

MAREK KWIEK*, ŁUKASZ SZYMULA**

Profesja akademicka w ujęciu globalnym: co Big Data mówią nam o udziale kobiet w nauce?

1. Wprowadzenie

Przedmiotem tej pracy jest zmieniająca się demografia globalnej kadry akademickiej z perspektywy wieku, płci, dyscypliny akademickiej i czasu. Nasze podejście jest zakrojone na dużą skalę, generacyjne, a także zarówno przekrojowe, jak i dynamiczne (wzdłużne): analizujemy 4,3 mln naukowców nieokazjonalnych (autorów co najmniej trzech artykułów) w ciągu ostatnich trzech dekad (1990–2021). Interesuje nas zmieniający się rozkład populacji mężczyzn i kobiet naukowców w czasie w różnych grupach wiekowych – zwłaszcza rozkład populacji młodych naukowców z maksymalnie 10-letnim doświadczeniem w publikowaniu – w 16 dyscyplinach STEMM (ściślych, technicznych, inżynierskich, matematycznych i medycznych). Badanie skupia się na obszarze OECD: ilekroć używany jest termin „globalny”, w sensie technicznym oznacza on „38 krajów OECD”.

Podejścia zakrojone na szeroką skalę i podejścia longitudinalne do badania różnic w karierze akademickiej między mężczyznami i kobietami ze względu na wiek i dyscyplinę są stosowane od niedawna, wraz z rosnącym dostępem do krajowych i globalnych, zarówno komercyjnych, jak i niekomercyjnych baz danych dotyczących kadry naukowej oraz baz administracyjnych i bibliometrycznych, takich jak Web of Science, Scopus i Microsoft Academic Graph (MAG), a także Academic Analytics i Digital Bibliography and Library Project (DBLP) w przypadku USA oraz CRISTIN w przypadku Norwegii czy POL-on w przypadku Polski (zob. Boekhout i in., 2021; Elsevier, 2017; Elsevier, 2020; King et al., 2017; Kwiek & Roszka, 2021a; Larivière et al., 2013; Nielsen & Andersen, 2021; Nygaard et al., 2022; Robinson-Garcia et al., 2020; Savage & Olejniczak, 2021; Way et al., 2017; Zhang et al., 2022). Postępy na polu metod dezambiguacji (czyli ujednolicania) autorów prac naukowych poczynione w ramach bibliometrii pozwalają na badanie karier akademickich w skali globalnej i w ujęciu czasowym. Globalna

* Prof. dr hab. Marek Kwiek (kwiekm@amu.edu.pl), Centrum Studiów nad Polityką Publiczną Uniwersytetu im. Adama Mickiewicza w Poznaniu; ORCID: orcid.org/0000-0001-7953-1063

** Łukasz Szymula (lukasz.szymula@amu.edu.pl), Centrum Studiów nad Polityką Publiczną Uniwersytetu im. Adama Mickiewicza w Poznaniu, Poznań; Wydział Matematyki i Informatyki UAM w Poznaniu; ORCID: orcid.org/0000-0001-8714-096X

kadra naukowa może być analizowana z zupełnie nowych perspektyw: naszym celem jest analiza młodych mężczyzn i kobiet naukowców. Podejście generacyjne do zmieniającej się kadry naukowej – zwłaszcza rozkład naukowców pod względem wieku w ramach poszczególnych grup wieku akademickiego (tj. lat od pierwszej publikacji) według płci (np. młodsze kobiety naukowcy vs. starsze kobiety naukowcy) – nie było dotąd stosowane w ilościowych badaniach nauki na poziomie globalnym.

Aktualnie uczestnictwo mężczyzn i kobiet w nauce można analizować w ujęciu czasowym na wcześniej nieosiągalnym poziomie szczegółowości w odniesieniu do krajów, instytucji i dyscyplin, a także w odniesieniu do grup wiekowych i etapów kariery zawodowej. Można badać publikacje i ich autorów za pomocą analiz czasowych, tematycznych, geograficznych i sieciowych, czyli wiązać je z czasem, tematyką badań, miejscami i innymi naukowcami (Börner, 2010, s. 62–63). Prowadząc nasze badania, poszliśmy śladem Huanga i współpracowników (2020), którzy zrekonstruowali pełne historie publikacyjne ponad 1,5 mln naukowców, aby zbadać nierówności między kobietami i mężczyznami w karierze naukowej na świecie (83 kraje, 13 dyscyplin), Boekhouta i współpracowników (2021), którzy prześledzili kariery publikacyjne około 6 mln mężczyzn i kobiet naukowców w latach 1996–2018, oraz Kinga i współpracowników (2017), którzy przeanalizowali 1,5 mln prac naukowych z bibliometrycznej bazy danych JSTOR, aby pokazać różnice między kobietami i mężczyznami w zakresie wskaźnika autocytowań w różnych dyscyplinach i w ujęciu czasowym. Żadna z tych prac nie koncentrowała się jednak na młodych naukowcach.

Celem naszego artykułu jest ustalenie, czego możemy dowiedzieć się o zmieniającej się demografii kadry naukowej na świecie oraz na przestrzeni czasu na podstawie dostępnych globalnych źródeł danych typu bibliometrycznego. Naszym celem było sprawdzenie, jak użyteczne mogą okazać się globalne źródła danych do analizy kadry naukowej w czterech wspomnianych, powiązanych ze sobą wymiarach: płci, wieku, dyscypliny i czasu. Przetestowaliśmy, w jaki sposób można mierzyć transformacje demograficzne globalnej profesji naukowej przy użyciu nowych źródeł danych, przekraczając tym samym tradycyjne podejście, w którym statystyki krajowe pochodzące z krajowych urzędów statystycznych są agregowane do wyższego poziomu, jak w przypadku baz danych dotyczących kadry naukowej tworzonych przez OECD, UNESCO i Unię Europejską (Eurostat).

Podejmując się tych badań, włączamy się do dyskusji na temat korzyści i niedogodności wynikających z wykorzystania globalnych publikacyjnych i cytowaniowych baz danych – czyli w proponowanym przez nas ujęciu „ustrukturyzowanych” Big Data (Holmes, 2017; Salganik, 2018; Selwyn, 2019) – do globalnych badań profesji akademickiej, w których dane dotyczące płci, wieku i dyscyplin tradycyjnie były dostępne niemal wyłącznie przekrojowo (pojedyncze punkty w czasie), głównie w małej skali krajowej

(przez studia przypadków), a ostatnio także w coraz większym stopniu w małej międzynarodowej skali porównawczej poprzez badania ankietowe profesji akademickiej. Badamy szczegółowo zmieniającą się kadrę naukową za pomocą dziesięciu 5-letnich grup wieku akademickiego w ramach każdej dyscypliny zarówno z perspektywy przekrojowej (2021 r.), jak i wzdłużnej (1990–2021).

1.1. Kobiety w obszarze STEM: kontekst teoretyczny

Globalny obraz młodych mężczyzn i kobiet w nauce to ogólny przegląd ich obecności w różnych dyscyplinach na całym świecie. Obraz ten pokazuje tendencje i wzorce w czasie dla różnych dyscyplin. Obecność ta różni się znacznie na poziomie krajowym ze względu na czynniki społeczne, ekonomiczne, polityczne i kulturowe. Istnieją kraje o silniejszej polityce i skuteczniejszych inicjatywach zachęcających kobiety do kształcenia w zakresie nauk ścisłych i przyrodniczych, z większą pulą absolwentek przystępujących do programów doktoranckich i rozpoczynających karierę naukową; są też kraje, w których postawy kulturowe i społeczne mogą zniechęcać kobiety do kariery naukowej. W związku z tym, chociaż różnice w zależności od kraju mogą być znaczące, przedmiotem naszego zainteresowania w tym artykule są globalne różnice między dyscyplinami, zmieniające się w czasie. Ukierunkowane interwencje i polityki mające na celu rozwiązanie problemu niedostatecznej reprezentacji kobiet w niektórych dyscyplinach, wynikającego zarówno z niskiego odsetka osób rozpoczynających karierę, jak i wysokiego odsetka osób przedwcześnie kończących karierę, muszą być rozwijane na poziomie krajowym. Badając kontekst krajowy, można uzyskać bardziej precyzyjne wyjaśnienie źródeł problemu reprezentacji kobiet w nauce, co z kolei pozwala na opracowywanie skuteczniejszych strategii na poziomie poszczególnych dyscyplin. W niniejszym badaniu nie bierzemy pod uwagę przerw w karierze, które mogą być częstsze wśród kobiet ze względu na obowiązki związane z pełnieniem opieki nad dziećmi czy osobami starszymi; nie rozpatrujemy też szerszego kontekstu dotyczącego równowagi między pracą a rodziną.

Młodzi naukowcy, a w szczególności młode kobiety naukowcy, stają przed wyjątkowymi wyzwaniami i barierami, aby rozpocząć, kontynuować i rozwijać karierę naukową. Oprócz niedostatecznej reprezentacji kobiet w nauce istnieją ukryte uprzedzenia (stereotypy i dyskryminacja kobiet w STEM); nieprzyjazna kultura miejsca pracy, zwłaszcza w dyscyplinach zdominowanych przez mężczyzn; a także wyzwania związane z równowagą między życiem zawodowym i życiem prywatnym (i obowiązkami macierzyńskimi), co może prowadzić do przerw w karierze i jej wolniejszego rozwoju. Jak wynika z raportu (Elsevier 2020), kobiety nadal borykają się z poważnymi problemami na każdym etapie kariery: są niedostatecznie reprezentowane na wyższych stanowiskach, rzadziej podejmują współpracę międzynarodową, częściej doświadczają przerw w karierze, rza-

dziej niż mężczyźni publikują artykuły w czasopismach o wysokim wskaźniku oddziaływania, a ich artykuły są średnio rzadziej cytowane. Do każdego z tych problemów istnieje rozbudowana literatura, również naszego autorstwa (zob. Kwiek 2022: 387–564; Kwiek i Roszka 2021a; Kwiek i Roszka 2021b; Kwiek i Roszka 2022a).

Podczas gdy zarówno mężczyźni, jak i kobiety odchodzą z nauki, wskaźnik odchodzenia dla kobiet w naukach ścisłych i inżynieryjnych jest wyższy od wskaźnika dla mężczyzn. Do najważniejszych teorii dotyczących porzucania przez kobiety kariery naukowej należą: teoria „przeciekającego rurociągu”, hipoteza „chłodnego klimatu” oraz hipoteza „autoselekcji”. Teoria „przeciekającego rurociągu” zakłada, że na każdym etapie kariery akademickiej dochodzi do znacznej utraty talentów wśród kobiet, począwszy od absolwentek, poprzez kobiety postdoków (czyli jeszcze przed pierwszym zatrudnieniem na etacie, ale już po uzyskaniu doktoratu), aż do kobiet asystentek i kobiet profesorów. Powodem mogą być bariery systemowe, takie jak choćby uprzedzenia i dyskryminacja (zob. np. Sheltzer i Smith 2014; Sexton et al. 2012; Shaw i Stanton 2012; oraz Wolfinger et al. 2008). Z kolei teoria chłodnego klimatu wskazuje, że nieprzyjazne środowisko pracy w dyscyplinach z obszaru STEM może zniechęcać kobiety do kontynuowania kariery (zob. np. Hall i Sandler 1982; Maranto i Griffin 2011; Morris i Daniel 2008; oraz Cornelius et al. 1988). Teoria autoselekcji dowodzi, że kobiety są niedoreprezentowane w dyscyplinach STEM, ponieważ są nimi mniej zainteresowane z powodu czynników społecznych i kulturowych, które je do nich zniechęcają (zob. np. Hyde et al. 1990; Britton 2017; Whitt et al. 1999). Natomiast metafora „szklanego sufitu” jest wykorzystywana do opisu nierówności rozkładu według płci w nauce z innej perspektywy: „niewidzialnej” tzn. trudno definiowalnej bariery, która uniemożliwia kobietom awans na wyższe poziomy zarządzania w organizacjach, w tym na uczelniach i w instytucjach akademickich. Istnieją systemowe bariery, które sprawiają, że kobiety nie są w stanie osiągnąć możliwości znajdujących się na wyższym poziomie zarządzania. Niewidzialna bariera ogranicza również uznanie zawodowe w nauce – niewiele kobiet w świecie zostaje profesorami zwyczajnymi (w Polsce poziom ten jest wyjątkowo wysoki) (zob. np. Tang 1997; Morrison et al. 1987).

