

Autoreferat

w postępowaniu o nadanie stopnia doktora habilitowanego w
dziedzinie nauk społecznych

w dyscyplinie ekonomia i finanse

Imię i nazwisko:

Wojciech Roszka

Tytuł osiągnięcia naukowego:

*Ekonomiczna analiza nauki w Polsce: produktywność i
nierówności w świetle zintegrowanych mikrodanych
administracyjnych i bibliometrycznych*

Miejsce zatrudnienia:

Uniwersytet Ekonomiczny w Poznaniu
Instytut Informatyki i Ekonomii Ilościowej
Katedra Statystyki

Profil ORCID: <https://orcid.org/0000-0003-4383-3259>

Poznań, 14 października 2025

Spis treści

1	Imię i nazwisko	1
2	Posiadane dyplomy i stopnie naukowe	1
3	Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych	1
4	Omówienie osiągnięcia naukowego, o którym mowa w art. 219 ust. 1 pkt 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce	2
4.1	Wprowadzenie do tematyki	6
4.2	Rozwój infrastruktury danych i metod analizy mikropoziomowej	9
4.2.1	Ograniczenia wcześniejszych rozwiązań	10
4.2.2	Mój wkład: autorskie podejście do integracji danych	11
4.2.3	Etapy integracji danych POL-on-Scopus	12
4.2.4	Rezultat i zastosowania	14
4.2.5	Znaczenie infrastrukturalne	16
4.2.6	Ograniczenia	18
4.2.7	Podsumowanie	19
4.3	Analiza struktury i dynamiki produktywności naukowej	20
4.3.1	Zakres i unikalność podejścia	21
4.3.2	Badania nad nierównościami w nauce	21
4.3.3	Główne wyniki i ich znaczenie	23
4.3.4	Precyzja i ograniczenia analiz	24
4.3.5	Zastosowania operatu poza analizą produktywności	26
4.3.6	Znaczenie dla badań międzynarodowych	27
4.3.7	Podsumowanie	28
4.4	Badania nad nierównościami w nauce	29
4.4.1	Znaczenie badań nad nierównościami w nauce	33
4.4.2	Nierówności płciowe w systemie nauki	35
4.4.3	Wiek, staż akademicki i nierówności pokoleniowe	36
4.4.4	Wzorce współpracy i dostęp do zasobów	37
4.4.5	Implikacje dla polityki naukowej	39
5	Aktywność naukowa poza jednostką organizacyjną zatrudnienia	43
5.1	Projekty realizowane na Uniwersytecie im. Adama Mickiewicza w Poznaniu	43
5.1.1	NCN MAESTRO „Program międzynarodowych badań porównawczych szkolnictwa wyższego” (2012–2017)	43

5.1.2	Doskonałość naukowa: konkurencyjność, mierzalność, umiędzynarodowienie (DIALOG, 2016–2019)	44
5.1.3	Uczelnia badawcza 2.0. Finansowanie, zarządzanie i polityka kadrowa (DIALOG, 2020–2023)	45
5.1.4	Polscy Naukowcy 2022	45
5.1.5	Ekspertyza bibliometryczna na zlecenie kierownictwa programu IDUB (marzec 2025)	47
5.2	Projekt regionalny na zlecenie administracji publicznej (2018 – 2019) . . .	48
5.3	Ekspertyza dla AQUANET S.A. w ramach Spółki Celowej Uniwersytetu Ekonomicznego w Poznaniu (2015)	49
5.4	Analiza publikacji w MDPI w jednostkach Polskiej Akademii Nauk (marzec 2025)	50
5.5	Działalność ekspercka na zlecenie Stowarzyszenia Metropolia Poznań (2016)	51
5.6	Ekspertyza dotycząca bazy danych RAD-on na zlecenie Ośrodka Przetwarzania Informacji – Państwowego Instytutu Badawczego (2024)	52
5.7	Wiceprzewodniczący Kapituły Programu Ariadna (od 2022)	53
5.8	Współpraca interdyscyplinarna w zakresie analizy danych biologicznych (2024–2025)	54
5.9	Staż naukowy na University of Southampton (2016)	55
5.10	Kierowanie komponentem międzynarodowym – Visegrad Fund	55
6	Odniesienie do kryteriów habilitacyjnych i dalsze plany badawcze	57
7	Pozostała działalność dydaktyczna akademicka i organizacyjna	59
7.1	Działalność dydaktyczna	59
7.2	Nagrody i wyróżnienia	60
7.3	Zaangażowanie organizacyjne w środowisku statystycznym	60
7.4	Wyjazdy szkoleniowe, konferencyjne i służbowe	61
8	Wykaz publikacji po uzyskaniu stopnia doktora (2013–2025)	62

1 Imię i nazwisko

Wojciech Roszka

2 Posiadane dyplomy i stopnie naukowe

2006 Tytuł zawodowy magistra nauk ekonomicznych w zakresie informatyki i statystyki. Praca magisterska pt. „*Analiza zdarzeń jako forma testowania średniej efektywności informacyjnej rynku kapitałowego*”; promotor: prof. dr hab. Witold Jurek; Wydział Ekonomii, Akademia Ekonomiczna w Poznaniu.

2013 Stopień naukowy doktora nauk ekonomicznych w dyscyplinie statystyka. Praca doktorska pt. „*Statystyczna integracja danych w badaniach społeczno-ekonomicznych*”; promotor: prof. dr hab. Elżbieta Gołata; recenzenci: prof. dr hab. Małgorzata Rószkiewicz (Szkoła Główna Handlowa), prof. dr hab. Marek Walesiak (Uniwersytet Ekonomiczny we Wrocławiu); Wydział Informatyki i Gospodarki Elektronicznej, Uniwersytet Ekonomiczny w Poznaniu. Dyplom z wyróżnieniem.

3 Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych

2021–obecnie Data scientist w Centrum Studiów nad Polityką Publiczną (Institute for Advanced Studies in Social Sciences and Humanities), Uniwersytet im. Adama Mickiewicza w Poznaniu.

2014–obecnie Adiunkt w Katedrze Statystyki, Uniwersytet Ekonomiczny w Poznaniu.

2013–2020 Post-doc w Centrum Studiów nad Polityką Publiczną (Institute for Advanced Studies in Social Sciences and Humanities), Uniwersytet im. Adama Mickiewicza w Poznaniu.

2013–2014 Asystent w Katedrze Statystyki, Uniwersytet Ekonomiczny w Poznaniu.

2009–2013 Doktorant w Katedrze Statystyki, Uniwersytet Ekonomiczny w Poznaniu.

2009–2011 Specjalista w Urzędzie Statystycznym w Poznaniu; członek zespołu badawczego ESSnet „Data Integration” oraz zespołu zajmującego się przygotowaniem rejestrów administracyjnych do wykorzystania w Narodowym Spisie Powszechnym Ludności i Mieszkań 2012.

4 Omówienie osiągnięcia naukowego, o którym mowa w art. 219 ust. 1 pkt 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce

Osiągnięcie naukowe lokuję w dyscyplinie *ekonomia i finanse*, w obszarze badań określanym międzynarodowo jako *economics of science* (Stephan, 1996), dalej określanym jako *ekonomiczna analiza nauki*. Nurt ten zajmuje się badaniem funkcjonowania nauki jako systemu produkcji wiedzy, w tym m.in. produktywności, współpracy, mobilności oraz nierówności w środowisku akademickim. Jest to podobzdar ekonomii analizujący mechanizmy alokacji zasobów, zachowania badaczy i instytucji oraz efektywność procesów generowania wiedzy.

W moim ujęciu *economics of science* obejmuje również analizę nierówności strukturalnych i równości szans w środowisku akademickim. Część badań wpisuje się w nurt *heterodox economics*, łącząc podejście ekonomiczne z perspektywą równości płci, różnorodności i sprawiedliwości instytucjonalnej. Analizowane zjawiska – w tym różnice płciowe w produktywności, strukturze współpracy i awansie zawodowym – sytuują się w obszarze badawczym na styku ekonomii, socjologii nauki i studiów nad równością płci (Etzkowitz and Kemelgor, 2001; Ceci and Williams, 2011). Nierówności w nauce mają w tym ujęciu nie tylko wymiar etyczny, ale również ekonomiczny – ograniczają efektywność wykorzystania kapitału ludzkiego i innowacyjnego potencjału systemu badawczego (Stephan and Levin, 1996).

Tematycznie osiągnięcie wpisuje się w międzynarodowy nurt badań określanych jako *science of science*, czyli ilościowych badań nad nauką i szkolnictwem wyższym (w polskim dyskursie spotyka się określenia: „naukoznawstwo ilościowe” lub „badania nad nauką”). Jest to dynamicznie rozwijający się kierunek łączący elementy ekonomii, zarządzania, informatyki i socjologii, którego celem jest zrozumienie dynamiki produktywności, współpracy i wpływu badań z wykorzystaniem danych dużej skali. W polskim systemie klasyfikacji badania tego typu sytuują się przede wszystkim w *ekonomii i finansach*, a częściowo także w socjologii (socjologii nauki). W moich pracach korzystam z metod naukometrycznych jako instrumentarium wspierającego analizy ekonomiczne, przy czym całość badań jednoznacznie lokuję w ekonomicznej analizie nauki, świadomie wykorzystując interdyscyplinarne podejście charakterystyczne dla *science of science*.

Istotnym kontekstem moich badań jest heterodoksyjny nurt ekonomii nauki, w którym obok efektywności analizuje się dystrybucję zasobów i równość dostępu do szans akademickich. Analizy różnic płciowych w produktywności naukowej, strukturze współpracy i

dynamice karier wpisują się w ten kierunek badań, łączący ilościowe metody ekonomiczne z krytycznym spojrzeniem na instytucjonalne uwarunkowania kariery naukowej. Zastosowane wskaźniki – takie jak *gender solo research gap* czy *index of gender homophily* – umożliwiają empiryczne uchwycenie mechanizmów opisanych w literaturze o stratyfikacji płciowej w nauce (Rossiter, 1993; Xie and Shauman, 2003; Caprile et al., 2012).

Mój wkład autorski

We wszystkich publikacjach wchodzących w skład osiągnięcia jestem współautorem wraz z prof. dr. hab. Markiem Kwiekiem. Mój wkład miał charakter koncepcyjny, metodologiczny i empiryczny. Odpowiadałem za zaprojektowanie i wdrożenie unikalnej w skali kraju procedury integracji danych administracyjnych (POL-on) z bibliometrycznymi (Scopus), która łączyła podejście deterministyczne, probabilistyczne oraz manualną weryfikację. Efektem była trwała i replikowalna infrastruktura mikrodanych obejmująca ponad 25 tysięcy naukowców, stanowiąca fundament dalszych analiz.

Na gruncie tej infrastruktury zastosowałem i zaadaptowałem zaawansowane techniki ilościowe, m.in. regresje logistyczne binarne i frakcyjne, modele panelowe z efektami stałymi, analizy kohortowe i sieciowe. Zaproponowałem wprowadzenie nowych kategorii i wskaźników, takich jak wiek akademicki, klasyfikacja 10/50, *gender solo research gap* czy *index of gender homophily*, które umożliwiły nowatorskie badania nad koncentracją dorobku, dynamiką karier oraz nierównościami w dostępie do kapitału akademickiego i sieci współpracy. Część analiz wpisywała się w nurt *ekonomicznej analizy nauki* (ang. *economics of science*) oraz w tradycję ekonomii heterodoksyjnej, w której kwestie równości płci i dystrybucji szans stanowią integralny element analizy efektywności.

Mój wkład miał również wymiar teoriiwórczy. Rozwinałem ramy analityczne ekonomicznej analizy nauki, wprowadzając m.in. rozróżnienie pomiędzy „produktywnością potencjalną” a „produktywnością zrealizowaną” oraz relacyjne ujęcie efektywności (nakłady–wyniki). Łączyłem uzyskane wyniki z literaturą międzynarodową i polskimi uwarunkowaniami instytucjonalnymi, formułując wnioski aplikacyjne dla polityki naukowej – dotyczące koncentracji dorobku, mobilności międzyklasowej, wsparcia młodszych kohort badaczy oraz równości płciowej w nauce.

Mój wkład we wszystkich publikacjach wynosi około 50%, obejmując zarówno stronę koncepcyjną i metodologiczną, jak i pełny proces analityczny – od formułowania hipotez, przez estymację modeli, po interpretację i rekomendacje.

Zintegrowane dane i metodologia

Kluczowym elementem mojego wkładu było zaprojektowanie i wdrożenie **metodologii integracji danych administracyjnych** z systemu POL-on (OPI) z danymi bibliometrycznymi z bazy Scopus z zastosowaniem podejścia *probabilistic record linkage*. Metodologia obejmowała etap koncepcyjny, dobór i adaptację algorytmów porównywania tekstów, implementację procedur dopasowania i walidacji oraz weryfikację jakości wyników – w

tym oszacowanie odsetka błędnych przyporządkowań (fałszywie pozytywnych) i odsetka braków w przyporządkowaniach (fałszywie negatywnych). Integracja była konieczna z uwagi na brak wspólnych identyfikatorów osobowych i wymagała zaawansowanych technik łączenia rekordów oraz logicznej walidacji.

Opracowana metodologia jest pierwszym tego typu rozwiązaniem zastosowanym w Polsce na tak dużą skalę. Jej efektem było powstanie **zintegrowanego operatu mikro danych o polskiej nauce (POL-on × Scopus)** obejmującego ponad 25 tysięcy naukowców, który ma charakter infrastruktury badawczej. Operat ten stanowił podstawę analiz zawartych w zgłoszonym cyklu i może być wykorzystywany w kolejnych projektach badawczych, w tym do porównań międzynarodowych oraz monitorowania polityki naukowej. Co istotne, baza wymaga **ciągłego uaktualniania i harmonizacji**, aby zachować aktualność i wartość analityczną. Proces ten – obejmujący przejrzystość, pełne pokrycie populacji docelowej, kompletność informacyjną, porównywalność w czasie, aktualność, kontrolę błędów oraz ochronę poufności – jest zgodny z rekomendacjami dotyczącymi jakości danych statystycznych (Eurostat, 2014, 2011; Zhang, 2012).

Przez określenie „pełnopopulacyjne” rozumiem pełne pokrycie **populacji badaczy widocznych w Scopus** (*Scopus-visible*) zlinkowanych z rejestrem POL-on na moment referencyjny bazy OPI. Wnioski odnoszę konsekwentnie do tego segmentu populacji. Ewentualne różnice pokrycia dziedzin (zwłaszcza w dziedzinach humanistyczno-społecznych; *Humanities and Social Sciences* (HSS)) ograniczam projektowo poprzez: (i) analizy warstwowe, (ii) ujęcie dyscyplin/okresów poprzez efekty stałe, (iii) testy wrażliwości raportowane w pracach z cyklu.

Zakres tematyczny cyklu

Cykl wnosi wkład do *ekonomii i finansów*, w szczególności do *ekonomicznej analizy nauki* (ang. *economics of science*) oraz nurtu *science of science*. Choć korzysta z teorii wypracowanych również w ramach socjologii, zastosowane podejścia mają charakter ekonomiczny i statystyczny, a wyniki rozwijają wiedzę w dyscyplinie *ekonomia i finanse*. Wykorzystują narzędzia ekonomii empirycznej (mikrodane, modele ekonometryczne, podejścia kohortowe) do badania mechanizmów sukcesu i selekcji w systemie nauki, odwołując się do literatury z *economics of science* oraz socjologii nauki (Merton, 1968; Bourdieu, 1975; Cole, 1973; Whitley, 2000).

Skoncentrowano się m.in. na zróżnicowaniu współpracy naukowej ze względu na płeć (McPherson et al., 2001; Bozeman and Corley, 2004), intensywności współpracy międzynarodowej i dostępie do globalnych sieci (Crane, 1972; Wagner, 2008), rozkładzie i trwałości produktywności (Lotka, 1926; de Solla Price, 1976; Allison et al., 1982), wpływie wieku biologicznego i akademickiego (Leahey, 2006; Hunter and Leahey, 2010), mobilności między klasami produktywności oraz roli publikacji indywidualnych i zespołowych.

Wyniki dostarczają *dowodów empirycznych* na działanie mechanizmów sukcesu akademickiego i *rozwijają* kategorie analityczne ekonomicznej analizy nauki (alokacja zasobów,

dystrybucja produktywności, efekty polityk publicznych) poprzez testowanie hipotez na pełnopolulacyjnych mikrodanych — *Scopus-visible* zintegrowanych z POL-on.

Zestawienie publikacji wchodzących w skład cyklu¹:

1. Kwiek, M., **Roszka, W.** (2021). Gender Disparities in International Research Collaboration: A Study of 25,000 University Professors. *Journal of Economic Surveys*, **35**(5), 1344–1380. (Q1; MN: 140)
2. Kwiek, M., **Roszka, W.** (2021). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, **15**(3), 1–38, art. no. 101171. (Q1; MN: 140)
- 2a. Kwiek, M., **Roszka, W.** (2021). Dlaczego w nauce dominuje współpraca z mężczyznami: homofilia ze względu na płeć na przykładzie 25 000 naukowców. *Nauka*, nr 2. (MN: 20)
3. Kwiek, M., **Roszka, W.** (2022). Are female scientists less inclined to publish alone? The gender solo research gap. *Scientometrics*, **127**(3), 1697–1735. (Q1; MN: 100)
4. Kwiek, M., **Roszka, W.** (2022). Academic vs. biological age in research on academic careers: A large-scale study with implications for scientifically developing systems. *Scientometrics*, **127**(6), 3543–3575. (Q1; MN: 100)
- 4a. Kwiek, M., **Roszka, W.** (2022). Globalny obieg naukowy a wiek w nauce: analiza 20 000 polskich naukowców. *Nauka*, nr 3. (MN: 20)
5. Kwiek, M., **Roszka, W.** (2024). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*, **87**(3), 519–549. (Q1; MN: 140)
- 5a. Kwiek, M., **Roszka, W.** (2023). Zawsze wysoce produktywni? Zawsze nisko produktywni? Dynamiczne ujęcie dorobku naukowego profesorów tytularnych wykorzystujące klasy produktywności. *Nauka*, nr 1. (MN: 20)
6. Kwiek, M., **Roszka, W.** (2025). Are scientists changing their research productivity classes when they move up the academic ladder? *Innovative Higher Education*, **50**(1), 329–367. (Q1; MN: 140)
7. Kwiek, M., **Roszka, W.** (2024). The young and the old, the fast and the slow: A large-scale study of productivity classes and rank advancement. *Studies in Higher Education*, **49**(11), 2036–2051. (Q1; MN: 140)

¹Objaśnienia: Q — kwartył w JCR 2025; MN — liczba punktów MEiN. Dla czasopism wielodyscyplinarnych przyjmuję kategorię wiodącą raportowaną przez JCR.

8. Kwiek, M., **Roszka, W.** (2024). Top research performance in Poland over three decades: A multidimensional micro-data approach. *Journal of Informetrics*, **18**(4), 1–16, art. no. 101595. (Q1; MN: 140)
- 8a. Kwiek, M., **Roszka, W.** (2024). Reguła 10/50. Analiza dorobku 150 000 polskich naukowców pokazuje trwałość systemowych nierówności w nauce. *Nauka*, nr 4. (MN: 20)

Uwaga dotycząca pozycji oznaczonych „a”. Publikacje z przyrostkiem „a” są polskojęzycznymi wersjami lub interpretacyjnymi rozwinięciami analiz opublikowanych po angielsku (czasopisma JCR). Ich celem jest **transfer wiedzy na polski rynek naukowy**. Pozycje te **nie wchodzą w skład zgłaszanego cyklu osiągnięcia naukowego i nie są wliczane** do zagregowanych wskaźników oddziaływania raportowanych dla rdzenia cyklu (np. Q1, IF, FWCI); w zestawieniach parametrycznych wykazywane są odrębnie (np. „Nauka”: 20 pkt MEiN).

Podsumowanie

Zgłoszony cykl publikacji ukazał się w prestiżowych czasopismach międzynarodowych; **wszystkie artykuły** opublikowano w tytułach należących do najwyższego kwartyłu (Q1) w rankingach dziedzinowych (JCR/Scopus). Artykuły cechują się ponadprzeciętnym oddziaływaniem badawczym, co potwierdza wysoki wskaźnik Field-Weighted Citation Impact (FWCI = 3,48 w latach 2020–2025, SciVal, stan na 3 września 2025 roku).

Cykl wnosi, w mojej ocenie, istotny wkład do *ekonomii i finansów* w obszarze *ekonomicznej analizy nauki* (ang. *economics of science*) — obejmującej zagadnienia alokacji zasobów, dystrybucji produktywności i konsekwencji polityk publicznych — oraz, komplementarnie, do *socjologii nauki*. Potwierdza i rozwija tezy o kumulatywnej przewadze i stratyfikacji w polu akademickim, identyfikuje *gender solo research gap*, ukazuje znaczenie homofilii i struktury sieci współpracy, dowodzi, że wiek akademicki trafniej niż biologiczny oddaje trajektorie produktywności, oraz ujawnia rozbieżności między logiką awansów instytucjonalnych a logiką produktywności.

Wnioski formułowane są dla populacji badaczy *widocznych w Scopus* i zidentyfikowanych w POL-on; różnice pokrycia dziedzin oraz zróżnicowanie instytucjonalne są jawnie kontrolowane (warstwowanie, efekty stałe), co nie podważa konkluzji dotyczących koncentracji dorobku (10/50), trwałości pozycji elity, mobilności między klasami produktywności, relacji produktywność–wiek akademicki oraz wzorców współpracy krajowej i międzynarodowej.

4.1 Wprowadzenie do tematyki

Zgłoszony cykl publikacji lokuje w dyscyplinie *ekonomia i finanse*. Dotyczy on analizy funkcjonowania systemu nauki w Polsce z wykorzystaniem metod ilościowych typo-

wych dla ekonomii empirycznej: analizy danych panelowych, modeli ekonometrycznych oraz pełnopopulacyjnych mikrodanych. Celem cyklu jest uchwycenie struktury i dynamiki produktywności naukowej, identyfikacja mechanizmów jej zróżnicowania oraz ocena efektywności instytucjonalnej. Badania wpisują się w nurt *science of science*, rozwijany w literaturze międzynarodowej, w którym szczególną uwagę zwraca się na alokację zasobów, dystrybucję produktywności i znaczenie polityki publicznej dla rozwoju kapitału ludzkiego w nauce (Stephan, 2012; Jones, 2005; Moed, 2005). Najbardziej naturalnym umiejscowieniem zgłoszonego cyklu publikacji byłaby *naukometria* lub *scientometria* (tzw. *science of science*), stanowiąca międzynarodowo wyodrębnione pole badawcze. W polskiej klasyfikacji dyscyplin nie występują jednak takie kategorie, dlatego badania tego typu lokowane są w ramach istniejących dyscyplin. W moim przypadku ich właściwym i zgodnym z porządkiem prawnym umiejscowieniem jest *ekonomia i finanse*.

Cele badawcze cyklu obejmowały:

- stworzenie trwałego operatu analitycznego poprzez integrację danych administracyjnych (POL-on) i bibliometrycznych (Scopus) przy użyciu metodologii *probabilistic record linkage* (Herzog et al., 2007; Christen, 2012), co umożliwiło analizy na poziomie mikrodanych,
- analizę zróżnicowań płciowych we współpracy międzynarodowej, homofilii płciowej i publikacjach jednoautorskich jako czynników różnicujących wykorzystanie kapitału ludzkiego w nauce,
- identyfikację wpływu wieku biologicznego i akademickiego na trajektorie produktywności, w nawiązaniu do koncepcji *life cycle* w badaniach kapitału ludzkiego,
- ocenę trwałości i mobilności *między klasami produktywności*, testując tezy o efektach kumulacji przewagi i hipotezy „sacred spark”,
- analizę koncentracji dorobku i jego korelatów instytucjonalno–demograficznych w świetle klasycznych rozkładów Lotki i Pareto (Lotka, 1926; de Solla Price, 1976).

