mediów społecznościowych w szkolnictwie wyższym” było przedstawienie wyników badania dotyczącego skali wykorzystania mediów społecznościowych przez uczelnie. W artykule opisano również metody pozyskiwania i przetwarzania danych, które można w dużym uproszczeniu przedstawiają się jak na rysunku poniżej.

### Rysunek 3. Proces pobierania i przetwarzania danych na potrzeby artykułu (Maślankowski, Brzezicki, 2020).



Źródło: (Maślankowski, Brzezicki, 2020).

Celem przeprowadzonych badań była kompleksowa ocena stopnia oraz charakteru wykorzystania mediów społecznościowych przez instytucje szkolnictwa wyższego w Polsce, z zastosowaniem zaawansowanych metod analitycznych osadzonych w paradygmacie big data. Założenia badawcze obejmowały zarówno pomiar zakresu obecności uczelni na platformach społecznościowych, jak i pogłębioną analizę tematyczną i funkcjonalną treści publikowanych w tych kanałach komunikacji. Tak skonstruowana perspektywa badawcza umożliwiła identyfikację dominujących wzorców komunikacyjnych oraz ocenę sposobów, w jakie szkoły

wyższe wykorzystują narzędzia mediów społecznościowych w procesach budowania wizerunku, podtrzymywania relacji ze społecznością akademicką oraz oddziaływania na otoczenie społeczno-gospodarcze.

Zrealizowane w pierwszym kwartale 2019 roku badania empiryczne wykazały wyraźną stratyfikację aktywności w mediach społecznościowych między uczelniami o zróżnicowanej skali działania. Jednostki duże, zarówno w wymiarze liczebności społeczności akademickiej, jak i wielkości organizacyjnej, cechują się najwyższym poziomem obecności i aktywności w środowisku mediów społecznościowych. Systematycznie publikują one rozbudowany zakres treści, obejmujący m.in. informacje o bieżących wydarzeniach, materiały promujące osiągnięcia naukowe i organizacyjne, a także komunikaty o charakterze relacji, wzmacniające identyfikację członków społeczności z uczelnią.

W przeciwieństwie do tego, uczelnie średnie i małe wykazują istotnie niższy poziom zaangażowania w wykorzystanie mediów społecznościowych, a odsetek jednostek posiadających oficjalne profile jest relatywnie ograniczony. Niski poziom obecności w tych kanałach prowadzi do niewykorzystania ich potencjału w zakresie komunikacji, promocji oraz dyfuzji informacji, co może negatywnie wpływać na postrzeganą widoczność i rozpoznawalność mniejszych instytucji szkolnictwa wyższego w przestrzeni publicznej.

Przeprowadzona analiza treści ujawniła ponadto istotne różnice między przekazami publikowanymi w mediach społecznościowych a materiałami umieszczanymi na oficjalnych stronach internetowych uczelni. Podczas gdy witryny www pełnią przede wszystkim funkcję informacyjno-administracyjną, a ich struktura i zawartość mają charakter stosunkowo statyczny, komunikacja w mediach społecznościowych jest zdecydowanie bardziej dynamiczna, zorientowana na promocję wydarzeń akademickich, inicjatyw naukowych i studenckich, a także na kształtowanie wizerunku uczelni w czasie rzeczywistym. Wyniki te wskazują, że media społecznościowe stanowią odrębny, wysoce specyficzny kanał komunikacji instytucjonalnej, którego funkcja wykracza poza prostą replikację treści publikowanych w innych źródłach informacyjnych i odgrywa coraz większą rolę w procesach komunikacyjnych szkół wyższych.

Warto zaznaczyć, że autor niniejszego wniosku jest twórcą pierwszej udostępnionej wersji oprogramowania do badania obecności przedsiębiorstw w mediach społecznościowych. Oprogramowanie to jest wykorzystywane obecnie w ramach różnych projektów Komisji Europejskiej oraz Eurostatu, i rozwijane jest przez osoby z całej Europy.

### **Syntetyczne ujęcie dorobku naukowego**

Podsumowując, moje publikacje naukowe koncentrują się głównie wokół przetwarzania i analizy danych. Od roku 2013 związane są przede wszystkim z przetwarzaniem dużych zbiorów danych w celu prowadzenia badań empirycznych. Poniższa tabela stanowi podsumowanie mojego dorobku naukowego.

**Tabela 4. Syntetyczne ujęcie aktywności naukowej**

Liczba recenzowanych publikacji naukowych	Przed uzyskaniem stopnia naukowego doktora	Po uzyskaniu stopnia naukowego doktora	Razem
samodzielnie	3	21	24
<i>w tym w języku angielskim</i>	2	9	11
współautorskie	9	25	34
<i>w tym w języku angielskim</i>	0	18	18
Razem (samodzielne i współautorskie)	12	46	58

**5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.**

**5.1. Współpraca w międzynarodowej sieci naukowej European Research Consortium for Information Systems (ERCIS), Niemcy**

European Research Consortium for Information Systems (ERCIS) stanowi dynamiczną, międzynarodową sieć badawczą skupiającą wiodące ośrodki naukowe oraz partnerów przemysłowych działających w obszarze systemów informatycznych (Information Systems). Jej siedziba jest przy Uniwersytecie Muenster. ERCIS zorganizowany jest jak typowa sieć badawcza, tj. posiada Steering Board, Advisory Board, Competence Centers, Research Clusters itp. Sieć obejmuje ponad 30 partnerskich instytucji akademickich z całego świata, w tym Uniwersytet Gdański oraz Politechnikę Wrocławską jako przedstawicieli z Polski. Celem ERCIS jest wspieranie współpracy naukowej, wymiany doświadczeń badawczych oraz rozwijania wspólnych inicjatyw projektowych i publikacyjnych w zakresie informatyki ekonomicznej.

W ramach współpracy z siecią ERCIS aktywnie uczestniczę w działaniach ukierunkowanych na rozwój badań naukowych, integrację środowiska międzynarodowego oraz promocję dorobku polskich badaczy w tej dziedzinie. Do najważniejszych inicjatyw, w które byłem zaangażowany, należą:

- 1. Udział w organizacji cyklu konferencji EuroSymposium on Systems Analysis and Design, później EuroSymposium on Digital Transformation, odbywających się pod oficjalnym patronatem sieci ERCIS.** W ostatnich latach konferencja ta stanowiła platformę prezentacji wyników badań oraz wymiany wiedzy między naukowcami i praktykami z całego świata.
- 2. Współkierowanie ścieżką tematyczną „Digital Transformation” w ramach prestiżowej międzynarodowej konferencji AIS International Conference on**

**Information Systems Development (ISD)** (140 pkt w krajowych kryteriach oceny), będącej kontynuacją inicjatywy EuroSymposium. Współprzewodniczącymi tej ścieżki są prof. Joerg Becker (ERCIS Academic Director, Uniwersytet w Münster) oraz autor niniejszego wniosku. Ścieżka została zrealizowana w ramach konferencji ISD 2024 w Sopocie (Polska) oraz ISD 2025 w Belgradzie (Serbia).

3. **Współredakcja monografii naukowej** opracowanej we współpracy z członkami sieci ERCIS:

Wrycza S., Maślankowski J. (red.), *Informatyka Ekonomiczna. Teoria i zastosowania*, Wydawnictwo Naukowe PWN, Warszawa 2019.

W publikacji tej uczestniczyli członkowie ERCIS m.in. prof. Joerg Becker (Uniwersytet w Münster, Niemcy) oraz prof. Reima Suomi (Uniwersytet w Turku, Finlandia), co stanowi przykład efektywnej współpracy międzynarodowej w ramach ERCIS.

### **Znaczenie współpracy dla dorobku habilitacyjnego**

Udział w sieci ERCIS w istotny sposób przyczynił się do rozwoju mojego dorobku naukowego i organizacyjnego. Współpraca z czołowymi badaczami z wiodących ośrodków akademickich (m.in. Uniwersytet w Münster, Uniwersytet w Turku, Uniwersytet w Agder) umożliwiła rozwinięcie międzynarodowych kontaktów badawczych, wspólne publikacje oraz inicjowanie projektów o zasięgu europejskim.

Aktywność w ramach ERCIS pozwoliła również na wzmocnienie pozycji polskiego środowiska naukowego w obszarze systemów informacyjnych oraz podniesienie rozpoznawalności badań prowadzonych na Uniwersytecie Gdańskim. Wkład w organizację konferencji o wysokim prestiżu (ISD, EuroSymposium) oraz współredakcję publikacji z udziałem uznanych naukowców z zagranicy stanowi wymierny dowód aktywności międzynarodowej i współpracy naukowej, istotny w kontekście osiągnięć wymaganych w postępowaniu habilitacyjnym.

### **Plany dalszej współpracy w ramach sieci ERCIS**

W kolejnych latach planuję kontynuować i rozwijać współpracę w ramach ERCIS, w szczególności poprzez inicjowanie wspólnych projektów badawczych finansowanych ze środków europejskich i współredagowanie publikacji naukowych w renomowanych wydawnictwach oraz dalszą organizację cyklu konferencji EuroSymposium i ścieżek tematycznych w ramach konferencji ISD lub innych.