1.2. Pytania badawcze

W tej pracy wprowadziliśmy nową jednostkę analizy: w centrum naszego zainteresowania znajduje się pojedynczy naukowiec (z jego unikalnymi cechami), a nie pojedyncza publikacja (z jej unikalnymi atrybutami wyprowadzanymi z metadanych). Zatem chociaż korzystamy z bibliometrycznego źródła danych (surowe mikrodane pochodzące z bazy Scopus, udostępnione nam przez ICSR Lab firmy Elsevier w ramach wieloletniej umowy o współpracy), skupiamy się na naukowcach i ich cechach, a nie na publikacjach i ich właściwościach.

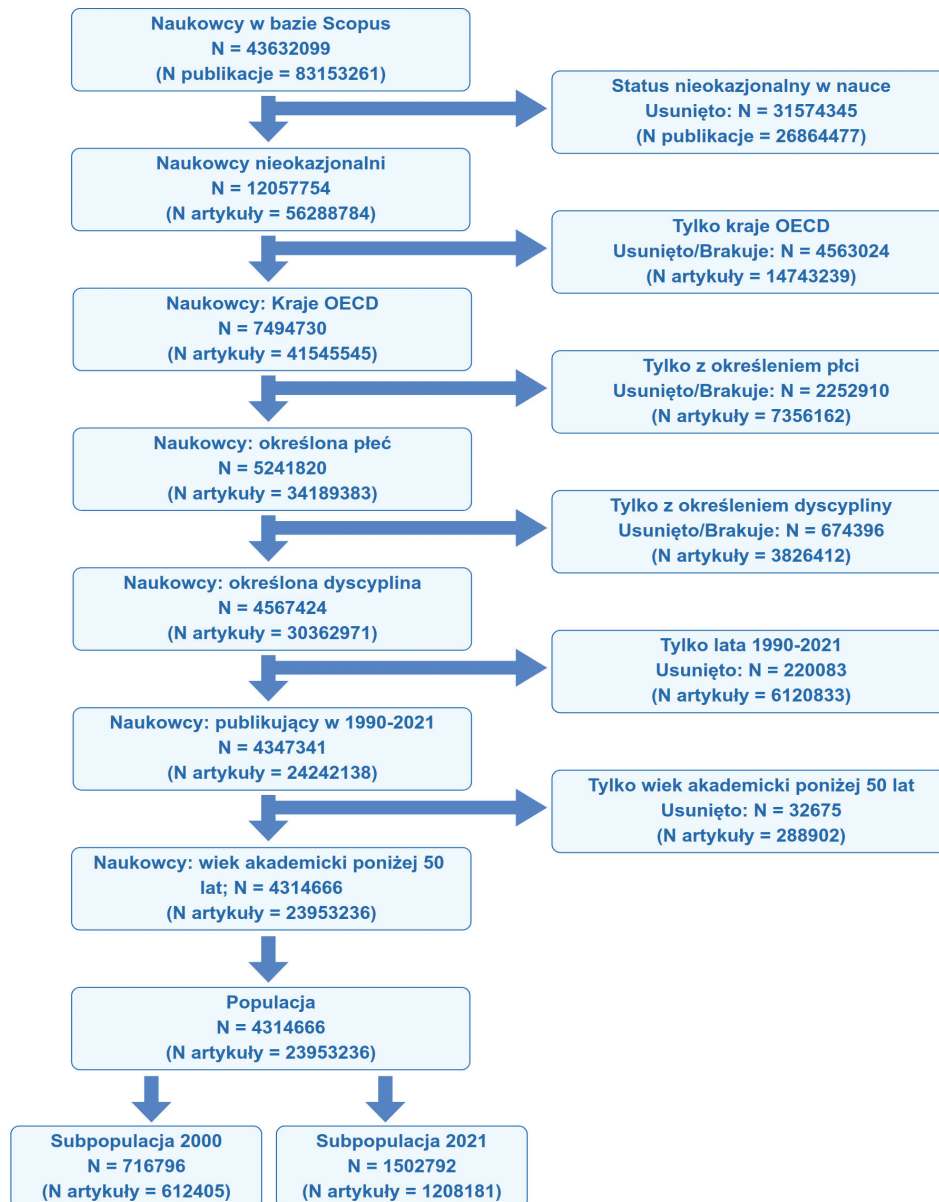
Posiadane i wytworzone przez nas mikrodane określają płeć, wiek akademicki czy doświadczenie akademickie, dyscyplinę, kraj, a także publikacje i ich typy (z okresu całej kariery naukowej); zamieniamy zatem bibliometryczne źródła danych o publikacjach w źródła danych o indywidualnych naukowcach.

Nasze trzy pytania badawcze są następujące: (1) Jaki jest globalny rozkład dyscyplinarny młodych kobiet i mężczyzn naukowców? (2) W jaki sposób rozkład ten zmienia się w czasie? (3) Jakie są trendy dotyczące obecności kobiet w nauce?

2. Dane i metody

Najważniejsze cechy populacji poddanej badaniu wzdłużnemu (dynamicznemu) dla lat 1990–2021 (4 314 666 naukowców, w tym 1 645 860, czyli 38,15% kobiet) przedstawiono w tabeli 1. Natomiast podstawowe cechy subpopulacji w badaniu przekrojowym dla roku 2021 (1 502 792 naukowców, w tym 579 399, czyli 38,55% kobiet) przedstawiono w tabeli 2. Nasza populacja została skonstruowana w następujący sposób (odnosimy się do populacji, a nie do próby, ponieważ dysponujemy ogółem naukowców, wraz z ich atrybutami, jako jednostkami analizy): po pierwsze, aby określić liczbę naukowców, wybrano unikalnych autorów publikacji (typ: artykuł w czasopiśmie, materiał konferencyjny w książce lub czasopiśmie), którzy opublikowali swoje prace w latach 1990–2021. Dla tej wybranej grupy autorów określono lata ich działalności naukowej. Otrzymany zbiór naukowców został następnie zawężony za pomocą pakietu pięciu ograniczeń: (1) kraj OECD, (2) dyscyplina STEMM, (3) płeć (ujęcie binarne: mężczyzna lub kobieta), (4) nieokazjonalny status w nauce: minimalny dorobek naukowy definiowany jako trzy artykuły naukowe w ciągu całej kariery naukowca oraz (5) wiek akademicki, czyli czas, jaki upłynął od pierwszej publikacji, w przedziale 1–50 lat. Dyscyplina naukowców została ustalona jako dyscyplina dominująca na podstawie wszystkich odnośników bibliograficznych w ich wszystkich publikacjach.

Wymóg posiadania co najmniej trzech publikacji w ciągu całego życia pozwolił nam na ograniczenie naszej populacji do naukowców nieokazjonalnych, czyli takich którzy funkcjonują w przestrzeni naukowej bardziej regularnie. Ponadto naukowcy posiadający jedną lub dwie publikacje w bazie Scopus z większym prawdopodobieństwem mogą być wynikiem błędów popełnianych przez algorytmy ujednolicania (tzw. dezambiguacji) nazwisk autorów (zob. Boekhout et al., 2021, s. 3). Warto zaznaczyć, że zasadniczo, pod względem dezambiguacji nazwisk autorów, baza Scopus działa skuteczniej niż Web of Science (Sugimoto & Larivière, 2018, s. 36). Następnie dla każdego naukowca określono doświadczenie akademickie wyrażone w pełnych latach, poczynawszy od roku pierwszej publikacji dowolnego typu. Dla każdego roku działalności badawczej naukowca określono długość jego doświadczenia akademickiego i przynależność do odpowiedniej akademickiej grupy wiekowej. Użyliśmy populacji dla lat 1990–2021 do analiz wzdłuż-



Ryc. 1. Schemat blokowy: etapy konstruowania populacji i dwóch subpopulacji

nych i subpopulacji dla 2021 roku do analizy przekrojowej. Rycina 1 podsumowuje konstrukcję populacji. Początkowo używaliśmy surowych danych dla roku 2020 i wcześniejszych na podstawie wersji bazy danych Scopus z dnia 18 sierpnia 2021 r. Ostatecz-

nie wykorzystano wersję bazy danych Scopus dla 2021 i wcześniejszych, datowaną na 21 października 2022 r.¹

Lista badanych dyscyplin STEM

Skupiliśmy się na wszystkich 16 dyscyplinach z obszaru STEMM, zgodnie z systemem klasyfikacji czasopism stosowanym w bazie Scopus (All Science Journal Classification, ASJC): AGRI, nauki rolnicze i biologiczne; BIO, biochemia, genetyka i biologia molekularna; CHEMENG, inżynieria chemiczna; CHEM, chemia; COMP, informatyka; EARTH, nauki o Ziemi i planetach; ENER, energia; ENG, inżynieria; ENVIR, nauka o środowisku; IMMU, immunologia i mikrobiologia; MATER, materiałoznawstwo; MATH, matematyka; MED medycyna, NEURO, neuronauka; PHARM, farmakologia, toksykologia i farmacja; oraz PHYS, fizyka i astronomia. Każdy naukowiec jest przypisany tylko do jednej dyscypliny.

Aby zbadać rozkład mężczyzn i kobiet w ramach kadry naukowej według grup wieku akademickiego, zastosowaliśmy dwa uzupełniające się podejścia, które nazwaliśmy „poziomym” i „pionowym”.

- (1) Podejście poziome: analiza rozkładu w ujęciu horyzontalnym w ramach tych samych grup wiekowych. Dla każdej dyscypliny, dla każdej z dziesięciu 5-letnich grup wiekowych, odsetek wszystkich naukowców (sumarycznie naukowców mężczyzn i kobiet) wynosi 100%.

¹ Aby uzyskać wyniki na poziomie zagregowanym, działania w ramach ICSR Lab opierały się na wykorzystaniu środowiska Databricks, które pozwalało na zarządzanie i wykonywanie obliczeń w chmurze na zasobach Amazon EC2. Skrypty do generowania wyników zostały napisane z wykorzystaniem biblioteki PySparkSQL. Prace nad uzyskaniem wyników przebiegały w dwóch etapach. Pierwszym etapem była praca na 1% danych bazy Scopus z datą udostępnienia zbioru 18 sierpnia 2021 roku (z ICSR Lab: 1% wolumenu danych na podstawie zbioru 20 000 publikacji z lat 2010–2018 i z uwzględnieniem wszystkich publikacji cytowanych przez te publikacje i powołujących się na nie) przy użyciu klastra w trybie standardowym z Databricks Runtime w wersji 11.2, wykorzystującym technologię Apache Spark w wersji 3.3.0, Scala 2.12 oraz instancję i3.2xlarge z 61 GB pamięci, 8 rdzeni, od jednego do czterech procesów roboczych dla węzła roboczego oraz instancji i3.xlarge z 30,5 GB pamięci, 4 rdzeni dla węzła sterownika. Testowe uruchomienia skryptów objęły 1% danych, a ich celem była optymalizacja czasu i kosztu wykonywanych obliczeń. Po sprawdzeniu poprawności skryptów wykonano ostateczne uruchomienie. Operacja została przeprowadzona na 100% bazy danych Scopus z datą udostępnienia zbioru 21 października 2022 roku przy użyciu klastra w trybie standardowym z Databricks Runtime w wersji 11.2 ML, wykorzystującym technologię Apache Spark w wersji 3.3.0, Scala 2.12 oraz instancję i3.2xlarge z 61 GB pamięci, 8 rdzeni, od jednego do sześciu procesów roboczych dla węzła roboczego oraz instancji c4.2xlarge z 15 GB pamięci, 4 rdzeni dla węzła sterownika. Czas wykonania całego skryptu zajął 1,13 godziny; operacja ta została uruchomiona 22 listopada 2022 roku.