Kluczowym etapem badań było stworzenie zintegrowanej bazy danych obejmującej ponad 25 tysięcy naukowców aktywnych publikacyjnie, widocznych w bazie Scopus i powiązanych z rejestrem POL-on (stan na listopad 2017 r.). Dla przejrzystości: *pełnopopulacyjność* odnosi się do całego segmentu badaczy *widocznych w Scopus* i obecnych w POL-on w dacie referencyjnej, a nie do całej populacji wszystkich badaczy w Polsce. Zróżnicowaną widoczność dziedzin w Scopus uwzględniam projektowo (warstwowanie, efekty stałe dla dyscyplin, alternatywne progi klasyfikacyjne). W odróżnieniu od większości badań opartych wyłącznie na danych bibliometrycznych, cykl wykorzystuje dane mikropoziomowe — łączące informacje o osobach, afiliacjach, dziedzinach i instytucjach. Operat ten ma charakter infrastrukturalny i wymaga stałego uaktualniania oraz harmonizacji, zgodnie

z rekomendacjami Eurostatu dotyczącymi jakości danych (Eurostat, 2011, 2014; Zhang, 2012).

Pod względem metodologicznym cykl wnosi oryginalny wkład w rozwój narzędzi badawczych i operatów danych. Połączenie źródeł administracyjnych z bibliometrycznymi umożliwiło analizy wcześniej niemożliwe na danych krajowych. Metody i operaty stanowią punkt wyjścia do dalszych badań międzynarodowych.

W części empirycznej cyklu:

- wykazano, że kobiety rzadziej niż mężczyźni angażują się we współpracę międzynarodową oraz częściej współpracują w ramach homofilii płciowej — co ogranicza dostęp do zróżnicowanych sieci akademickich,
- zidentyfikowano *gender solo research gap*, interpretowaną jako nowy czynnik ograniczający pełne wykorzystanie potencjału badaczek,
- ukazano, że wiek akademicki (liczony od daty pierwszej publikacji) trafniej niż wiek biologiczny oddaje trajektorie produktywności, co ma znaczenie dla oceny inwestycji w kapitał ludzki,
- potwierdzono trwałość pozycji najbardziej produktywnych naukowców (analizy kohortowe), zgodnie z teorią kumulatywnej przewagi,
- zaproponowano typologie ścieżek kariery („the young and the old, the fast and the slow”), pokazując znaczenie timingu awansów instytucjonalnych,
- przeanalizowano relację między awansem a zmianami klas produktywności — wskazując, że logika instytucjonalna nie zawsze pokrywa się z logiką efektywności badawczej,
- wykazano, że 10% najbardziej produktywnych naukowców odpowiada za ponad 50% całego dorobku, potwierdzając ciężkoogonowe rozkłady produktywności.

Zgłoszone publikacje można podzielić na dwie grupy: (1) rozwój metodologii i infrastruktury danych; (2) analizy empiryczne dotyczące produktywności, nierówności i koncentracji dorobku.

Pozycjonowanie teoretyczne (ekonomia i finanse). Pracę lokuję w *ekonomii i finansach*, w obszarze badawczym *ekonomicznej analizy nauki* (ang. *economics of science*). Analizuję mechanizmy dystrybucji produktywności badawczej, koncentracji dorobku oraz instytucjonalnych uwarunkowań efektywności systemu nauki. Odwołuję się do klasycznych praw rozkładów produktywności (Lotka, Pareto, Price), teorii kapitału ludzkiego oraz nurtu badań nad endogenicznym wzrostem gospodarczym, w którym aktywność naukowa stanowi kluczowy czynnik rozwoju (Stephan, 2012; Jones, 2005; Auerbach, 2019).

Zintegrowane mikrodane umożliwiają analizy na poziomie mikro (badacze), mezo (instytucje) i makro (kohorty), co pozwala empirycznie testować hipotezy dotyczące koncentracji dorobku, nierówności płciowych i wiekowych oraz ich wpływu na efektywność i konkurencyjność systemu nauki.

4.2 Rozwój infrastruktury danych i metod analizy mikropoziomowej

Pierwszy obszar mojego osiągnięcia naukowego obejmuje zaprojektowanie, przeprowadzenie i wdrożenie metodologii integracji danych administracyjnych z systemu POL-on (OPI) z danymi bibliometrycznymi z bazy Scopus, opartej na podejściu probabilistycznym (*probabilistic record linkage*). Celem było stworzenie trwałego, zintegrowanego operatu mikrodanych, który umożliwia analizy ekonomiczne dotyczące funkcjonowania systemu nauki, w tym produktywności badawczej, efektywności instytucjonalnej oraz dystrybucji zasobów w szkolnictwie wyższym.

Dotychczasowy brak takiego narzędzia istotnie ograniczał zakres i jakość analiz ilościowych prowadzonych na krajowych danych. Badania nad produktywnością, współpracą czy mobilnością były oparte na pojedynczych źródłach — administracyjnych lub bibliometrycznych — co uniemożliwiało jednoczesne uwzględnienie kluczowych uwarunkowań demograficznych (wiek, płeć), instytucjonalnych (typ uczelni, stanowisko) i publikacyjnych (liczba i jakość publikacji, współautorstwo). W efekcie wiele analiz miało charakter fragmentaryczny, a ich wyniki nie oddawały złożoności mechanizmów wpływających na efektywność i konkurencyjność polskiego systemu nauki (Fortin and Currie, 2013; Gingras, 2008; Sugimoto et al., 2013).

Opracowana przeze mnie metodologia integracyjna adresuje ten problem poprzez zastosowanie zaawansowanych technik łączenia danych bez wspólnego identyfikatora — tzw. *probabilistic record linkage* (Fellegi and Sunter, 1969; Winkler, 2006; Christen, 2012). Cały proces został zaprojektowany, przetestowany i zaimplementowany w środowisku R (m.in. `RecordLinkage`, `fastLink`, `stringdist`, `data.table`, `dplyr`), z uwzględnieniem specyfiki polskich danych oraz typowych wyzwań tekstowych (zmiennosc zapisu nazwisk, aliasy jednostek, transliteracje, brak diakrytyków, niejednoznaczność afiliacji). Dzięki tej integracji możliwe stało się zbudowanie zbioru obejmującego ponad 25 tysięcy osób aktywnych publikacyjnie w polskim szkolnictwie wyższym, dla których połączono informacje administracyjne (zatrudnienie, stanowiska, dyscypliny) z dorobkiem publikacyjnym i wskaźnikami cytowalności.

Powstały operat stanowi pierwszy w Polsce tak szeroki i metodycznie udokumentowany zasób w obszarze badań nad systemem nauki z perspektywy ekonomii empirycznej (Fortunato et al., 2018; Lane et al., 2011a; Glänzel and Moed, 2013). Jego zastosowanie umożliwiło oryginalne analizy dotyczące m.in. trwałości produktywności, trajektorii

karier akademickich, różnic płciowych czy zróżnicowań instytucjonalnych. Operat zaprojektowano jako trwałą infrastrukturę badawczą o architekturze modularnej (zgodnej ze standardami FAIR), z możliwością rozszerzania o kolejne źródła (np. finansowanie grantowe, ORCID, Crossref/OpenAlex) oraz ciągłej aktualizacji i harmonizacji zgodnie z rekomendacjami jakości danych (Eurostat, 2014, 2011; Zhang, 2012).

W mojej ocenie zaprojektowana metodologia integracji oraz zbudowany na jej podstawie zbiór danych stanowią najważniejszy element mojego wkładu w rozwój metod analizy ekonomicznej nauki w Polsce. Ich znaczenie wykracza poza ramy pojedynczych publikacji — tworzą fundament analityczny umożliwiający systemowe monitorowanie, modelowanie i diagnozowanie procesów zachodzących w środowisku akademickim na poziomie mikro (badacze), mezo (instytucje) i makro (system krajowy).

4.2.1 Ograniczenia wcześniejszych rozwiązań

Dotychczasowe analizy prowadzone na danych krajowych opierały się najczęściej na pojedynczych źródłach – danych administracyjnych lub bibliometrycznych – co znacząco ograniczało ich użyteczność w badaniach ekonomicznych nad nauką. Dane administracyjne z systemu POL-on (OPI) zawierają informacje o statusie zatrudnienia, stopniach i tytułach naukowych, afiliacji instytucjonalnej, a także podstawowe dane demograficzne (płeć, data urodzenia). Jednak nie obejmują one żadnych informacji na temat dorobku publikacyjnego, cytowań czy współpracy międzynarodowej, które są kluczowe dla analizy produktywności i efektywności badawczej.

Z kolei bazy bibliometryczne – takie jak Scopus czy Web of Science – oferują szczegółowe dane o publikacjach, cytowaniach, metrykach czasopism (Impact Factor, percentyle Scopus) oraz typach współautorstwa (krajowe, międzynarodowe, indywidualne). Jednak źródła te nie zawierają informacji o przebiegu kariery zawodowej autora, jego statusie zatrudnienia, dyscyplinie czy typie instytucji naukowej. W rezultacie, analizy oparte wyłącznie na jednym z tych zasobów nie pozwalały na kompleksową ocenę produktywności i alokacji zasobów w systemie nauki (Moed, 2005; Glänzel and Moed, 2013; Huang et al., 2020; Fortunato et al., 2018).

Dodatkowym ograniczeniem było to, że dane bibliometryczne przypisują aktywność publikacyjną do identyfikatorów autorskich tworzonych automatycznie. Choć Scopus stosuje wewnętrzne Author ID, w praktyce jedna osoba może być reprezentowana przez kilka profili, szczególnie w przypadku niejednoznacznego zapisu nazwisk lub zmian afiliacyjnych. Dane administracyjne z POL-on nie zawierają informacji bibliometrycznych ani powiązań z systemami zewnętrznymi, co uniemożliwiało bezpośrednie przypisanie aktywności publikacyjnej do konkretnej osoby. Mimo rosnącej popularności systemów takich jak ORCID, ich pokrycie dla badaczy afiliowanych w Polsce w analizowanym okresie było niewystarczające. Konieczne było zatem zastosowanie metod probabilistycznego dopasowywania rekordów.

W efekcie, nawet najbardziej zaawansowane dotąd badania nad polską nauką były prowadzone w ramach sztucznego podziału: na analizy instytucjonalne (w oparciu o POL-on) oraz analizy bibliometryczne (w oparciu o Scopus lub WoS), bez możliwości ich wzajemnego powiązania. Taka separacja skutkowałą uproszczonym obrazem procesów i utrudniała testowanie hipotez ekonomicznych dotyczących produktywności, efektywności instytucjonalnej czy dystrybucji zasobów (Costas et al., 2010).

W Polsce brakowało dotąd infrastruktury danych umożliwiającej prowadzenie mikro-poziomowych analiz ilościowych w sposób porównywalny z rozwiązaniami stosowanymi w takich krajach jak Kanada, Holandia, Stany Zjednoczone czy Niemcy, gdzie integracja danych personalnych, finansowych i bibliometrycznych stała się standardem (Larivière and Gingras, 2010; van Leeuwen, 2006; Lane et al., 2011b; Sugimoto et al., 2013). W krajach tych rozwijane są zintegrowane operaty danych łączące informacje o zatrudnieniu, finansowaniu badań, dorobku publikacyjnym oraz cechach demograficznych, co umożliwia prowadzenie analiz wysokiej jakości i wspiera procesy decyzyjne w polityce publicznej (*evidence-informed policy making*) (Hicks et al., 2015; Ioannidis et al., 2016).

W tym kontekście luka analityczna w Polsce była szczególnie dotkliwa. Mimo istnienia dwóch dużych zasobów (POL-on i Scopus), brak wspólnego identyfikatora i niedostateczny rozwój metod integracji danych powodowały, że nie istniały narzędzia pozwalające na prowadzenie kompleksowych analiz produktywności i efektywności systemu nauki. Opracowana przeze mnie procedura integracji stanowi odpowiedź na te wyzwania – w kolejnych podrozdziałach przedstawiam jej założenia, implementację oraz rezultaty.

4.2.2 Mój wkład: autorskie podejście do integracji danych

Odpowiadając na przedstawione wcześniej ograniczenia, zaprojektowałem i samodzielnie wdrożyłem procedurę *probabilistycznego łączenia rekordów* (*probabilistic record linkage*), która umożliwiła integrację danych z dwóch niezależnych źródeł: systemu POL-on (OPI) oraz bazy Scopus. Byłem autorem zarówno koncepcji metodologicznej, jak i implementacji całego procesu. Obejmowało to przygotowanie danych wejściowych, konstrukcję miar podobieństwa, walidację wyników i budowę zbioru docelowego. Celem było stworzenie operatu mikrodanych o jakości pozwalającej na prowadzenie systematycznych badań empirycznych nad produktywnością i efektywnością systemu nauki w Polsce.

Zaprojektowane przeze mnie rozwiązanie nie wymagało obecności wspólnego identyfikatora (np. PESEL, ORCID, ID pracownika czy ID publikacyjnego), lecz identyfikowało rekordy tej samej osoby na podstawie zestawu cech tekstowych, logicznych i kontekstowych: nazwiska i imienia (w różnych wariantach), afiliacji, inicjałów kolejnych imion, informacji o współautorach, dyscypliny nauki, instytucji zatrudnienia czy historii publikacji. Kluczowym elementem było zastosowanie **blokowania według instytucji**, co istotnie zwiększyło precyzję dopasowań. Dodatkowo opracowałem mechanizm ważenia cech — największą wagę przypisywano zgodności współautorów oraz afiliacji, co pozwoliło lepiej

uchwycić kontekst badawczy.

Dopasowania klasyfikowano na podstawie zagregowanego wskaźnika prawdopodobieństwa, a przypadki niejednoznaczne rozstrzygano według reguł warunkowych łączących kilka cech. Dla obserwacji, w których automatyczne procedury nie były wystarczające, zastosowałem dodatkową walidację manualną. Dzięki temu udało się osiągnąć bardzo wysoką precyzję dopasowań przy akceptowalnym poziomie niedopasowań, co potwierdziła walidacja ekspercka i obliczone wskaźniki jakości (precision, recall, F1-score). Relatywnie wyższy wskaźnik fałszywych niedopasowań wynikał z częstych przypadków relacji jeden–do–wielu po stronie Scopus — tej samej osobie odpowiadało kilka profili autorских.

Efektom procedury był operat obejmujący ponad 25 tysięcy naukowców zatrudnionych w polskich uczelniach akademickich ze stopniem doktora lub wyższym. Jego jakość została potwierdzona w procesie walidacji (m.in. poprzez ocenę wskaźników błędów klasyfikacji FDR i FNR oraz zgodność z dodatkowymi źródłami informacji). Zbiór ten zaprojektowałem jako **trwałą, replikowalną i skalowalną** infrastrukturę analityczną, umożliwiającą dalsze badania nad produktywnością, współpracą, nierównościami i mobilnością w nauce, a także zapewniającą możliwość porównań międzynarodowych. Dokumentacja kodu, parametrów i decyzji algorytmicznych sprawia, że operat może być rozwijany i ponownie wykorzystywany.

Zaproponowane rozwiązanie było pierwszym w Polsce przykładem **transparentnej integracji danych administracyjnych i bibliometrycznych** w skali całego systemu nauki. Odpowiadałem za koncepcję, implementację, testy, walidację oraz dalsze wykorzystanie operatu w analizach ekonometrycznych (m.in. regresje panelowe, modele frakcyjne, analizy kohortowe) prezentowanych w publikacjach cyklu.

Mój wkład miał również wymiar teoriiotwórczy: procedura integracji nie była wyłącznie operacją techniczną, lecz **nową koncepcją metodologiczną** w badaniach nad nauką w Polsce. Po raz pierwszy stworzono spójną ramę łączenia danych administracyjnych z bibliometrycznymi w skali całego systemu. Zdefiniowałem kryteria jakości dopasowań, zaproponowałem sposób ważenia cech oraz mechanizmy blokowania i walidacji, które mają charakter uniwersalny i mogą być stosowane w innych krajach i dyscyplinach. W ten sposób integracja danych stała się nie tylko narzędziem technicznym, ale **fundamentem nowego podejścia badawczego**, umożliwiającego systematyczne analizy produktywności i nierówności akademickich w perspektywie mikropoziomowej.

4.2.3 Etapy integracji danych POL-on–Scopus

Zaprojektowana i przeprowadzona przeze mnie procedura integracji składała się z czterech głównych etapów. Jej celem było uzyskanie wysokiej jakości dopasowań między rekordami z dwóch niezależnych źródeł — systemu POL-on (OPI) oraz bazy Scopus — przy uwzględnieniu ograniczeń typowych dla danych administracyjnych i bibliometrycznych

(brak wspólnych identyfikatorów, błędy zapisu, transliteracje, aliasy instytucji). Integracja stanowiła warunek niezbędny do prowadzenia analiz ekonomicznych nad produktywnością naukową, efektywnością instytucjonalną i alokacją zasobów w systemie szkolnictwa wyższego. Całość zaimplementowano w języku R (rdzeniowo `fastLink`; uzupełniająco m.in. `RecordLinkage`, `stringdist`, `data.table`).

Etap 1: Przygotowanie danych wejściowych: Znormalizowano i oczyszczono dane z POL-on i Scopusa. Ujednolicono zapis nazwisk (warianty z i bez diakrytyków) oraz nazw instytucji (np. „UAM” → „Uniwersytet im. Adama Mickiewicza”). Usunięto obserwacje niekompletne. Każdej osobie w POL-on nadano unikalny identyfikator wewnętrzny. Zastosowano reguły harmonizacji pól tekstowych i kalendarzowych (formaty dat, skróty, transliteracje).

Etap 2: Generowanie par kandydatów i dopasowanie probabilistyczne: Wykorzystano klasyczny model Fellegiego–Suntera z *blokowaniem* według instytucji (porównania wyłącznie w obrębie tej samej uczelni). Obliczano miary podobieństwa łańcuchów znaków (m.in. odległość Jaro–Winklera) dla nazwisk, pierwszych i — jeśli występowały — drugich imion. Poszczególnym cechom przypisano wagi, a następnie wyliczono zagregowany wskaźnik dopasowania (*score*) w przedziale $[0, 1]$. Blokowanie znacząco ograniczyło przestrzeń porównań bez utraty pokrycia.

Etap 3: Klasyfikacja i selekcja dopasowań: Na podstawie progu prawdopodobieństwa $\xi_{ij} \geq 0,85$ pary rekordów uznawano za dopasowane; pozostałe — za niedopasowane. Przypadki niejednoznaczne wykluczono zgodnie z konserwatywnymi regułami decyzyjnymi. Dla rekordów z wieloma potencjalnymi dopasowaniami zastosowano reguły rozstrzygające (ważona zgodność współautorów, dyscypliny i afiliacji).

Etap 4: Walidacja ekspercka i ocena jakości: Ręczna ocena losowej próby 500 par potwierdziła wysoką jakość dopasowań: *precision* przekroczyła 97%, *recall* osiągnął 94%, a *F1-score* wyniósł 0,96. Dla progu $\xi_{ij} \geq 0,85$ oszacowano również wskaźniki populacyjne: **FDR** (ang. *false discovery rate*, fałszywe dopasowania) = 0,21% oraz **FNR** (ang. *false non-match rate*, fałszywe niedopasowania) = 39,91%. Relatywnie wyższy FNR wynikał z częstych przypadków relacji jeden–do–wielu po stronie Scopusa: ta sama osoba z POL-on mogła odpowiadać kilku profilom (Scopus Author ID) wskutek wariantów zapisu („ó”/„o”, „ż”/„z”), przestawień liter, różnic afiliacyjnych czy transliteracji. Najczęściej obserwowano *profil główny* (z większością publikacji) oraz *profile poboczne* (z pojedynczymi pozycjami). Uzyskane wskaźniki jakości są porównywalne z rezultatami najlepszych projektów integracyjnych w krajach OECD (Winkler,

2006; Herzog et al., 2007; Christen, 2012; Enamorado et al., 2019), co potwierdza ich przydatność dla analiz porównawczych.

Wszystkie operacje przeprowadzono w środowisku R z pełną dokumentacją i możliwością replikacji (kontrola wersji kodu, *seed* dla komponentów losowych, dziennik decyzji algorytmicznych). Końcowy operat — obejmujący 32 937 unikalnych osób i 38 750 połączeń z profilami Scopus — zapisano w formacie zgodnym ze standardami FAIR (Findable, Accessible, Interoperable, Reusable) i przystosowano do dalszych analiz w R, SPSS, Stata oraz Excelu. Na potrzeby analiz badawczych wyodrębniono naukowców, którzy opublikowali **co najmniej jeden artykuł w czasopiśmie w latach 2009–2018**. Ostateczny operat objął 25 463 osoby, które w tym okresie opublikowały łącznie 158 743 artykuły. Populacja ta została **celowo zredukowana** do badaczy, których rekordy udało się zintegrować oraz którzy posiadali **stopień doktora lub wyższy**, co zapewniało spójność porównań i koncentrację analiz na pracownikach akademickich aktywnych w polskim szkolnictwie wyższym. Zbiór został przygotowany z myślą o zastosowaniach w analizach ekonometrycznych produktywności, dystrybucji dorobku i efektywności instytucjonalnej.

4.2.4 Rezultat i zastosowania

W wyniku przeprowadzonej procedury integracyjnej powstał zintegrowany zbiór danych obejmujący 25 463 naukowców zatrudnionych w polskich uczelniach akademickich według stanu na listopad 2017 roku (moment referencyjny bazy OPI), którzy zostali zidentyfikowani jako autorzy artykułów w czasopismach indeksowanych w Scopus. Oznacza to, że operat obejmuje *pełnopopulację badaczy widocznych w Scopus* wśród pracowników instytucji akademickich w Polsce, przy czym populacja ta została celowo zredukowana do osób, których rekordy udało się zintegrować oraz które posiadały stopień doktora lub wyższy. Takie zawężenie zapewniało spójność porównań i koncentrację analiz na pracownikach naukowo-dydaktycznych aktywnych w badaniach, a tym samym — największe znaczenie z punktu widzenia *ekonomicznej analizy nauki* (ang. *economics of science*).

Każdy rekord w zintegrowanym zbiorze łączy informacje pochodzące z dwóch źródeł — systemu POL-on (OPI) oraz bazy Scopus — w trzech głównych wymiarach:

- **Dane demograficzne:** płeć (na podstawie klasyfikacji imienia lub przypisania w OPI), wiek biologiczny (wyliczany z daty urodzenia), wiek akademicki (liczony jako liczba lat od daty pierwszej publikacji w Scopus).
- **Dane instytucjonalne:** typ uczelni (akademicka, zawodowa, publiczna, niepubliczna), nazwa i kod jednostki zatrudniającej, stanowisko (adiunkt, profesor uczelni, profesor tytularny), stopień i tytuł naukowy, przypisana dyscyplina według klasyfikacji ASJC.

- **Dane bibliometryczne:** liczba publikacji, liczba cytowań (całkowita i średnia), liczba współautorów, udział publikacji powstałych *we współpracy międzynarodowej*, metryki czasopism (Impact Factor, percentylu Scopus), typ współautorstwa (indywidualne, krajowe, międzynarodowe).

Zbiór — będący efektem autorskiego podejścia do integracji — charakteryzuje się wysoką jakością, spójnością logiczną oraz gotowością do *analiz longitudinalnych*. Zaprojektowano go modułowo, z myślą o rozszerzeniach o kolejne zmienne i źródła (np. finansowanie grantowe, dane patentowe) oraz o *cyklicznej aktualizacji* w kolejnych okresach sprawozdawczych.