Ponadto planowane jest zacieśnienie współpracy dydaktycznej i badawczej z wybranymi ośrodkami ERCIS (m.in. Münster, Agder, Turku) w zakresie transformacji cyfrowej, zrównoważonych systemów informacyjnych oraz sztucznej inteligencji w gospodarce. Działania te mają na celu dalsze umacnianie międzynarodowej pozycji Uniwersytetu Gdańskiego w obszarze badań nad systemami informacyjnymi i cyfryzacją organizacji.

Współpraca z ERCIS pozwoliła mi również wizytować wiele ośrodków naukowych w Europie, m.in. Uniwersytet w Lund (Niderlandy), czy Uniwersytet w Turku (Finlandia). Warto podkreślić, że obecnie sieć ta posiada dwóch dyrektorów akademickich. Jednym z nich jest wspomniany prof. J. Becker, dwukrotny doktor honoris causa, natomiast drugim jest prof. J. vom Brocke, obecny Prezes-elekt towarzystwa naukowego AIS (Association for Information Systems). Świadczy to o randze tej instytucji w świecie nauki.

## **5.2. Współpraca z SNSPA, National University of Political Studies and Public Administration, Rumunia**

Współpraca ze Școala Națională de Studii Politice și Administrative (SNSPA - angielska nazwa National University of Political Studies and Public Administration) w Rumunii została zainicjowana w roku 2015. Wśród najważniejszych elementów tej współpracy wymienić należy:

- 1) Współpraca naukowa podczas corocznej konferencji SCIC – Smart Cities International Conference (SCIC) – jestem członkiem ścisłego grona Komitetu Naukowego począwszy od 2017 roku;
- 2) Współpraca w projektach naukowych, m.in. w ramach projektu SIPOCA.
- 3) Wyjazdy dydaktyczne Erasmus STA+.

W ramach współpracy naukowej przy organizacji konferencji SCIC, zostałem zaproszony do ścisłego grona naukowego w roku 2015. W ramach tej współpracy intensywnie współpracuję z przedstawicielami SNSPA, w szczególności z prof. C. Vrabie. Za współpracę z SNSPA w tym zakresie zostałem wyróżniony m.in. **nagrodą za wybitny wkład w rozwój konferencji naukowej SCIC**. Pełna treść: “Recognition Award proudly presented to Jacek Maślankowski for his distinguished contribution to the Development of Smart(er) Communities and his commitment to Excellency in Public Administration Research. Thank you for your service as our Smart Cities Ambassador!”

Dzięki ścisłej współpracy z SNSPA zostałem zaproszony do udziału w projekcie EU SIPOCA 20, w ramach programu “Setting up the development framework for e-government instruments”. Moja aktywność polegała na prezentowaniu rozwiązań e-government przed przedstawicielami środowiska naukowego oraz administracji publicznej, również z udziałem tamtejszych ministrów. W latach 2018-2021 wygłosiłem trzy tzw. Keynote Speech. Ponadto pełniłem w projekcie funkcję konsultacyjną, gdzie zarówno przedstawiciele świata nauki, jak i administracji publicznej, mogli czerpać wiedzę z mojego doświadczenia w zakresie współpracy z jednostkami administracji publicznej.

Pełnię również rolę członka Komitetu Naukowego Smart Edu Hub, utworzonego na SNSPA (<https://www.smart-edu-hub.eu/home/team>).

## **5.3. Współpraca z TuE – Technical University of Eindhoven, Niderlandy**

Współpraca z prof. Pietem Daasem z Eindhoven University of Technology (TU/e, Niderlandy) dotyczyła przede wszystkim badań naukowych w zakresie wykorzystania danych

masowych (big data) w statystyce publicznej. Wspólne działania obejmowały zarówno prace badawcze, jak i opracowywanie raportów naukowych realizowanych na zlecenie Komisji Europejskiej, dotyczących wykorzystania danych pochodzących z sieci Internet w badaniach statystycznych oraz rozwoju metod analizy danych wielkoskalowych.

W wyniku współpracy powstały liczne opracowania i raporty naukowe, a do najważniejszych osiągnięć zaliczyć należy:

- 1) Artykuł naukowy indeksowany w bazie Web of Science: Daas, P., Maślankowski, J. (2023). Current challenges and possible big data solutions for the use of web data as a source for official statistics. *Wiadomości Statystyczne. The Polish Statistician*, 68(12), 49–64. <https://doi.org/10.59139/ws.2023.12.3>
- 2) Organizację wraz z prof. Pietem Daasem wspólnej ścieżki naukowej w ramach konferencji (International Conference on Advanced Research Methods and Analytics) CARMA 2022, zrealizowanej na zaproszenie prof. Josepa Domenecha z Politechniki w Walencji (Hiszpania).

W ramach pierwszego obszaru współpracy, trwającego blisko dwa lata, opracowano publikację naukową poświęconą analizie wyzwań związanych z wykorzystaniem danych pochodzących z sieci Web jako źródła informacji dla statystyki publicznej. W pracy zaprezentowano problemy metodologiczne oraz możliwe rozwiązania w zakresie masowego pozyskiwania danych (web scrapingu) i ich przetwarzania w celach badawczych.

Drugi obszar współpracy obejmował działania organizacyjne i popularyzatorskie, w tym wystąpienia konferencyjne realizowane na zaproszenie prof. Josepa Domenecha. Wieloletnia współpraca z tym badaczem, od dwóch dekad zajmującym się analizą i eksploracją zasobów internetowych, umożliwiła rozwinięcie badań nad zastosowaniem metod big data do analizy treści stron internetowych organizacji i instytucji.

Współpraca z prof. Pietem Daasem miała istotne znaczenie dla rozwoju mojej działalności naukowej, zwłaszcza w zakresie wykorzystania danych webowych w badaniach statystycznych. Prof. Daas uważany jest za autorytet w dziedzinie metodologii big data dla statystyki publicznej, a jego dorobek naukowy wielokrotnie był prezentowany i cytowany w publikacjach międzynarodowych. Trzeba również podkreślić, że prof. Piet Daas zaprezentował mnie w jednym ze swoich wystąpień na Technical University of Eindhoven jako bliskiego współpracownika, mającego wpływ na jego badania, jak również umieścił mnie jako współautora badań realizowanych przez niego na uczelni TU/e.

#### **5.4. Inne instytucje naukowe**

W ramach działalności naukowej podjąłem współpracę z pracownikami Uniwersytetu Warszawskiego, która zaowocowała m.in. prowadzeniem zajęć dydaktycznych na Uniwersytecie Gdańskim w ramach szkoły letniej big data. Zajęcia, realizowane w języku angielskim, obejmowały problematykę zastosowań baz danych oraz wizualizacji danych opartych na dużych zbiorach danych.

Aktualnie z mojej strony prowadzona jest współpraca z prof. Vilną Nekrašaitė-Liegė z Politechniki w Wilnie (Vilnius Gediminas Technical University). W wyniku tej współpracy przygotowano m.in. artykuł naukowy, który został zaakceptowany do publikacji w czasopiśmie

posiadającym współczynnik Impact Factor. Celem publikacji jest przedstawienie metodologii dotyczącej prób nielosowych stosowanych w badaniach realizowanych z wykorzystaniem danych pochodzących ze stron internetowych. Obecnie trwają prace nad wprowadzeniem poprawek redakcyjnych do wskazanego artykułu.

W 2006 roku rozpocząłem członkostwo w organizacji Association for Information Systems (AIS). W tym samym roku współtworzyłem polski oddział stowarzyszenia – The Polish Chapter of AIS (PLAIS) – w ramach którego przez pierwsze lata pełniłem funkcję sekretarza, a następnie skarbnika. Jestem członkiem założycielem (Founding Member) tego polskiego oddziału AIS.

### 5.5. Konferencje naukowe

Byłem członkiem komitetów naukowych i programowych konferencji organizowanych na różnych uczelniach w Polsce i za granicą. W poniższej tabeli zaprezentowano wybrane konferencje, w których wniosłem znaczący wkład.