Tabela 1. Populacja dla lat 1990–2021: najważniejsze cechy

		Kobiety			Mężczyźni			Razem		
		<i>n</i>	% wier- szowy	% kolu- mno- wy	<i>n</i>	% wier- szowy	% kolu- mno- wy	<i>n</i>	% wier- szowy	% kolu- mno- wy
Dyscypliny	Razem	1 645 860	38,15	100	2 668 806	61,85	100	4 314 666	100	100
	AGRI	104 805	39,98	6,37	157 318	60,02	5,89	262 123	100	6,08
	BIO	328 806	46,26	19,98	381 963	53,74	14,31	710 769	100	16,47
	CHEM	87 608	30,16	5,32	202 843	69,84	7,60	290 451	100	6,73
	CHEMENG	4294	23,06	0,26	14 330	76,94	0,54	18 624	100	0,43
	COMP	16 191	16,59	0,98	81 414	83,41	3,05	97 605	100	2,26
	EARTH	34 042	27,62	2,07	89 221	72,38	3,34	123 263	100	2,86
	ENER	3255	19,09	0,20	13 793	80,91	0,52	17 048	100	0,40
	ENG	24 992	11,52	1,52	191 978	88,48	7,19	216 970	100	5,03
	ENVIR	35 867	38,35	2,18	57 661	61,65	2,16	93 528	100	2,17
	IMMU	26 805	53,24	1,63	23 547	46,76	0,88	50 352	100	1,17
	MATER	26 227	26,16	1,59	74 043	73,84	2,77	100 270	100	2,32
	MATH	11 915	20,15	0,72	47 206	79,85	1,77	59 121	100	1,37
	MED	836 890	45,44	50,85	1 005 040	54,56	37,66	1 841 930	100	42,69
	NEURO	40 961	47,20	2,49	45 819	52,80	1,72	86 780	100	2,01
	PHARM	15 641	41,35	0,95	22 183	58,65	0,83	37 824	100	0,88
	PHYS	47 561	15,44	2,89	260 447	84,56	9,76	308 008	100	7,14
Kraje OECD (TOP 10)	USA	540 501	39,73	32,84	819 882	60,27	30,72	1 360 383	100	31,53
	Japonia	92 601	19,28	5,63	387 599	80,72	14,52	480 200	100	11,13
	Niemcy	118 509	33,49	7,20	235 312	66,51	8,82	353 821	100	8,20
	Wielka Bryt.	116 285	39,49	7,07	178 187	60,51	6,68	294 472	100	6,82
	Włochy	119 688	50,36	7,27	117 960	49,64	4,42	237 648	100	5,51
	Francja	93 770	42,07	5,70	129 110	57,93	4,84	222 880	100	5,17
	Kanada	68 983	42,75	4,19	92 393	57,25	3,46	161 376	100	3,74
	Hiszpania	71 656	48,13	4,35	77 233	51,87	2,89	148 889	100	3,45
	Australia	50 652	44,79	3,08	62 425	55,21	2,34	113 077	100	2,62
	Korea Płd.	19 886	19,32	1,21	83 038	80,68	3,11	102 924	100	2,39

- (2) Podejście pionowe: analiza rozkładu w ujęciu pionowym – oddzielnie mężczyźni i oddzielnie kobiety – we wszystkich grupach wiekowych. W każdej dyscyplinie występuje 100% mężczyzn i 100% kobiet naukowców, różnie rozmieszczonych w 10 grupach wiekowych.

Tabela 2. Subpopulacja dla roku 2021: najważniejsze cechy

		Kobiety			Mężczyźni			Razem		
		<i>n</i>	% wier- szowy	% kolu- mno- wy	<i>n</i>	% wier- szowy	% kolu- mno- wy	<i>n</i>	% wier- szowy	% kolu- mno- wy
Grupy wiekowe	Razem	579 399	38,55	100	923 393	61,45	100	1 502 792	100	100
	5 lat i mniej	148 749	46,26	25,67	172 795	53,74	18,71	321 544	100	21,40
	6–10	149 875	43,47	25,87	194 936	56,53	21,11	344 811	100	22,94
	11–15	102 419	40,52	17,68	150 366	59,48	16,28	252 785	100	16,82
	16–20	71 335	36,73	12,31		63,27	13,31	194 213	100	12,92
	21–25	45 297	32,74	7,82	93 052	67,26	10,08	138 349	100	9,21
	26–30	30 302	28,86	5,23	74 698	71,14	8,09	105 000	100	6,99
	31–35	17 736	24,83	3,06	53 682	75,17	5,81	71 418	100	4,75
	36–40	8 432	20,58	1,46	32 541	79,42	3,52	40 973	100	2,73
	41–45	3 833	17,27	0,66	18 357	82,73	1,99	22 190	100	1,48
	46–50	1 421	12,35	0,25	10 088	87,65	1,09	11 509	100	0,77
Dyscypliny	AGRI	42 657	40,13	7,36	63 645	59,87	6,89	106 302	100	7,07
	BIO	92 185	43,27	15,91	120 854	56,73	13,09	213 039	100	14,18
	CHEM	22 450	30,21	3,87	51 862	69,79	5,62	74 312	100	4,94
	CHEMENG	1 287	24,98	0,22	3 865	75,02	0,42	5 152	100	0,34
	COMP	6 449	18,20	1,11	28 986	81,80	3,14	35 435	100	2,36
	EARTH	14 446	27,87	2,49	37 390	72,13	4,05	51 836	100	3,45
	ENER	1 527	20,28	0,26	6 004	79,72	0,65	7 531	100	0,50
	ENG	9 029	13,82	1,56	56 326	86,18	6,10	65 355	100	4,35
	ENVIR	14 688	40,15	2,54	21 892	59,85	2,37	36 580	100	2,43
	IMMU	6 949	50,03	1,20	6 940	49,97	0,75	13 889	100	0,92
	MATER	10 257	27,09	1,77	27 601	72,91	2,99	37 858	100	2,52
	MATH	4 653	20,02	0,80	18 590	79,98	2,01	23 243	100	1,55
	MED	318 792	46,14	55,02	372 166	53,86	40,30	690 958	100	45,98
	NEURO	1 3873	43,76	2,39	17 833	56,24	1,93	31 706	100	2,11
	PHARM	3 190	45,98	0,55	3 748	54,02	0,41	6 938	100	0,46
	PHYS	16 967	16,53	2,93	85 691	83,47	9,28	102 658	100	6,83
Kraje OECD (TOP 10)	USA	176 646	40,63	30,49	258 155	59,37	27,96	434 801	100	28,93
	Japonia	22 331	18,15	3,85	100 695	81,85	10,90	123 026	100	8,19
	Niemcy	36 659	32,19	6,33	77 212	67,81	8,36	113 871	100	7,58
	Włochy	51 171	49,21	8,83	52 821	50,79	5,72	103 992	100	6,92
	Wielka Bryt.	40 328	38,88	6,96	63 392	61,12	6,87	103 720	100	6,90
	Francja	31 657	39,74	5,46	47 996	60,26	5,20	79 653	100	5,30
	Hiszpania	29 067	46,89	5,02	32 925	53,11	3,57	61 992	100	4,13
	Kanada	24 022	42,36	4,15	32 685	57,64	3,54	56 707	100	3,77
	Australia	21 160	44,49	3,65	26 396	55,51	2,86	47 556	100	3,16
	Korea Płd.	7 903	19,31	1,36	33 034	80,69	3,58	40 937	100	2,72

Tabela 3. Subpopulacja dla roku 2021 według dyscypliny i płci

Dyscyplina	Razem	Mężczyźni	Kobiety	Odsetek mężczyzn	Odsetek kobiet
MED	690 958	372 166	318 792	53,86	46,14
BIO	213 039	120 854	92 185	56,73	43,27
AGRI	106 302	63 645	42 657	59,87	40,13
PHYS	102 658	85 691	16 967	83,47	16,53
CHEM	74 312	51 862	22 450	69,79	30,21
ENG	65 355	56 326	9 029	86,18	13,82
EARTH	51 836	37 390	14 446	72,13	27,87
MATER	37 858	27 601	10 257	72,91	27,09
ENVIR	36 580	21 892	14 688	59,85	40,15
COMP	35 435	28 986	6 449	81,80	18,20
NEURO	31 706	17 833	13 873	56,24	43,76
MATH	23 243	18 590	4 653	79,98	20,02
IMMU	13 889	6 940	6 949	49,97	50,03
ENER	7 531	6 004	1 527	79,72	20,28
PHARM	6 938	3 748	3 190	54,02	45,98
CHEMENG	5 152	3 865	1 287	75,02	24,98
TOTAL	1 502 792	923 393	579 399	61,45	38,55

Część naszych badań opiera się na konstrukcji badań wzdłużnych (dynamicznych) w szerszym sensie, co wymaga krótkiego wyjaśnienia metodologicznego. W badaniach wzdłużnych w wąskim sensie dane są zbierane w wielu punktach w czasie od tej samej grupy uczestników; wykorzystaliśmy to wąskie podejście w najnowszym badaniu 2326 polskich profesorów tytularnych, śledząc ich awanse, publikacje i klasy produktywności w okresie 40 lat (Kwiek i Roszka 2023). W klasycznych definicjach badania wzdłużne dotyczą zbierania i analizowania danych w czasie i jest to szeroki termin opisujący rodzinę metod: badania wzdłużne obejmują powtarzane badania przekrojowe, badania prospektywne i badania retrospektywne (Menard 2002: 2–3). W punkcie wyjścia projekt wzdłużny pozwala na pomiar różnic lub zmian zmiennych pomiędzy jednym okresem a okresem kolejnym. Badania wzdłużne w tym szerszym znaczeniu to badania, w których (a) dane są zbierane dla każdej cechy czy zmiennej dla dwóch lub więcej odrębnych okresów; (b) analizowane osoby czy przypadki są te same (lub przynajmniej porównywalne) pomiędzy jednym a drugim okresem; oraz (c) analiza obejmuje porównanie danych pomiędzy okresami (Menard 2002: 2).