Zintegrowany operat został wykorzystany w szeregu analiz ekonometrycznych i statystycznych, które stały się podstawą artykułów [1]–[7], m.in.:

- estymacja modeli regresji logistycznej i frakcyjnej (prawdopodobieństwo udziału we współpracy międzynarodowej, prawdopodobieństwo publikowania solo, zależność wieku akademickiego i produktywności),
- analizy kohortowe — grupowanie według roku uzyskania doktoratu i badanie trajektorii produktywności w ujęciu pokoleniowym,
- rekonstrukcja indywidualnych ścieżek karier i identyfikacja typologii rozwoju (np. kariery intensywne i szybkie vs. rozciągnięte w czasie),
- analiza mobilności między klasami produktywności, ocena stabilności pozycji w czasie oraz wpływu zmian instytucjonalnych na produktywność,
- testowanie hipotez o związku między typem instytucji, płcią, wiekiem, dyscypliną naukową a poziomem dorobku publikacyjnego, w tym z wykorzystaniem *modeli wielopoziomowych*.

Operat umożliwia modelowanie zjawisk dotąd trudno uchwytnych — np. różnic w produktywności pomiędzy naukowcami w tym samym wieku biologicznym, lecz o różnym wieku akademickim, różnic w zachowaniach publikacyjnych między uczelniami publicznymi i niepublicznymi, czy identyfikacji ścieżek rozwoju odbiegających od liniowego wzorca awansu. Dzięki połączeniu danych z dwóch komplementarnych źródeł możliwe jest jednoczesne uchwycenie wielu poziomów analizy: mikro (naukowcy), mezo (instytucje) i makro (kohorty, system jako całość). Pozwala to osadzać zjawiska w pełniejszym kontekście oraz testować modele zgodne z hierarchiczną strukturą danych.

Co szczególnie istotne, powstały zbiór okazał się przydatny w tworzeniu i walidacji *nowych wskaźników ewaluacyjnych*, np.:

- produktywność skorygowana o wiek biologiczny lub wiek akademicki,

- produktywność w percentylach (np. top 20%, top 10%, top 5%, top 1%),
- udział publikacji w czasopismach z 90. lub 95. *percentylu Scopus*,
- wskaźniki widoczności międzynarodowej skorygowane o dziedzinę,
- indeksy nierówności dorobku (np. indeks Giniego, wykładnik rozkładu Pareto).

Wyniki uzyskane na podstawie operatu ujawniły *strukturalne* zróżnicowania w polskim środowisku naukowym — zarówno w dostępie do współpracy międzynarodowej, jak i w koncentracji dorobku w wąskiej elicie. Umożliwiły też identyfikację „niewidocznych” barier awansu, wynikających nie z dorobku, lecz z cech instytucjonalnych, płci czy typu uczelni. Tego typu wnioski nie byłyby możliwe do sformułowania na podstawie samego Scopusa lub samego POL-on, analizowanych oddzielnie.

Z perspektywy ekonomii i polityki nauki zintegrowany operat może pełnić funkcję infrastruktury wspierającej:

- konstrukcję bardziej sprawiedliwych i trafnych wskaźników ewaluacyjnych,
- monitorowanie efektów polityk instytucjonalnych (reformy ewaluacji, programy grantowe, działania projakościowe),
- planowanie działań wspierających rozwój różnych grup naukowców (zależnie od etapu kariery, instytucji, dyscypliny),
- porównania międzynarodowe z wykorzystaniem analogicznych zbiorów z Kanady, Holandii czy USA.

Dzięki swojej trwałości, możliwości aktualizacji i elastyczności analitycznej opracowany zbiór może stać się trwałym elementem infrastruktury wiedzy o polskiej nauce. Jego przygotowanie nie tylko umożliwiło realizację zgłoszonego cyklu badań, lecz także otworzyło nowe możliwości dalszej eksploracji zjawisk takich jak produktywność, nierówności, mobilność i efektywność systemu akademickiego z perspektywy ekonomii i finansów.

4.2.5 Znaczenie infrastrukturalne

Stworzony przeze mnie operat danych ma charakter **trwały, replikowalny i skalowalny**. Jego konstrukcja opiera się na przejrzystej, w pełni udokumentowanej metodologii, w której każdy etap — od oczyszczania i harmonizacji, przez klasyfikację dopasowań, po walidację jakości — został opisany w sposób umożliwiający ponowne zastosowanie w analogicznych środowiskach danych. Ze względu na charakter danych administracyjnych i ograniczenia dostępowe sam operat nie jest publicznie udostępniany; **publiczna i odtwarzalna** pozostaje natomiast metodologia, logika algorytmów i dokumentacja techniczna.

Operat pełni więc równocześnie funkcję *narzędzia analitycznego, infrastruktury danych* oraz *platformy rozwojowej* dla badań nad efektywnością i funkcjonowaniem systemu nauki w Polsce — i jest jednym z pierwszych w kraju tak systematycznie i metodycznie udokumentowanych połączeń danych administracyjnych i bibliometrycznych w skali całego systemu.

Dzięki rozbudowanemu zakresowi zmiennych — demograficznych (płeć, wiek), instytucjonalnych (typ uczelni, stanowisko, dyscyplina) oraz bibliometrycznych (produktywność, cytowalność, współautorstwo, widoczność międzynarodowa) — operat znajduje zastosowanie przede wszystkim w badaniach *ekonomicznej analizy nauki* (ang. *economics of science*), a także w praktyce polityki publicznej. Przykładowe aplikacje obejmują:

- ocenę skuteczności i selektywności systemów grantowych (związek finansowania z dorobkiem, widzialnością i współpracą),
- analizę efektywności alokacji zasobów i zwrotu z inwestycji w kapitał ludzki,
- śledzenie mobilności poziomej i pionowej (przejścia między instytucjami, awanse, zmiany dyscypliny),
- identyfikację barier strukturalnych (związanych z płcią, wiekiem, typem uczelni),
- analizę wpływu reform ewaluacyjnych i finansowych na wzorce publikacyjne i współautorstwo,
- prognozowanie dynamiki zatrudnienia, produktywności i współpracy.

Zastosowane podejście odpowiada na jedno z kluczowych wyzwań współczesnej *science of science* — **łączenie danych z wielu źródeł przy braku wspólnego identyfikatora** — szeroko dyskutowane w literaturze (Fortunato et al., 2018; Ziman, 2000; Lane et al., 2011a; Larivière and Gingras, 2010; Glänzel and Moed, 2013). Wyniki pokazują, że dzięki metodom probabilistycznym i wnikliwej walidacji można skutecznie przezwyciężyć to ograniczenie także w warunkach systemów w transformacji.

Operat zaprojektowano *modularnie* i zgodnie z zasadami FAIR (Findable, Accessible, Interoperable, Reusable): posiada słowniki zmiennych i kodów, historię zmian, reguły wersjonowania i plan aktualizacji. Struktura danych umożliwia łączenie z rejestrami grantów (NCN, NCBR), danymi o projektach europejskich (Horyzont), bazami cytowań (Crossref, OpenAlex) oraz identyfikatorami ORCID — otwierając drogę do rozwoju nowoczesnych ekosystemów danych wspierających *evidence-informed policy making*.

Z perspektywy międzynarodowej operat jest porównywalny metodologicznie z rozwiązaniami stosowanymi m.in. przez CWTS (Holandia), OST (Kanada) czy projektami STAR METRICS/UMETRICS (USA) (Larivière and Gingras, 2010; van Leeuwen, 2006; Lane et al., 2011b). Może także stanowić punkt odniesienia dla partnerstw w regionie Europy Środkowo-Wschodniej, gdzie podobne infrastruktury danych dopiero się kształtują.

Istotnym zastosowaniem operatu, wykraczającym poza analizę dorobku jednostek, jest wsparcie **zarządzania uczelniami jako organizacjami wiedzy**. Dostęp do rzetelnych mikro danych umożliwia projektowanie polityk instytucjonalnych, budowę narzędzi monitoringu aktywności, identyfikację potencjałów badawczych, modelowanie mobilności kadr oraz konstruowanie bardziej sprawiedliwych, przejrzystych wskaźników efektywności. W tym sensie operat wspiera przechodzenie od zarządzania intuicyjnego do zarządzania *opartego na danych* (evidence-informed institutional governance), zwiększając efektywność i konkurencyjność polskiego systemu szkolnictwa wyższego.

4.2.6 Ograniczenia

Przeprowadzone analizy obejmują populację naukowców posiadających dorobek zarejestrowany w bazie Scopus. Jest to rozwiązanie logiczne z perspektywy międzynarodowej porównywalności i dostępności metryk, jednak pewna część środowiska naukowego (publikująca głównie w języku polskim lub w dyscyplinach słabiej reprezentowanych w Scopus) nie jest bezpośrednio ujęta. Wnioski odnoszą się więc przede wszystkim do segmentu nauki zorientowanego na publikacje międzynarodowe.

Zakres danych nie wynika z arbitralnych decyzji badawczych, lecz z obiektywnych ograniczeń dostępu i jakości źródeł. Dane administracyjne z systemu POL-on, choć sukcesywnie uaktualniane, charakteryzują się niewielkim pokryciem artykułów naukowych, a ich pozyskanie obarczone jest barierami administracyjno-prawnymi (czasochłonne procedury i konieczność uzyskania zgód instytucjonalnych). Z kolei dane bibliometryczne z bazy Scopus wiążą się z bardzo wysokimi kosztami (jednorazowy zakup, płatne licencje na API) bądź koniecznością długotrwałego pobierania ręcznego, co znacząco utrudnia budowę i aktualizację operatu mikro danych.

Ograniczenia dotyczą także zakresu merytorycznego dostępnych danych. W operacie brak jest bezpośrednich informacji o finansowaniu badań (granty NCN, NCBR, programy europejskie), o środowisku instytucjonalnym czy o politykach kadrowych poszczególnych jednostek. Dane te są trudne do pozyskania i zintegrowania, zwłaszcza że polskie szkolnictwo wyższe i nauka obejmują zróżnicowane segmenty instytucjonalne (uczelnie publiczne i prywatne, jednostki PAN, instytuty badawcze), działające w odmiennych warunkach prawnych i finansowych. Główną zaletą stworzonego operatu jest możliwość łącznego badania zmiennych, które dotąd nie występowały razem w pojedynczych źródłach, jednak nie wyczerpuje on całego spektrum czynników kształtujących produktywność i nierówności.

Należy także mieć świadomość ograniczeń interpretacyjnych. Analizy ilościowe, nawet prowadzone na pełnych zbiorach, nie są w stanie uchwycić całego kontekstu społecznego i kulturowego procesów akademickich. Formułowane wnioski traktują jako identyfikację mechanizmów i korelacji, a nie jako deterministyczne prawa. Na przykład fakt, że kobiety rzadziej awansują do elity publikacyjnej, jest statystycznym faktem, którego przyczyny

mogą być różnorakie (od czynników rodzinnych, przez wybory życiowe, po bariery systemowe). Rekomendacje polityczne mają więc charakter *evidence-informed*, a nie ostateczny. W przyszłości planuję uzupełnienie analiz ilościowych o komponent jakościowy (np. wywiady z naukowcami), co pozwoli pełniej wyjaśnić obserwowane zależności.

W przyszłości istotnym rozwinięciem badań byłoby także objęcie analizą dorobku lokalnego (np. poprzez integrację z krajowymi bazami publikacji), co pozwoliłoby ocenić, czy mechanizmy nierówności obserwowane w segmentach widocznych w Scopus występują również w kontekście krajowym.

4.2.7 Podsumowanie

Zintegrowany operat mikrodanych łączący POL-on i Scopus stanowi **fundament analityczny** całego zgłaszanego cyklu badań nad produktywnością, efektywnością i nierównościami w polskim systemie nauki. Po raz pierwszy w Polsce umożliwił on analizę pełnego kontekstu funkcjonowania naukowców w skali całego środowiska — łącząc dorobek publikacyjny z wiekiem, płcią, stanowiskiem, instytucją i dyscypliną — oraz badanie mobilności, współpracy i koncentracji dorobku z wykorzystaniem podejścia mikropoziomowego.

Opracowana metodologia cechuje się wysoką precyzją i odpornością na niedoskonałości danych, a zarazem elastycznością umożliwiającą przeniesienie do innych kontekstów i cykliczną aktualizację. Może być stosowana w analizach instytucjonalnych i sektorowych oraz w porównaniach międzynarodowych — do monitorowania trendów produktywności, modelowania mobilności i oceny skuteczności polityk finansowania badań. Choć sam operat, ze względów dostępowych, nie jest publiczny, **procedury integracyjne i dokumentacja** są transparentne i w pełni odtwarzalne w warunkach analogicznych.

Z perspektywy krajowej operat stanowi trwałą infrastrukturę wiedzy na styku *ekonomicznej analizy nauki* (ang. *economics of science*), polityki publicznej i praktyki zarządzania szkolnictwem wyższym. Służy nie tylko badaniom akademickim, ale także praktyce decyzyjnej — jako podstawa konstrukcji wskaźników efektywności, oceny skuteczności polityk projakościowych oraz narzędzie *evidence-informed policy making*. Uwzględniając wskazane ograniczenia dotyczące zakresu danych i ich interpretacji, uważam, że operat — jako *narzędzie, metodologia i infrastruktura* — stanowi kluczowe osiągnięcie mojego dorobku: łączy nowatorstwo koncepcyjne, zaawansowanie metodologiczne i znaczenie aplikacyjne, a przy tym tworzy solidny punkt wyjścia dla dalszych krajowych i międzynarodowych badań ilościowych nad *ekonomią i finansami szkolnictwa wyższego*.

W całym procesie kierowałem się zasadami przejrzystości: dokumentacją etapów algorytmicznych i parametrów dopasowania, dziennikiem decyzji oraz polityką wersjonowania. Uważam, że rozwój kultury otwartej nauki powinien obejmować również badania ekonomiczne nad szkolnictwem wyższym; budowa trwałych, rzetelnych i metodycznie udokumentowanych operatów — takich jak ten — jest ważnym krokiem ku większej replikowalności i jakości analiz w polityce naukowej i edukacyjnej.

4.3 Analiza struktury i dynamiki produktywności naukowej

Drugim obszarem mojego osiągnięcia była pogłębiona analiza *struktury* i *dynamiki* produktywności naukowej w Polsce. Produktywność — rozumiana jako intensywność, jakość i trwałość aktywności publikacyjnej — stanowi kluczowy wymiar aktywności badawczej, odgrywający zasadniczą rolę w systemach ewaluacyjnych oraz w *ekonomicznej analizie nauki* (ang. *economics of science*) (Stephan, 2012; Fortunato et al., 2018).

W literaturze ekonomicznej podkreśla się jej *wielowymiarowość*: oprócz liczby publikacji ważne są cytowalność, prestiż czasopism, zakres i układ współpracy (krajowej i międzynarodowej), charakter współautorstwa oraz tempo akumulacji dorobku (Abramo et al., 2011; Larivière and Gingras, 2010; Sugimoto et al., 2013). Dopiero łączne ujęcie tych aspektów pozwala uchwycić rzeczywistą wartość i dynamikę aktywności badawczej.

W Polsce wcześniejsze badania koncentrowały się głównie na analizach przekrojowych, opartych na danych agregowanych lub jednostkowych, ale pozbawionych komponentu dynamicznego i mikropowiązania demograficzno-instytucjonalnych (Kwiek, 2018; Laudel, 2006; Kwiek, 2016). Rzadko uwzględniano czynniki pokoleniowe oraz relację produktywności do przebiegu kariery, co utrudniało testowanie hipotez o mechanizmach kumulacji dorobku i efektach cyklu życia w nauce.

W zgłoszonym cyklu wdrożyłem program badawczy łączący pełnopopulacyjne mikro-dane *Scopus-visible* z zaawansowanymi metodami modelowania, aby zmapować wzorce produktywności w polskim szkolnictwie wyższym. Wykorzystany operat — opisany w poprzednim obszarze — obejmuje ponad 25 tysięcy aktywnych badaczy i łączy informacje demograficzne, instytucjonalne i bibliometryczne, co umożliwiło analizy na poziomie jednostek, instytucji i kohort.

Skoncentrowałem się na czterech wymiarach:

- **Struktura rozkładu** dorobku i jego koncentracja (ciężkoogonowe rozkłady, elita produktywności);
- **Dynamika rozwoju** w ujęciu kohortowym i longitudinalnym (trajektorie, tempo akumulacji);
- **Trwałość pozycji** oraz **mobilność** między klasami produktywności;
- **Relacja** między ścieżką instytucjonalną a faktyczną efektywnością badawczą.

Na potrzeby analiz zaproponowałem i wdrożyłem oryginalne narzędzie klasyfikacyjne — **regułę 10/50** — pozwalające na jednoznaczną identyfikację elity (top 10%) i analizę koncentracji oraz mobilności. Klasyfikacja ta powstała w dwóch wariantach: (a) na danych *Scopus-only* (niezależny komponent analityczny, przydatny do analiz przekrojów czasowych i porównań międzynarodowych) oraz (b) na danych zintegrowanych POL-on-Scopus (umożliwiających testy hipotez z udziałem zmiennych demograficznych i

instytucjonalnych). Tak zaprojektowane narzędzie posłużyło do badania dynamiki karier, strukturalnych nierówności i do oceny potencjalnych skutków polityk publicznych.

Dzięki integracji danych oraz wykorzystaniu bogatego zestawu zmiennych możliwe było testowanie hipotez o mechanizmach akumulacji przewagi, efektach ścieżki kariery i selekcji instytucjonalnej w polskim szkolnictwie wyższym. Wyniki mają znaczenie poznawcze — jako wkład do międzynarodowej literatury nad produktywnością naukową — oraz aplikacyjne, ponieważ wspierają projektowanie polityk publicznych ograniczających nadmierną koncentrację dorobku, wspierających mobilność i młodsze kohorty badaczy. Analizy osadzam w ramach *ekonomicznej analizy nauki* i teorii kapitału ludzkiego, w której inwestycje w badania i rozwój traktowane są jako kluczowy element wzrostu gospodarczego i konkurencyjności.

4.3.1 Zakres i unikalność podejścia

W odróżnieniu od wcześniejszych prac krajowych, opartych na fragmentarycznych danych instytucjonalnych albo wyłącznie na źródłach bibliometrycznych, moje podejście umożliwiło analizę *pełnej populacji* naukowców zatrudnionych w polskich uczelniach akademickich (wg OPI), widocznych w Scopus, z mikropoziomowym powiązaniem informacji demograficznych, instytucjonalnych i bibliometrycznych. Integracja POL-on-Scopus pozwoliła jednocześnie kontrolować wiek biologiczny i akademicki, płeć, typ uczelni, jednostkę zatrudniającą, stanowisko i dyscyplinę, a także uwzględniać metryki publikacyjne (liczba i cytowalność prac, prestiż czasopism, współautorstwo krajowe i międzynarodowe). W literaturze ekonomicznej tego typu zbiory określa się jako warunek rozwoju *evidence-informed policy making* i analizy efektywności systemu szkolnictwa wyższego (Lane et al., 2011a; Fortunato et al., 2018; Moed, 2005).

Operat posłużył do estymacji modeli regresji logistycznej i frakcyjnej regresji logistycznej, a także do konstrukcji klasyfikacji produktywności (reguła 10/50) oraz analizy mobilności między klasami w ujęciu kohortowym i longitudinalnym. W przeciwieństwie do analiz przekrojowych, modelowanie dynamiczne pozwala wykrywać mechanizmy akumulacji dorobku i trajektorii produktywności, co umożliwia testowanie hipotez o cyklu życia, kapitale ludzkim i efektach kumulacji zasobów w nauce (Long, 1992; Sugimoto et al., 2013).

4.3.2 Badania nad nierównościami w nauce

W analizach dotyczących nierówności w nauce odpowiadałem za całość podejścia metodologicznego i implementację badań empirycznych. Mój wkład obejmował zarówno opracowanie narzędzi metodologicznych, jak i pełny proces analityczny — od formułowania hipotez, przez konstrukcję wskaźników i modeli, po interpretację wyników i rekomendacje. W szczególności odpowiadałem za:

- **Dobór ram analitycznych:** podstawą moich analiz były koncepcje *ekonomicznej analizy nauki* (ang. *economics of science*) oraz teorii kapitału ludzkiego, zwłaszcza dotyczące efektywności wykorzystania zasobów i cyklu życia kariery naukowej. Ramy te uzupełniałem o inspiracje z socjologii nauki (np. teoria kumulatywnej przewagi Mertona), traktując je jako komplementarne do ujęcia ekonomicznego. W praktyce operacjonalizowałem zmienne trudne do uchwycenia w danych administracyjnych, takie jak wiek akademicki, typy współpracy czy klasy produktywności.
- **Konstrukcję nowych wskaźników:** zaprojektowałem i wdrożyłem dwie oryginalne miary nierówności płciowych:
 1. *gender solo research gap* — różnicę w częstości publikacji jednoautorskich między kobietami i mężczyznami, interpretowaną jako miarę indywidualnego kapitału badawczego i widoczności naukowej,
 2. *gender homophily index* — mierzącą skłonność do współpracy w zespołach jedнопłciowych, pozwalającą uchwycić ograniczenia w dostępie do zróżnicowanych sieci badawczych.

Wskaźniki te zostały zaprojektowane jako narzędzia *ekonomicznej analizy nauki*, możliwe do zastosowania w badaniach porównawczych na poziomie międzynarodowym.

- **Projektowanie i estymację modeli statystycznych:** odpowiadałem za konstrukcję modeli logistycznych (np. prawdopodobieństwo udziału we współpracy międzynarodowej), frakcyjnych (udział publikacji solo, udział prac w zespołach jedнопłciowych) oraz logistycznych z efektami stałymi dla dyscyplin i okresów, co pozwoliło kontrolować heterogeniczność i porównywać wyniki między grupami naukowców. Modele implementowałem w R, stosując stratyfikację danych i kontrolę efektów instytucjonalnych.
- **Analizy sieciowe współpracy:** wykorzystałem dane o współautorstwie do badania konfiguracji zespołów badawczych, mierząc m.in. homofilię płciową i różnorodność składu zespołów. W (Kwiek and Roszka, 2021a) wykazałem, że wzorce współpracy istotnie wpływają na dostęp do zasobów kapitału ludzkiego i mogą sprzyjać bądź ograniczać długookresową produktywność.
- **Interpretację i rekomendacje:** łączyłem uzyskane wyniki z literaturą międzynarodową, wskazując na konsekwencje ekonomiczne dla polityki naukowej. Na tej podstawie formułowałem rekomendacje dotyczące m.in. wspierania udziału kobiet w badaniach międzynarodowych, uwzględniania wieku akademickiego przy ocenie dorobku oraz monitorowania wzorców współpracy jako czynnika efektywnej alokacji zasobów i równości szans.

Całość podejścia była **nowatorska w Polsce**: po raz pierwszy pełnopopulacyjne mikro-dane zostały użyte do systematycznej analizy nierówności w nauce w ujęciu ekonomicznym. Opracowane wskaźniki i modele pozwoliły na testowanie hipotez o zróżnicowaniach płciowych, wiekowych i instytucjonalnych w sposób porównywalny z międzynarodowymi standardami.

Mój wkład miał także wymiar teoriiwórczy: wprowadziłem do badań nad nierównościami oryginalne ramy analityczne, takie jak *gender solo research gap* i *gender homophily index*. Ich zastosowanie w analizach pełnopopulacyjnych stanowi nie tylko narzędzie empiryczne, ale także **wkład do ekonomicznej analizy stratyfikacji nauki**, rozwijającej klasyczne ujęcia (np. kumulatywnej przewagi) w kierunku badań nad efektywnością, kapitałem ludzkim i polityką publiczną. Socjologia nauki pełniła w tym kontekście rolę dyscypliny komplementarnej, dostarczającej dodatkowych inspiracji, natomiast główne osadzenie analiz miało charakter ekonomiczny.