**Tabela 5. Lista konferencji naukowych ze znaczącym wkładem autora wniosku**

Nazwa konferencji	Funkcja
Seria corocznej konferencji EuroSymposium, lata 2010-2023, Sopot, Polska	Członek Komitetu Organizacyjnego, Recenzent, Redaktor LNBIP (od 2017 roku)
ISD 2026, AIS 34th International Conference on Information Systems Development, Association for Information Systems, Praga, Czechy	Track Co-Chair, Członek Komitetu Programowego, Recenzent
ISD 2025, AIS 33rd International Conference on Information Systems Development, Association for Information Systems, Belgrad, Serbia	Track Co-Chair, Członek Komitetu Programowego, Recenzent
ISD 2024, AIS 32nd International Conference on Information Systems Development, Association for Information Systems, Sopot, Polska	Track Co-Chair, Członek Komitetu Programowego, członek Komitetu Organizacyjnego, Recenzent
SAC 2026, Track on Lean and Agile Software Development, The 41st ACM/SIGAPP Symposium On Applied Computing, Thessaloniki, Greece, March 23-27, 2026	Członek Komitetu Programowego, Recenzent
ICTERI 2025, International Conference on ICT in Education, Research and Industrial Applications, Nicea, Francja	Członek Komitetu Programowego, Recenzent

<b>Nazwa konferencji</b>	<b>Funkcja</b>
ICTERI 2024, International Conference on ICT in Education, Research and Industrial Applications, Lwów, Ukraina	Członek Komitetu Programowego, Recenzent
ICTERI 2023, International Conference on ICT in Education, Research and Industrial Applications, Iwano Frankiwnsk, Ukraina	Członek Komitetu Programowego, Recenzent
18th Conference on Computer Science and Intelligence Systems, FedCSIS 2023 (IEEE #57573), Warsaw, Poland, 17–20 September, 2023	Członek Komitetu Programowego, Recenzent
PERS@ADBIS, 3rd Workshop on Personalization and Recommender Systems at the 28th European Conference on Advances in Databases and Information Systems, August 28, 2024, Bayonne, France	Członek Komitetu Programowego, Recenzent
SMARTGREENS 2020: 9th International Conference on Smart Cities and Green ICT Systems, online	Członek Komitetu Programowego
ICIM 2009, International Conference on Information Management, Gdańsk, Polska	Członek Komitetu Organizacyjnego
ICIM 2015, International Conference on Information Management, Gdańsk, Polska	Członek Komitetu Organizacyjnego, Redaktor materiałów konferencyjnych
ICEP 2010, The 2nd International Conference on eCommerce, ePayments and New Entrepreneurship, Sopot, Polska	Członek Komitetu Organizacyjnego
SCIC 2017-2025 (coroczna konferencja) – Smart Cities International Conference, Bukareszt, Rumunia	Członek Komitetu Naukowego
ICTM 2016, International Conference on ICT Management for Global Competitiveness and Economic Growth in Emerging Economies, Wrocław, Polska	Członek Komitetu Naukowego i Recenzującego

## **6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę.**

### **6.1.Osiągnięcia dydaktyczne**

Jako najważniejsze osiągnięcie dydaktyczne uznaję:

- współautorstwo książek dydaktycznych (w tym nagrodzona przez PTI), autorstwo rozdziałów w książkach dydaktycznych oraz współredakcja monografii akademickiej, sugerowanej studentom jako zasób wiedzy naukowej dotyczący informatyki ekonomicznej,
- uzyskanie dwóch wyróżnień w konkursie Nauczyciel Roku im. Celestyna Mrongowiusza, organizowanego w Uniwersytecie Gdańskim,
- uzyskanie Medalu Komisji Edukacji Narodowej,
- współdziałanie w tworzeniu programu studiów kierunku Informatyka i Ekonometria, w szczególności na specjalności Aplikacje informatyczne w biznesie,
- prowadzenie licznych zajęć dydaktycznych w języku polskim i angielskim, w tym na studiach podyplomowych, wysoko ocenianych przez studentów.

**Współautorstwo książek dydaktycznych, autorstwo rozdziałów w książkach dydaktycznych oraz redagowanie monografii naukowej będącej podręcznikiem akademickim**

#### **Monografia naukowa – podręcznik akademicki**

Najważniejsze dzieło to monografia naukowa, która pełni rolę podręcznika akademickiego dla studentów:

- Wrycza, S., Maślankowski, J. (red.), (2019) Informatyka ekonomiczna. Teoria i zastosowania., Wydawnictwo Naukowe PWN, Warszawa, 863 strony

Monografia ta zdobyła wyróżnienie w Konkursie PTI (Polskie Towarzystwo Informatyczne) na **Najlepszą Polską Książkę Informatyczną Roku** (nagrodę przyznano w 2021 r.).

#### **Współautorstwo książek dydaktycznych:**

- Wrycza, S., Marcinkowski, B., Maślankowski, J. (2012) UML 2.x: ćwiczenia zaawansowane, Helion, Gliwice

#### **Autorstwo rozdziałów w książkach dydaktycznych:**

- Maślankowski, J. (2016) Pozyskiwanie danych na potrzeby analiz biznesowych [w:] Buchnowska D. (eds.): Kokpity menedżerskie w analizie i prezentacji danych

biznesowych z wykorzystaniem MS Excel, 2016, Sopot, Wydział Zarządzania Uniwersytetu Gdańskiego, str. 11-30.

- Maślankowski, J. (2016) Podstawy tworzenia kokpitu menedżerskiego z wykorzystaniem Power Pivot [w:] Buchnowska D. (eds.): Kokpity menedżerskie w analizie i prezentacji danych biznesowych z wykorzystaniem MS Excel, 2016, Sopot, Wydział Zarządzania Uniwersytetu Gdańskiego, str. 119-132.
- Maślankowski, J. (2008) Projektowanie sieci komputerowych [w:] Sieci komputerowe w Biznesie (seria Informatyka Ekonomiczna), Wydawnictwo UG, 2008.
- Maślankowski, J., Lenartowicz, T. (2008) Bezpieczeństwo transmisji danych [w:] Sieci komputerowe w Biznesie (seria Informatyka Ekonomiczna), Wydawnictwo UG, 2008.
- Maślankowski, J. (2006), Diagramy maszyny stanowej, [w:] Wrycza, S. (red), UML 2.1. Ćwiczenia, Helion, Gliwice, str. 81-88
- Maślankowski, J. (2006), Diagramy wdrożeniowe (komponentów i rozlokowania), [w:] Wrycza, S. (red), UML 2.1. Ćwiczenia, Helion, Gliwice, str.123-132
- Maślankowski, J. (2006), UML a programowanie, [w:] Wrycza, S. (red), UML 2.1. Ćwiczenia, Helion, Gliwice, str. 217-236

### **Konkurs Nauczyciel Roku im. Celestyna Mrongowiusza**

W ramach działalności dydaktycznej prowadziłem intensywną i wysoko ocenianą działalność nauczycielską, czego wyrazem są uzyskane wyróżnienia i nagrody przyznawane przez Uniwersytet Gdański.

Zostałem dwukrotnie uhonorowany wyróżnieniem w konkursie im. Celestyna Mrongowiusza, stanowiące jedno z najwyższych wyróżnień dydaktycznych przyznawanych nauczycielom akademickim Uniwersytetu Gdańskiego:

- Nauczyciel Roku im. Celestyna Mrongowiusza – Gdańsk, 20 marca 2013 r. (za rok akademicki 2011/2012),
- Nauczyciel Roku im. Celestyna Mrongowiusza – Gdańsk, 20 marca 2015 r. (za rok akademicki 2013/2014).

Nagroda im. Celestyna Mrongowiusza przyznawana jest nauczycielom akademickim wyróżniającym się szczególnymi osiągnięciami dydaktycznymi, wysokim poziomem zaangażowania w proces kształcenia oraz uznaniem wśród studentów. Laureaci wyłaniani są na podstawie wyników ocen i opinii studentów. Z każdego wydziału Uniwersytetu Gdańskiego nominowana jest corocznie tylko jedna osoba, która uzyskuje najlepsze oceny studentów.

Otrzymane wyróżnienia stanowią potwierdzenie wysokich kompetencji dydaktycznych wnioskującego, jego zaangażowania w proces dydaktyczny oraz uznania w środowisku akademickim Uniwersytetu Gdańskiego.

**Tabela 6. Wyróżnienia w konkursie Nauczyciel Roku im. Celestyna Mrongowiusza**

<b>Data otrzymania wyróżnienia</b>	<b>Rodzaj konkursu</b>
20 marca 2013 r. (za rok akademicki 2011/2012)	Nagroda Nauczyciel Roku im. Celestyna Mrongowiusza (Uniwersytet Gdański)
20 marca 2015 r. (za rok akademicki 2013/2014)	Nagroda Nauczyciel Roku im. Celestyna Mrongowiusza (Uniwersytet Gdański)

Pragnę podkreślić, że podobne, lecz mniej znaczące wyróżnienia uzyskiwałem również na innych uczelniach z którymi współpracowałem. Jest to rezultatem wysokiej oceny zajęć prowadzonych przeze mnie, przekazywanej w ankietach wypełnianych przez studentów na temat wykładowców prowadzących dane przedmioty.

#### **Medal Komisji Edukacji Narodowej**

W sierpniu 2022 roku zostałem odznaczony Medalem Komisji Edukacji Narodowej w uznaniu osiągnięć w działalności naukowej i dydaktycznej. Wyróżnienie to zostało przyznane na wniosek złożony przez Wydział Zarządzania Uniwersytetu Gdańskiego, który został pozytywnie zaopiniowany i przekazany do dalszego procedowania.

**Tabela 7. Informacje nt. uzyskania Medalu Komisji Edukacji Narodowej**

<b>Rodzaj</b>	<b>Data i miejsce</b>	<b>Przyznający nagrodę/wyróżnienie</b>
Medal Komisji Edukacji Narodowej	Sierpień 2022 r., Warszawa	Minister właściwy do spraw oświaty i wychowania.