Nasze badanie reprezentuje zarówno konstrukcję przekrojową (w analizach pojedynczego punktu w czasie, 2021), jak i szeroko rozumianą konstrukcję wzdłużną, w jej

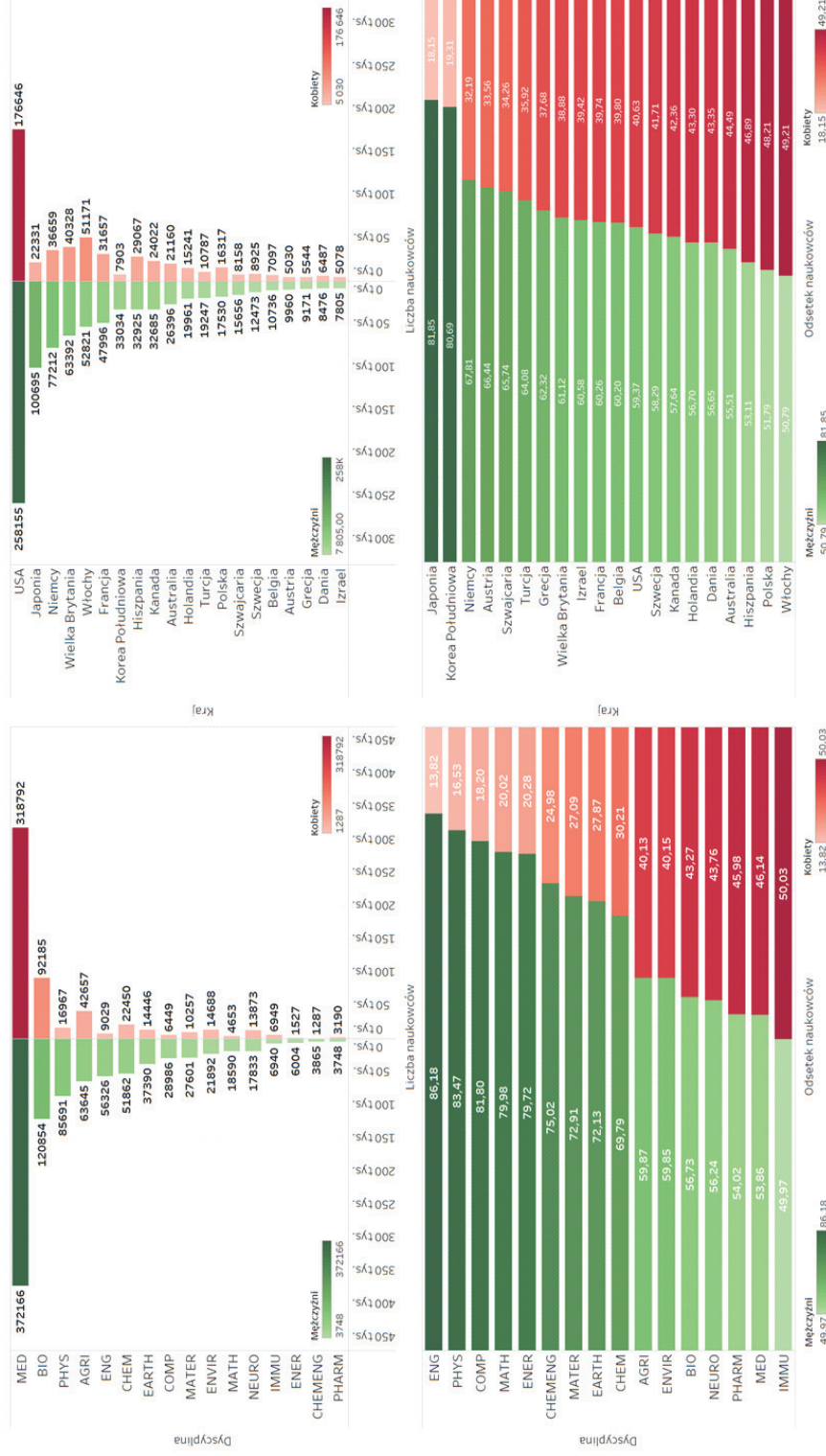
wariancie powtórzonego badania przekrojowego (analizując dwa punkty w czasie, 2000 i 2021, oraz trend w latach 1990–2021, zgodnie z założeniem, że dane przekrojowe mogą być powtarzane w czasie przy zachowaniu wysokiego poziomu spójności między pytaniami, Ruspini 1999). Nasze zestawy przypadków – naukowcy z ich mikrodanymi – dla poszczególnych okresów nie są całkowicie różne: do pewnego stopnia pokrywają się (dla naukowców aktywnych przez dłuższy czas). Nasze mikrodane obejmują indywidualnych naukowców, co oznacza, że ich zapisy na poziomie indywidualnym zawierają te same zmienne mierzone w kilku różnych punktach czasowych. Z tego powodu zostały one połączone w jeden zbiór danych: zwiększyło to wielkość próby, a także wprowadziło wymiar czasowy, tak jak to sugeruje literatura przedmiotu (Ruspini 1999: 222).

3. Wyniki badań

3.1. Wyniki ogólne

Jakkolwiek analiza zmian liczby mężczyzn i kobiet naukowców w czasie może być zniekształcona przez niemożność odróżnienia ekspansji liczby naukowców od ekspansji liczby czasopism indeksowanych w dużych bazach bibliometrycznych, to można prześledzić zmieniającą się względną obecność kobiet w nauce. Chociaż rosnąca liczba publikujących naukowców w czasie jest skorelowana z rosnącym wskaźnikiem pokrycia czasopism w bazie Scopus, odsetek publikujących mężczyzn i kobiet naukowców jest niezależny od wskaźnika pokrycia czasopism. Co za tym idzie, o ile zmieniająca się w czasie *liczba* publikujących naukowców nie stanowi rzetelnej miary zmieniającego się udziału kobiet w globalnej nauce, o tyle *odsetek* publikujących mężczyzn i kobiet naukowców adekwatnie odzwierciedla zmiany zachodzące w globalnej kadrze naukowej w czasie.

W 2021 roku 45,98% globalnej kadry naukowej w STEMM (zgodnie z definicją przyjętą w tym badaniu: w 38 krajach OECD i o nieokazjonalnym statusie publikowania w bazie Scopus) zajmowało się badaniami medycznymi (MED: 690 958 naukowców), a drugą co do wielkości dyscypliną była biochemia, genetyka i biologia molekularna (BIO: 213 039 naukowców). W 2021 r. w naszej populacji znalazło się 1,5 mln naukowców zgodnie z definicjami, w tym 923 tys. mężczyzn i 579 tys. kobiet (38,55%). Większość kobiet naukowców (63,09%) koncentrowała się w sześciu krajach: w USA, Włoszech, Wielkiej Brytanii, Niemczech, Francji i Hiszpanii. Ponad 70% kobiet naukowców zajmowało się medycyną oraz biochemią, genetyką i biologią molekularną. Immunologia i mikrobiologia (IMMU) wykazywały najwyższy odsetek kobiet naukowców (50,03%), a następnie znalazło się kilka dziedzin z ponad 40% udziałem kobiet (np. AGRI, ENVIR, BIO, NEURO, PHARM i MED). Z kolei w inżynierii, fizyce i astronomii, informatyce i matematyce odsetek kobiet wynosił nie więcej niż 20%.



Ryc. 2. Liczba publikujących naukowców w dziedzinie STEMM w 38 krajach OECD według dyscypliny i płci (po lewej u góry) oraz według kraju (tylko 20 największych systemów) i płci (po prawej u góry). Udział według dyscypliny i płci (po lewej na dole) oraz według kraju (tylko 20 największych systemów OECD) i płci (po prawej na dole) (w %), 2021 ($N = 1502792$)

3.2. Wyniki: podejście poziome

3.2.1. Ujęcie przekrojowe (2021): wszystkie grupy wiekowe poziomo

Dyscypliny w jednym punkcie czasowym (rok 2021) są reprezentowane przez naukowców z różnych grup wieku akademickiego. Rycina 3 pokazuje odsetek kobiet naukowców w poszczególnych dyscyplinach według grup wieku akademickiego. Zasadniczo obserwujemy skutki znacznego napływu kobiet naukowców (obecnych w 2021 r.) do większości dyscyplin w ostatnich latach i dekadach: dla młodszych pokoleń pracujących w 2021 r. odsetek kobiet naukowców jest zdecydowanie wyższy niż dla starszych pokoleń.

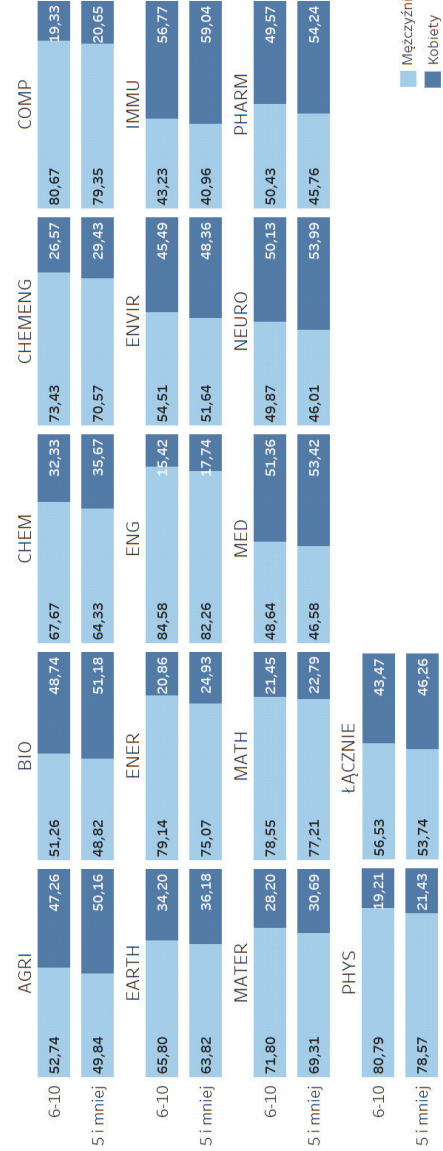
Ogólnie rzecz biorąc, oczekując rosnącej liczby kobiet naukowców we wszystkich dyscyplinach STEMM w miarę przesuwania się w dół w ramach grup wieku akademickiego, badaliśmy bieżące zmiany w oparciu o ujęcie przekrojowe (jeden rok – 2021), zwłaszcza badając najmłodsze grupy wiekowe. MED i BIO wykazują strukturę, w której dla każdej kolejnej młodszej grupy wiekowej w 2021 r. obserwowano wyższy udział kobiet naukowców. PHYS, COMP, ENG i MATH, zwane dalej wielką czwórką, z matematyzowane dyscypliny tradycyjnie zdominowane przez mężczyzn, skupiające 226 691 naukowców w naszej populacji (15,09%; w tym zaledwie 37 098, czyli 16,37% kobiet), wykazują natomiast stabilną strukturę, w której dla każdej kolejnej młodszej grupy wiekowej w 2021 r. obserwowany jest podobny (lub tylko nieznacznie wyższy) udział kobiet.

Te dwa przeciwstawne wzorce demograficzne pokazują różny poziom napływu młodych kobiet naukowców do dyscyplin w przeszłości: duży i rosnący oraz mały i stabilny. Wystarczy porównać matematykę MATH i biochemię, genetykę i biologię molekularną BIO na ryc. 3: w jednym interesującym nas roku, przy najnowszych dostępnych danych, udział kobiet bardzo młodych, młodych i w średnim wieku w MATH jest praktycznie taki sam; z kolei w przypadku BIO udział kobiet w tych samych grupach wieku akademickiego stale rośnie w każdej kolejnej młodszej grupie.

Obecny globalny rozkład dyscyplinarny młodych kobiet w nauce ma znaczenie dla kwestii rozkładu mężczyzn i kobiet w nauce w przyszłości, pomimo wysokiego wskaźnika odchodzenia z nauki wśród młodych naukowców ogółem i młodych kobiet naukowców w szczególności (około 10% rocznie; zob. Boothby i in. 2022). Obecne młode kohorty będą w ciągu dekady kohortami w średnim wieku, a obecne najstarsze kohorty przestaną publikować, wycofując się z pracy akademickiej, co pociągnie za sobą nowe wyzwania dla wielkiej czwórki stale silnie zdominowanej przez mężczyzn i coraz wyraźniej odstającej od pozostałych dyscyplin z obszaru STEMM (nie wspominając o dyscyplinach społecznych i humanistycznych, pomijanych w tym badaniu z powodu niedokładności wybranego źródła danych – bibliometrycznej bazy Scopus).