4.3.3 Główne wyniki i ich znaczenie

Przeprowadzone analizy ilościowe i kohortowe, oparte na zintegrowanym zbiorze danych, dostarczyły szeregu istotnych rezultatów empirycznych, które rzucają nowe światło na strukturę i dynamikę produktywności naukowej w Polsce. Wyniki te mają znaczenie nie tylko dla poznania mechanizmów funkcjonowania krajowego systemu nauki, lecz także dla projektowania polityk publicznych i oceny efektywności instytucjonalnej w szkolnictwie wyższym.

- **Wysoka koncentracja produktywności**: zgodnie z regułą 10/50, 10% najbardziej produktywnych naukowców odpowiada za ponad 50% wszystkich publikacji z afiliacją polską w bazie Scopus. Zjawisko to — typowe dla rozkładów ciężkoogonowych (Lotka, Pareto) — potwierdzono empirycznie na poziomie pełnopopulacyjnym. Koncentracja dorobku utrzymuje się w czasie, niezależnie od reform systemowych, co wskazuje na trwałość strukturalnych mechanizmów akumulacji i ograniczoną elastyczność redystrybucji zasobów.
- **Trwałość przynależności do elity**: analiza macierzy przejść między klasami produktywności wykazała, że badacze, którzy we wczesnych etapach kariery wchodzą do grupy top 10%, w większości utrzymują tę pozycję przez wiele lat. Jest to przejaw silnej zależności od ścieżki (path dependence) oraz efektów akumulacji kapitału ludzkiego i reputacyjnego, które zwiększają szanse na dalsze osiągnięcia i dostęp do zasobów.
- **Produktywność zależna od wieku akademickiego**: szczyt aktywności publikacyjnej przypada zazwyczaj na pierwsze 10–15 lat kariery naukowej, po czym obserwuje się stabilizację lub spadek. Analiza z wykorzystaniem wieku akademickiego

— obliczanego jako liczba lat od pierwszej publikacji w Scopus — pozwoliła precyzyjnie odróżnić cykl życia produktywności od efektów starzenia się biologicznego, co wpisuje się w literaturę ekonomii pracy i teorii cyklu życia (Kwiek and Roszka, 2022a).

- **Ograniczona mobilność w rozkładzie produktywności:** zarówno analizy kohortowe, jak i macierze przejść wskazują na stosunkowo niską mobilność naukowców w obrębie rozkładu produktywności. Awans do grupy top 10% jest rzadki, zwłaszcza w późniejszych fazach kariery. Taka struktura ogranicza konkurencyjność systemu i zmniejsza skuteczność redystrybucji zasobów poprzez mechanizmy instytucjonalne.
- **Rozbieżność między ścieżką instytucjonalną a produktywnością:** awans na wyższe stanowiska akademickie nie zawsze wiąże się z dalszym wzrostem produktywności. W wielu przypadkach obserwowano stabilizację lub spadek dorobku po uzyskaniu tytułu profesora. Zjawisko to, znane również z badań międzynarodowych (Zuckerman and Merton, 1977; Diamond, 1986), wskazuje na niedoskonałości systemu bodźców i potrzebę projektowania bardziej efektywnych mechanizmów wynagradzania i awansów.
- **Przydatność klasyfikacji produktywnościowej opartej na regule 10/50:** zaproponowany podział na klasy produktywności — z wyodrębnieniem elity (top 10%) oraz pozostałych badaczy (bottom 90%) — pozwolił na identyfikację struktury koncentracji dorobku w polskiej nauce. Klasy te okazały się użyteczne zarówno jako zmienne zależne w modelach ekonometrycznych, jak i narzędzie do monitorowania koncentracji zasobów oraz możliwości awansu. Dzięki swojej przejrzystości mogą być stosowane także w praktyce polityki publicznej jako wskaźniki wspierające projektowanie systemów oceny i finansowania.

4.3.4 Precyzja i ograniczenia analiz

Wszystkie modele statystyczne estymowano z zachowaniem wysokich standardów jakości danych i poprawności metodologicznej. Przed analizą przeprowadzono kontrolę kompletności, imputację braków w zmiennych demograficznych (np. płci) oraz odfiltrowano rekordy o niskim poziomie dopasowania. Modele testowano w alternatywnych specyfikacjach (zestawy zmiennych kontrolnych, poziomy agregacji), a w analizach regresyjnych kontrolowano efekty instytucjonalne (typ uczelni) i dziedzinowe (grupy ASJC). W części modeli zastosowano efekty stałe dla dyscyplin i okresów (modelowanie „*within variation*”), a w przypadku zmiennych frakcyjnych (np. udział publikacji solo, udział we współpracy międzynarodowej) uwzględniono specyfikę rozkładów ograniczonych.

Pomimo zachowanej staranności, analizy napotykają pewne ograniczenia metodologiczne i strukturalne:

- **Brak danych o finansowaniu zewnętrznym.** Zbiór nie obejmuje grantów krajowych (NCN, NCBR) ani europejskich (Horyzont, ERC), co ogranicza możliwość analizy relacji między produktywnością a pozyskiwaniem zasobów finansowych.
- **Brak informacji o rolach autorskich.** Ograniczenia bazy Scopus uniemożliwiają rozróżnienie ról autorskich (pierwszy, korespondencyjny, ostatni autor), co utrudnia analizę wkładu indywidualnego i hierarchii zasług w publikacjach wieloautorskich.
- **Niejednoznaczności w dopasowaniu rekordów.** Mimo wysokich wskaźników jakości (precyzja 97%, czułość 94%, $F1=0,96$) istnieje niewielkie ryzyko błędów w przypadku nazwisk częstych, transliterowanych lub zmienionych (np. po ślubie), zwłaszcza przy niewielkim dorobku i nietypowych afiliacjach.
- **Pokrycie bazy Scopus i heterogeniczność dziedzin.** Pokrycie STEM jest lepsze niż HSS; ponadto pierwsza publikacja w Scopus nie zawsze odpowiada faktycznej pierwszej publikacji badacza, co może powodować przesunięcia w pomiarze wieku akademickiego w niektórych dziedzinach.
- **Endogeniczność i kolejność zdarzeń.** Relacje produktywność–awans czy produktywność–współpraca międzynarodowa mogą być dwukierunkowe; zastosowane modele ograniczają, lecz nie eliminują ryzyka odwróconej przyczynowości.
- **Trunkacja i selekcja.** Analizy obejmują populację badaczy widocznych w Scopus; wnioski nie muszą być w pełni reprezentatywne dla aktywności publikacyjnej poza czasopismami indeksowanymi.

Jednocześnie konstrukcja operatu — integracja dwóch niezależnych źródeł, walidacja ekspercka i wykorzystanie pełnej populacji *Scopus-visible* — minimalizuje wpływ tych ograniczeń na wnioski. W szczególności:

- oczyszczono zbiory z duplikatów i rekordów niekompletnych,
- próg dopasowania POL-on–Scopus wyznaczono empirycznie ($\xi_{ij} \geq 0,85$),
- zweryfikowano spójność cech demograficznych i instytucjonalnych (płeć, wiek, typ uczelni) z danymi GUS i POL-on, w zakresie ich dostępności.

W rezultacie uzyskane wyniki można uznać za **wiarygodne i stabilne**, a identyfikowane ograniczenia wskazują jednocześnie kierunki dalszego rozwoju operatu — w szczególności włączenie informacji o finansowaniu badań, rolach autorskich oraz jakościowych zmiennych z badań ankietowych. Dzięki temu baza może stanowić punkt wyjścia do coraz bardziej kompletnych modeli ekonometrycznych w obszarze *science of science*.

4.3.5 Zastosowania operatu poza analizą produktywności

Zintegrowany zbiór danych POL-on-Scopus posłużył nie tylko do badań nad strukturą i dynamiką produktywności, ale także jako *fundament* eksploracji innych aspektów funkcjonowania systemu nauki w Polsce. Jego konstrukcja — łącząca cechy demograficzne, instytucjonalne, kohortowe i bibliometryczne — umożliwia testowanie hipotez wykraczających poza wąsko rozumiane miary aktywności publikacyjnej (Moed, 2005; Fortunato et al., 2018). W szczególności:

- **Mobilność między klasami produktywności (ujęcie kohortowe).** Analizowano, w jakim stopniu badacze zmieniają pozycję w rozkładzie produktywności, w tym awansują do elity top 10% (generującej ok. 50% dorobku — reguła 10/50). Mobilność ta różnicuje się w zależności od kohorty, typu uczelni, dyscypliny i płci (Cole, 1973; Long, 1992; Fox, 2005).
- **Produktywność a ścieżka awansu instytucjonalnego.** Porównano rytm rozwoju dorobku z formalnym przebiegiem kariery (stopień doktora, stopień doktora habilitowanego, tytuł profesora). W wielu przypadkach obserwowano brak pełnej synchronizacji tych ścieżek, co prowokuje pytania o efektywność i meritokratyczność awansów (Zuckerman and Merton, 1977; Diamond, 1986).
- **Zróżnicowania instytucjonalne.** Dane umożliwiły porównania dorobku w zależności od typu uczelni i uczestnictwa w programie IDUB. Badacze z uczelni IDUB częściej należeli do top 10%, co sugeruje, że alokacja dodatkowych zasobów finansowych przekłada się na większą koncentrację produktywności.
- **Różnice płciowe i pokoleniowe.** Modelowano wpływ płci, wieku biologicznego i akademickiego na produktywność oraz współpracę międzynarodową. Zidentyfikowano m.in. *gender solo research gap* i homofilię płciową (Kwiek and Roszka, 2021b, 2022b).
- **Typologie trajektorii karier.** W (Kwiek and Roszka, 2024c) wyróżniono typy „early fast”, „late bloomers”, „declining after promotion”, „volatile”, łącząc analizy kohortowe z longitudinalnymi — co pozwoliło uchwycić heterogeniczność ścieżek i zaproponować alternatywne kryteria oceny (Long, 1992; Fortin and Currie, 2013).

Dzięki tym zastosowaniom operat działa jako narzędzie *badawcze, diagnostyczne i polityczne*. Jego potencjał obejmuje:

- wsparcie projektowania polityk zwiększających efektywność alokacji zasobów (mobilność między klasami produktywności, między instytucjami i dyscyplinami) (Fox and Colatrella, 2001; UNESCO, 2021),

- monitorowanie skutków reform (np. systemu ewaluacji, regulacji awansów, programów projakościowych) (Hicks et al., 2015),
- diagnozowanie obszarów wymagających interwencji (młodsze kohorty, kobiety, badacze z mniejszych ośrodków) (van den Besselaar and Sandström, 2011),
- budowę bardziej kontekstowych wskaźników produktywności i efektywności (uwzględniających wiek akademicki, typ kariery, dynamikę rozwoju) (San Francisco Declaration on Research Assessment, 2012; Sugimoto et al., 2013).

Kolejnym kierunkiem jest analiza **interdyscyplinarności**. Struktura danych pozwala mierzyć przekrojowość dorobku na poziomie publikacji (kombinacje kodów ASJC) i zespołów (współpraca autorów z różnych dyscyplin), a także śledzić zmienność tematyczną w czasie — co wspiera badania nad efektywnością inwestycji w badania interdyscyplinarne.

Mam świadomość ryzyk *metrics-driven science*: nadmiernego podporządkowania badań prostym wskaźnikom (liczba publikacji, cytowań, punkty), strategicznych zachowań (np. *salami slicing*), unikania ryzykownych tematów czy ograniczania interdyscyplinarności. Operat — dzięki szczegółowości i elastyczności — umożliwia konstruowanie wskaźników bardziej kontekstowych i różnicujących, co sprzyja polityce nastawionej na długofalowy rozwój kapitału ludzkiego i wzrost gospodarczy oparty na wiedzy.

W tym sensie prezentowane badania stanowią wkład do bardziej zniuansowanej refleksji nad rolą wskaźników w zarządzaniu nauką — nie tylko poprzez doskonalenie pomiaru, lecz także przez analizę konsekwencji systemowych i instytucjonalnych, istotnych z perspektywy ekonomii i finansów.

4.3.6 Znaczenie dla badań międzynarodowych

Zaproponowane podejście metodologiczne wpisuje się w aktualne standardy i rekomendacje międzynarodowe dotyczące analizy systemów nauki oraz odpowiedzialnego wykorzystania danych w procesach ewaluacyjnych i zarządczych. Kluczowe elementy — integracja danych administracyjnych i bibliometrycznych, uwzględnienie wieku akademickiego, podejście kohortowe oraz analiza longitudinalna — pozostają zgodne z postulatami:

- **Leiden Manifesto for Research Metrics** (Hicks et al., 2015): ocena dorobku w kontekście czasowym, instytucjonalnym i dyscyplinarnym oraz ograniczanie roli prostych wskaźników;
- **San Francisco Declaration on Research Assessment (DORA)** (San Francisco Declaration on Research Assessment, 2012): różnicowanie miar oceny dorobku i krytyczne podejście do metryk czasopism;
- **Ram UNESCO dotyczących otwartej nauki** (UNESCO, 2021): transparentność metod, replikowalność analiz i inkluzywność projektowania systemów oceny.

Opracowany operat oraz wdrożona klasyfikacja percentylowa (w tym reguła *10/50*) mają *potencjał adaptacyjny* w analizach porównawczych. W krajach dysponujących rozbudowanymi rejestrami administracyjnymi (np. Holandia, Kanada, USA, Australia) możliwe jest pełne zastosowanie procedury integracyjnej (łączenie danych kadrowych z bibliometrycznymi). W państwach o ograniczonej dostępności danych administracyjnych podejście percentylowe może być implementowane na podstawie samych danych bibliometrycznych (Scopus, Dimensions, OpenAlex, Crossref), przy zachowaniu porównywalnych definicji i horyzontów czasowych.

Zjawiska zaobserwowane w polskim systemie — koncentracja dorobku, ograniczona mobilność między klasami produktywności, trwałość pozycji elity oraz częściowa niespójność między awansem instytucjonalnym a aktywnością publikacyjną — mają charakter szerzej potwierdzony w literaturze międzynarodowej (USA, Kanada, Niemcy, Holandia, Wielka Brytania) (Ioannidis et al., 2016; Fortunato et al., 2018; Larivière and Gingras, 2010). Proponowana metodologia umożliwia:

- porównania udziału elit badawczych w całkowitej produkcji naukowej,
- analizę trwałości i mobilności elity w długim okresie (ujęcia kohortowe i longitudinalne),
- identyfikację barier awansu i różnic w alokacji zasobów w systemach o odmiennym stopniu selektywności,
- ocenę wpływu reform ewaluacyjnych i polityk projakościowych na dynamikę karier oraz efektywność wykorzystania kapitału ludzkiego.

Dzięki transparentności algorytmicznej, zgodności z zasadami otwartej nauki i możliwości dostosowania do różnych środowisk danych, operat POL-on-Scopus może stanowić wzorzec dla budowy narodowych operatów mikrodanych o nauce — zwłaszcza w krajach reformujących politykę naukową i systemy oceny dorobku. Z perspektywy *ekonomii i finansów* stanowi on podstawę do analiz dystrybucji produktywności, efektywności instytucjonalnej oraz wpływu polityk publicznych na rozwój kapitału ludzkiego i konkurencyjność systemów szkolnictwa wyższego.

4.3.7 Podsumowanie

Przedstawione analizy wnoszą oryginalny wkład do badań nad produktywnością naukową w Polsce i wpisują się w międzynarodowy nurt badań nad strukturą oraz dynamiką karier akademickich (Cole, 1973; Fortunato et al., 2018; Sugimoto et al., 2013). Na podstawie danych o pełnej populacji naukowców *widocznych w Scopus* i zatrudnionych w polskich uczelniach akademickich wykazano znaczną koncentrację dorobku — zgodną z klasycznymi rozkładami Lotki-Price’a — oraz ograniczoną mobilność między klasami produktywności (Lotka, 1926; de Solla Price, 1976). Około 10% najbardziej produktywnych badaczy

odpowiada za blisko 50% całkowitej liczby publikacji z afiliacją polską (reguła 10/50), co wskazuje na silną koncentrację sukcesu i zasobów badawczych. Struktura ta pozostaje względnie stabilna w czasie, co sugeruje działanie trwałych procesów akumulacji przewagi i efektów skali (Merton, 1968; Ioannidis et al., 2020).

Połączenie klasyfikacji 10/50, ujęcia kohortowego i modeli longitudinalnych pozwoliło precyzyjnie uchwycić te zjawiska oraz zbadać ich uwarunkowania (wiek akademicki, dyscyplina, typ uczelni, status instytucjonalny). Wykazano istotne różnice pokoleniowe i heterogeniczność ścieżek kariery, niedostrzegalne w analizach statycznych (Kwiek and Roszka, 2022a, 2024c).

Implikacje dla polityki są dwojakie. Po pierwsze, wskazują na potrzebę odchodzenia od punktowych ocen przekrojowych na rzecz monitorowania długookresowych trajektorii, z włączeniem wieku akademickiego do konstrukcji wskaźników. Po drugie, sugerują projektowanie instrumentów sprzyjających mobilności w rozkładzie produktywności i ograniczaniu nadmiernej koncentracji, co z perspektywy ekonomii oznacza bardziej efektywną alokację kapitału ludzkiego i finansowego (San Francisco Declaration on Research Assessment, 2012; Hicks et al., 2015).

Z metodologicznego punktu widzenia analiza pokazuje potencjał zintegrowanych mikrodanych: odpowiednio przygotowany i zwalidowany operat umożliwia badania o dużej mocy wyjaśniającej, łączące uwarunkowania indywidualne, instytucjonalne i systemowe (Herzog et al., 2007; Winkler, 2006; Christen, 2012). Zestaw zastosowanych narzędzi — integracja administracyjno-bibliometryczna, regresje logistyczne z efektami stałymi, frakcyjne modele logistyczne oraz klasyfikacje percentylowe — może być z powodzeniem adaptowany w badaniach międzynarodowych.

Łącząc nowoczesne metody analizy danych z empiryczną diagnozą złożonych procesów alokacji zasobów w systemie nauki, cykl ten pokazuje, że mikrodane mogą służyć nie tylko opisowi, ale i projektowaniu polityk publicznych (Lane et al., 2011a; UNESCO, 2021). Rezultaty oparte na pełnopopulacyjnym, zintegrowanym zbiorze POL-on–Scopus stanowią trzon zgłaszanego cyklu; uzupełnia je publikacja oparta na danych *Scopus-only*, rozwijająca perspektywę porównawczą i testującą uniwersalność wniosków.

4.4 Badania nad nierównościami w nauce

Trzeci obszar mojego osiągnięcia naukowego dotyczył systematycznej analizy nierówności w systemie nauki w Polsce — ze szczególnym uwzględnieniem wymiaru płci, wieku biologicznego i akademickiego, instytucji zatrudnienia oraz wzorców współpracy. Opracowany przeze mnie operat danych umożliwił — po raz pierwszy w Polsce — przeprowadzenie badań nad nierównościami z wykorzystaniem danych pełnopopulacyjnych, obejmujących ponad 25 tysięcy naukowców widocznych międzynarodowo. Dzięki integracji danych administracyjnych z systemu POL-on z bibliometrycznymi danymi z bazy Scopus możliwe

było jednoczesne uchwycenie indywidualnych, instytucjonalnych i strukturalnych uwarunkowań nierówności w systemie nauki.

Badania koncentrowały się na czterech kluczowych wymiarach:

1. nierównościach płciowych w zakresie produktywności, współpracy i publikacji jednoautorskich (artykuły [1]–[3]), analizowanych jako różnice w wykorzystaniu kapitału ludzkiego i dostępie do zasobów,
2. wpływie wieku biologicznego i akademickiego na produktywność i przebieg kariery (artykuł [4]), w tym na tempo zwrotu z inwestycji w kapitał ludzki,
3. dostępie do współpracy międzynarodowej i struktur o wysokim prestiżu (artykuły [1], [2], [8]) — kluczowych dla budowy przewagi konkurencyjnej systemu,
4. implikacjach dla projektowania polityk naukowych wspierających inkluzywność, mobilność i efektywność alokacji zasobów badawczych.

Zaproponowane podejście łączyło zaawansowaną analizę ilościową z interpretacją wyników w świetle literatury międzynarodowej z zakresu gender studies, stratyfikacji akademickiej, teorii kumulatywnej przewagi oraz *ekonomicznej analizy nauki* (ang. *economics of science*) (Merton, 1968; Fox, 2005; Sugimoto et al., 2013; Allison and Stewart, 1974). Odwoływano się m.in. do ustaleń West et al. (2013), wskazujących na systematyczne różnice w autorskiej widoczności kobiet i mężczyzn, oraz Elsevier (2020), podkreślających ograniczony udział kobiet w międzynarodowych sieciach współpracy. Wyniki te mają bezpośrednie znaczenie dla oceny efektywności wykorzystania kapitału ludzkiego w systemie nauki.

Zidentyfikowane w analizach zróżnicowania instytucjonalne i geograficzne — zwłaszcza w dostępie do projektów o wysokim prestiżu oraz umiędzynarodowienia dorobku — wpisują się w ustalenia dotyczące stratyfikacji systemu nauki (van den Besselaar and Sandström, 2011). Mechanizmy reprodukcji przewag są zakorzenione nie tylko w działaniach jednostkowych, lecz także w logice funkcjonowania zamkniętych środowisk akademickich (tzw. *invisible colleges* (Crane, 1972)) oraz w utrwalonych schematach oceny dorobku. Z ekonomicznego punktu widzenia prowadzi to do nieefektywnej alokacji zasobów, marnotrawienia potencjału oraz obniżenia produktywności całego systemu.

Podjęte badania uwzględniały zalecenia inicjatyw międzynarodowych w zakresie odpowiedzialnej ewaluacji nauki, takich jak Deklaracja DORA (San Francisco Declaration on Research Assessment, 2012) czy Leiden Manifesto (Hicks et al., 2015), akcentujących dostosowanie wskaźników do zróżnicowanych ścieżek kariery, promowanie równości szans oraz ograniczenie stosowania jednowymiarowych metryk. W tym ujęciu nierówności traktowane są nie tylko jako problem etyczny, lecz także jako czynnik ograniczający efektywność systemową i konkurencyjność międzynarodową.

Mój wkład autorski

W analizach zawartych w tej części cyklu odpowiadałem za całość podejścia metodologicznego i jego implementację. Mój wkład obejmował w szczególności:

- **Dobór ram teoretycznych i ich adaptację do danych ilościowych.** Podstawą analiz były koncepcje *ekonomicznej analizy nauki* (ang. *economics of science*) oraz teorii kapitału ludzkiego (w tym ujęcie *life cycle*), które uzupełniałem o inspiracje z socjologii nauki (np. teorie kumulatywnej przewagi Mertona, badania nad sieciami współautorskimi). W praktyce oznaczało to operacjonalizację zmiennych trudnych do uchwycenia w klasycznych danych administracyjnych, takich jak wiek akademicki, typy współpracy czy klasy produktywności. Wiek akademicki — liczony od daty pierwszej publikacji w Scopus — okazał się szczególnie przydatny do testowania hipotez o nierównościach pokoleniowych i efektywności inwestycji w kapitał ludzki.
- **Konstrukcję nowych wskaźników.** Zaprojektowałem i wdrożyłem miary nierówności płciowych dotąd niewykorzystywane w badaniach krajowych:
 1. *gender solo research gap* — różnicę w częstości publikowania jednoautorskiego między kobietami i mężczyznami, interpretowaną jako wskaźnik indywidualnego kapitału badawczego i widoczności naukowej,
 2. *gender homophily index* — mierzącą skłonność do współpracy w zespołach jednopłciowych, co pozwalało uchwycić bariery w dostępie do zróżnicowanych sieci współpracy.

Oba wskaźniki mają charakter uniwersalny i nadają się do replikacji w badaniach międzynarodowych, co czyni je narzędziem analizy porównawczej w ramach *ekonomicznej analizy nauki*.