#### **Opiekun przedmiotów na kierunku Informatyka i Ekonometria**

W ramach realizowanej działalności dydaktycznej na kierunku Informatyka i Ekonometria prowadzę zajęcia z przedmiotów w języku angielskim: Implementation of Big Data Solutions oraz Introduction to Business Intelligence, wraz z ich polskimi odpowiednikami: Wdrażanie rozwiązań big data oraz Wprowadzenie do Business Intelligence. Dodatkowo prowadzone są przeze mnie zajęcia z przedmiotu Bazy danych. Zajęcia te obejmują zarówno komponent teoretyczny, jak i praktyczny, zgodnie z obowiązującymi standardami kształcenia na poziomie studiów wyższych.

## Prowadzenie zajęć dydaktycznych

W ciągu ponad 20-letniego okresu zatrudnienia na Uniwersytecie Gdańskim prowadziłem zajęcia obejmujące szeroki zakres przedmiotów, w tym między innymi następujące grupy tematyczne:

- Programowanie – Visual Basic, C#, Java, Python, ze szczególnym uwzględnieniem projektów praktycznych i ćwiczeń laboratoryjnych, mających na celu rozwój kompetencji programistycznych studentów.
- Bazy danych – SQL i NoSQL, obejmujące zarówno teorię projektowania baz danych, jak i praktyczne zastosowania w systemach informatycznych.
- Sieci komputerowe – w tym certyfikowane kursy CCNA, w ramach których studenci zdobywali zarówno wiedzę teoretyczną, jak i umiejętności praktyczne w zakresie konfiguracji i zarządzania sieciami komputerowymi.
- Analiza danych i Business Intelligence – przedmioty takie jak Hurtownie danych, czy powiązane z Business Intelligence oraz big data, w ramach których realizowane były zadania obejmujące modelowanie danych, przetwarzanie dużych zbiorów danych oraz wizualizację wyników analitycznych.

Zajęcia były prowadzone zarówno w języku polskim, jak i angielskim, obejmując komponent teoretyczny oraz praktyczny, zgodnie ze standardami kształcenia na poziomie studiów wyższych. Równocześnie uczestniczyłem w opracowywaniu i aktualizacji programów nauczania, wprowadzając innowacyjne metody dydaktyczne, w tym elementy projektowe i laboratoria praktyczne, dostosowane do zmieniających się wymogów branży IT oraz potrzeb studentów. W efekcie podejmowane działania przyczyniły się do podniesienia jakości kształcenia oraz przygotowania studentów do samodzielnej pracy zawodowej w środowisku informatycznym. Szczegółowa lista prowadzonych w Uniwersytecie Gdańskim przedmiotów została zaprezentowana w dwóch poniższych tabelach, obejmujących odpowiednio przedmioty prowadzone w języku polskim, jak i angielskim.

**Tabela 8. Lista przedmiotów prowadzonych w języku polskim**

Przedmiot	Rodzaj zajęć
ABC IT	Ćwiczenia
Administrowanie bazami danych	Ćwiczenia
Algorytmy i struktury danych	Wykład
Analiza danych OLAP	Ćwiczenia
Aplikacje rozproszone w języku Java	Ćwiczenia
Bazy danych	Wykład, ćwiczenia
Bezpieczeństwo serwisów internetowych	Ćwiczenia

Bezpieczeństwo systemów	Wykład, ćwiczenia
Big data w biznesie	Wykład, ćwiczenia
Hurtownie danych	Wykład, ćwiczenia
Implementowanie rozwiązań big data	Wykład, ćwiczenia
Informatyczne wspomaganie decyzji inwestycyjnych	Ćwiczenia
Informatyka w administracji	Wykład
Informatyka w zarządzaniu	Ćwiczenia
Język SQL	Ćwiczenia
Komputerowa analiza danych	Ćwiczenia
Nierelacyjne rozwiązania bazodanowe	Wykład, ćwiczenia
Programowanie aplikacji bazodanowych	Ćwiczenia
Programowanie komputerów	Ćwiczenia
Programowanie obiektowe	Wykład, ćwiczenia
Programowanie sieciowe	Wykład, ćwiczenia
Programowanie w języku Python	Ćwiczenia
Seminarium licencjackie	Seminarium
Seminarium magisterskie	Seminarium
Serwerowe systemy operacyjne	Ćwiczenia
Serwery aplikacyjne	Ćwiczenia
Serwisy internetowe	Ćwiczenia
Standardy XML	Ćwiczenia
System operacyjny Linux	Ćwiczenia
System operacyjny MS Windows Server	Ćwiczenia
Systemy baz danych	Wykład, ćwiczenia

Systemy big data	Wykład, ćwiczenia
Systemy Business Intelligence	Wykład, ćwiczenia
Technologia baz danych	Wykład, ćwiczenia
Technologia Informacyjna	Ćwiczenia
Technologie sieci WAN	Ćwiczenia
Telemedycyna i e-zdrowie	Wykład
Wprowadzenie do Business Intelligence	Wykład, ćwiczenia

**Tabela 9. Lista przedmiotów prowadzonych w języku angielskim**

Przedmiot	Rodzaj zajęć
ABC of IT	Wykład
Big Data Systems	Ćwiczenia
Business Intelligence systems	Ćwiczenia
Business Outcomes of Big Data Analysis	Wykład
Business Reports in SQL	Wykład
Challenges of Business Informatics	Wykład
Data warehouses	Wykład, ćwiczenia
E-government and Smart Cities	Wykład
Fundamentals of Business Informatics	Wykład
Fundamentals of Business Intelligence	Wykład, ćwiczenia
Implementation of Big Data Solutions	Wykład, ćwiczenia
Information Processing for Investment Professionals	Wykład, ćwiczenia
Information Technologies	Ćwiczenia
Introduction to Linux in Business	Wykład

Master Seminar	Seminarium
Open Source Applications in Business	Wykład

### **Prowadzenie zajęć na studiach podyplomowych**

W ramach prowadzonej działalności dydaktycznej zostałem zaproszony do realizacji zajęć na studiach podyplomowych, obejmujących następujące przedmioty:

- Analiza danych. Big Data – współprowadzenie zajęć z przedmiotu Wprowadzenie do big data (12 godzin) oraz prowadzenie zajęć z komponentu Hurtownie danych (12 godzin).
- Efektywne zarządzanie jednostkami ochrony zdrowia – prowadzenie zajęć z przedmiotu Informatyka medyczna (12 godzin).
- Data Science – analiza danych z wykorzystaniem AI – prowadzenie zajęć z przedmiotu Programowanie w języku Python (66 godzin).

Zajęcia te obejmowały zarówno komponent teoretyczny, jak i praktyczny, przy czym szczególny nacisk położono na rozwój umiejętności praktycznych uczestników w zakresie analizy danych, projektowania rozwiązań informatycznych oraz zastosowania narzędzi i technologii stosowanych w nowoczesnych systemach informatycznych. Równocześnie uczestniczyłem w opracowywaniu i aktualizacji programów nauczania poszczególnych studiów podyplomowych, wprowadzając innowacyjne metody dydaktyczne, w tym elementy projektowe i praktyczne laboratoria, w celu dostosowania treści kształcenia do wymogów rynku pracy i potrzeb zawodowych słuchaczy.

### **Uczestniczenie w programie ProUG – Program Rozwoju Uniwersytetu Gdańskiego**

W latach 1.10.2018 r. – 31.12.2023 r. uczestniczyłem w projekcie ProUG „PROgram Rozwoju Uniwersytetu Gdańskiego (ProUG)”, nr POWR.03.05.00-IP.09-00-PZ3/17. Projekt otrzymał dofinansowanie w ramach Programu Operacyjnego Wiedza Edukacja Rozwój, Osi Priorytetowej III Szkolnictwo wyższe dla gospodarki i rozwoju, Działanie 3.5 Kompleksowe programy szkół wyższych w kwocie ogółem na poziomie 18 626 672,02 zł.

W ramach projektu odpowiadałem za utworzenie i prowadzenie zajęć dydaktycznych z zakresu informatyki na wielu wydziałach Uniwersytetu Gdańskiego, m.in. na Wydziale Prawa i Administracji (anglojęzyczny kierunek Criminology and Criminal Justice) oraz Wydziale Zarządzania (kierunki Informatyka i ekonometria, specj. Aplikacje Informatyczne w Biznesie oraz Finance and Accounting).

### **6.2.Osiągnięcia organizacyjne**

Jako najważniejsze osiągnięcia organizacyjne, związane z Uniwersytetem Gdańskim, wskazuję wymienione w poniższej tabeli. Spośród nich chciałbym szczególną uwagę zwrócić nad stałą opieką nad kołami naukowymi, kierowanie praktykami studenckimi na kierunku Informatyka i ekonometria, pełnienie funkcji redaktora tematycznego oraz statystycznego

czasopism, jak również zasiadanie w zarządzie Stowarzyszenia Absolwentów. Dużo zaangażowania wymaga ode mnie również opieka administracyjno-organizacyjna oraz prowadzenie certyfikowanych kursów Cisco, w ramach Akademii Cisco Uniwersytetu Gdańskiego.