		AGRI			BIO			CHEM			CHEMENG		
46-50	89,08	10,92	83,96	16,04	90,91	92,73	100,00	9,09	92,73	100,00	9,09	92,73	7,27
41-45	84,58	15,42	76,50	23,50	85,67	87,27	92,73	14,33	87,27	92,73	14,33	87,27	7,27
36-40	80,68	19,32	72,90	27,10	82,45	85,67	90,20	17,55	85,67	90,20	17,55	85,67	9,80
31-35	76,02	23,98	68,96	31,04	80,11	85,39	85,39	19,89	85,39	85,39	19,89	85,39	14,61
26-30	70,25	29,75	65,22	34,78	75,56	78,85	78,85	24,44	78,85	78,85	24,44	78,85	21,15
21-25	66,80	33,20	62,15	37,85	72,39	76,47	76,47	27,61	76,47	76,47	27,61	76,47	23,53
16-20	61,52	38,48	58,58	41,42	70,95	76,86	76,86	29,05	76,86	76,86	29,05	76,86	23,14
11-15	57,41	42,59	54,40	45,60	68,60	72,97	72,97	31,40	72,97	72,97	31,40	72,97	27,03
6-10	52,74	47,26	51,26	48,74	67,67	73,43	73,43	32,33	73,43	73,43	32,33	73,43	26,57
5 i mniej	49,84	50,16	48,82	51,18	64,33	70,57	70,57	35,67	70,57	70,57	35,67	70,57	29,43
		EARTH			ENER			ENG			ENVIR		
46-50	93,66	6,34	100,00		98,19	92,67	1,81	92,67	98,19	92,67	1,81	92,67	7,83
41-45	90,06	9,94	96,77	3,23	95,50	87,92	4,50	87,92	95,50	87,92	4,50	87,92	12,08
36-40	84,65	15,35	94,29	5,71	94,58	83,31	5,42	83,31	94,58	83,31	5,42	83,31	16,69
31-35	83,42	16,58	88,65	1,35	92,27	78,44	7,73	78,44	92,27	78,44	7,73	78,44	21,56
26-30	81,01	18,99	89,37	10,63	90,75	72,20	9,25	72,20	90,75	72,20	9,25	72,20	27,80
21-25	76,19	23,81	83,48	16,52	88,67	66,97	11,33	66,97	88,67	66,97	11,33	66,97	33,03
16-20	72,18	27,82	84,52	15,48	87,22	62,91	12,78	62,91	87,22	62,91	12,78	62,91	37,09
11-15	68,77	31,23	80,09	19,91	85,67	57,87	14,33	57,87	85,67	57,87	14,33	57,87	42,13
6-10	65,80	34,20	79,14	20,86	84,58	54,51	15,42	54,51	84,58	54,51	15,42	54,51	45,49
5 i mniej	63,82	36,18	75,07	24,93	82,26	51,64	17,74	51,64	82,26	51,64	17,74	51,64	48,36
		MATER			MATH			MED			NEURO		
46-50	89,74	10,26	95,50	4,50	84,06	82,17	15,94	82,17	84,06	82,17	15,94	82,17	17,83
41-45	91,74	8,26	90,36	9,64	79,73	76,13	20,27	76,13	79,73	76,13	20,27	76,13	23,87
36-40	84,14	15,86	89,21	10,79	75,42	72,15	24,58	72,15	75,42	72,15	24,58	72,15	27,85
31-35	82,77	17,23	85,99	14,01	70,14	68,23	29,86	68,23	70,14	68,23	29,86	68,23	31,77
26-30	80,15	19,85	79,29	20,71	65,33	60,40	34,67	60,40	65,33	60,40	34,67	60,40	34,31
21-25	77,10	22,90	77,49	22,51	60,40	55,58	39,60	55,58	60,40	55,58	39,60	55,58	39,14
16-20	73,03	26,97	79,08	20,92	55,58	51,89	44,42	51,89	55,58	51,89	44,42	51,89	44,70
11-15	72,71	27,29	78,98	21,02	51,89	48,11	48,11	48,11	51,89	48,11	48,11	48,11	50,13
6-10	71,80	28,20	78,55	21,45	48,64	49,87	51,36	49,87	48,64	49,87	51,36	49,87	50,13
5 i mniej	69,31	30,69	77,21	22,79	46,58	46,01	53,42	46,01	46,58	46,01	53,42	46,01	53,99
		PHYS			ŁĄCZNIE								
46-50	94,96	5,04	87,65	12,35									
41-45	92,20	7,80	82,73	17,27									
36-40	90,39	9,61	79,42	20,58									
31-35	89,55	10,45	75,17	24,83									
26-30	87,02	12,98	71,14	28,86									
21-25	85,10	14,90	67,26	32,74									
16-20	83,12	16,88	63,27	36,73									
11-15	82,81	17,19	59,48	40,52									
6-10	80,79	19,21	56,53	43,47									
5 i mniej	78,57	21,43	53,74	46,26									

Ryc. 3. Rosnący udział kobiet w młodszych pokoleniach naukowców, z nielicznymi wyjątkami (np. COMP, MATH). Ujęcie poziome: rozkład publikujących naukowców z grupy STEMM według dyscypliny, grupy wieku akademickiego i płci (procenty wierszowe: 100% poziomo), 2021 ($N = 1\,502\,792$)



Ryc. 4. Zbliżenie tylko na młodych naukowców. Większa obecność młodych mężczyzn niż młodych kobiet we wszystkich dyscyplinach STEM z wyjątkiem sześciu (np. MED). Ujęcie poziome: tylko młodzi naukowcy (wiek akademicki 10 lat i mniej). Rozkład młodych naukowców według dyscypliny, grupy wieku akademickiego i płci (procenty wierszowe): 100% poziomo, 2021 ($N = 666\,355$)

Tradycyjne dane zagregowane pod względem płci i wieku dotyczące ogółu naukowców w poszczególnych dyscyplinach, państwach i instytucjach przesłaniają znacznie bardziej subtelny obraz zmieniającej się dynamiki zmian w obrębie poszczególnych dyscyplin i grup wieku akademickiego i pomiędzy nimi. W tym badaniu analizujemy subpopulację „młodych” naukowców (wiek akademicki 10 i mniej lat, ryc. 4).

3.2.2. Porównawcze ujęcie poziome (2000 vs. 2021)

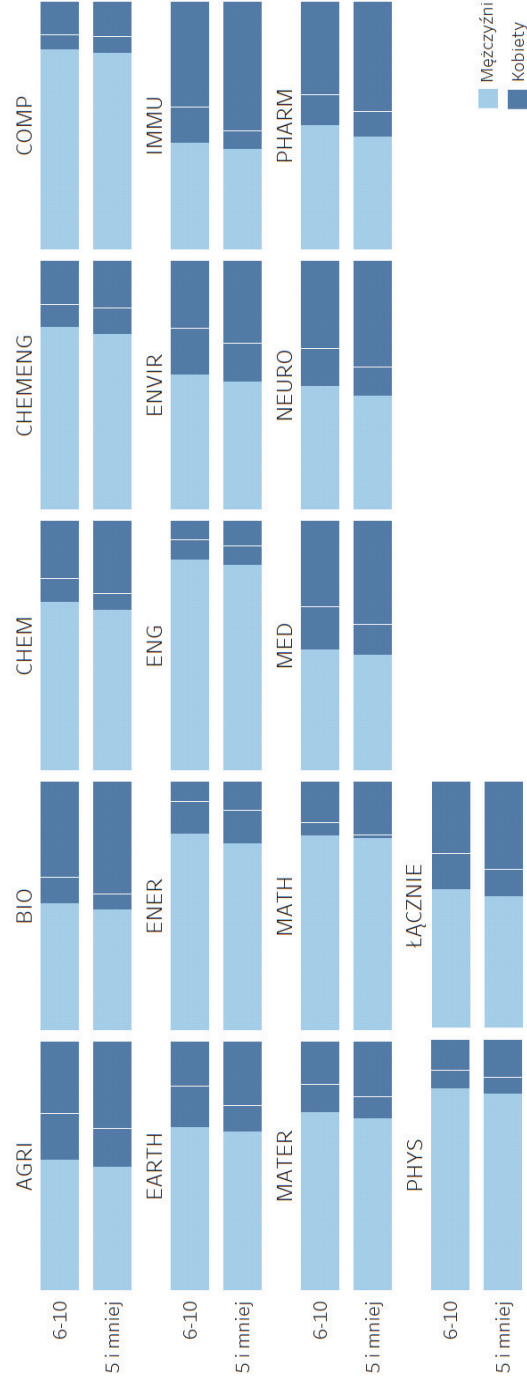
Porównując udział kobiet w dyscyplinach STEMM z jeszcze innej perspektywy dwóch ujęć z lat 2000 i 2021 (ryc. 5): w przypadku wszystkich dyscyplin ich udział wzrósł, choć w różnym stopniu. Białe linie pokazują udział kobiet naukowców dla roku 2000, natomiast ciemnoniebieskie słupki po prawej stronie pokazują go dla roku 2021. W przypadku najmłodszej grupy wiekowej, dla wszystkich dyscyplin łącznie, udział ten zwiększył się z jednej trzeciej do połowy (z 34,93% do 50,16%), a udział mężczyzn naukowców zmniejszył się z dwóch trzecich do połowy (z 65,07% do 49,84%). Porównując starszą kategorię wiekową 31–35 lat (wieku akademickiego czyli lat od pierwszej publikacji, jak w całej pracy), udział kobiet naukowców wzrósł trzykrotnie, z 8,12% do 23,98%. Z perspektywy dwóch dekad zmiany są zauważalne we wszystkich dyscyplinach – choć w większości przypadków można je określić jako niewielkie.

3.2.3. Malejąca izolacja kobiet naukowców w wielkiej czwórce dyscyplin zmatematyzowanych

Porównaliśmy udział młodych i starszych kobiet i mężczyzn naukowców w poszczególnych dyscyplinach w jednym roku – 2021. Wśród młodych naukowców udział kobiet w kilku dyscyplinach wynosił około połowy (np. BIO i MED), natomiast wśród starszych naukowców udział ten był znacznie niższy (tab. 4). W niektórych dyscyplinach udział starszych kobiet naukowców wynosił 10% lub mniej, co oznacza, że różnice między kobietami i mężczyznami były co najmniej 10-krotne (np. ENG i PHYS: odpowiednio 6,31% i 9,21%).

W wielu instytucjach starsze kobiety naukowcy nie tyle stanowiły mniejszość, co raczej były symbolami (pojedynczymi, wzorcowymi naukowcami reprezentującymi wszystkie kobiety naukowców; zob. specyfikę owych *tokens* w Kanter, 1977; na temat roli klimatu panującego na wydziałach zob. Fox & Nikivincze, 2021). Jednak izolacja młodych kobiet naukowców w COMP, ENG, MATH i PHYS z perspektywy międzykohortowej maleje, czyli ich widoczność wśród najmłodszych kohort rośnie.

Młodsze grupy wiekowe wykazują większą liczbę kobiet naukowców i ich wyższe odsetki we wszystkich dyscyplinach, również w dyscyplinach zdominowanych przez mężczyzn (ENG, MATH, PHYS) i w dyscyplinach bliższych równomiernemu rozkładowi mężczyzn i kobiet (MED, AGRI, BIO). Kobiety są liczniejsze i stanowią wyższe odsetki naukowców w młodszych grupach wieku akademickiego. We wszystkich dyscyplinach,



Ryc. 5. Rosnący udział młodych kobiet naukowców we wszystkich dyscyplinach w czasie. Przegląd kierunku zmian procentowych, 2000 vs. 2021: podejście poziome. Zbliżenie tylko na młodych naukowców (wiek akademicki 10 lat lub mniej). Rozkład młodych publikujących naukowców STEM według dyscypliny, grupy wieku akademickiego i płci; ciemnoniebieski odsetek kobiet naukowców 2021, białe linie odsetek kobiet naukowców 2000 (procent wierszo-
wy: 100% poziomo) $N_{2021} = 666\ 355$, $N_{2000} = 437\ 113$)

Tabela 4. Częstości i odsetki kobiet naukowców wśród publikujących naukowców STEM w dwóch kohortach (młodzi i starsi), 2021

Dyscyplina	Młodzi naukowcy (10 lat i mniej doświadczenia)			Starsi naukowcy (31–50 lat doświadczenia)			
	Wszyscy młodzi naukowcy	Młodzi naukowcy – kobiety	% młodych naukowców – kobiet	Dyscyplina	Wszyscy starsi naukowcy (31–50)	Starsi naukowcy – kobiety	% Starszych naukowców – kobiet
AGRI	41 954	20 389	48,60	AGRI	10 799	2206	20,43
BIO	89 295	44 533	49,87	BIO	23 377	6422	27,47
CHEM	36 368	12 394	34,08	CHEM	7582	1313	17,32
CHEMENG	2523	707	28,02	CHEMENG	455	51	11,21
COMP	12 678	2518	19,86	COMP	2642	353	13,36
EARTH	18 168	6363	35,02	EARTH	7205	1026	14,24
ENER	4420	1013	22,92	ENER	252	21	8,33
ENG	28 808	4745	16,47	ENG	4864	307	6,31
ENVIR	16 557	7758	46,86	ENVIR	2545	458	18,00
IMMU	5651	3270	57,87	IMMU	1587	430	27,1
MATER	20 664	6103	29,53	MATER	2097	323	15,4
MATH	8327	1835	22,04	MATH	3481	386	11,09
MED	324 524	170 004	52,39	MED	60 685	15 775	25,99
NEURO	14 260	7400	51,89	NEURO	2903	758	26,11
PHARM	3341	1741	52,11	PHARM	744	223	29,97
PHYS	38 817	7851	20,23	PHYS	14 872	1370	9,21
RAZEM	666 355	298 624	44,81	RAZEM	146 090	31 422	21,51

z wyjątkiem sześciu (AGRI, BIO, IMMU, MED, NEURO i PHARM), jest więcej najmłodszych mężczyzn niż najmłodszych kobiet naukowców, a także, co poniekąd oczywiste, znacznie więcej starszych mężczyzn niż starszych kobiet naukowców.