- **Projektowanie i estymację modeli statystycznych.** Odpowiadałem za konstrukcję i implementację modeli regresji:
 - *logistycznych* (np. prawdopodobieństwo udziału we współpracy międzynarodowej),
 - *frakcyjnych* (udział publikacji jednoautorskich, udział publikacji w zespołach jednopłciowych),
 - *logistycznych z efektami stałymi*, gdy wymagała tego struktura danych i konieczność kontroli heterogeniczności (efekty dyscyplinowe i okresowe).

Modele estymowałem w R, uwzględniając czynniki instytucjonalne i dziedzinowe poprzez stratyfikację danych oraz zastosowanie efektów stałych.

- **Analizy sieciowe współpracy.** Wykorzystałem strukturę współautorstwa do badania konfiguracji zespołów, mierząc m.in. homofilię płciową i różnorodność składu. W (Kwiek and Roszka, 2021a) pokazałem, że wzorce doboru partnerów publikacyjnych istotnie różnicują dostęp do zasobów i sieci o zróżnicowanym kapitale, a tym samym wpływają na trajektorie karier i efektywność wykorzystania kapitału ludzkiego.
- **Interpretację i rekomendacje.** Łączyłem uzyskane wyniki z literaturą międzynarodową, wskazując na podobieństwa i różnice w stosunku do systemów zachodnich. Na tej podstawie formułowałem rekomendacje dla polityki naukowej: m.in. wspieranie udziału kobiet w badaniach międzynarodowych, uwzględnianie wieku akademickiego w ocenie dorobku, a także monitorowanie wzorców współpracy jako czynnika równości szans i efektywnej alokacji zasobów.

Całość podejścia miała charakter **nowatorski w Polsce**: po raz pierwszy wykorzystano pełnopopulacyjne mikrodane do systematycznej analizy nierówności w nauce w ujęciu ekonomicznym. Opracowane wskaźniki i modele pozwalają na testowanie hipotez o zróżnicowaniach płciowych, wiekowych i instytucjonalnych w sposób porównywalny z międzynarodowymi standardami. Mój wkład polegał więc zarówno na *innowacji metodologicznej*, jak i na *przekładzie wyników na praktyczne rekomendacje polityczne*, co łączy perspektywę naukometryczną z ekonomią i polityką publiczną.

Mój wkład miał również wymiar teoriiotwórczy: wprowadziłem do badań nad nierównościami w nauce oryginalne kategorie i ramy interpretacyjne, takie jak *gender solo research gap* czy *gender homophily index*, które pozwalają konceptualizować różnice płciowe w dostępie do kapitału akademickiego i widoczności naukowej. Ich operacjonalizacja i zastosowanie w analizach pełnopopulacyjnych stanowią nie tylko narzędzie empiryczne, lecz także **wkład do ekonomicznej analizy stratyfikacji w nauce**, rozwijającej i adaptującej klasyczne ujęcia (Mertonowska kumulatywna przewaga, teorie kapitału społecznego) do współczesnych realiów polskiego i europejskiego systemu akademickiego. Socjologia nauki pełniła w tym kontekście rolę dyscypliny komplementarnej, dostarczającej dodatkowych inspiracji teoretycznych, natomiast zasadnicze osadzenie analiz miało charakter ekonomiczny.

Nowatorstwo podejścia

Na tle wcześniejszych badań nad nierównościami — często opartych na ograniczonych próbach, danych sondażowych lub fragmentarycznych analizach bibliometrycznych — zastosowane podejście wyróżniało się kilkoma cechami:

- **pełnopopulacyjnym zasięgiem danych** (25 463 naukowców widocznych w Scopus i zatrudnionych w polskich uczelniach), co redukuje ryzyko błędu selekcji i umożliwia formułowanie wniosków systemowych,

- **mikropoziomową strukturą** — indywidualne profile publikacyjne i ścieżki kariery pozwalały uchwycić heterogeniczność rozwoju naukowców, niedostępną w danych zagregowanych,
- **innowacyjną integracją** danych administracyjnych i bibliometrycznych mimo braku wspólnego identyfikatora, co umożliwiło jednoczesne uwzględnienie cech demograficznych, instytucjonalnych i publikacyjnych,
- **konstrukcją zmiennych dynamicznych** (wiek akademicki, rok uzyskania doktoratu, kohorty publikacyjne), które pozwoliły badać procesy w ujęciu longitudinalnym i kohortowym,
- **szerokim zakresem analiz nierówności** — nie tylko między płciami, ale także między kohortami, typami uczelni, regionami i klasami produktywności, co dało podstawę do bardziej kompleksowych wniosków o alokacji kapitału ludzkiego.

Takie podejście umożliwiło identyfikację istotnych luk — m.in. niższego udziału kobiet w publikacjach jednoautorskich, rzadszej obecności w zespołach międzynarodowych czy ograniczonej pozycji w strukturach współautorskich. Wyniki te mają znaczenie nie tylko etyczne, ale także ekonomiczne: wskazują na niewykorzystany potencjał kapitału ludzkiego i ograniczenia w efektywności systemu nauki.

Sformułowane na tej podstawie rekomendacje pozostają spójne z międzynarodowymi standardami odpowiedzialnej ewaluacji (DORA, Leiden Manifesto, ramy UNESCO dotyczące otwartej nauki) (San Francisco Declaration on Research Assessment, 2012; Hicks et al., 2015; UNESCO, 2021), a zarazem podkreślają znaczenie polityk wspierających inkluzywność i mobilność jako warunku zwiększania konkurencyjności i długofalowej efektywności systemu szkolnictwa wyższego.

4.4.1 Znaczenie badań nad nierównościami w nauce

Badania nad nierównościami odgrywają kluczową rolę w zrozumieniu działania współczesnych systemów akademickich. Nierówności te dotyczą zarówno cech indywidualnych — takich jak płeć, wiek, instytucja zatrudnienia czy etap kariery — jak i mechanizmów systemowych. Te ostatnie sprzyjają kumulacji przewag jednych grup i marginalizacji innych. Wpływają one na sprawiedliwość systemu, jego efektywność oraz zdolność do zrównoważonej produkcji wiedzy. (Fox and Colatrella, 2001; van den Besselaar and Sandström, 2011; Sugimoto et al., 2013). Z perspektywy ekonomii oznacza to zarówno niewykorzystanie kapitału ludzkiego, jak i niższą stopę zwrotu z publicznych inwestycji w badania.

W literaturze podkreśla się, że sukces naukowy jest także efektem mechanizmów społecznych, instytucjonalnych i sieciowych, takich jak:

- *efekt Matyldy* — niedoszacowanie dorobku kobiet i niższa widoczność ich publikacji (Holman et al., 2018),

- *efekt Mertona* — kumulatywna przewaga badaczy odnoszących wcześniejsze sukcesy (Merton, 1968),
- homofilia instytucjonalna i płciowa — preferencja współpracy z podobnymi partnerami (McPherson et al., 2001; Sugimoto et al., 2013),
- stratyfikacja zasobów — koncentracja grantów, infrastruktury i mobilności wokół elity (Leišytė, 2016; Fortin and Currie, 2013),
- wykluczenie z nieformalnych sieci wsparcia, mentorstwa i informacji (Fox, 2005; Chavarro et al., 2018).

W dłuższej perspektywie mechanizmy te ograniczają różnorodność i sprzyjają trwałej nierównowadze między badaczami różniącymi się nie tylko płcią, ale także kohortą, typem instytucji czy pozycją geograficzną. Skutkiem bywa niższa innowacyjność, zubożenie perspektyw badawczych oraz utrata potencjału intelektualnego w skali systemowej (Hicks et al., 2015; UNESCO, 2021). Z ekonomicznego punktu widzenia oznacza to także zmniejszenie konkurencyjności gospodarki opartej na wiedzy.

Dlatego temat nierówności należy traktować jako *pełnoprawny komponent* analiz produktywności i dynamiki karier, a nie ich uzupełnienie. Strukturalne różnice w dostępie do zasobów i sieci współpracy współkształtują rozkłady produktywności i mechanizmy akumulacji przewagi. Rekomendacje organizacji takich jak **DORA** (San Francisco Declaration on Research Assessment, 2012), **Leiden Manifesto** (Hicks et al., 2015) czy **UNESCO** (UNESCO, 2021) wskazują jednoznacznie na potrzebę:

- odejścia od prostych metryk i wskaźników agregowanych,
- różnicowania kryteriów oceny względem wieku, ścieżki kariery i dostępności zasobów,
- promowania inkluzywności, równości szans i przejrzystości procedur.

Zgłoszony cykl wnosi do tej debaty unikalny wkład empiryczny: po raz pierwszy w Polsce zastosowano podejście łączące dane administracyjne (zatrudnienie, wiek, stanowisko, instytucja) z danymi bibliometrycznymi (publikacje, współpraca, widoczność), co umożliwiło pogłębioną analizę nierówności w systemie nauki. Zintegrowane podejście pozwoliło badać zarówno nierówności poziome (np. różnice w liczbie publikacji), jak i strukturalne (np. dostęp do sieci współpracy, przywództwa autorskiego, mobilności międzynarodowej), z wykorzystaniem zaawansowanych metod analizy danych (m.in. regresji frakcyjnej i modeli z efektami stałymi). Wyniki te dostarczają podstaw do projektowania polityk publicznych, które łączą perspektywę równości szans z dążeniem do zwiększenia efektywności i konkurencyjności systemu.

4.4.2 Nierówności płciowe w systemie nauki

W ramach cyklu podjąłem analizę nierówności płciowych, wykorzystując połączone dane administracyjne i bibliometryczne, które umożliwiły jednoczesne uwzględnienie zmiennych demograficznych, instytucjonalnych oraz cech dorobku publikacyjnego i współpracy. Taka perspektywa pozwoliła na identyfikację trwałych różnic między kobietami i mężczyznami w zakresie:

- udziału w publikacjach jednoautorskich (*gender solo research gap*),
- wzorców współautorstwa (*gender homophily*),
- zaangażowania we współpracę międzynarodową,
- dostępu do cytowanych, prestiżowych publikacji oraz do centralnych pozycji w strukturach współpracy.

W artykułach (Kwiek and Roszka, 2021b, 2022b) wykazano, że kobiety rzadziej niż mężczyźni publikują samodzielnie, a ich udział w publikacjach międzynarodowych oraz zespołach o szerokim zasięgu jest istotnie niższy — także po kontrolowaniu poziomu produktywności. Skonstruowany przeze mnie wskaźnik *gender solo research gap*, oszacowany modelami frakcyjnej regresji logistycznej, wskazuje na uśrednioną różnicę 6–10 punktów procentowych na niekorzyść kobiet (w zależności od dyscypliny i wieku akademickiego). Z ekonomicznego punktu widzenia oznacza to niewykorzystanie części kapitału ludzkiego oraz mniejszą efektywność inwestycji publicznych w naukę.

Analiza wzorców współautorskich wykazała, że kobiety częściej tworzą zespoły jedнопłciowe, co ogranicza ich dostęp do zróżnicowanych sieci i zasobów akademickich. Mężczyźni natomiast częściej funkcjonują w bardziej heterogenicznych sieciach współpracy, co zwiększa ich centralność i widoczność (Sugimoto et al., 2013; West et al., 2013). Nierówności te są szczególnie widoczne w obszarze *STEM* (Science, Technology, Engineering, Mathematics; w niektórych badaniach rozszerzonym do *STEMM*, obejmującym dodatkowo medycynę), natomiast w *HSS* ich skala jest mniejsza (Sugimoto et al., 2013). W ujęciu intersekcyjnym kobiety zatrudnione w mniejszych instytucjach lub w ośrodkach spoza elity badawczej mają dodatkowo ograniczony dostęp do współpracy międzynarodowej i publikacji w czasopiśmie o najwyższym prestiżu (van den Besselaar and Sandström, 2011; Holman et al., 2018; Huang et al., 2020).

Zidentyfikowane zjawiska odzwierciedlają mechanizmy strukturalne: dodatkowe obciążenia pozanaukowe, słabszy dostęp do mentoringu i sieci wsparcia, a także większy udział kobiet w „niewidzialnej pracy” akademickiej (dydaktyka, administracja) (Fox, 2005; Elsevier, 2020; UNESCO, 2021). W dłuższej perspektywie prowadzi to do obniżenia konkurencyjności systemu nauki, ponieważ ogranicza pełne wykorzystanie potencjału badaczek.

Rekomendacje:

- wdrożenie transparentnego raportowania ról autorskich (np. CRediT) (Larivière et al., 2016), co pozwoli lepiej mierzyć indywidualny wkład i ograniczy niedoszacowanie dorobku kobiet,
- uwzględnianie przerw w karierze (np. urlopy opiekuńcze) w systemach ewaluacji i awansu, co zwiększy porównywalność ścieżek,
- monitorowanie udziału kobiet w prestiżowych aktywnościach (granty, redakcje, keynote talks), aby przeciwdziałać koncentracji zasobów wśród jednej grupy,
- promowanie różnorodności zespołów w projektach grantowych i ocenie instytucji, co sprzyja zarówno inkluzywności, jak i większej innowacyjności.

Zastosowane podejście — integracja danych oraz estymacja modeli frakcyjnych — pozwoliło uchwycić nie tylko różnice „poziome” (liczba publikacji), lecz także asymetrie w strukturze współpracy i pozycji w sieci. Wskazuje to na konieczność traktowania nierówności płciowych nie tylko jako kwestii równościowej, ale również jako problemu *ekonomicznej efektywności* systemu nauki.

4.4.3 Wiek, staż akademicki i nierówności pokoleniowe

Istotnym, a niedostatecznie uwzględnianym wymiarem nierówności jest wpływ wieku biologicznego i akademickiego na produktywność, widoczność i szanse awansu. W artykule (Kwiek and Roszka, 2022a) wprowadzam rozróżnienie między *wiekami biologicznym* a *wiekami akademickim* (liczonym jako liczba lat od pierwszej publikacji w Scopus), co pozwala uchwycić rzeczywisty staż badawczy oraz zredukować zakłócenia wynikające z opóźnionego wejścia do systemu, przerw w karierze czy nielinearności ścieżek. Takie podejście zwiększa trafność analiz produktywności i pozwala na bardziej sprawiedliwe porównania między badaczami.

Modele regresyjne pokazały, że wiek akademicki jest silniejszym predyktorem produktywności niż wiek biologiczny — zarówno w modelach z efektami stałymi dla dyscyplin, jak i w modelach frakcyjnych. Efekt ten utrzymywał się także w alternatywnych specyfikacjach (różne progi klasyfikacji produktywności, alternatywne miary publikacyjnej intensywności), co wskazuje na jego stabilność. Warstwowanie według kohort i wieku akademickiego ujawnia różnice trajektorii niezależne od wieku kalendarzowego (Sugimoto et al., 2016; Larivière and Sugimoto, 2015).

W analizach kohortowych i porównaniach produktywności szczególną uwagę poświęcono różnicom płciowym, które w wielu krajach stanowią trwały element stratyfikacji akademickiej. Wskaźniki *gender solo research gap* oraz *index of gender homophily* pozwalają ilościowo uchwycić nierówności w zakresie współautorstwa i samodzielności badawczej, tradycyjnie analizowane w ujęciu heterodoksyjnej ekonomii nauki. Zastosowanie pełnopopulacyjnych danych administracyjno-bibliometrycznych umożliwiło weryfikację hipotez o

zróźnicowaniu produktywności i trwałości elity w zależności od płci i wieku akademickiego (Kwiek, 2020).

Z perspektywy ekonomicznej nieuwzględnianie wieku akademickiego w ocenie dorobku prowadzi do nieefektywnej alokacji zasobów...

W (Kwiek and Roszka, 2024c) analizuję różnice międzykohortowe w tempie wzrostu dorobku oraz mobilności między klasami produktywności (z użyciem klasyfikacji 10/50). Młodsze kohorty (po 2005 r.) częściej osiągają pozycję w top 10% we wcześniejszych fazach kariery i cechują się większą dynamiką wzrostu; starsze kohorty wykazują większą stabilność, ale mają mniejszą szansę wejścia do elity. Zjawisko to odzwierciedla *bias of acceleration*: system premiuje szybkie, wczesne sukcesy, co sprzyja kumulacji przewag (Merton, 1968; Fortunato et al., 2018). W grupach przecinających się (np. kobiety o długim wieku akademickim w mniejszych ośrodkach) efekty nierówności są szczególnie nasilone (Sugimoto et al., 2013; UNESCO, 2021).

Z perspektywy ekonomicznej nieuwzględnianie wieku akademickiego w ocenie dorobku prowadzi do nieefektywnej alokacji zasobów: niedoszacowania potencjału młodszych kohort, pomijania *late bloomers*, a także do utrwalenia niesprawiedliwych ścieżek awansów. Podejście oparte na wieku akademickim umożliwia projektowanie polityk bardziej spójnych z realnym przebiegiem karier, a tym samym zwiększających zwrot z inwestycji publicznych w naukę.

Rekomendacje:

- wprowadzenie wieku akademickiego jako standardowego wymiaru oceny dorobku (Kwiek and Roszka, 2022a),
- formalne uwzględnianie *career breaks* i korekt w ewaluacji (np. modele progowe/tempa rozwoju stosowane w Australii i Kanadzie) (Sugimoto et al., 2016; Larivière and Sugimoto, 2015),
- projektowanie alternatywnych ścieżek kariery (eksperymentalnych, dydaktycznych, interdyscyplinarnych), aby ograniczyć negatywne skutki „bias of acceleration” (UNESCO, 2021),
- systemowe monitorowanie kohortowych trajektorii w ujęciu longitudinalnym i powiązanie ich z instrumentami wsparcia (granty startowe, mentoring, polityki projaściowe) (Fortunato et al., 2018; Moed, 2005).

4.4.4 Wzorce współpracy i dostęp do zasobów

Współpraca naukowa jest jednym z głównych mechanizmów akumulacji przewagi w nauce: sprzyja wzrostowi produktywności, zwiększa cytowalność, ułatwia dostęp do finansowania grantowego i wzmacnia prestiż instytucjonalny. Jednak dostęp do międzynarodowych sieci

badawczych i prestiżowych kanałów publikacji jest silnie zróżnicowany i niejednokrotnie ograniczony przez uwarunkowania instytucjonalne, geograficzne oraz demograficzne.

W artykułach (Kwiek and Roszka, 2021b) (dane POL-on-Scopus) oraz (Kwiek and Roszka, 2024b) (analiza *Scopus-only*) przeanalizowałem wzorce współautorstwa i udział we współpracy międzynarodowej, identyfikując mechanizmy koncentracji i wykluczenia.

Najważniejsze ustalenia empiryczne:

- **Elita badawcza (top 10%)** tworzy zwarte i silnie powiązane sieci współautorskie, w których współpraca międzynarodowa jest standardem. Sieci te są zlokalizowane głównie w największych ośrodkach akademickich i dyscyplinach STEM, co potwierdza tezę o przestrzennej i dziedzinowej koncentracji zasobów (Fortunato et al., 2018).
- **Badacze z mniejszych uczelni i jednostek regionalnych** mają istotnie ograniczony dostęp do międzynarodowych partnerstw, co przekłada się na mniejszą widoczność publikacyjną i ograniczoną obecność w prestiżowych czasopismach (Chavarro et al., 2018). Efekt ten utrzymuje się także po kontrolowaniu produktywności i dyscypliny.
- **Kobiety i badacze o późniejszym wieku akademickim** są niedoreprezentowani w strukturach współpracy o wysokim prestiżu (Q1, wysoko cytowane prace, projekty z partnerami zagranicznymi). Modele regresji wskazują, że nawet przy tej samej produktywności kobiety mają przeciętnie mniejszą liczbę unikalnych współautorów i węższy zasięg geograficzny współpracy (Fox, 2005; Sugimoto et al., 2013; West et al., 2013; Holman et al., 2018).
- **Silna koncentracja współpracy międzynarodowej:** w części dyscyplin 10–15% najbardziej aktywnych badaczy odpowiada za ponad połowę publikacji z udziałem autorów zagranicznych. Zjawisko to zwiększa dywergencję między elitą a resztą systemu.

Wnioski analityczne: Modele logistyczne i frakcyjne (z kontrolą dyscypliny, produktywności, płci i instytucji) pokazują, że kluczowymi barierami są:

- *lokalizacja instytucjonalna* — przewaga centralnych, dużych ośrodków nad jednostkami regionalnymi,
- *brak wczesnych doświadczeń we współpracy zagranicznej* — co prowadzi do efektu ścieżki zależnej i ogranicza późniejsze szanse wejścia do prestiżowych sieci.

Rekomendacje dla polityki naukowej:

- wzmocnienie instrumentów wspierających mobilność międzynarodową w początkowych fazach kariery (np. granty startowe, staże postdocowe),

- monitorowanie i ograniczanie koncentracji współpracy w elitarnych sieciach poprzez polityki inkluzywne (uwzględniające mniejsze uczelnie i regiony),
- promowanie różnorodnych i interdyscyplinarnych zespołów w ocenie grantów i ewaluacji instytucjonalnej,
- integracja perspektywy płci i wieku akademickiego w ocenie dostępu do międzynarodowych zasobów i sieci.

Uzyskane wyniki potwierdzają, że współpraca naukowa w Polsce — podobnie jak w innych krajach (Fortunato et al., 2018; Larivière and Sugimoto, 2015) — jest silnie stratyfikowana, a jej elitarna koncentracja utrwała nierówności i ogranicza pełne wykorzystanie potencjału badawczego całej populacji.

4.4.5 Implikacje dla polityki naukowej

Wnioski płynące z przeprowadzonych przeze mnie analiz nierówności w systemie nauki mają bezpośrednie znaczenie dla projektowania polityki opartej na danych (*evidence-informed policy*). Zidentyfikowane zjawiska – takie jak *gender solo research gap*, nierówności we współpracy, efekty wieku akademickiego, ograniczona mobilność międzyklasowa oraz strukturalne wykluczenie instytucjonalne – wskazują, że obecne mechanizmy oceny, finansowania i awansu mogą nieświadomie reprodukować istniejące asymetrie (Fox, 2005; Sugimoto et al., 2013; Ioannidis et al., 2016).

Analizy oparte na pełnej populacji badaczy zatrudnionych w polskich uczelniach i widocznych w Scopus pokazują jednoznacznie, że system nauki w Polsce funkcjonuje w warunkach wysokiej selektywności. Raz zdobyta przewaga (np. prestiżowa afiliacja, wczesna produktywność, współpraca międzynarodowa) przekłada się na kolejne sukcesy – zgodnie z mechanizmem kumulatywnej przewagi (Merton, 1968; Fortunato et al., 2018). Jednocześnie badacze o nielinearnych karierach, późnym wejściu do systemu lub bez wsparcia instytucjonalnego napotykaają trwałe bariery w dostępie do zasobów i prestiżu.

W oparciu o przeprowadzone badania rekomenduję działania w trzech priorytetowych obszarach:

Redefinicja wskaźników produktywności i dorobku naukowego. Obecnie stosowane metryki (np. liczba publikacji, liczba cytowań, punkty MEiN) nie uwzględniają istotnych różnic w trajektorii kariery, co prowadzi do systematycznej stronniczości. Pomijane są m.in.:

- wiek akademicki i dynamika dorobku w czasie (Kwiek and Roszka, 2022a),
- przerwy w karierze naukowej (career breaks) związane z opieką, zdrowiem lub migracją,

- zróżnicowane role autorskie (lider, autor korespondencyjny, członek zespołu),
- nierówny dostęp do zasobów (granty, infrastruktura, sieci współpracy).

Rekomenduję wdrożenie *relatywnych* wskaźników produktywności – skorygowanych o wiek akademicki, typ instytucji oraz rolę autorską – które lepiej oddają rzeczywisty wkład badacza. Proponowane narzędzia, takie jak produktywność percentylowa, udział publikacji jednoautorskich czy miary centralności w sieci współpracy, mogą stać się podstawą bardziej kontekstowej i sprawiedliwej oceny dorobku (Sugimoto et al., 2013).