**Tabela 10. Najważniejsze osiągnięcia organizacyjne związane z Uniwersytetem Gdańskim**

<b>Data</b>	<b>Funkcja</b>
2020 r. – nadal	Kierownik Praktyk Studenckich kierunku Informatyka i Ekonometria na Wydziale Zarządzania Uniwersytetu Gdańskiego
2016 r. – nadal	V-ce Prezes Stowarzyszenia Absolwentów WSHM (Wyższa Szkoła Handlu Morskiego), WSE (Wyższa Szkoła Ekonomiczna) oraz Wydziałów Ekonomicznych Uniwersytetu Gdańskiego
2014 r. – nadal	Opiekun Koła Naukowego Informatyki Lider
2008-2013 r.	Opiekun Koła Naukowego E-xpert
2016-2020 r.	Członek Rady Dziekana Wydziału Zarządzania Uniwersytetu Gdańskiego
2021-2024 r.	Członek Rady Wydziału Zarządzania Uniwersytetu Gdańskiego
2011-2020 r.	Redaktor tematyczny czasopisma <i>Zarządzanie i Finanse. Journal of Management and Science</i> . ISSN 2084-5189. Wydawnictwo Wydziału Zarządzania UG.
2016 r. – nadal	Redaktor statystyczny czasopisma <i>Financial Law Review</i> , ISSN 2299-6834, Wydawnictwo UG.
2011 r. – nadal	Komitet Organizacyjny Konferencji <i>EuroSymposium</i>
2017-2023 r.	Redakcja naukowa czasopisma <i>LNBIP (Lecture Notes in Business Information Processing)</i> wydawnictwa Springer
2024-2025 r.	Przewodniczący tracka Digital Transformation (Track Chair) konferencji Information Systems Development – edycje 2024 oraz 2025 (140 pkt MNiSW).
2009-2015 r.	Członek Komitetu Organizacyjnego Konferencji <i>ICIM International Conference on Information Management</i> (General Chair prof. dr hab. B.F.Kubiak)
2010 r.	Członek Komitetu Organizacyjnego Konferencji <i>International Conference on eCommerce, ePayments and New Entrepreneurship</i>

2011-2024 r.	Członek Komisji ds. Nagród i Odznaczeń na Wydziale Zarządzania Uniwersytetu Gdańskiego
2012 r. – nadal	Funkcja administracyjno-organizacyjne w Akademii Cisco Uniwersytetu Gdańskiego (administrator Akademii Cisco)

W ramach udziału w Komitecie Organizacyjnym Konferencji EuroSymposium, byłem współredaktorem następujących materiałów konferencyjnych, wydawanych przez wydawnictwo Springer, zgodnie z tabelą poniżej.

**Tabela 11. Lista współredagowanych materiałów konferencyjnych LNBIP Springer**

Rok	Tytuł
2017	10th SIGSAND/PLAIS EuroSymposium 2017: S. Wrycza, J. Maślankowski (eds), Information Systems: Development, Research, Applications, Education, Series: Lecture Notes in Business Information Processing 300, Springer, 2017, DOI: 10.1007/978-3-319-66996-0
2018	11th SIGSAND/PLAIS EuroSymposium 2018: S. Wrycza, J. Maślankowski (eds), Information Systems: Research, Development, Applications, Education, Series: Lecture Notes in Business Information Processing 333, Springer, 2018, DOI: 10.1007/978-3-030-00060-8
2019	12th SIGSAND/PLAIS EuroSymposium 2019: S. Wrycza, J. Maślankowski (eds), Information Systems: Research, Development, Applications, Education, Series: Lecture Notes in Business Information Processing 359, Springer, 2019, DOI: 10.1007/978-3-030-29608-7
2021	13th SIGSAND/PLAIS EuroSymposium 2021: S. Wrycza, J. Maślankowski (eds), Digital Transformation, Series: Lecture Notes in Business Information Processing 429, Springer, 2021, DOI: 10.1007/978-3-030-85893-3
2022	14th PLAIS EuroSymposium 2022: J. Maślankowski, B. Marcinkowski, P. Rupino da Cunha (eds), Digital Transformation, Series: Lecture Notes in Business Information Processing 465, Springer, 2022, DOI: 10.1007/978-3-031-23012-7
2023	15th PLAIS EuroSymposium 2023: J. Maślankowski, B. Marcinkowski, P. Rupino da Cunha (eds), Digital Transformation, Series: Lecture Notes in Business Information Processing 495, Springer, 2023, DOI: 10.1007/978-3-031-43590-4

W ramach konferencji AIS ISD 2024 oraz ISD 2025 odpowiadałem za przydzielanie recenzentów do artykułów naukowych, pisanie metarecenzji oraz współpraca z General Chair oraz moimi co-chairami, wśród których był również prof. J. Becker z ERCIS/Uniwersytetu w Monastyrze. W ramach konferencji serii *International Conference on Information*

*Management* byłem w Komitecie Organizacyjnym w latach 2010-2015 r. W ramach tej współpracy opublikowane zostały m.in. materiały konferencyjne *Information Management in Practice*, które współredagowałem (poniższa tabela). Byłem również członkiem komitetów organizacyjnych innych konferencji, jak np. *ICeP 2010*, która odbyła się w Sopocie w 2010 r.

**Tabela 12. Współredakcja materiałów konferencyjnych krajowej konferencji ICIM**

Rok	Tytuł
2015	Kubiak Bernard, Maślankowski Jacek (eds.), <i>Information Management in Practice</i> , 2015, Sopot, Faculty of Management University of Gdańsk, 373 p., ISBN 978-83-64669-05-7

Zasiadałem również w komitetach redakcyjnych czasopism naukowych. Byłem **redaktorem tematycznym** z zakresu informatyki ekonomicznej w czasopiśmie *Zarządzanie i Finanse. Journal of Management and Science*. ISSN 2084-5189. Wydawnictwo Wydziału Zarządzania UG. Pełnię również rolę **redaktora statystycznego** czasopisma *Financial Law Review*, ISSN 2299-6834, Wydawnictwo UG.

W ramach działalności wychowawczej i popularyzatorskiej wspierałem studentów pełniąc rolę **Opiekuna nad Kołem Naukowym Lider (od 2014 r.)**, w ramach którego współorganizowałem coroczną konferencję Liderreferencja, mającą na celu popularyzację informatyki ekonomicznej wśród studentów. Prelegentami konferencji byli przedstawiciele firm informatycznych. Najważniejszym osiągnięciem tej działalności było uzyskiwanie licznych wyróżnień przez studentów pod moją opieką, w tym w międzynarodowym konkursie MS Imagine Cup. Istotną częścią działalności Koła była także promocja Wydziału Zarządzania Uniwersytetu Gdańskiego poprzez organizację licznych wydarzeń, w tym charytatywnych, w których współuczestniczyłem wraz ze studentami, takich jak Liderzy planszówek w Ergo Arenie oraz stanowiska promujące naukę w Galerii Bałtyckiej. W latach 2008-2013 pełniłem funkcję opiekuna Koła Naukowego E-xpert, działającego również na Wydziale Zarządzania Uniwersytetu Gdańskiego.

Od 2016 roku pełnię funkcję **Wiceprezesa Stowarzyszenia Absolwentów WSHM, WSE** oraz Wydziałów Ekonomicznych Uniwersytetu Gdańskiego. W ramach tej działalności odpowiadam m.in. za organizację prelekcji absolwentów osiągających znaczące sukcesy zawodowe powiązane z kierunkami studiów. W tym miejscu chciałbym poinformować, że WSHM (Wyższa Szkoła Handlu Morskiego) oraz WSE (Wyższa Szkoła Ekonomiczna) to uczelnie, które w roku 1970 zostały przekształcone w Uniwersytet Gdański.

Od 2020 roku jestem **Kierownikiem Praktyk Studenckich** na kierunku Informatyka i Ekonometria na Wydziale Zarządzania Uniwersytetu Gdańskiego, co obejmuje m.in. podpisywanie umów z przedsiębiorstwami oraz rozliczanie studentów z odbytych praktyk.

W latach 2011-2024 byłem członkiem Komisji ds. Nagród i Odznaczeń na Wydziale Zarządzania Uniwersytetu Gdańskiego. Do moich zadań należało opiniowanie wniosków wpływających do Komisji, jak również uczestniczenie w głosowaniach nad przyznawanymi nagrodami oraz odznaczeniami.

Od 2012 roku pełnię funkcję administracyjno-organizacyjną w Akademii Cisco Uniwersytetu Gdańskiego, odpowiadając za organizację Akademii, w tym rekrutację uczestników, planowanie zajęć we współpracy z instruktorami oraz współpracę z Rektorem Uniwersytetu Gdańskiego w zakresie rozliczania kursów. Posiadam również uprawnienia instruktorskie i prowadzę zajęcia w ramach Akademii Cisco.