Tabela 4 pokazuje wzrost odsetka kobiet wśród młodszych grup wieku akademickiego (10 lat lub mniej doświadczenia) w porównaniu ze starszymi grupami (31–50 lat doświadczenia) we wszystkich dyscyplinach. Sugeruje to wyraźnie rosnący trend uczestnictwa kobiet w nauce. Jednak w kontekście naukowców z różnych grup wieku akademickiego pracujących w tym samym czasie (2021) w wielkiej czwórce dyscyplin z matematyzowanych izolacja młodych kobiet znacznie się zmniejszyła w porównaniu z izolacją starszych kobiet. W 2021 r. dla tych czterech dyscyplin odsetek kobiet w młodszych pokoleniach był co najmniej dwukrotnie wyższy niż w starszych pokoleniach: na przykład w dyscyplinie inżynieria młode kobiety stanowiły 16,47% w porównaniu z zaledwie 6,31% dla starszych kohort (tab. 4).

Ten trend większej obecności kobiet w młodszych kohortach jest silniejszy w dyscyplinach bliższych osiągnięcia równego rozkładu mężczyzn i kobiet. Na przykład w medycynie w 2021 r. młode kobiety stanowiły 52,39% (w porównaniu z 25,99% w przypadku starszych kobiet naukowców; a w biochemii odpowiednio 49,87% i 27,47%).

Dzięki naszym mikrodanym na poziomie indywidualnym możemy dokładniej przeanalizować, co w praktyce oznacza izolacja kobiet w dyscyplinach STEMM. Jak pokazano w tabeli 5, kobiety są bardziej reprezentowane w ujęciu liczbowym i procentowym, gdy przechodzimy od starszych do młodszych pokoleń w dziesięciu grupach wieku akademickiego (w tym samym roku 2021). Wskazuje to na malejącą izolację kobiet z każdą kolejną młodszą grupą wiekową. Szczegółowe przykłady dodatkowo uwydatniają różnice pomiędzy obecnością kobiet w młodszych i starszych pokoleniach.

Na przykład w przypadku inżynierii, w grupie wieku akademickiego 36–40 lat, były tylko 84 (publikujące) kobiety w porównaniu z 1486 mężczyznami. Obrazuje to zdecydowaną różnicę pod względem reprezentacji – mężczyźni mają ponad 17-krotną przewagę liczebną nad kobietami. Jednak w młodszej (5 lat i mniej) grupie wieku akademickiego różnica ta znacznie maleje, z 2316 kobietami inżynierami i 10 739 mężczyznami inżynierami. W tym przypadku liczba mężczyzn jest już tylko pięciokrotnie większa od liczby kobiet. Ten przykład ilustruje, jak radykalnie izolacja kobiet w inżynierii maleje w młodszych pokoleniach. W dyscyplinie fizyka i astronomia (PHYS), w grupie wieku akademickiego 46–50 lat, było tylko 79 kobiet w porównaniu z 1489 mężczyznami (19-krotna różnica). Natomiast w młodszej (5 lat i mniej) grupie wieku akademickiego było 3817 kobiet fizyków i 13 998 mężczyzn fizyków (zaledwie 4-krotna różnica).

Akademickie światy młodych kobiet z dyscyplin wielkiej czwórki dziś i 20–30 lat temu są zadziwiająco różne, a dzisiejsze starsze kobiety są tymi młodymi kobietami sprzed kilkudziesięciu lat, które przetrwały w środowisku zdominowanym przez mężczyzn.

Tabela 5. Zbliżenie na liczbę młodych vs. starszych naukowców: dane zdezagregowane ze względu na płeć i grupy wieku akademickiego, rozkład publikujących naukowców z dziedziny STEMM według wybranych grup wieku akademickiego i płci, 2021 r.

Dyscyplina	Płeć	5 lat i mniej	6–10 lat	31–35 lat	36–40 lat
AGRI	Kobiety	9714	10 675	1238	647
	Mężczyźni	9652	11 913	3925	2702
BIO	Kobiety	21 139	23 394	3463	1757
	Mężczyźni	20 161	24 601	7692	4726
CHEM	Kobiety	6793	5601	693	380
	Mężczyźni	12 253	11 721	2792	1785
CHEMENG	Kobiety	377	330	32	15
	Mężczyźni	904	912	187	138
COMP	Kobiety	1049	1469	231	76
	Mężczyźni	4030	6130	1344	648
EARTH	Kobiety	2732	3631	534	335
	Mężczyźni	4820	6985	2686	1848
ENER	Kobiety	557	456	16	4
	Mężczyźni	1677	1730	125	66
ENG	Kobiety	2316	2429	198	84
	Mężczyźni	10 739	13 324	2362	1466
ENVIR	Kobiety	3807	3951	277	130
	Mężczyźni	4065	4734	1008	649
IMMU	Kobiety	1617	1653	249	104
	Mężczyźni	1122	1259	557	305
MATER	Kobiety	3397	2706	193	98
	Mężczyźni	7670	6891	927	520
MATH	Kobiety	829	1006	193	112
	Mężczyźni	2808	3684	1185	926
MED	Kobiety	86 100	83 904	9217	4005
	Mężczyźni	75 065	79 455	21 655	12 289
NEURO	Kobiety	3520	3880	369	227
	Mężczyźni	3000	3860	968	588
PHARM	Kobiety	985	756	128	62
	Mężczyźni	831	769	229	159
PHYS	Kobiety	3817	4034	705	396
	Mężczyźni	13 998	16 968	6040	3726
RAZEM	Kobiety	148 749	149 875	17 736	8432
	Mężczyźni	172 795	194 936	53 682	32 541

3.3. Wyniki: podejście pionowe

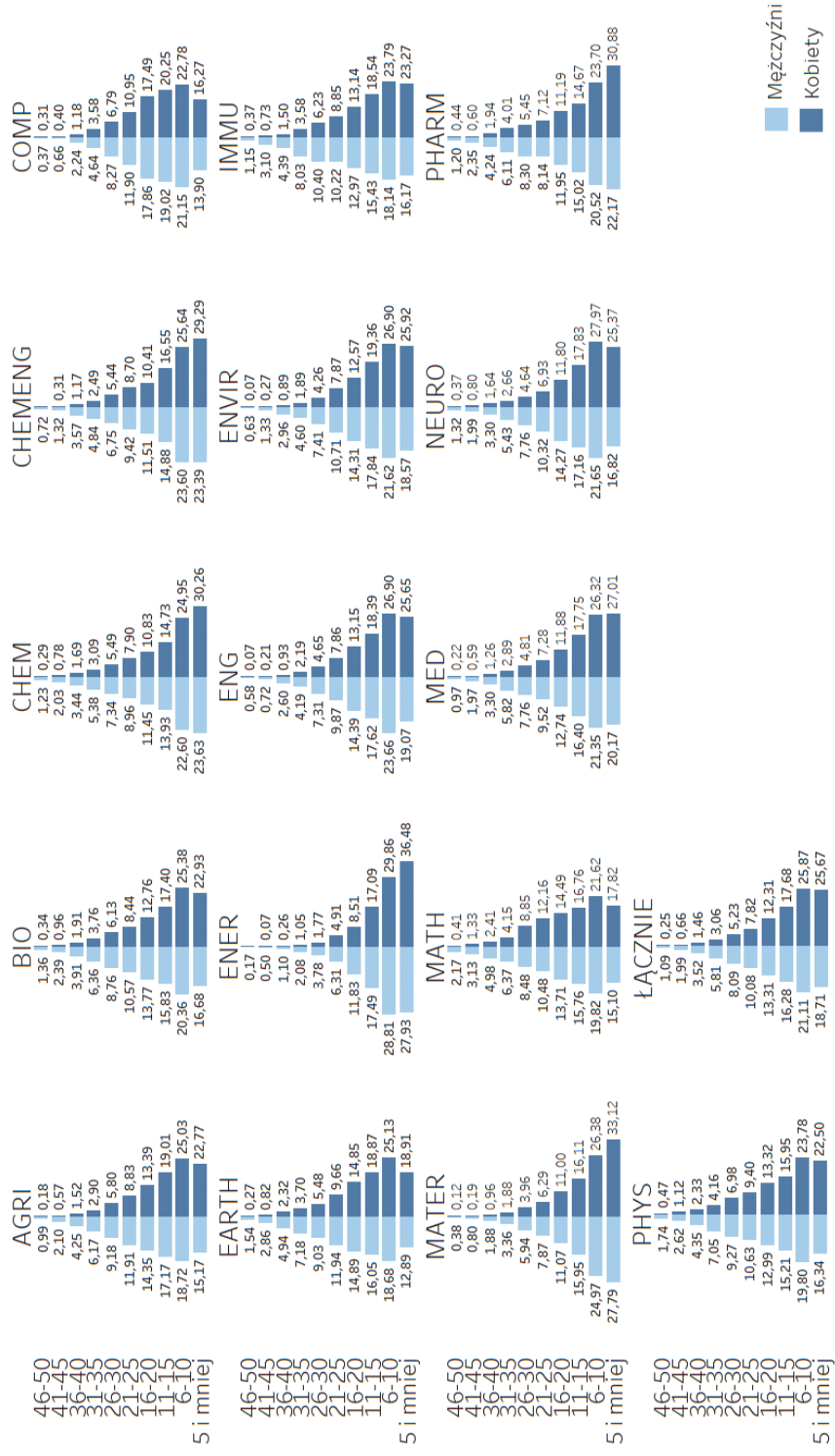
3.3.1. Ujęcie przekrojowe (2021): wszystkie grupy wieku akademickiego pionowo

Badając rozkład mężczyzn i kobiet w obrębie dyscyplin, stwierdziliśmy, że w większości dyscyplin (dziewięć) największa część kobiet znajdowała się w dwóch najmłodszych grupach wieku akademickiego. To znaczy w grupach z nie więcej niż 10-letnim doświadczeniem naukowym (ryc. 6). Młode kobiety dominowały (> 50%) wśród wszystkich kobiet w takich dyscyplinach, jak CHEM, ENG czy MED. W tych dyscyplinach napływ (publikujących) kobiet w ostatnich kilkunastu latach był ogromny. Najniższy odsetek młodych kobiet wśród wszystkich kobiet naukowców – czyli ich najmniejszy napływ (< 40%) – dotyczył dyscyplin COMP i MATH. We wszystkich dyscyplinach łącznie udział młodych kobiet wśród wszystkich kobiet naukowców wyniósł 51,54%, a udział młodych mężczyzn wśród wszystkich mężczyzn naukowców był znacznie niższy i wyniósł 39,82%. Wyłaniający się z naszych badań obraz wspiera narrację o coraz młodszych kobietach w nauce: spośród wszystkich kobiet obecnych aktualnie w globalnej nauce, ponad połowa ma nie więcej niż 10 lat doświadczenia w publikowaniu (patrz szczegóły na ryc. 7).