Zwiększenie równości w dostępie do zasobów, współpracy i szans. Moje analizy pokazują, że kobiety, osoby z mniejszych uczelni oraz badacze z późniejszych kohort napotykają istotne bariery w dostępie do:

- prestiżowych sieci współpracy i zespołów badawczych,
- projektów międzynarodowych i mobilności akademickiej,
- ról liderek w publikowaniu i grantach.

Zaleca się wprowadzenie mechanizmów wyrównawczych, takich jak:

- priorytety w grantach mobilnościowych i partnerskich dla grup niedoreprezentowanych (młodsze kohorty, kobiety, uczelnie regionalne),
- dodatkowa punktacja za różnorodność zespołów (np. w ocenie wniosków NCN),
- instrumenty wspierające powroty po przerwach w karierze (granty powrotowe),
- programy mentorskie i konsorcyjne łączące badaczy z różnych typów instytucji.

Są to rozwiązania o szybkim i mierzalnym wpływie, redukujące nierówności już na etapie dostępu do zasobów.

Transparentność i systematyczne monitorowanie nierówności. Równość w systemie nauki wymaga aktywnego monitorowania i przejrzystości procesów oceny. Rekomenduję:

- wdrożenie systematycznego monitorowania nierówności w dostępie do zasobów, współpracy i prestiżu (np. przez MEiN, NCN),
- upowszechnienie standardu **CRediT** jako narzędzia raportowania ról autorskich,
- obowiązek raportowania składu zespołów pod względem płci, wieku akademickiego i typu instytucji,

- włączenie wskaźników inkluzywności i otwartości (np. udział kobiet w rolach liderkich, interdyscyplinarność) do oceny jednostek i projektów,
- rozwój praktyk otwartej nauki i otwartych danych (zgodnie z UNESCO 2021), zwiększających przejrzystość i replikowalność ocen.

Podsumowanie. Zgłoszone przeze mnie analizy stanowią nie tylko diagnozę systemu, lecz także podstawę realistycznych i spójnych z najlepszymi praktykami międzynarodowymi działań reformatorskich. Pokazują one, że reforma systemu oceny i finansowania, oparta na pełnopopulacyjnych mikrodanych i wskaźnikach relatywnych, może zwiększyć efektywność, sprawiedliwość i innowacyjność polskiej nauki. Moje propozycje łączą unikalny wkład empiryczny (integracja POL-on–Scopus) z jasnymi wskazówkami politycznymi, co pozwala na lepsze wykorzystanie kapitału ludzkiego w nauce i ograniczenie strat wynikających z niewykorzystanego potencjału.

Wnioski z podstawowych osiągnięć naukowych

Zgłoszone przeze mnie osiągnięcie naukowe – cykl publikacji poświęconych analizie struktury, dynamiki i uwarunkowań produktywności naukowej w Polsce – stanowi próbę kompleksowego opisu funkcjonowania systemu akademickiego z wykorzystaniem danych pełnopopulacyjnych oraz technik statystycznych umożliwiających analizę zjawisk w ujęciu mikropoziomym, kohortowym i longitudinalnym. Punktem wyjścia była potrzeba lepszego zrozumienia procesów sukcesu, mobilności oraz nierówności w nauce, a także stworzenia infrastruktury analitycznej mogącej służyć dalszym badaniom i diagnozie systemowej.

W ramach cyklu zaprojektowałem i wdrożyłem **operat danych integrujący źródła administracyjne (POL-on) i bibliometryczne (Scopus)**, z wykorzystaniem podejścia probabilistycznego (Christen, 2012). Operat ten może stanowić trwałe i replikowalne narzędzie analityczne, umożliwiające modelowanie ścieżek karier naukowych, analizę rozkładów produktywności oraz identyfikację uwarunkowań demograficznych i instytucjonalnych. Opracowane procedury oraz przyjęta metodologia są zgodne z podejściami stosowanymi w *science of science* w krajach o bardziej rozwiniętej infrastrukturze danych, a zarazem stanowią **pierwszy w Polsce przykład pełnopopulacyjnej integracji administracyjno-bibliometrycznej**.

Analizy przeprowadzone na tej bazie pozwoliły uchwycić zjawiska takie jak: koncentracja dorobku publikacyjnego (reguła 10/50), trwałość pozycji elity badawczej (Kwiek and Roszka, 2024a), ograniczona mobilność między klasami produktywności (Kwiek and Roszka, 2025) czy też rozbieżności między karierą instytucjonalną a rzeczywistą aktywnością naukową (Kwiek and Roszka, 2025). Wyniki wskazują również na występowanie nierówności związanych z płcią, wiekiem akademickim i typem instytucji, w tym m.in. *gender solo research gap* (Kwiek and Roszka, 2022b), homofilii płciowej (Kwiek and Rosz-

ka, 2021a) oraz zróżnicowanego dostępu do współpracy międzynarodowej i prestiżowych kanałów publikacji. Kobiety napotykają także dodatkowe bariery w dostępie do najbardziej prestiżowych sieci współpracy i są rzadziej obecne w strukturach współautorstwa o wysokiej centralności, co ogranicza ich szanse rozwoju kariery (Kwiek and Roszka, 2021a).

W zgłoszonych pracach podjęto również próbę sformułowania **rekomendacji dla polityki naukowej** – dotyczących m.in. konstrukcji alternatywnych wskaźników produktywności, uwzględnienia wieku akademickiego w ewaluacji dorobku oraz potrzeby systematycznego monitorowania nierówności. Wnioski te mogą być pomocne w projektowaniu bardziej zróżnicowanych, inkluzywnych i opartych na danych modeli zarządzania wiedzą i kapitałem ludzkim w nauce.

Mam świadomość, że przedstawione podejścia nie wyczerpują złożoności analizowanych zjawisk, a wiele ustaleń wymaga dalszej weryfikacji i pogłębienia. Jednym z kluczowych wniosków jest jednak wysoka **stabilność pozycji w strukturze produktywności badawczej**, co zostało po raz pierwszy wykazane w analizie pełnopopulacyjnej (Kwiek and Roszka, 2024a). Uważam, że zaproponowana infrastruktura danych, zastosowane metody oraz uzyskane wyniki stanowią istotny krok w kierunku rozwijania badań nad funkcjonowaniem systemu nauki w Polsce – zarówno w wymiarze akademickim, jak i aplikacyjnym – i mogą znaleźć zastosowanie także w badaniach porównawczych oraz działaniach instytucjonalnych.

Warto podkreślić, że rosnącą rolę w integracji danych pełni obecnie **deterministyczne dopasowywanie rekordów** z wykorzystaniem identyfikatorów takich jak ORCID. Choć w okresie realizacji prac pokrycie ORCID wśród badaczy afiliowanych w Polsce było zbyt niskie, aby stanowić główny klucz integracyjny, obserwowany wzrost popularności tego standardu może w przyszłości umożliwić przejście do integracji deterministycznej lub hybrydowej. Zaprojektowana przeze mnie procedura probabilistyczna, choć powstała w warunkach braku jednoznacznych identyfikatorów, pozostaje **elastycznym fundamentem**, który może być adaptowany do nowych realiów i stanowić podstawę nowoczesnych operatów danych łączących podejścia deterministyczne i probabilistyczne – zależnie od dostępności i jakości źródeł.

Przeprowadzone analizy nierówności zrealizowano w oparciu o zintegrowany operat danych POL-on-Scopus (stanowiący podstawę zgłaszanego cyklu) i uzupełniono analizą bibliometryczną opartą wyłącznie na danych Scopus w jednym z artykułów cyklu. Dzięki temu osiągnięcie łączy **oryginalny wkład metodologiczny z empirycznym testem teorii socjologii nauki i ekonomii szkolnictwa wyższego**, lokując je w dyscyplinie *ekonomia i finanse*, a zarazem w międzynarodowym nurcie *science of science*.

5 Aktywność naukowa poza jednostką organizacyjną zatrudnienia

W latach 2016–2025 prowadziłem systematyczną i zróżnicowaną działalność naukową poza macierzystą jednostką organizacyjną zatrudnienia – Uniwersytetem Ekonomicznym w Poznaniu. Obejmowała ona zarówno długookresową współpracę z Uniwersytetem im. Adama Mickiewicza w Poznaniu, jak i realizację ekspertyz oraz projektów analitycznych na zlecenie instytucji publicznych, instytutów naukowych oraz podmiotów gospodarczych. Jej istotnym komponentem był również trzymiesięczny staż naukowy na University of Southampton w Wielkiej Brytanii.

Aktywność ta miała charakter interdyscyplinarny, aplikacyjny i teoriiotwórczy. W projektach realizowanych poza UEP odpowiadałem nie tylko za analizy danych i dobór metod ilościowych, lecz także za rozwój autorskich koncepcji wskaźnikowych, konstrukcję operatów mikrodanych, modelowanie zjawisk z wykorzystaniem metod regresyjnych oraz interpretację wyników w kontekście instytucjonalnym i systemowym. Moja rola obejmowała również współprojektowanie struktury raportów i artykułów naukowych oraz formułowanie wniosków i rekomendacji politycznych, co miało znaczenie zarówno dla debaty akademickiej, jak i praktyki instytucjonalnej.

Działania te były prowadzone w pełni niezależnie od zatrudnienia w UEP – instytucjonalnie, formalnie i finansowo – i stanowiły ważny element mojego rozwoju naukowego. Rozszerzyły zakres tematyczny badań, umożliwiły testowanie i wdrażanie nowych metod statystycznych oraz pogłębiły doświadczenie w pracy zespołowej w środowiskach badawczych o zróżnicowanej specyfice dziedzinowej. W efekcie przyczyniły się zarówno do powstania publikacji zgłoszonych jako osiągnięcie naukowe, jak i do transferu wiedzy do polityk publicznych i praktyki gospodarczej, wpisując się w nurt *science of science* oraz *ekonomicznej analizy nauki* (ang. *economics of science*).

5.1 Projekty realizowane na Uniwersytecie im. Adama Mickiewicza w Poznaniu

W latach 2013–2024 byłem wykonawcą czterech ogólnopolskich projektów badawczych realizowanych w Centrum Studiów nad Polityką Publiczną UAM pod kierownictwem prof. Marka Kwieka. Poniżej przedstawiam ich zwięzły opis i mój wkład.

5.1.1 NCN MAESTRO „Program międzynarodowych badań porównawczych szkolnictwa wyższego” (2012–2017)

Projekt MAESTRO realizowano w latach 2012–2017 w Centrum Studiów nad Polityką Publiczną UAM. Był to wówczas największy w Polsce projekt poświęcony badaniom

porównawczym szkolnictwa wyższego (NCN, nr 2011/02/A/HS6/00183). Celem było zrozumienie przemian systemów akademickich w Europie i na świecie – mechanizmów stratyfikacji w nauce, wzorców publikowania i współpracy oraz dynamiki karier.

Mój wkład (część ilościowa projektu):

- **Przygotowanie i harmonizacja danych:** udział w tworzeniu pierwszych wersji operatu mikro danych o polskich naukowcach; czyszczenie, harmonizacja i wstępne dopasowanie rekordów między źródłami.
- **Operacjonalizacja pojęć i wskaźników:** współudział w konstrukcji wczesnych miar produktywności, umiędzynarodowienia i współpracy (punkt wyjścia dla późniejszych, bardziej zaawansowanych wskaźników).
- **Analizy porównawcze:** prace nad ujęciami kohortowymi i porównaniami Polska–Europa Zachodnia w zakresie ścieżek kariery i produktywności.
- **Wkład metodologiczny:** rozwijanie ram łączących perspektywy socjologiczną i ekonomiczną, które następnie rozwinąłem w kolejnych projektach.

Projekt MAESTRO był **fundamentem metodologicznym i koncepcyjnym** dla późniejszej integracji danych administracyjnych i bibliometrycznych oraz moich badań w obszarze *ekonomicznej analizy nauki*. Wnioski i doświadczenia uzyskane w ramach tego projektu stanowiły punkt wyjścia do opracowania autorskiego podejścia integracyjnego zastosowanego w osiągnięciu habilitacyjnym.

5.1.2 Doskonałość naukowa: konkurencyjność, mierzalność, umiędzynarodowienie (DIALOG, 2016–2019)

Projekt „Doskonałość naukowa” miał na celu pogłębioną diagnozę funkcjonowania polskiego szkolnictwa wyższego ze szczególnym naciskiem na finansowanie, politykę kadrową i awanse. Badania oparto na integracji danych administracyjnych (POL-on), bibliometrycznych (Scopus) oraz informacji finansowo–organizacyjnych na poziomie uczelni.

Mój wkład:

- **Integracja i rekonstrukcja danych:** zaprojektowanie procedur łączenia źródeł (POL-on, Scopus, dane finansowe), dopasowanie rekordów instytucjonalnych, harmonizacja kategorii i walidacja kompletności.
- **Wskaźniki efektywności:** opracowanie wskaźników relacyjnych (nakłady–wyniki) dla oceny efektywności instytucjonalnej; wykorzystanie ich w rankingach i analizach porównawczych.

- **Ścieżki kariery i awanse:** analizy zależności między produktywnością naukową a tempem awansów; identyfikacja rozbieżności między awansem instytucyjnym a realną produktywnością.
- **Zróżnicowanie instytucjonalne:** klasyfikacja typów uczelni (badawcze, klasyczne, zawodowe, niskoproduktywne) oparta na strukturze publikacji i umiędzynarodowieniu.

Wkład teoriiwórczy: rozróżnienie „produktywności potencjalnej” (wynikającej z zasobów) i „zrealizowanej” (mierzonej bibliometrycznie), rozwinięcie podejścia relacyjnego (nakłady–wyniki) oraz zastosowanie ujęcia kohortowego do analizy awansów.

5.1.3 Uczelnia badawcza 2.0. Finansowanie, zarządzanie i polityka kadrowa (DIALOG, 2020–2023)

Projekt „Uczelnia badawcza 2.0” dostarczył zintegrowaną diagnozę systemu szkolnictwa wyższego z wykorzystaniem danych administracyjnych (POL-on), bibliometrycznych (Scopus) oraz finansowo–organizacyjnych.

Mój wkład:

- **Integracja i rekonstrukcja danych:** pełna integracja wieloźródłowa (POL-on, Scopus, finanse), w tym dopasowanie instytucji, harmonizacja kategorii i walidacja kompletności; przygotowanie nowatorskiego operatu mikrodanych o uczelniach.
- **Wskaźniki efektywności instytucjonalnej:** zaprojektowanie zestawu wskaźników względem rezultatów (publikacje w górnych percentylach, umiędzynarodowienie, udział publikacji zagranicznych); konstrukcja rankingów i profili instytucjonalnych.
- **Wzorce awansu naukowego:** modelowanie zależności między produktywnością (ilościową i jakościową) a prawdopodobieństwem oraz tempem awansów.
- **Zróżnicowanie instytucjonalne:** klasyfikacja typów uczelni według struktury publikacji, koncentracji dorobku i udziału badań międzynarodowych; analiza skutków IDUB i implikacji dla finansowania.

Wkład teoriiwórczy: ugruntowanie rozróżnienia „produktywność potencjalna” vs „zrealizowana”, rozwinięcie podejścia relacyjnego (nakłady–wyniki) oraz pogłębienie ujęcia kohortowego i longitudinalnego w analizie karier.

5.1.4 Polscy Naukowcy 2022

Projekt *Polscy Naukowcy 2022* był największym w Polsce przedsięwzięciem badawczym dotyczącym szkolnictwa wyższego, zrealizowanym w ramach programu „Nauka dla Społeczeństwa” w CSnPP UAM. Badanie objęło ponad 60 tys. polskich naukowców, do których

skierowano zaproszenia do udziału w ankiecie internetowej (CAWI); **kwestionariusz w pełni wypełniło około 11 tys. osób**. Jest to bardzo dobry wynik – odpowiadający zwrotności na poziomie blisko 20% – biorąc pod uwagę, że typowe wskaźniki odpowiedzi w badaniach internetowych mieszczą się zazwyczaj w przedziale 5–7% (Manfreda et al., 2008).

Innowacją projektu było rozszerzenie wyników ankiety o integrację dwóch źródeł: danych administracyjnych z systemu POL-on (zatrudnienie, awanse, afiliacje) oraz danych bibliometrycznych z bazy Scopus (publikacje, cytowania, współpraca międzynarodowa). Integracja ta miała charakter **deterministyczny**, ponieważ respondenci byli identyfikowani nie tylko na podstawie imienia, nazwiska, afiliacji i adresu e-mail, lecz także dzięki unikalnym identyfikatorom *Scopus ID*.

Mój wkład:

- udział w projektowaniu i realizacji procedury deterministycznej integracji danych ankietowych z bazami POL-on i Scopus przy wykorzystaniu *Scopus ID*, co zapewniło jednoznaczne powiązanie odpowiedzi respondentów z ich dorobkiem publikacyjnym i karierą akademicką;
- wsparcie koncepcyjne i techniczne procesu integracji CAWI z danymi administracyjnymi i bibliometrycznymi, w tym walidacja powstałego operatu;
- rozwój infrastruktury danych o nauce poprzez włączenie do operatu zmiennych **subiektywnych** (postawy, aspiracje, motywacje, oceny polityk), których nie obejmowały wcześniejsze badania pełnopopulacyjne;
- analiza jakości i zwrotności badania CAWI w populacji ponad 60 tys. naukowców – z rekordowym odsetkiem pełnych odpowiedzi (ok. 11 tys.), co znacząco przewyższa typowe wskaźniki odpowiedzi w badaniach internetowych;
- interpretacja wyników w kontekście umiędzynarodowionej literatury o karierach akademickich, aspiracjach płacowych i uwarunkowaniach produktywności, w duchu *ekonomicznej analizy nauki* (ang. *economics of science*) i *naukoznawstwa ilościowego* (*science of science*).

Projekt ten **poszerza operat o zmienne subiektywne** (postawy, motywacje, oceny polityk), tworząc nową infrastrukturę danych o polskiej nauce. **Wyniki są obecnie w drugiej rundzie recenzji i nie wchodzi w skład zgłaszanego cyklu habilitacyjnego**, stanowią jednak ważne uzupełnienie i zaplecze metodologiczne moich dalszych badań w zakresie *ekonomicznej analizy nauki*.

5.1.5 Ekspertyza bibliometryczna na zlecenie kierownictwa programu IDUB (marzec 2025)

Na zlecenie kierownictwa programu IDUB² przygotowałem w marcu 2025 roku **samodzieloną ekspertyzę bibliometryczną** dotyczącą skali, dynamiki i charakteru publikowania w czasopismach wydawnictwa MDPI. Celem opracowania było zdiagnozowanie znaczenia MDPI w dorobku pracowników UAM oraz porównanie pozycji uczelni z innymi ośrodkami uczestniczącymi w programie Inicjatywa Doskonałości – Uczelnia Badawcza (IDUB).

Zakres ekspertyzy obejmował:

- identyfikację autorów i jednostek organizacyjnych UAM publikujących w MDPI (na podstawie danych Scopus),
- analizę liczby publikacji, cytawalności oraz jakości czasopism (CiteScore, SNIP, SJR),
- porównania międzyuczelniane (udział publikacji MDPI w dorobku, relacje cytowań do IF czasopism),
- klasyfikację jednostek uczestniczących w IDUB pod względem intensywności publikowania w MDPI,
- ocenę potencjalnych ryzyk związanych z koncentracją dorobku w czasopismach o zróżnicowanym poziomie selektywności.

Znaczenie instytucjonalne. Ekspertyza została wykorzystana wewnętrznie przez władze programu IDUB jako materiał wspierający politykę wydawniczą, w tym rozważania nad kształtowaniem zasad publikowania i wspierania jakościowego rozwoju dorobku. Stanowiła narzędzie do podejmowania decyzji strategicznych w zakresie polityki naukowej i ewaluacji.

Działalność ta była realizowana całkowicie poza jednostką mojego zatrudnienia (UEP) i miała charakter niezależnej, eksperckiej analizy bibliometrycznej o wysokim znaczeniu instytucjonalnym.

Uwagi ogólne dotyczące operatu mikrodanych

W wielu projektach badawczych opisanych w niniejszym autoreferacie wykorzystano ten sam zintegrowany operat mikrodanych obejmujący 25 463 naukowców zatrudnionych w

²Inicjatywa Doskonałości – Uczelnia Badawcza (IDUB); Program Ministerstwa Nauki i Szkolnictwa Wyższego ustanowiony w 2019 r., którego celem jest podniesienie międzynarodowej pozycji wybranych uczelni polskich poprzez dodatkowe finansowanie badań naukowych i umiędzynarodowienie działalności akademickiej.

polskich uczelniach akademickich (stan na listopad 2017 roku) posiadających stopień doktora lub wyższy. Operat ten powstał w ramach projektu *Transformacje pracy akademickiej: produktywność badawcza, sukces i status w nauce* (NCN MAESTRO, 2019–2022), gdzie odpowiadałem za jego opracowanie, walidację oraz przygotowanie do analiz longitudinalnych.

W kolejnych projektach (m.in. *Uczelnia badawcza 2.0* oraz *Polscy Naukowcy 2022*) operat ten był rozszerzany o dodatkowe źródła danych (np. dane finansowe, wyniki ankiet, dane o grantach i współpracy międzynarodowej) i aktualizowany zgodnie z potrzebami analitycznymi. Tym samym stanowił on trwałą infrastrukturę analityczną, będącą podstawą wielu publikacji naukowych zgłoszonych jako osiągnięcie habilitacyjne.

Warto podkreślić, że wykorzystanie tego samego operatu w różnych projektach nie oznacza powielania analiz, lecz budowę **komplementarnych modułów badawczych**. Każda z publikacji wchodzących w skład cyklu odpowiada na odrębne pytania badawcze, stosuje własne zestawy wskaźników i modele (logistyczne, frakcyjne, kohortowe, longitudinalne), a wspólna infrastruktura danych zapewnia jedynie spójność i porównywalność wyników. Dzięki temu możliwe było uchwycenie zarówno zjawisk punktowych (np. *gender solo research gap*), jak i procesów długookresowych (np. trwałości pozycji w elitach produktywnościowych).

Rozwój operatu miał charakter iteracyjny: począwszy od integracji POL-on i Scopus w projekcie MAESTRO, poprzez rozszerzenie o dane finansowe i organizacyjne w projekcie *Uczelnia badawcza 2.0*, aż po połączenie z wynikami badania ankietowego CAWI w projekcie *Polscy Naukowcy 2022*. Każdy etap wносił nowe źródła danych i metody analityczne, co zwiększało wartość operatu i pozwalało na formułowanie coraz bardziej zaawansowanych wniosków o funkcjonowaniu systemu nauki.

Uwaga: Ze względu na wspólną genezę operatu, jego wykorzystanie w różnych projektach nie jest powieleniem osiągnięć, lecz świadomym i metodycznie uzasadnionym podejściem do budowy komplementarnych analiz w ramach jednej infrastruktury danych.

Rezultatem współpracy w ramach powyższych projektów są wszystkie publikacje zgłoszone jako osiągnięcie naukowe w niniejszym postępowaniu habilitacyjnym. Powstały one w ramach prac badawczych realizowanych w UAM i stanowią efekt mojego wkładu analitycznego, metodologicznego oraz interpretacyjnego. Działalność ta była całkowicie niezależna od mojego zatrudnienia na Uniwersytecie Ekonomicznym w Poznaniu.