### **6.3.Osiągnięcia popularyzujące naukę**

Do osiągnięć popularyzujących naukę zaliczę:

- 1) Zasiadanie w Jury polskich i międzynarodowych hackathonów oraz mentoring,
- 2) Organizację konferencji popularnonaukowych wraz z Kołem Naukowym Informatyki L!der, jako Opiekun Koła,
- 3) Wystąpienia na licznych wykładach popularyzujących naukę, promujących Wydział,
- 4) Organizacja spotkań popularnonaukowych w ramach Stowarzyszenia Absolwentów WSHM, WSE oraz Wydziałów Ekonomicznych Uniwersytetu Gdańskiego.

W swojej działalności popularyzatorskiej uczestniczyłem w pracach jury licznych hackathonów, zarówno krajowych, takich jak Hackathon SDG (2018, organizator: Główny Urząd Statystyczny), jak i prestiżowych wydarzeniach międzynarodowych, w tym EU Datathon (2021, organizator: Komisja Europejska). W ramach wybranych hackathonów pełniłem również funkcję mentora, wspierając uczestników w rozwiązywaniu problemów praktycznych.

W latach 2017–2022, we współpracy z Kołem Naukowym L!der, organizowałem cykl otwartych konferencji popularnonaukowych L!derferencja, skierowanych m.in. do uczniów klas maturalnych. Konferencje te miały na celu pogłębianie wiedzy z zakresu informatyki ekonomicznej i prezentowanie praktycznych aspektów tej dziedziny, w tym zarządzania projektami informatycznymi oraz budowy systemów informatycznych. Prelegentami konferencji byli praktycy, którzy przedstawiali zagadnienia z perspektywy realizowanych projektów w przedsiębiorstwach. Podsumowując, studenci Koła organizowali m.in.:

- cykliczną konferencję L!derferencja,
- warsztaty informatyczne zapraszając prelegentów – praktyków z firm, m.in. w zakresie big data, np. nauka języka Scala w Apache Spark,
- charytatywne wydarzenie o nazwie L!derzy planszówek, odbywające się w Ergo Arenie na granicy Gdańska i Sopotu, gdzie zbierano datki na rzecz chorych dzieci (pieniądze zostały przekazane przez członków Koła w Uniwersyteckim Centrum Klinicznym, co zostało odpowiednio udokumentowane).

Do wystąpień popularyzujących naukę w szczególności zaliczyć należy moją aktywność w zakresie wydarzeń takich jak:

- dni otwarte szkół średnich, np. wystąpienie nt. big data w języku angielskim dla Zespołu Szkół Akademickich w Elblągu, gdzie uczestniczyło kilkadziesiąt uczniów klas maturalnych,
- w ramach wydarzenia Uniwersalna Strefa Nauki odbywającego się w Galerii Bałtyckiej w Gdańsku, w postaci "Warsztaty na temat bezpieczeństwa w sieci.", gdzie m.in. został przeprowadzony wywiad ze mną przez dziennikarzy lokalnego radia.

W ramach pełnienia funkcji Wiceprezesa Stowarzyszenia Absolwentów WSHM, WSE oraz Wydziałów Ekonomicznych Uniwersytetu Gdańskiego organizowałem liczne wykłady popularnonaukowe, w których prelegentami byli absolwenci WSHM, WSE oraz Wydziałów Ekonomicznych Uniwersytetu Gdańskiego, prezentujący swoje doświadczenia zawodowe i osiągnięcia, związane z kierunkiem studiów.

### **Promotorstwo prac magisterskich i licencjackich**

Posiadam doświadczenie w zakresie opieki naukowej nad pracami dyplomowymi, pełniąc funkcję promotora łącznie 36 prac magisterskich, spośród których 11 zostało napisanych w języku angielskim, oraz 23 prac licencjackich, w tym 3 w języku angielskim, zgodnie z poniższą tabelą.

**Tabela 13. Promotorstwo prac magisterskich i licencjackich**

<b>Rodzaj prac</b>	<b>Obronione prace</b>	<b>w tym w języku angielskim</b>
magisterskie	36	11
licencjackie	23	3

Szczególne znaczenie w mojej działalności dydaktycznej miała rola pierwszego promotora prac magisterskich realizowanych w ramach anglojęzycznej specjalności Business Informatics na kierunku Informatyka i Ekonometria, co stanowiło istotny wkład w rozwój oferty edukacyjnej o charakterze międzynarodowym na wydziale. Ponadto odegrałem pionierską rolę w zakresie promotorstwa pomocniczego. Zostałem wyznaczony na pierwszego promotora pomocniczego na Wydziale, w rezultacie czego aktywnie uczestniczyłem w opracowaniu uchwały regulującej zasady sprawowania funkcji promotora pomocniczego, przyczyniając się tym samym do ugruntowania formalnych podstaw tej roli w strukturze jednostki akademickiej.

**7. Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej.**

**Współpraca z instytucjami w ramach prowadzenia prac naukowych**

W toku mojej działalności zawodowej nawiązałem współpracę z wieloma instytucjami międzynarodowymi i krajowymi oraz przedsiębiorstwami, w tym w szczególności:

- OECD – Organizacja Współpracy Gospodarczej i Rozwoju (Organisation for Economic Co-operation and Development) – pełniłem istotne funkcje, m.in. przewodniczącego sieci INES LSO oraz grupy roboczej ESO, a także uczestniczyłem w opracowywaniu licznych raportów i prac metodologicznych;
- Eurostat, Komisja Europejska – współpracowałem jako konsultant w zakresie doskonalenia metodologii pozyskiwania danych z Internetu, w szczególności w projektach dotyczących analizy ofert pracy w sieci;
- UNECE – United Nations Economic Commission for Europe, EKG ONZ – jestem współautorem dwóch raportów, w tym opracowań o charakterze metodologicznym;
- Urząd Statystyczny w Gdańsku, Główny Urząd Statystyczny – pełnię funkcję konsultanta;
- REDD AI – uczestniczyłem w projekcie finansowanym ze środków europejskich, którego celem było opracowanie narzędzia predykcji cen nieruchomości biurowych.

Działalność ta obejmowała zarówno wkład merytoryczny w projekty badawcze i raporty metodologiczne, jak i współpracę przy wdrażaniu rozwiązań praktycznych, co stanowiło istotny element mojej aktywności naukowo-zawodowej.

**Współpraca z OECD**

W latach 2017–2021 w ramach współpracy z OECD pełniłem funkcję przewodniczącego grupy roboczej ESO (Education and Social Outcomes) w ramach sieci INES LSO (Indicators of Education Systems, Network for Data Development on Labour Market and Social Outcomes of Education). Sieć OECD INES LSO zajmuje się rozwojem wskaźników dotyczących uwarunkowań rynku pracy w zależności od poziomu i rodzaju edukacji, a także analizą społecznych i ekonomicznych efektów kształcenia. W skład sieci wchodzi blisko 100 członków reprezentujących ponad 30 krajów OECD oraz krajów partnerskich. Spotkania sieci odbywają się dwa razy w roku w różnych lokalizacjach na świecie.

Do kluczowych etapów mojej działalności w OECD należy zaliczyć:

- Pełnienie funkcji przewodniczącego Sieci OECD INES LSO,
- Pełnienie funkcji przewodniczącego Grupy Roboczej ESO.

Jako przewodniczący Grupy Roboczej ESO byłem odpowiedzialny za inicjowanie prac badawczych, nadzór nad ich przebiegiem oraz opracowywanie zestawu wskaźników

dotyczących społecznych efektów kształcenia, co stanowiło istotny wkład w rozwój metodologii OECD w zakresie edukacji i rynku pracy.

**Tabela 14. Działalność w ramach Organizacji Współpracy Gospodarczej i Rozwoju (OECD)**

<b>Organizacja</b>	<b>Funkcja</b>	<b>Data</b>
OECD	Członek sieci	2010 r. – nadal
OECD	Przewodniczący grupy roboczej ESO	2017- 2021 r., 2025 r. – nadal
OECD	Przewodniczący sieci INES LSO	2022 – 2024 r.

Za jedno z najważniejszych wyróżnień w mojej karierze zawodowej i naukowej uważam fakt, że w 2021 roku zostałem nominowany przez sześć krajów członkowskich OECD i jednogłośnie wybrany na przewodniczącego sieci INES LSO na trzyletnią kadencję, bez możliwości jej przedłużenia, zgodnie ze statutem sieci. Pełnienie tej funkcji wiązało się z odpowiedzialnością za kierowanie międzynarodową grupą ekspertów, koordynowanie prac nad wskaźnikami dotyczącymi społecznych i ekonomicznych efektów kształcenia oraz inicjowanie projektów badawczych, które mają zastosowanie w polityce edukacyjnej krajów członkowskich OECD i krajów partnerskich.