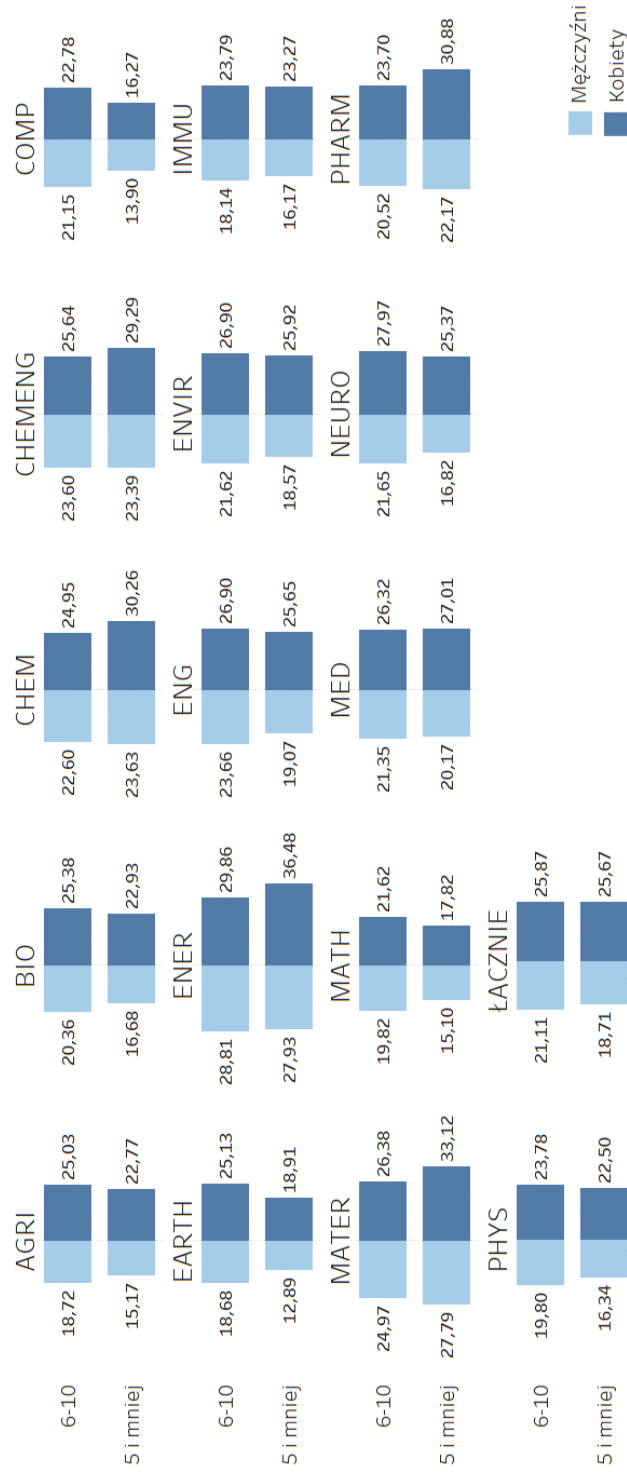
3.3.2. Porównawczy ujęcie pionowe (2000 vs. 2021)

W tej części pracy omawiamy zmiany zachodzące w piramidach wieku na przestrzeni dwóch dekad, porównując piramidy wieku naukowców z roku 2021 i 2000 w ujęciu płci. Piramida wieku składa się z połączonych wykresów słupkowych dla mężczyzn i kobiet, przy czym na osi pionowej przedstawiony jest wiek. W naszym przypadku, przypomnijmy, to wiek akademicki, czyli liczba lat, jaka upłynęła od pierwszej publikacji dowolnego typu indeksowanej w bazie Scopus. Piramidy wieku z 2021 roku (jasnoniebieskie) zostały nałożone na piramidy z 2000 roku (ciemnoniebieskie). Dla każdej z 10 grup wieku akademickiego w naszej populacji słupek odchodzący od osi w prawo przedstawia udział kobiet w tej grupie, a słupek znajdujący się po lewej stronie – pokazuje udział mężczyzn (zob. Wachter 2014: 218–221). Obie piramidy wieku obejmują inną populację (w każdym przypadku są naukowcy przychodzący i odchodzący); jednak niektóre kohorty naukowców są wspólne. Naukowcy włączeni do naszego badania publikowali w latach 1970–2021 (dla danych z 2021 roku) oraz w latach 1950–2000 (dla danych z 2000 roku).

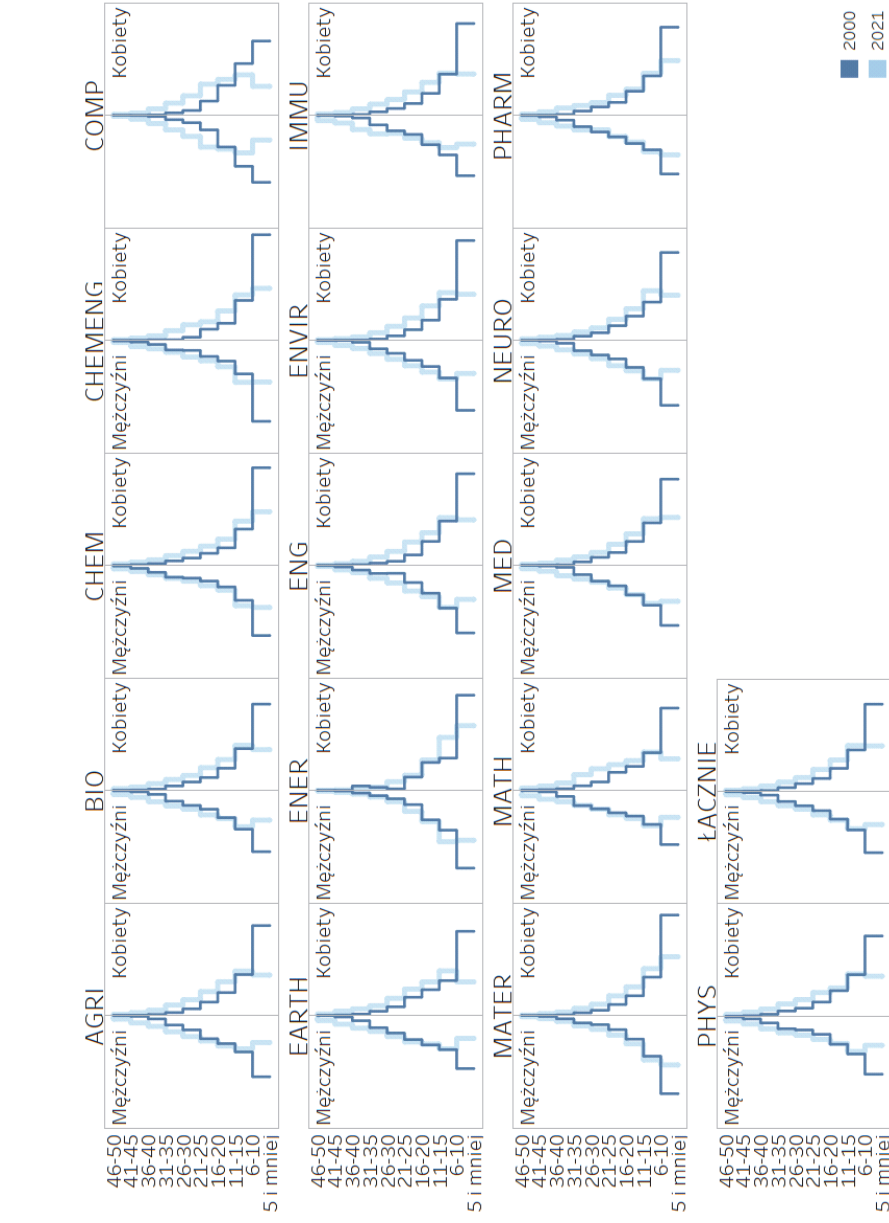
Na rycinie 8 przedstawiamy odsetki mężczyzn i kobiet naukowców wśród autorów publikujących w dwóch punktach czasowych (pomijając liczbę autorów) według dyscyplin. Korzystając z tych samych zasad doboru próby, podejście to pozwala nam porównać demografię naukowców i skupić się na młodych (i starszych) naukowcach. Rycina 8 pokazuje ujęcia z 2021 i 2000 roku według grup wieku akademickiego, dyscypliny i płci, wskazując na rozkład mężczyzn i kobiet naukowców w każdej dyscyplinie i ilustrując dynamikę zmian.



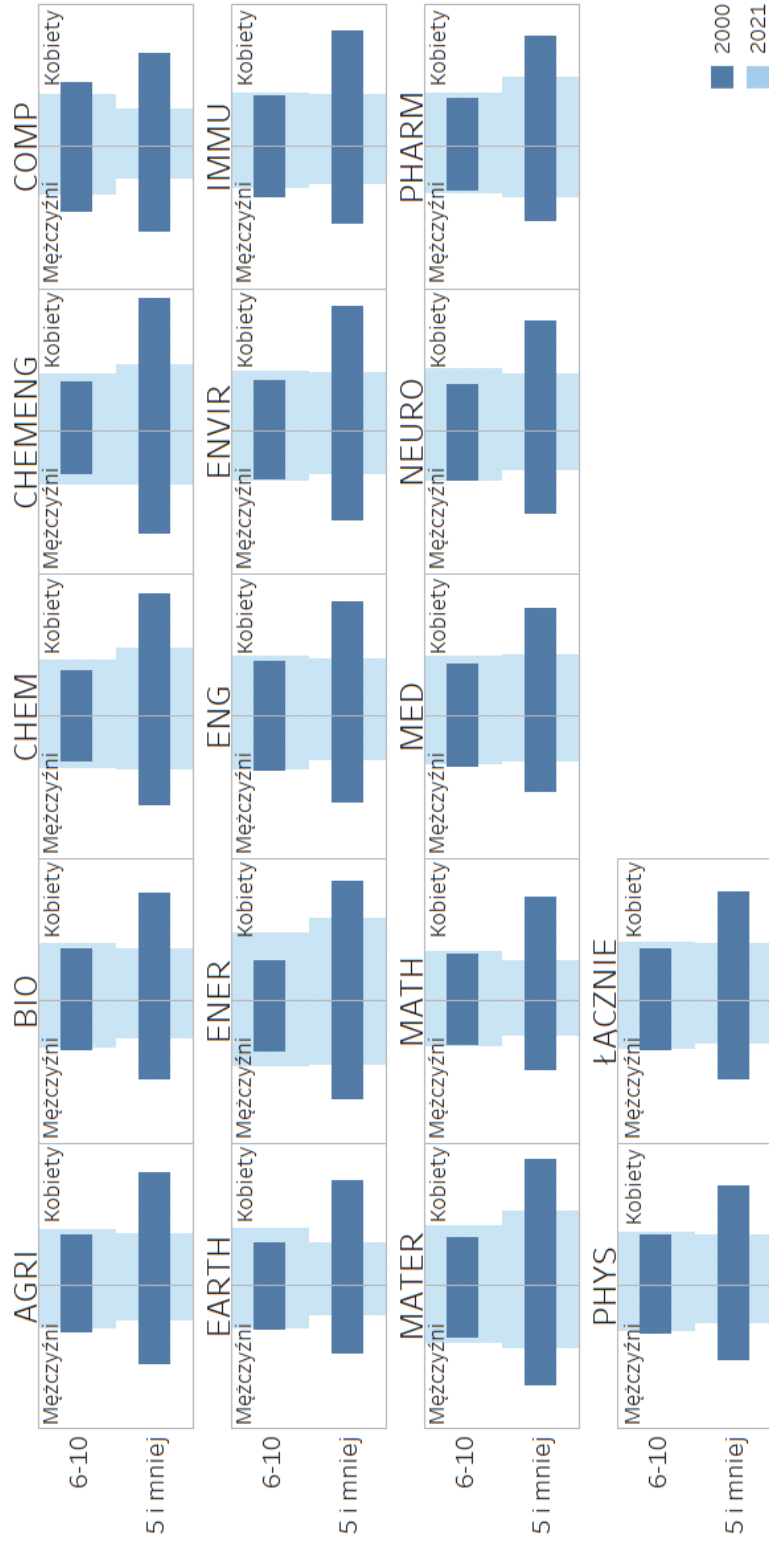
Ryc. 6. Młode kobiety w obszarze STEM: w większości dyscyplin większość kobiet należy do dwóch najmłodszych grup wieku akademickiego. Ujęcie pionowe: rozkład publikujących naukowców w obszarze STEM według dyscypliny, grupy wieku akademickiego i płci (procenty kolumnowe: 100% pionowo, dla wszystkich grup łącznie), 2021 ($N = 1502792$)