5.2 Projekt regionalny na zlecenie administracji publicznej (2018 – 2019)

W latach 2018–2019 uczestniczyłem jako ekspert w projekcie realizowanym na zlecenie Ministerstwa Inwestycji i Rozwoju oraz urzędów marszałkowskich trzech województw: opolskiego, świętokrzyskiego i warmińsko-mazurskiego. Projekt nosił tytuł: *Analiza przewidy-*

wanych konsekwencji sytuacji demograficznej w relacji do systemu społeczno-gospodarczego i układu przestrzennego badanych obszarów. Określenie scenariuszy rozwoju społeczno-gospodarczego i przestrzennego do 2030 roku i dotyczył skutków depopulacji oraz starzenia się ludności dla polityk publicznych na poziomie regionalnym i lokalnym.

Działiałem w interdyscyplinarnym zespole eksperckim, którego członkowie pochodzili z różnych instytucji naukowych (poza UEP). Moim zadaniem było przygotowanie części ilościowej i prognostycznej raportów, obejmującej analizy danych demograficznych i społeczno-gospodarczych, a także scenariusze rozwojowe dla poziomu powiatowego i gminnego. W szczególności odpowiadałem za:

- konstrukcję wskaźników ryzyka demograficznego (m.in. indeksy starości, współczynniki obciążenia demograficznego, trendy migracyjne),
- analizy porównawcze między jednostkami samorządu terytorialnego,
- opracowanie alternatywnych scenariuszy przyszłych zmian ludnościowych i ich wpływu na dostęp do usług publicznych oraz rozwój infrastruktury,
- interpretację wyników w kontekście planowania przestrzennego i regionalnej polityki społeczno-gospodarczej.

Efektom pracy były trzy raporty końcowe opublikowane przez urzędy marszałkowskie poszczególnych województw:

- *Województwo opolskie* (2019): Raport.
- *Województwo świętokrzyskie* (2019): Raport.
- *Województwo warmińsko-mazurskie* (2019): Raport.

Raporty te posłużyły jako dokumenty eksperckie wykorzystywane w opracowaniu strategii demograficznych i programów rozwoju regionalnego. Moja rola koncentrowała się na przełożeniu zaawansowanych analiz statystycznych i prognoz demograficznych na rekomendacje praktyczne dla polityki regionalnej. Działalność ta miała charakter naukowo-analityczny i została zrealizowana całkowicie poza jednostką organizacyjną mojego zatrudnienia. Jednocześnie stanowiła istotne uzupełnienie mojego programu badawczego, pozwalając na zastosowanie metod statystycznych rozwijanych w naukach ekonomicznych i społecznych w praktyce polityki publicznej.

5.3 Ekspertyza dla AQUANET S.A. w ramach Spółki Celowej Uniwersytetu Ekonomicznego w Poznaniu (2015)

W 2015 roku uczestniczyłem jako ekspert w opracowaniu wyników badania ankietowego zrealizowanego przez AQUANET S.A. — jednego z największych operatorów sieci wodno-kanalizacyjnych w Polsce, obsługującego Poznań oraz kilkanaście sąsiednich gmin. Projekt

został zlecony w ramach działalności Spółki Celowej Uniwersytetu Ekonomicznego w Poznaniu i miał na celu diagnozę opinii klientów, ocenę jakości usług oraz identyfikację obszarów wymagających usprawnień.

W ramach opracowania odpowiadałem za:

- przygotowanie planu analitycznego i założeń metodologicznych analizy wyników, w tym dobór technik statystycznych adekwatnych do struktury danych i celów badania,
- analizę struktury próby z uwzględnieniem zróżnicowania terytorialnego i demograficznego respondentów (podział na miasto Poznań i gminy aglomeracyjne),
- przeprowadzenie analiz przekrojowych i testów niezależności dla danych kategori-
cznych (m.in. testów chi-kwadrat i analizy częstości),
- opracowanie interpretacji wyników w układzie tematycznym (ocena jakości usług, preferencje dotyczące kontaktu z firmą, świadomość działań inwestycyjnych i proekologicznych, wizerunek spółki),
- współtworzenie dokumentu końcowego zawierającego rekomendacje dla zarządu spółki w zakresie komunikacji, rozwoju usług oraz polityki informacyjnej.

Analiza została przeprowadzona we współpracy z dr. Marcinem Szymkowiakiem, jednak większość obliczeń statystycznych, a także interpretacja wyników i struktura raportu, były moją samodzielną odpowiedzialnością. Ekspertyza miała charakter praktyczny i została wykorzystana przez zarząd AQUANET S.A. w pracach nad nową strategią komunikacyjną i modernizacyjną.

Działalność ta stanowi przykład zaangażowania w zastosowanie wiedzy statystycznej i metodologicznej w kontekście biznesowym i społecznym, poza jednostką organizacyjną zatrudnienia. Pokazuje również możliwość transferu narzędzi analizy danych, rozwijanych przeze mnie w badaniach akademickich, do zastosowań praktycznych w sektorze gospodarki komunalnej. Dzięki temu doświadczeniu mogłem testować skuteczność metod ilościowych w diagnozowaniu potrzeb klientów i przekładaniu wyników na rekomendacje dla polityki usługowej przedsiębiorstwa.

5.4 Analiza publikacji w MDPI w jednostkach Polskiej Akademii Nauk (marzec 2025)

W marcu 2025 roku, na zlecenie pracowników Polskiej Akademii Nauk, przygotowałem niezależną analizę dotyczącą skali i charakterystyki publikacji afiliowanych w instytutach PAN, które ukazały się w czasopismach wydawnictwa MDPI (Multidisciplinary Digital Publishing Institute). Celem opracowania było dostarczenie podstaw do oceny kierunków

i selektywności polityki wydawniczej instytutów PAN, a także zidentyfikowanie specyfiki współpracy międzynarodowej oraz profilu publikacyjnego w tej grupie czasopism.

W ramach opracowania przeanalizowałem:

- strukturę publikacji MDPI według instytutów PAN i obszarów nauki (zgodnie z klasyfikacją OECD i dyscyplinami obowiązującymi w ewaluacji),
- dynamikę liczby publikacji w latach 2016–2022, w tym przyrosty roczne i zmiany w udziale MDPI w całkowitym dorobku publikacyjnym instytutów,
- wskaźniki bibliometryczne (Impact Factor, CiteScore, SJR, SNIP) oraz ich relację do cytowalności artykułów afiliowanych przez PAN,
- udział współautorstwa międzynarodowego oraz stopień dopasowania tematyki publikacji do profilu badawczego danego instytutu,
- indeksy intensywności publikacyjnej (liczba artykułów MDPI na pracownika naukowego) i ich zróżnicowanie między instytutami PAN.

Opracowanie miało charakter eksploracyjno-analityczny i zostało wykonane z pełnym zachowaniem autonomii badawczej. Dane pochodziły z bazy Scopus oraz ogólnodostępnych źródeł informacyjnych PAN. Raport końcowy został wykorzystany w kontekście wewnętrznych dyskusji o strategii publikacyjnej i ocenie jakości publikacji w czasopiśmie otwartego dostępu, ze szczególnym uwzględnieniem zjawiska masowego publikowania w MDPI.

Działalność ta miała charakter ekspercki, realizowany całkowicie poza jednostką organizacyjną zatrudnienia (UEP), i była oparta na samodzielnie zaprojektowanej procedurze badawczej. Jej wyniki stanowią przykład zastosowania metod analizy bibliometrycznej do oceny jakości i selektywności dorobku naukowego w instytucjach badawczych. Wpisuje się to w mój szerszy program badawczy dotyczący diagnozy systemowych wzorców produktywności i nierówności w nauce, a zarazem dostarcza praktycznych rekomendacji dla polityki instytucjonalnej w zakresie publikowania.

5.5 Działalność ekspercka na zlecenie Stowarzyszenia Metropolia Poznań (2016)

W 2016 roku współpracowałem ze Stowarzyszeniem Metropolia Poznań w ramach umowy cywilnoprawnej, której celem było opracowanie spójnych, porównywalnych i możliwych do aktualizacji raportów tabelarycznych dotyczących sytuacji społeczno-demograficznej oraz przestrzennej miasta Poznania i gmin aglomeracji poznańskiej.

Zamówienie to miało charakter ekspercki i zostało zrealizowane całkowicie poza moją jednostką organizacyjną zatrudnienia (UEP). Obejmowało wsparcie analityczne dla

polityki miejskiej i metropolitalnej w zakresie wykorzystania danych statystycznych pochodzących z różnych źródeł (m.in. GUS, BDL, dane lokalne). Zakres moich działań obejmował:

- harmonizację pojęć i definicji wykorzystywanych w statystyce publicznej, tak aby zapewnić porównywalność wskaźników między gminami aglomeracji oraz w ujęciu czasowym,
- opracowanie metodyki budowy zestandaryzowanych tabel statystycznych, które umożliwiały syntetyczne i przystępne prezentowanie charakterystyk społecznych, demograficznych i przestrzennych poszczególnych jednostek terytorialnych,
- analizę spójności danych i wsparcie merytoryczne w zakresie interpretacji obserwowanych zjawisk oraz konstrukcji wskaźników do celów raportowania,
- opracowanie raportu końcowego służącego jako baza wiedzy dla dalszych prac planistycznych i strategicznych Stowarzyszenia Metropolia Poznań.

Współpraca ta miała na celu zapewnienie władzom lokalnym i metropolitalnym narzędzi do prowadzenia polityki opartej na danych, zgodnie z ideą *evidence-based policy*. Zastosowane przeze mnie podejścia statystyczne i metody harmonizacji danych zostały wykorzystane w późniejszych opracowaniach metropolitalnych i miały znaczenie w debacie publicznej dotyczącej rozwoju aglomeracji poznańskiej.

Działalność ta stanowi przykład praktycznego zastosowania mojej wiedzy metodologicznej poza uczelnią i wpisuje się w szerszy program badawczy ukierunkowany na tworzenie i adaptację narzędzi analitycznych umożliwiających diagnozę zjawisk społecznych, demograficznych i instytucjonalnych na poziomie regionalnym i lokalnym.

5.6 Ekspertyza dotycząca bazy danych RAD-on na zlecenie Ośrodka Przetwarzania Informacji – Państwowego Instytutu Badawczego (2024)

W 2024 roku, na zlecenie Ośrodka Przetwarzania Informacji – Państwowego Instytutu Badawczego (OPI PIB), przeprowadziłem ekspercką analizę zasobów bazy danych RAD-on – największego publicznego repozytorium danych o nauce i szkolnictwie wyższym w Polsce, prowadzonego na mocy ustawy *Prawo o szkolnictwie wyższym i nauce*.

Celem ekspertyzy było określenie potencjału analitycznego bazy oraz wskazanie barier utrudniających jej wykorzystanie do analiz naukometycznych, ewaluacyjnych i instytucjonalnych. Moja praca miała charakter całościowy i obejmowała wszystkie główne komponenty RAD-on, w tym:

- ocenę struktury logicznej i organizacyjnej danych (relacje między tabelami, identyfikatory, zakresy czasowe, poziomy agregacji),
- analizę jakości danych pod względem kompletności, spójności i aktualności, z uwzględnieniem problemów wynikających z niejednolitego raportowania przez jednostki naukowe,
- identyfikację ograniczeń analitycznych, takich jak brak zmiennych kluczowych, trudności w harmonizacji danych w przekroju lat czy niejednoznaczność przypisania podmiotów i osób do dyscyplin,
- ocenę potencjału wykorzystania danych RAD-on w obszarach takich jak monitorowanie kariery naukowej, produktywność badawcza, mobilność akademicka, uczestnictwo w ewaluacji i finansowaniu działalności badawczej,
- opracowanie szczegółowych rekomendacji dotyczących dalszego rozwoju systemu, w tym: rozszerzenia dokumentacji, stworzenia znormalizowanych widoków dla użytkowników akademickich, poprawy interoperacyjności i możliwości eksportu danych.

Wnioski z analizy zawarte zostały w szczegółowym raporcie końcowym, który przekazano kierownictwu OPI PIB. Zostały one wykorzystane do oceny stanu obecnego systemu i jako wkład do planowanych działań rozwojowych.

Ekspertyza miała charakter niezależny i została wykonana poza jednostką organizacyjną zatrudnienia (UEP). Stanowi przykład mojego zaangażowania w rozwój ogólnopolskiej infrastruktury informacyjnej o nauce oraz doradztwa instytucjonalnego na poziomie krajowym. Jest także naturalnym rozszerzeniem mojego programu badawczego opartego na integracji i analizie mikrodanych, łącząc wymiar metodologiczny z aplikacyjnym.

5.7 Wiceprzewodniczący Kapituły Programu Ariadna (od 2022)

Od 2022 roku pełnię funkcję **Wiceprzewodniczącego Kapituły Programu *Ariadna***, prowadzonego przez firmę Predictive Solutions – polskiego dystrybutora oprogramowania IBM SPSS³. Program *Ariadna* to ogólnopolska inicjatywa wspierająca środowisko akademickie poprzez udostępnianie licencji edukacyjnych, popularyzację narzędzi analizy danych oraz rozwój kompetencji dydaktycznych i badawczych w zakresie wykorzystania oprogramowania SPSS.

Jako wiceprzewodniczący Kapituły uczestniczę w kształtowaniu polityki merytorycznej programu, opiniowaniu działań jego organizatorów oraz wyznaczaniu kierunków wsparcia dla instytucji akademickich. Moja działalność obejmuje także promowanie wysokich standardów dydaktycznych w nauczaniu statystyki i analizy danych oraz łączenie doświadczenia akademickiego z praktycznym zastosowaniem narzędzi IBM SPSS. Ważnym

³<https://ariadna.predictivesolutions.pl/>

elementem programu jest certyfikacja studentów ze znajomości metod ilościowych i środowiska IBM SPSS, prowadzona zarówno na uczelniach, jak i w siedzibie Predictive Solutions; uzyskanie tego certyfikatu stanowi dla wielu pracodawców wiarygodne potwierdzenie wysokich kompetencji analitycznych absolwentów. Dodatkowo, w ramach *Ariadny* organizowana jest cyklicznie konferencja naukowa *Analityczne Wyzwania*⁴, gromadząca czołowych statystyków, socjologów ilościowych i analityków z różnych dziedzin, stanowiąca ważne forum wymiany doświadczeń między nauką a praktyką. Funkcja ta pozwala mi łączyć perspektywę badacza i dydaktyka z działaniami na rzecz całego środowiska akademickiego.

5.8 Współpraca interdyscyplinarna w zakresie analizy danych biologicznych (2024–2025)

W latach 2024–2025 uczestniczyłem w interdyscyplinarnym projekcie badawczym z zakresu biologii molekularnej roślin, którego rezultatem była publikacja:

Gzyl, J., Żywicki, M., Deckert, J., Lichocki, M., Skafe, Y., Bąk, J., **Roszka, W.**, Shen, Y., Sun, X., Chmielowska-Bąk, J. (2025). *Oxytranscriptome of soybean seedlings under short-term cadmium treatment*. *Scientific Reports*, 15, 25308. 10.1038/s41598-025-09324-0

Artykuł dotyczył analizy zmian w ekspresji genów (RNA-seq) u siewek soi poddanych krótkotrwałemu stresowi kadmowemu. Moja rola polegała na opracowaniu i wdrożeniu procedur analizy danych eksperymentalnych, w szczególności:

- przetwarzaniu i statystycznej analizie danych transkryptomicznych RNA-seq (pipeline bioinformatyczny),
- doborze i zastosowaniu metod detekcji różnic w ekspresji genów (DESeq2),
- interpretacji wyników na poziomie klas funkcjonalnych genów (analiza GO),
- współredagowaniu części metodologicznej i dyskusji statystycznej.

Działalność ta miała charakter pozauniwersytecki i interdyscyplinarny. Współpraca potwierdziła możliwość skutecznego transferu zaawansowanych metod statystycznych do analizy danych biologicznych oraz pokazała uniwersalność narzędzi ilościowych rozwijanych w moim programie badawczym. Wpisuje się ona w szerszą perspektywę integracji i analizy danych wielowymiarowych, stanowiąc przykład praktycznego zastosowania kompetencji statystycznych poza obszarem ekonomii i nauk społecznych.

⁴<https://analitycznewyzwania.predictivesolutions.pl/>

5.9 Staż naukowy na University of Southampton (2016)

W okresie od 1 października do 31 grudnia 2016 roku odbyłem trzymiesięczny staż naukowy na University of Southampton (Wielka Brytania), w jednej z czołowych jednostek akademickich w Europie w dziedzinie statystyki stosowanej. Staż został zrealizowany w School of Social Sciences (Social Statistics and Demography), w zespole specjalizującym się w analizie danych społecznych i demograficznych, pod opieką prof. Li-Chun Zhan-ga, uznanego międzynarodowo eksperta w zakresie statystyki małych obszarów i metod integracji danych.

Jego głównym celem było rozwinięcie kompetencji z zakresu:

- estymacji dla małych obszarów (small area estimation),
- zastosowania technik dopasowywania rekordów (record linkage) i integracji danych,
- zaawansowanej analizy danych demograficznych i regionalnych.

Podczas stażu uczestniczyłem w seminariach naukowych, konsultacjach z zespołem badawczym oraz pracach nad zastosowaniem modeli bayesowskich i mieszanych do szacowania zmiennych społecznych na poziomie lokalnym. Nabyte doświadczenie miało charakter praktyczny i bezpośrednio przełożyło się na późniejsze projekty badawcze – zwłaszcza w zakresie metod integracji danych administracyjnych i bibliometrycznych oraz analiz regionalnych.

Staż ten stanowi przykład międzynarodowej aktywności naukowej wykraczającej poza ramy jednostki zatrudnienia i był ważnym etapem przygotowania do opracowania zintegrowanego operatu mikrodanych (POL-on-Scopus), który stał się fundamentem zgłaszanego osiągnięcia habilitacyjnego.

5.10 Kierowanie komponentem międzynarodowym – Visegrad Fund

W okresie 21.09.2021 – 18.02.2023 byłem koordynatorem (Principal Investigator po stronie polskiej) w projekcie *International Visegrad Fund* (ID 2212009) pt. *Measuring university excellence in the V4 region; eliminating invisibility and get recognition*⁵, realizowanym w konsorcjum ośmiu uczelni z państw Grupy Wyszehradzkiej (m.in. Óbuda University – koordynator, University of Pécs, Tomas Bata University in Zlín, Slovak University of Technology in Bratislava, Politechnika Wrocławska, Politechnika Warszawska).

Projekt badał specyfikę uczelni z Europy Środkowej, które najczęściej plasują się w rankingach międzynarodowych poniżej pierwszego tysiąca, a ich pozycja pozostaje trudna do jednoznacznej identyfikacji. W ramach trzech pakietów roboczych (*literature review*,

⁵<https://uni-obuda.hu/v4-uni-visibility/>

provider survey, stakeholder survey) analizowano uwarunkowania prawne i instytucjonalne, kulturę organizacyjną uczelni oraz opinie absolwentów na temat jakości kształcenia i postrzegania „doskonałości akademickiej”.

Byłem szczególnie zaangażowany w *Provider Survey WP* oraz *Stakeholder Survey WP*, odpowiadając za adaptację kwestionariuszy, integrację danych z Polski oraz analizę porównawczą. Jako PI po stronie polskiej koordynowałem prace w krajowym zespole, odpowiadając zarówno za stronę merytoryczną, jak i organizacyjną. Uczestniczyłem również w międzynarodowej konferencji *Central European University Visibility, Excellence and Ranking Forum*, która odbyła się w listopadzie 2022 r. w Budapeszcie. Rezultatem projektu był kompleksowy raport porównawczy dla uczelni V4, analizy rankingowe i bibliometryczne identyfikujące mocne i słabe strony uczelni, a także rekomendacje dla instytucji i decydentów dotyczące zwiększania widoczności uczelni w skali międzynarodowej. Wyniki udostępniono również w językach narodowych partnerów, co zwiększyło ich przydatność dla lokalnych środowisk akademickich.

Dzięki uczestnictwu w powyższych projektach i inicjatywach prowadziłem działalność naukową o charakterze interdyscyplinarnym, w innej uczelni akademickiej niż moja jednostka zatrudnienia. Publikacje powstałe w ich ramach stanowią zgłoszone osiągnięcie naukowe w niniejszym postępowaniu habilitacyjnym. Moja rola w każdym z projektów miała charakter samodzielny i merytorycznie istotny, koncentrowała się na stronie analitycznej, metodologicznej oraz interpretacyjnej.

6 Odniesienie do kryteriów habilitacyjnych i dalsze plany badawcze

Zgłoszone osiągnięcie obejmuje osiem publikacji opublikowanych w uznanych czasopiśmie międzynarodowych (łącznie 1040 pkt. MEiN, wszystkie w JCR), opartych na pełnopopulacyjnych danych zintegrowanych metodą probabilistycznego dopasowania rekordów. Prace dotyczą struktury, dynamiki i nierówności w polskiej nauce, stanowiąc wkład w rozwój badań empirycznych z zakresu *ekonomicznej analizy nauki* (ang. *economics of science*) oraz *science of science*.

W cyklu zastosowano oryginalne rozwiązania metodologiczne: probabilistyczne łączenie danych administracyjnych i bibliometrycznych, analizy kohortowe z uwzględnieniem wieku akademickiego, klasyfikację produktywności opartą na regule 10/50, modele regresji frakcyjnej i z efektami stałymi oraz analizy sieciowe współpracy. Efektem jest zbudowanie trwałego, replikowalnego operatu mikrodanych, który umożliwia analizy mikropoziome, kohortowe i longitudinalne, a także konstrukcję nowych wskaźników, m.in. *gender solo research gap* i *gender homophily index*. Wyniki badań stanowią zarówno **wkład metodologiczny i analityczny**, jak i **podstawę aplikacyjną** dla projektowania polityk naukowych zgodnych z zasadami odpowiedzialnej ewaluacji (DORA, *Leiden Manifesto*).

W realizacji badań odpowiadałem za koncepcję, integrację danych, dobór i implementację metod analitycznych, interpretację wyników oraz przygotowanie głównej części tekstów. Współautor skupiał się na analizie literatury, osadzeniu wyników i redakcji końcowej. Zakres zadań potwierdza moją samodzielność badawczą, rozwijaną również w ramach współpracy międzynarodowej (komponent polski projektu *Visegrad Fund*, ID 2212009, 2021–2023).

Działalność naukowa poza macierzystą uczelnią stanowi istotny element mojego dorobku. W latach 2016–2025 prowadziłem liczne projekty badawcze i analityczne w ramach współpracy z Uniwersytetem im. Adama Mickiewicza w Poznaniu, Ośrodkiem Przetwarzania Informacji – Państwowym Instytutem Badawczym (OPI PIB) oraz instytucjami publicznymi i gospodarczymi. Obejmowały one m.in. udział w projektach MAESTRO, DIALOG i *Nauka dla Społeczeństwa*, przygotowanie ekspertyz dotyczących infrastruktury danych naukowych, opracowywanie wskaźników i modeli statystycznych dla potrzeb polityki publicznej oraz popularyzację idei odpowiedzialnej ewaluacji. Działalność ta miała charakter interdyscyplinarny i aplikacyjny, łącząc analizy empiryczne z projektowaniem narzędzi wspierających zarządzanie nauką.

Wyniki cyklu stanowią punkt wyjścia do dalszego programu badawczego poświęconego analizie struktury, dynamiki i nierówności w systemie nauki, w tym rozwoju i aktualizacji operatu danych, badań porównawczych w Europie Środkowo-Wschodniej, ewaluacji reform systemowych oraz opracowania nowych wskaźników produktywności, współpracy i

mobilności.

W planowanych badaniach zamierzam rozwijać wątki związane z ekonomią heterodoksyjną i analizą nierówności w nauce, w szczególności badanie zróżnicowanych ścieżek karier akademickich kobiet i mężczyzn (*gendered career paths*), różnic w dostępie do finansowania oraz w sieciach współpracy międzynarodowej. Celem jest pogłębienie zrozumienia mechanizmów dystrybucji szans i barier w rozwoju naukowym w ujęciu ilościowym, instytucjonalnym i porównawczym.