Sieć OECD INES LSO odpowiada za przygotowanie i opracowanie treści dotyczących społecznych i ekonomicznych uwarunkowań kształcenia, które są prezentowane w prestiżowej publikacji OECD Education at a Glance (EaG), stanowiącej jedno z najbardziej uznanych źródeł danych i analiz w zakresie edukacji na świecie. Jestem stałym współpracownikiem tej publikacji, wymienianym wśród współautorów począwszy od edycji EaG 2011, co pozwalało mi nie tylko na systematyczne monitorowanie trendów edukacyjnych, ale również na aktywne kształtowanie metodologii i standardów pomiaru wskaźników edukacyjnych w międzynarodowej perspektywie.

### **Współpraca z Eurostat**

W ramach współpracy z Eurostatem istotnym elementem mojej działalności naukowej był rozwój metod pozyskiwania informacji z Internetu. Od 2016 roku uczestniczę jako członek głównych projektów grantowych dotyczących pozyskiwania i przetwarzania danych internetowych, w tym projektów ESSnet Big Data I oraz ESSnet Big Data II. W ramach tych projektów opracowywałem i rozwijałem narzędzia służące do identyfikacji obecności przedsiębiorstw w mediach społecznościowych, wraz z walidacją danych i weryfikacją rodzaju aktywności w tych serwisach. Narzędzia te były następnie rozwijane i adaptowane przez współpracowników z różnych krajów, co pozwoliło na ich zastosowanie w szerokim, międzynarodowym kontekście. Pozwalało mi to na stosowanie w praktyce rozwiązań, o których pisałem w pracach naukowych wymienianych w niniejszym autoreferacie jako osiągnięcia dodatkowe.

Wyniki współpracy z Eurostatem znalazły również odzwierciedlenie w mojej monografii, w której w rozdziałach teoretycznych cytowane są raporty i opracowania przygotowane również przez mnie na potrzeby Eurostatu. Opracowane narzędzia i metody pozwoliły na standaryzację procesu pozyskiwania danych internetowych, umożliwiając tworzenie wiarygodnych wskaźników ekonomicznych i społecznych w skali międzynarodowej.

**Tabela 15. Współpraca z Komisją Europejską (Eurostat)**

<b>Organizacja</b>	<b>Funkcja</b>	<b>Data</b>
Eurostat, Komisja Europejska	Członek zespołu w ramach grantu ESSnet Big Data I	2016-2018 r.
Eurostat, Komisja Europejska	Członek zespołu w ramach grantu ESSnet Big Data II	2018-2020 r.
Eurostat, Komisja Europejska	Koordinator pakietu roboczego 2 w ramach grantu Web Intelligence Network	2020-2025 r.
Eurostat, Komisja Europejska	Koordinator grantu Web Data in Statistical Production Process. Web Intelligence Hub.	2025 r. - nadal

W latach 2020–2025 pełniłem funkcję koordynatora pakietu roboczego 2 (Work Package 2: OBEC and OJA) w projektach realizowanych we współpracy z Eurostatem, którego celem było wdrożenie do produkcji statystyk eksperymentalnych w ramach dwóch inicjatyw: Online Based Enterprise Characteristics (charakterystyka przedsiębiorstw obecnych online) oraz Online Job Advertisements (oferty pracy online). W ramach tego pakietu byłem odpowiedzialny za opracowanie i wdrożenie metodologii pozyskiwania danych z Internetu, definiowanie standardów oceny jakości danych oraz przygotowanie i nadzór nad skryptami wykorzystującymi infrastrukturę chmurową Eurostatu, co umożliwiło automatyczne pozyskiwanie danych, ich przetwarzanie oraz udostępnianie wyników analiz w formie statystyk eksperymentalnych. Realizacja tego pakietu obejmowała współpracę z zespołami z wielu krajów, koordynację prac badawczo-analitycznych oraz monitorowanie postępów w zakresie metodologii i jakości danych.

Od czerwca 2025 roku kontynuuję współpracę z Eurostatem jako koordynator po stronie polskiej w projekcie grantowym Web Data in Statistical Production Process, gdzie odpowiadam za koordynację działań badawczo-wdrożeniowych w Polsce, w tym za nadzór nad pozyskiwaniem, przetwarzaniem i weryfikacją danych internetowych w kontekście statystyk oficjalnych. Działalność ta przyczynia się do rozwoju metodologii statystyki eksperymentalnej oraz wdrażania innowacyjnych rozwiązań w zakresie big data w instytucjach statystycznych, zarówno na poziomie krajowym, jak i międzynarodowym.

Jako jedyna osoba spoza Komisji Europejskiej (Eurostat, CEDEFOP), zostałem zaproszony do udziału w podcaście nt. rozwiązań big data w Komisji Europejskiej. Planowana data publikowania podcastu to drugi kwartał 2026 r. (Stats in a Wrap. Episode 29).

W ramach współpracy z Komisją Europejską jestem współautorem kilkudziesięciu raportów, z czego większość powiązana jest z big data. Wybrane raporty, których jestem współautorem i są powiązane z big data, znajdują się poniżej.

**Tabela 16. Wybrane współautorskie raporty Komisji Europejskiej związane z big data**

<b>Rola</b>	<b>Tytuł raportu</b>
Współautor	Maślankowski, J., Salgado, D., Quaresma, S., Ascari G., Brancato, G., Consiglio, L., Righi, P., Tuoto, T., Daas, P., Six, M., Kowarik, A., (2019) Deliverable K1:First Draft of the Quality Guidelines for the Acquisition and Usage of Big Data, Work package K, Deliverable K.1.
Współautor	Stateva, G., ten Bosch, O., Maślankowski, J., Righi, A., Scannapieco, M., Greenaway, M., Swier, N., Jansson, I., Wu, D. (2017) Legal aspects related to Web scraping of Enterprise Web Sites. ESSnet Big Data I, Work package 2, Deliverable 2.1.
Współautor	Stateva, G., ten Bosch, O., Maślankowski, J., Barcaroli, G., Scannapieco, M., Greenaway, M., Jansson, I., Wu, D. (2017) Methodological and IT Issues and Solutions. ESSnet Big Data I, Work package 2, Deliverable 2.4.
Współautor	Consten, A., Chavdarov, V., Daas, P., Horvat, V., Maslankowski, J., Quaresma, S., Six, M., Tuoto, T. (2018) Report describing the methodology of using Big Data for official statistics and the most important questions for future studies. ESSnet Big Data I, Work package 8, Deliverable 8.4.
Współautor	Consten, A., Chavdarov, V., Daas, P., Horvat, V., Maślankowski, J., Quaresma, S., Six, M., Tuoto, T. (2018) Report describing the quality aspects of big data for official statistics. ESSnet Big Data, Deliverable 8.2. Eurostat, European Commission.
Kierujący pracami, współautor	Maślankowski, J., Bis, M., Blatt, D. D. P., Brnot, N., Columbano, A., Eisinias, Ž., Georgiev, K., Gussenbauer, J., Kuehnemann, H., Lucarelli, A., Magistro, R., Nekrašaitė-Liegė, V., Niewiadomska, E., Ruocco, G., Summa, D. (2025), Web Intelligence Network. Deliverable 2.4 Suggested tables with experimental statistics and metadata including quality assessment, Eurostat, European Commission.
Kierujący pracami, współautor	Maślankowski, J., Bis, M., Blatt, D. D. P., Brnot, N., Columbano, A., Eisinias, Ž., Georgiev, K., Gussenbauer, J., Kuehnemann, H., Lucarelli, A., Magistro, R., Nekrašaitė-Liegė, V., Niewiadomska, E., Ruocco, G., Summa, D. (2024), Deliverable 2.6: Technical guidelines on the implementation of the scripts produced for OJA and OBEC statistics, Eurostat, European Commission.

## Współpraca z UNECE

W ramach współpracy z Europejską Komisją Gospodarczą Organizacji Narodów Zjednoczonych (UNECE) istotnym elementem mojej działalności naukowej było uczestnictwo w przygotowaniu dwóch kluczowych opracowań międzynarodowych:

- A Suggested Framework for the Quality of Big Data (2014) oraz
- Guide on Measuring Human Capital (2016).

W obu publikacjach zostałem wymieniony z imienia i nazwiska jako współautor, co stanowi formalne potwierdzenie mojego wkładu w prace badawcze i metodologiczne.

**Tabela 17. Wybrane współautorskie raporty ONZ**

Organizacja	Funkcja	Data
UNECE	Współautor: A Suggested Framework for the Quality of Big Data, UNECE, 2014	2012-2014 r.
UNECE	Współautor: Guide on Measuring Human Capital, United Nations Economic Commission for Europe, wydawnictwo: United Nations Publications. New York and Geneva, 2016, ISBN 978-92-1-117120-4, e-ISBN 978-92-1-059866-8	2014-2016 r.

Pierwsze opracowanie stanowiło jedno z pierwszych kompleksowych badań nad jakością danych big data i wprowadzało wielowymiarowe podejście do oceny jakości danych, obejmujące zarówno proces przetwarzania danych, jak i różnorodność źródeł danych. Opracowanie to nie tylko wyznaczyło nowe standardy metodologiczne w zakresie jakości danych, ale zostało również przeze mnie cytowane w monografii naukowej, którą przedstawiam jako istotny wkład w rozwój nauk o zarządzaniu i jakości.

W drugim opracowaniu, Guide on Measuring Human Capital, wraz z zespołem podejmowałem próby opracowania zestawu mierników umożliwiających systematyczną ocenę rozwoju kapitału ludzkiego. Wypracowane w tym zakresie narzędzia i wskaźniki stanowią fundament dalszych prac badawczych prowadzonych przez OECD w zakresie pomiaru kapitału ludzkiego i mają istotne znaczenie dla rozwoju międzynarodowej metodologii statystycznej. Moja działalność w tym obszarze obejmowała zarówno udział w tworzeniu koncepcji wskaźników, jak i analizę ich przydatności w praktyce, co przyczyniło się do popularyzacji oraz ustandaryzowania metod oceny kapitału ludzkiego w ujęciu międzynarodowym.

## Współpraca z Urzędem Statystycznym w Gdańsku (Główny Urząd Statystyczny)

Poza pracą w Uniwersytecie Gdańskim, jestem również konsultantem zatrudnionym w niepełnym wymiarze czasu pracy w Urzędzie Statystycznym w Gdańsku, chociaż moja współpraca ma miejsce głównie na szczeblu centralnym, tj. w Głównym Urzędzie Statystycznym. W latach wieloletniej współpracy z Urzędem Statystycznym w Gdańsku,

realizowanej głównie na szczeblu centralnym w Głównym Urzędzie Statystycznym, pełnię funkcję konsultanta w niepełnym wymiarze czasu pracy. W ramach tej współpracy uczestniczyłem w licznych projektach metodologicznych dotyczących różnych aspektów życia społecznego, obejmujących zarówno opracowanie, jak i nadzór nad wdrażaniem innowacyjnych metod badawczych. Jestem współautorem wielu publikacji wydanych przez GUS.

W ostatnich latach szczególnie nacisk położyłem na wykorzystanie źródeł danych typu big data, w tym na inicjowanie, rozwijanie oraz udoskonalanie metodologii pozyskiwania i przetwarzania danych. Praca ta umożliwiła implementację wielu koncepcji naukowych w praktyce statystycznej, co przyczyniło się do uzyskania wymiernych efektów w zakresie jakości i rzetelności analiz statystycznych oraz pogłębienia mojego rozumienia potencjału big data w badaniach społeczno-ekonomicznych.

### **Współpraca z REDD AI**

W ramach praktycznego zastosowania metod big data uczestniczyłem w projekcie finansowanym ze środków europejskich, którego celem było opracowanie narzędzi do szacowania cen nieruchomości komercyjnych, ze szczególnym uwzględnieniem nieruchomości biurowych, obszaru wskazywanego w literaturze jako luka badawcza. W projekcie pełniłem funkcję głównego inżyniera ds. uczenia maszynowego, odpowiadając za projektowanie, implementację, testowanie oraz przekazanie algorytmów predykcyjnych do zastosowań praktycznych.

Wyniki tej współpracy zaowocowały opracowaniem metod, które pozwalają na dokładne szacowanie wartości nieruchomości komercyjnych z wykorzystaniem danych big data, a także przyczyniły się do przygotowania dwóch wysokopunktowanych publikacji naukowych, których wydanie miało miejsce w listopadzie 2025 roku oraz będzie miało miejsce w marcu 2026 r. Działalność ta łączy w sobie elementy badań podstawowych i stosowanych, jednocześnie przyczyniając się do rozwoju metod analizy danych oraz praktycznych zastosowań big data w obszarze ekonomii i zarządzania nieruchomościami.

### **Nagrody i wyróżnienia**

W swojej karierze zawodowej zostałem uhonorowany wieloma nagrodami i wyróżnieniami, spośród których najważniejsze wypisałem w poniższej tabeli.

**Tabela 18. Wybrane nagrody i wyróżnienia związane z działalnością naukową i dydaktyczną**

<b>Rodzaj</b>	<b>Data i miejsce</b>	<b>Przyznający nagrodę/wyróżnienie</b>
Medal Komisji Edukacji Narodowej	Sierpień 2022 r., Warszawa	Minister właściwy do spraw oświaty i wychowania.
Indywidualna Nagroda Rektora II stopnia za działalność naukową	Październik 2017 r., Gdańsk	Rektor Uniwersytetu Gdańskiego
Indywidualna Nagroda Rektora II stopnia za działalność naukowo-organizacyjną	Marzec 2020 r., Gdańsk	Rektor Uniwersytetu Gdańskiego
Nagroda Nauczyciel Roku im. Celestyna Mrongowiusza (Uniwersytet Gdański)	20 marca 2013 r. (za rok akademicki 2011/2012)	Rektor Uniwersytetu Gdańskiego
Nagroda Nauczyciel Roku im. Celestyna Mrongowiusza (Uniwersytet Gdański)	20 marca 2015 r. (za rok akademicki 2013/2014)	Rektor Uniwersytetu Gdańskiego
Nagroda za osiągnięcia naukowe, organizacyjne i dydaktyczne	Czerwiec 2019 r., Sopot	Dziekan Wydziału Zarządzania Uniwersytetu Gdańskiego
Najlepsza Polska Książka Informatyczna w roku 2020	Czerwiec 2021 r., Warszawa	Polskie Towarzystwo Informatyczne (PTI)
Wyróżnienie (Recognition Award) za wkład w rozwój konferencji SCIC (Smart Cities International Conference)	Grudzień 2022 r., Bukareszt	Władze SNSPA (National University of Political Studies and Public Administration)

## Nagrody na hackathonach

W ramach działalności dodatkowej uczestniczyłem w licznych hackathonach, pełniąc funkcje zarówno uczestnika odpowiedzialnego za programowanie rozwiązań, jak i koordynatora zespołów. Do najbardziej znaczących sukcesów zespołowych należy zaliczyć:

- I miejsce – UNECE HLG Hackathon (organizator: Europejska Komisja Gospodarcza ONZ), 5-7 września 2017 r.
- I miejsce – Europejski Hackathon Big Data, marzec 2019 r. (organizator: Komisja Europejska)
- II miejsce – Europejski Hackathon Big Data, marzec 2021 r. (organizator: Komisja Europejska)
- VI miejsce – Europejski Hackathon Big Data, marzec 2023 r. (organizator: Komisja Europejska)

Uważam, że istotnym wkładem w rozwój technologii big data była moja aktywność w ramach tych hackathonów, gdyż rozwiązania i pomysły wypracowane podczas wydarzeń były następnie rozwijane i wdrażane przez różne organizacje, m.in. CEDEFOP (Europejskie Centrum Rozwoju Kształcenia Zawodowego, Komisja Europejska).

**Tabela 19. Lista zdobytych nagród na hackathonach**

Organizator	Nazwa	Data	Miejsce
Organizacja Narodów Zjednoczonych, Europejska Komisja Gospodarcza	Hackathon UNECE HLG SDG	5-7 września 2017 r. (64,5 godziny)	I miejsce (równorzędne)
Komisja Europejska, Eurostat	EU Big Data Hackathon	marzec 2019 r.	I miejsce
Komisja Europejska, Eurostat	EU Big Data Hackathon	marzec 2021 r.	II miejsce
Komisja Europejska, Eurostat	EU Big Data Hackathon	marzec 2023 r.	VI miejsce (nagrodzone trofeum)

## Recenzje artykułów naukowych

W ramach działalności naukowej od wielu lat pełnię funkcję recenzenta artykułów naukowych, uczestnicząc w procesie oceny merytorycznej i metodologicznej prac w czasopiśmie naukowych o zasięgu krajowym i międzynarodowym. Do tej pory wykonałem ponad 100 recenzji, obejmujących zarówno periodyki indeksowane w bazach Web of Science

i SCOPUS, jak również czasopisma z przyznanym Impact Factor. Recenzowane przeze mnie artykuły dotyczyły szerokiego spektrum zagadnień badawczych, w tym m.in. informatyki, analizy danych, metodologii badań empirycznych, zastosowań big data w naukach ekonomicznych, zarządzaniu danymi oraz zastosowań technologii informacyjnych w badaniach społecznych i ekonomicznych.

Systematyczne pełnienie funkcji recenzenta umożliwiło mi nie tylko krytyczną ocenę jakości naukowej i metodologicznej publikacji, lecz także aktywny udział w procesie podnoszenia standardów publikacyjnych oraz weryfikacji innowacyjności podejścia badawczego autorów. Moja działalność recenzencka przyczyniła się do promowania najlepszych praktyk w nauce, kształtowania jakości badań empirycznych oraz popularyzacji nowatorskich metod badawczych w społeczności międzynarodowej. Recenzowanie artykułów naukowych traktuję również jako integralny element mojej aktywności naukowej, który pozwala mi na ciągłe aktualizowanie wiedzy, monitorowanie trendów badawczych oraz identyfikowanie kierunków rozwoju w obszarze nauk ekonomicznych, informatycznych i społecznych.



.....  
(podpis wnioskodawcy)