Kierunek ten łączy refleksję nad efektywnością z problematyką równości szans i różnorodności akademickiej, wpisując się w nurt *heterodox economics* i empirycznych badań nad polityką naukową.

W perspektywie kilkuletniej planuję rozwijać trzy komplementarne linie badawcze:

- **Rozszerzenie operatu danych** o źródła dotąd niedostępne (ORCID, OpenAlex, CRIS, krajowe bazy grantowe NCN/NCBR), co pozwoli na pełniejsze analizy przepływów finansowych, współpracy międzynarodowej i mobilności naukowców. Docelowo operat ma stać się otwartą infrastrukturą analityczną dla badań nad nauką w Polsce i w regionie.
- **Badania porównawcze i longitudinalne** w ujęciu międzynarodowym – zwłaszcza Polska–Czechy–Węgry–Niemcy – dotyczące dynamiki produktywności, mobilności i nierówności w kontekście różnych modeli finansowania nauki. Ujednolicone definicje i metody analityczne umożliwią identyfikację regionalnych wzorców rozwoju systemów akademickich.
- **Modelowanie polityk naukowych** w duchu podejścia *evidence-informed policy making* – testowanie zależności między parametrami finansowania, strukturą współpracy i wynikami naukowymi, z wykorzystaniem modeli panelowych i symulacji scenariuszowych. Celem jest rozwój narzędzi wspierających projektowanie bardziej efektywnych, zrównoważonych i inkluzywnych polityk naukowych.

Równolegle planuję upowszechniać wyniki badań poprzez publikacje w czołowych czasopismach z zakresu naukometrii, *economics of science* i polityki szkolnictwa wyższego, a także poprzez działania dydaktyczne i szkoleniowe. Zamierzam kontynuować współpracę z krajowymi i zagranicznymi ośrodkami badawczymi (m.in. UAM, OPI PIB, University of Southampton), rozwijać komponenty międzynarodowe operatu danych oraz tworzyć warunki do transferu wyników do praktyki instytucjonalnej.

Uważam, że zgłoszone osiągnięcie i planowana kontynuacja badań wnoszą oryginalny, interdyscyplinarny i trwały wkład do rozwoju wiedzy w dyscyplinie *ekonomia i finanse*, a opracowana metodologia i wnioski mają znaczenie zarówno poznawcze, jak i praktyczne – dla krajowej i międzynarodowej analizy systemów akademickich.

7 Pozostała działalność dydaktyczna akademicka i organizacyjna

7.1 Działalność dydaktyczna

Od ponad 15 lat prowadzę zajęcia dydaktyczne na Uniwersytecie Ekonomicznym w Poznaniu. Od 5 lat uczę także w Collegium Da Vinci. Realizuję kursy z zakresu statystyki, metod ilościowych oraz analizy danych — w języku polskim i angielskim, na poziomie studiów licencjackich i magisterskich.

W okresie po uzyskaniu stopnia doktora (od 2013 roku) zrealizowałem łącznie ponad **3300 godzin dydaktycznych**. Do najważniejszych należą:

- Statystyka opisowa (871 godz.), Statystyka (564 godz.) i Statystyka matematyczna (228 godz.),
- Kursy komputerowe: Analiza danych z R, Wprowadzenie do pakietu statystycznego R oraz R Programming (łącznie 238 godz.),
- Zajęcia w języku angielskim: R Programming, Modern Statistical Methods (180 godz.), Statistics (185 godz.), Statistics for Business and Economics (30 godz.), Real Estate Market Analysis (70 godz.),
- Zajęcia praktyczne: Analiza statystyczna w badaniach rynku (486 godz.), Ilościowe metody wyceny przedsiębiorstw (120 godz.), Statystyczna integracja danych (156 godz.), Statystyczna analiza ankiet (18 godz.),
- Inne: Klasa Akademicka, Metoda reprezentacyjna, SeminaRIA dyplomowe (131 godz.).

W Collegium Da Vinci prowadzę zajęcia z takich przedmiotów jak:

- Zaawansowana eksploracja danych,
- Statystyka w zastosowaniach,
- Podstawy języka R w statystyce opisowej,
- Statystyka opisowa,
- Wstęp do Big Data w ekonomii,
- Statystyka matematyczna,
- Storytelling danych,
- SeminaRIA licencjackie.

Byłem promotorem **34 prac licencjackich**, zarówno na studiach dziennych, jak i niestacjonarnych, oraz recenzentem szeregu prac dyplomowych. Uzyskałem bardzo wysoką ocenę dydaktyczną w ostatniej ocenie okresowej na Uniwersytecie Ekonomicznym w Poznaniu.

7.2 Nagrody i wyróżnienia

Za działalność naukową i publikacyjną pięciokrotnie otrzymałem **Nagrodę Rektora Uniwersytetu Ekonomicznego w Poznaniu** za oryginalne i twórcze osiągnięcia naukowe. Wysoka jakość prowadzonych badań oraz publikacje w renomowanych czasopismach przyczyniły się do uzyskania przez jednostkę bardzo dobrej oceny parametrycznej.

Dzięki prowadzonym analizom statystycznym i pracom nad integracją danych przyczyniłem się również do rozwoju metodologii stosowanej w nauce o nauce. Zostało to dostrzeżone zarówno w projektach badawczych, jak i publikacjach zgłoszonych jako osiągnięcie habilitacyjne.

7.3 Zaangażowanie organizacyjne w środowisku statystycznym

Moja działalność akademicka obejmuje również aktywność organizacyjną oraz członkostwo w gremiach naukowych i towarzystwach branżowych:

- **Członek Komisji Rewizyjnej Oddziału Polskiego Towarzystwa Statystycznego w Poznaniu (od 2020)** — uczestniczę w bieżących działaniach Oddziału, w tym w zakresie nadzoru wewnętrznego nad działalnością statutową oraz organizacyjną.
- **Członek Sekcji Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego (od 2010)** — biorę udział w pracach środowiska skupionego wokół zastosowań statystyki w klasyfikacji, eksploracji i analizie danych, w tym w corocznych konferencjach naukowych SKAD.
- **Członek komitetu organizacyjnego XXX Konferencji Naukowej Sekcji Klasyfikacji i Analizy Danych PTS (2021)** — brałem udział w przygotowaniu programu konferencji, ocenie zgłoszonych wystąpień oraz organizacji technicznej wydarzenia, które odbyło się w formule online.
- **Przewodniczący jury oraz koordynator kategorii „Statystyk/Badacz” w I i II edycji Ogólnopolskiego Konkursu „Ignacy Jan Paderewski” (2021–2022)** — odpowiadałem za przygotowanie i realizację zadań konkursowych w tej kategorii oraz przewodniczyłem pracom jury dla uczniów szkół podstawowych i ponadpodstawowych. Kategoria miała na celu rozwijanie umiejętności badawczych i statystycz-

nych młodzieży, kształtowanie kompetencji analitycznych oraz popularyzację metod ilościowych w kontekście postaci i działalności Ignacego Jana Paderewskiego.

- **Wiceprzewodniczący Kapituły Programu Ariadna (od 2022)** — współodpowiadam za kierunki rozwoju programu edukacyjnego *Ariadna* prowadzonego przez firmę Predictive Solutions (polskiego dystrybutora IBM SPSS). Program wspiera uczelnie wyższe w zakresie dydaktyki statystycznej i analizy danych, udostępniając oprogramowanie IBM SPSS oraz materiały dydaktyczne. W ramach Kapituły uczestniczę w ocenie inicjatyw akademickich, rekomendacjach dotyczących rozwoju programu oraz popularyzacji zastosowań statystyki w edukacji i badaniach.

7.4 Wyjazdy szkoleniowe, konferencyjne i służbowe

W ramach działalności naukowej i dydaktycznej uczestniczyłem w licznych wyjazdach szkoleniowych, konferencyjnych i służbowych, które umożliwiły mi systematyczny rozwój kompetencji metodologicznych oraz prezentację wyników badań. Brałem udział zarówno w szkoleniach specjalistycznych organizowanych przez międzynarodowe firmy technologiczne (SAS, IBM), jak i w konferencjach naukowych o zasięgu ogólnopolskim i branżowym, takich jak *Kongres Statystyki Polskiej*, konferencje Sekcji Klasyfikacji i Analizy Danych PTS czy spotkania *Analityczne Wyzwania*.

Udział w tych wydarzeniach sprzyjał podnoszeniu kwalifikacji warsztatowych, nawiązywaniu kontaktów naukowych oraz aktywnej wymianie doświadczeń z przedstawicielami środowiska statystycznego, ekonomicznego i społecznego.

Równolegle prowadziłem także działania popularyzatorskie – w ramach programu *Klasa akademicka* realizowałem zajęcia z zakresu statystyki i analizy danych w liceach Wielkopolski, przybliżając uczniom praktyczne zastosowania metod ilościowych.

Aktywność ta stanowiła integralną część mojego rozwoju zawodowego, łącząc aspekty badawcze, dydaktyczne i środowiskowe oraz przyczyniając się do upowszechniania wiedzy o statystyce i analizie danych.

8 Wykaz publikacji po uzyskaniu stopnia doktora (2013–2025)

Poniższy wykaz obejmuje wszystkie publikacje po uzyskaniu stopnia doktora (2013–2025), w tym także osiem artykułów wskazanych jako cykl stanowiący podstawę osiągnięcia habilitacyjnego.

1. Gzyl, J., Żywicki, M., Deckert, J., Lichocki, M., Skafe, Y., Bąk, J., **Roszka, W.**, Shen, Y., Sun, X., & Chmielowska-Bąk, J. (2025). Oxytranscriptome of soybean seedlings under short-term cadmium treatment. *Scientific Reports*, 15, 25308. <https://doi.org/10.1038/s41598-025-09324-0>
2. Kwiek, M., & **Roszka, W.** (2025). Are scientists changing their research productivity classes when they move up the academic ladder? *Innovative Higher Education*, 50(1), 329–367. <https://doi.org/10.1007/s10755-024-09735-3>
3. Kwiek, M., & **Roszka, W.** (2024). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*, 87(3), 519–549. <https://doi.org/10.1007/s10734-023-01022-y>
4. Kwiek, M., & **Roszka, W.** (2024). The young and the old, the fast and the slow: A large-scale study of productivity classes and rank advancement. *Studies in Higher Education*, 49(11), 2036–2051. <https://doi.org/10.1080/03075079.2023.2288172>
5. Kwiek, M., & **Roszka, W.** (2024). Top research performance in Poland over three decades: A multidimensional micro-data approach. *Journal of Informetrics*, 18(4), art. no. 101595. <https://doi.org/10.1016/j.joi.2024.101595>
6. Kwiek, M., & **Roszka, W.** (2024). Reguła 10/50. Analiza dorobku 150 tysięcy polskich naukowców pokazuje trwałość systemowych nierówności w nauce. *Nauka*, 4, 47–70. <https://doi.org/10.24425/nauka.2024.152910>
7. Kwiek, M., & **Roszka, W.** (2023). Zawsze wysoce produktywni? Zawsze nisko produktywni? Dynamiczne ujęcie dorobku naukowego profesorów tytularnych wykorzystujące klasy produktywności. *Nauka*, 1, 7–30. <https://doi.org/10.24425/nauka.2023.144944>
8. Kwiek, M., & **Roszka, W.** (2022). Are female scientists less inclined to publish alone? The gender solo research gap. *Scientometrics*, 127(3), 1697–1735. <https://doi.org/10.1007/s11192-022-04308-7>

9. Kwiek, M., & **Roszka, W.** (2022). Academic vs. biological age in research on academic careers: A large-scale study with implications for scientifically developing systems. *Scientometrics*, 127(6), 3543–3575. <https://doi.org/10.1007/s11192-022-04363-0>
10. Kwiek, M., & **Roszka, W.** (2022). Globalny obieg naukowy a wiek w nauce: analiza 20 000 polskich naukowców. *Nauka*, 3, 89–110. <https://doi.org/10.24425/nauka.2022.140330>
11. Kwiek, M., & **Roszka, W.** (2021). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, 15(3), art. no. 101171. <https://doi.org/10.1016/j.joi.2021.101171>
12. Kwiek, M., & **Roszka, W.** (2021). Gender disparities in international research collaboration: A study of 25,000 university professors. *Journal of Economic Surveys*, 35(5), 1344–1380. <https://doi.org/10.1111/joes.12395>
13. Kwiek, M., & **Roszka, W.** (2021). Dlaczego w nauce dominuje współpraca z mężczyznami: homofilia ze względu na płeć na przykładzie 25 000 naukowców. *Nauka*, 2, 51–74. <https://doi.org/10.24425/nauka.2021.136304>
14. **Roszka, W.** (2019). Spatial microsimulation of personal income in Poland. *Statistics in Transition*, 20(3), 1–20. <https://doi.org/10.21307/stattrans-2019-028>
15. Szymkowiak, M., & **Roszka, W.** (2016). Potencjał gospodarczy gmin aglomeracji poznańskiej w ujęciu taksonomicznym. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 427, 271–282. <https://doi.org/10.15611/pn.2016.427.23>
16. **Roszka, W.** (2015). Konstrukcja syntetycznych zbiorów danych na potrzeby estymacji dla małych domen. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 384, 303–313. <https://doi.org/10.15611/pn.2015.384.27>
17. **Roszka, W.** (2015). Some practical issues related to the integration of data from sample surveys. *Statistika*, 95(1), 60–73. <https://www.czso.cz/documents/10180/20550323/32019715q1060.pdf/3c31037b-cecb-47a2-8044-7acf40e413ae?version=1.0>
18. Soniewicki, M., & **Roszka, W.** (2014). Competitiveness of manufacturing enterprises – The role of information and communication technologies supporting knowledge management. In L. Yang, J. Kantola, A. de Hoyos, K. Keminishi, & G. Duysters (Eds.), *Proceedings of the 11th International Conference on Innovation and Management*, 563–573. <http://icim.vamk.fi/2014/uploads/UploadPaperDir/11thICIM2014.pdf>

19. **Roszka, W.**, & Szymkowiak, M. (2014). Podejście kalibracyjne w statystycznej integracji danych. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 328, 280–289. <http://www.dbc.wroc.pl/publication/28006>
20. **Roszka, W.** (2013). Szacowanie łącznych charakterystyk cech nieobserwowanych łącznie. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 279, 237–246. <https://www.dbc.wroc.pl/dlibra/publication/27297/edition/24672>

Literatura

- Abramo, G., D'Angelo, C. A., and Costa, F. D. (2011). Research productivity: Are higher academic ranks more productive than lower ones? *Scientometrics*, 88(3):915–928.
- Allison, P. D., Long, J. S., and Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*, 47(5):615–625.
- Allison, P. D. and Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*, 39(4):596–606.
- Auerbach, P. (2019). *The Future of Academic Capitalism*. Routledge, London.
- Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, 14(6):19–47.
- Bozeman, B. and Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4):599–616.
- Caprile, M., Meulders, D., O'Dorchai, S., and Zimmermann, K. (2012). *Meta-analysis of Gender and Science Research: Synthesis Report*. European Commission, DG Research and Innovation.
- Ceci, S. J. and Williams, W. M. (2011). *Understanding Current Causes of Women's Underrepresentation in Science*. Proceedings of the National Academy of Sciences, 108(8): 3157–3162.
- Chavarro, D., Ràfols, I., and Tang, P. (2018). Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. *Research Policy*, 47(9):1651–1663.
- Christen, P. (2012). *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Springer, Berlin, Heidelberg.
- Cole, J. R. (1973). *Social Stratification in Science*. University of Chicago Press, Chicago.

- Costas, R., van Leeuwen, T. N., and van Raan, A. F. J. (2010). Is scientific performance a function of age? a bibliometric analysis of se and nl researchers. *Scientometrics*, 82(3):517–532.
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. University of Chicago Press, Chicago.
- de Solla Price, D. J. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292–306.
- Diamond, Arthur M., J. (1986). The life-cycle research productivity of mathematicians and scientists. *Journal of Gerontology*, 41(4):520–525.
- Elsevier (2020). The researcher journey through a gender lens. *Elsevier*.
- Enamorado, T., Fifield, B., and Imai, K. (2019). Using a probabilistic model to assist merging of large-scale administrative records. *American Political Science Review*, 113(2):353–371.
- Etzkowitz, H. and Kemelgor, C. (2001). *Gender Inequality in Science: A Universal Condition?* *Journal of Women and Minorities in Science and Engineering*, 7(3): 255–281.
- Eurostat (2011). Essnet on data integration: Final report. Technical report, Eurostat.
- Eurostat (2014). European statistics code of practice and the quality assurance framework (ess qaf). Technical report, Eurostat.
- Fellegi, I. P. and Sunter, A. B. (1969). A theory for record linkage. *Journal of the American Statistical Association*, 64(328):1183–1210.
- Fortin, J.-M. and Currie, D. (2013). Big science vs. little science: How scientific impact scales with funding. *PLOS ONE*, 8(6):e65263.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., and Barabási, A.-L. (2018). Science of science. *Science*, 359(6379):eaao0185.
- Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, 35(1):131–150.
- Fox, M. F. and Colatrella, C. (2001). Participation, performance, and advancement of women in academic science and engineering: What is at issue and why. *Journal of Technology Transfer*, 26(4):377–386.
- Gingras, Y. (2008). Trends in the evolution of metrics used to evaluate science and scientists. *History of Science*, 46(3):335–355.

- Glänzel, W. and Moed, H. F. (2013). Thoughts and facts on bibliometric indicators. *Scientometrics*, 96(1):381–394.
- Herzog, T. N., Scheuren, F. J., and Winkler, W. E. (2007). Data quality and record linkage techniques. *Springer*.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I. (2015). Bibliometrics: The leiden manifesto for research metrics. *Nature*, 520(7548):429–431.
- Holman, L., Stuart-Fox, D., and Hauser, C. E. (2018). The gender gap in science: How long until women are equally represented? *PLOS Biology*, 16(4):e2004956.
- Huang, J., Gates, A. J., Sinatra, R., and Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117(9):4609–4616.
- Hunter, L. A. and Leahey, E. (2010). Parenting and research productivity: New evidence and methods. *Social Studies of Science*, 40(3):433–451.
- Ioannidis, J. P. A., Boyack, K. W., and Baas, J. (2016). Citation metrics: A primer on how (not) to normalize. *PLOS Biology*, 14(9):e1002542.
- Ioannidis, J. P. A., Boyack, K. W., and Baas, J. (2020). Updated science-wide author databases of standardized citation indicators. *PLOS Biology*, 18(10):e3000918.
- Jones, B. F. (2005). The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317.
- Kwiek, M. (2016). The european research elite: A cross-national study of highly productive academics in 11 countries. *Higher Education*, 71(3):379–397.
- Kwiek, M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*, 115(1):415–462.
- Kwiek, M. (2020). *Academic Top Earners: Research Productivity, Prestige and the Role of Gender*. *Studies in Higher Education*, 45(1): 1–18.
- Kwiek, M. and Roszka, W. (2021a). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, 15(3):1–38. Art. no. 101171.
- Kwiek, M. and Roszka, W. (2021b). Gender disparities in international research collaboration: A study of 25,000 university professors. *Journal of Economic Surveys*, 35(5):1344–1380.

- Kwiek, M. and Roszka, W. (2022a). Academic vs. biological age in research on academic careers. *Scientometrics*, 127(6):3543–3575.
- Kwiek, M. and Roszka, W. (2022b). Are female scientists less inclined to publish alone? the gender solo research gap. *Scientometrics*, 127(3):1697–1735.
- Kwiek, M. and Roszka, W. (2024a). Once highly productive, forever highly productive? full professors’ research productivity from a longitudinal perspective. *Higher Education*, 87(3):519–549.
- Kwiek, M. and Roszka, W. (2024b). Top research performance in poland over three decades: A multidimensional micro-data approach. *Journal of Informetrics*, 18(4):1–16. Art. no. 101595.
- Kwiek, M. and Roszka, W. (2024c). The young and the old, the fast and the slow: a large-scale study of productivity classes and rank advancement. *Studies in Higher Education*, 49(11):2036–2051.
- Kwiek, M. and Roszka, W. (2025). Are scientists changing their research productivity classes when they move up the academic ladder? *Innovative Higher Education*, 50(1):329–367.
- Lane, J., Fealing, K. H., III, J. H. M., and Shipp, S. S. (2011a). *The Science of Science Policy: A Handbook*, chapter Introduction, pages 1–16. Stanford University Press, Stanford, CA.
- Lane, J. I., Husbands-Fealing, K., and King, J. (2011b). The science of science policy. In Bainbridge, W. S., editor, *Leadership in Science and Technology: A Reference Handbook*, pages 1–10. SAGE Publications.
- Larivière, V. and Gingras, Y. (2010). The impact factor’s matthew effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, 61(2):424–427.
- Larivière, V. and Sugimoto, C. R. (2015). Age stratification and scientific impact. *Journal of the American Society for Information Science and Technology*, 66(10):1954–1963.
- Larivière, V., Sugimoto, C. R., Macaluso, B., Milojević, S., Cronin, B., and Thelwall, M. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3):417–435.
- Laudel, G. (2006). The art of getting funded: How scientists adapt to their funding conditions. *Science and Public Policy*, 33(7):489–504.

- Leahey, E. (2006). Gender differences in productivity: Research specialization as a missing link. *Gender & Society*, 20(6):754–780.
- Leišytė, L. (2016). New public management and research productivity: A precarious state of affairs of academic work in the netherlands. *Studies in Higher Education*, 41(5):828–846.
- Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71(1):159–178.
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323.
- Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, 50(1):79–104.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.
- Merton, R. K. (1968). The matthew effect in science. *Science*, 159(3810):56–63.
- Moed, H. F. (2005). *Citation Analysis in Research Evaluation*. Springer, Dordrecht.
- Rossiter, M. W. (1993). *The Matthew Matilda Effect in Science*. Social Studies of Science, 23(2): 325–341.
- San Francisco Declaration on Research Assessment (2012). San francisco declaration on research assessment (dora). *American Society for Cell Biology*.
- Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature*, 34(3):1199–1235.
- Stephan, P. E. (2012). How economics shapes science. *Harvard University Press*. Book.
- Stephan, P. E. and Levin, S. G. (1996). Exceptional contributions to u.s. science by the foreign-born and foreign-educated. *Population Research and Policy Review*, 15(1):55–72.
- Sugimoto, C. R., Larivière, V., Ni, C., Gingras, Y., and Cronin, B. (2013). Global gender disparities in science. *Nature*, 504(7479):211–213.
- Sugimoto, C. R., Ni, C., and Larivière, V. (2016). On the relationship between gender disparities in scholarly communication and country-level indicators. *PLOS ONE*, 11(4):e0148259.

- UNESCO (2021). Unesco science report: The race against time for smarter development.
- van den Besselaar, P. and Sandström, U. (2011). Gender differences in research performance and its impact on careers: a longitudinal case study. *Scientometrics*, 87(3):483–499.
- van Leeuwen, T. N. (2006). The application of bibliometric analyses in the evaluation of social science research. *Who Benefits from Social Science Research?*, pages 133–151.
- Wagner, C. S. (2008). *The New Invisible College: Science for Development*. Brookings Institution Press, Washington, DC.
- West, J. M., Jacquet, J., King, M., MacDonnell, S., and Neylon, C. (2013). The role of gender in scholarly authorship. *PLOS ONE*, 8(7):e66212.
- Whitley, R. (2000). The intellectual and social organization of the sciences. *Oxford University Press*. Book.
- Winkler, W. E. (2006). Overview of record linkage and current research directions. *Research Report Series*, 2006-2.
- Xie, Y. and Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Harvard University Press.
- Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. *Statistical Journal of the IAOS*, 28(3-4):219–233.
- Ziman, J. (2000). *Real Science: What it is, and what it means*. Cambridge University Press, Cambridge.
- Zuckerman, H. and Merton, R. K. (1977). Age, aging, and age structure in science. *Sociological Inquiry*, 47(1):85–113.