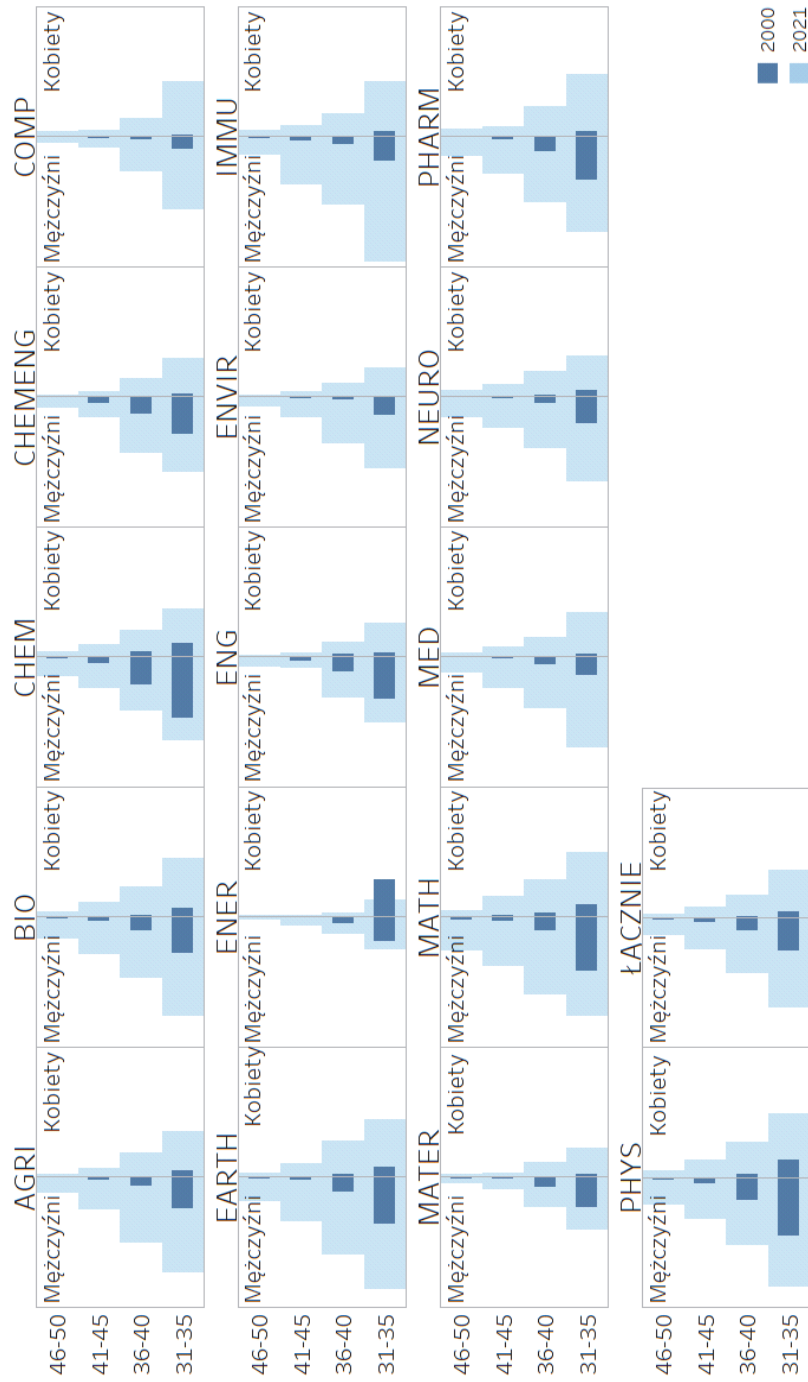
Ryc. 7. Zbliżenie tylko na młodych naukowców. Wyższa koncentracja młodych kobiet niż młodych mężczyzn we wszystkich dyscyplinach. Podejście pionowe: rozkład publikujących naukowców z dziedziny STEMM według dyscyplin, grup wieku akademickiego i płci (procenty kolumnowe, pionowo: odsetek młodych kobiet wśród wszystkich kobiet i młodych mężczyzn wśród wszystkich mężczyzn; kobiety w kolorze ciemnoniebieskim), 2021 ($N = 666\,355$)



Ryc. 8. Malejące odsetki najmłodszych mężczyzn i kobiet naukowców wśród wszystkich mężczyzn i kobiet naukowców w czasie, we wszystkich dyscyplinach. Podsumowanie kierunku zmian w procentach, 2000 vs. 2021: ujęcie pionowe. Rozkład publikujących naukowców z obszaru STEM według dyscypliny, grupy wieku akademickiego i płci (procenty kolumnowe: 100% pionowo dla wszystkich grup wieku akademickiego łącznie, ciemnoniebieski rok 2000, jasnoniebieski rok 2021) ($N_{2021} = 1\,502\,792$, $N_{2000} = 716\,796$)



Ryc. 9. Malejąca w czasie baza młodych naukowców, zarówno mężczyzn, jak i kobiet. Przegląd kierunku zmian procentowych, 2000 vs. 2021: ujęcie pionowe. Zbliżenie tylko na młodych naukowców (wiek akademicki 10 lat lub mniej). Rozkład młodych, publikujących naukowców w obszarze STEM według dyscypliny, grupy wieku akademickiego i płci, 2000 (kolor ciemnoniebieski) i 2021 (kolor jasnoniebieski) (w oparciu o procenty kolumnowe) ($N_{2021} = 666\,355$, $N_{2000} = 437\,113$)



Ryc. 10. Poszerzająca się w czasie baza starszych naukowców, zarówno mężczyzn, jak i kobiet. Przegląd kierunku zmian, 2000 vs. 2021: ujęcie pionowe. Zbliżenie tylko na starszych naukowców: wiek akademicki 31 – 50 lat. Rozkład starszych, publikujących naukowców z obszaru STEM według dyscypliny, grupy wieku akademickiego i płci, 2000 (kolor ciemnoniebieski) i 2021 (kolor jasnoniebieski) (w oparciu o procenty kolumnowe) ($N_{2000} = 17\,463$)

Podczas gdy w kolejnej części 3.4 wykorzystano analizę trendów, aby pokazać zmiany w odsetkach kobiet naukowców według dyscyplin, ta część dodaje do analizy wiek (akademicki).

Najogólniej rzecz ujmując, każda dyscyplina ma specyficzną strukturę w formie piramidy demograficznej, w której wiek biologiczny zastępowany jest wiekiem akademickim (czy zawodowym). W przypadku każdej dyscypliny piramida wieku w różnym stopniu zwęża się u góry i rozszerza u dołu. Podstawa piramidy odzwierciedla odsetek młodych naukowców wśród wszystkich naukowców, a jej wierzchołek reprezentuje odsetek starszych naukowców wśród wszystkich naukowców. Szersza podstawa wskazuje na wyższy odsetek młodych naukowców.

W 2021 r. we wszystkich dyscyplinach zaobserwowano podobny wzorzec: podstawa piramidy wieku (pierwsza grupa wieku akademickiego, 5 lat i mniej) była węższa w porównaniu z okresem sprzed dwóch dekad zarówno w przypadku mężczyzn, jak i kobiet naukowców. Udział młodych kobiet wśród wszystkich kobiet znacznie się obniżył w porównaniu z mniejszymi spadkami w przypadku młodych mężczyzn (zob. ryc. 9). Ten spadek może również wskazywać na to, że młode kobiety, które trafiły do środowiska nauki akademickiej dwie dekady wcześniej, pozostały w systemie w 2021 r., zwiększając obecność kobiet w starszych grupach wieku akademickiego. Kurcząca się podstawa piramidy u kobiet naukowców w 2021 r. w porównaniu z 2000 r. jest również widoczna dla wszystkich dyscyplin łącznie. Pod względem struktur wiekowych w demografii (Rowland, 2014, s. 98–107) nasze struktury z 2000 r. można klasyfikować jako „bardzo młode”, a struktury z 2021 r. jako „młode” lub „dojrzałe”.

Natomiast porównując odsetki starszych naukowców mężczyzn i kobiet w 2000 i 2021 r. w ramach dyscyplin (ryc. 10), wzorzec jest jednoznaczny: odsetki w przypadku obu płci w czterech starszych grupach wieku akademickiego były znacznie wyższe w 2021 r. niż w 2020 r. W każdej starszej kategorii naukowców, dla każdej dyscypliny, bez wyjątków, odnotowano wyższy odsetek starszych naukowców w 2021 r. niż w 2020 r., co wskazuje na kolejny wymiar starzenia się kadry naukowej (innym jest mediana wieku naukowców). Kadra naukowa w świecie starzeje się systemowo, wraz ze starzeniem większości społeczeństw – co jednak analizujemy szczegółowo na potrzeby innego badania.

3.4. Wyniki: udział kobiet naukowców według dyscyplin (trendy 1990–2021)

W tej części pracy, na koniec, analizujemy zmieniające się w czasie uczestnictwo kobiet w nauce, aby przetestować założenie, że napływ kobiet do nauki w ciągu ostatnich trzech dekad był silnie zróżnicowany i uzależniony od dyscypliny.

Liczba indywidualnych naukowców wykorzystanych tutaj do zbadania trendu w czasie wynosiła 4,3 mln (61,85% mężczyzn i 38,15% kobiet, tab. 1). Badaliśmy trend dotyczący odsetka kobiet naukowców obecnych w globalnej nauce (w postaci publikowania

artykułów naukowych indeksowanych w bazie Scopus) w latach 1990–2021. W naszej analizie wykorzystaliśmy trend liniowy w postaci $y = at + b$. W równaniu tym b oznacza miejsce przecięcia linii trendu przez oś y , natomiast a oznacza współczynnik nachylenia linii trendu. Współczynnik nachylenia opisuje jak stroma jest ta linia, wykorzystując wartość dodatnią lub ujemną. Wartość współczynnika nachylenia a wskazuje na średnią zmianę rok do roku, a wartość b wskazuje na poziom zjawiska w okresie zerowym (poprzedzającym pierwszy rok analizy).

W niektórych dyscyplinach uczestnictwo kobiet w nauce było wysokie przy silnym wzroście (MED i PHARM) lub wysokie przy słabym wzroście (BIO); natomiast w innych było ono niskie przy silnym wzroście (AGRI, CHEMENG). Wielka czwórka, czyli grupa czterech dyscyplin intensywnie zmatematyzowanych, charakteryzowała się niskim poziomem uczestnictwa kobiet i jego słabym wzrostem (COMP, ENG, MATH i PHYS). Dla wszystkich dyscyplin łącznie wzrost był znaczny – z 22,16% do 38,55%. Odsetek kobiet naukowców rósł we wszystkich dyscyplinach, choć w różnym tempie. Dyscypliny MATH, COMP, PHYS i ENG charakteryzowały się najniższym wzrostem, ze współczynnikami nachylenia równymi lub mniejszymi niż 0,33 (tab. 6). Wszystkie współczynniki nachylenia były dodatnie, wskazując na tendencję wzrostową odsetka kobiet naukowców we wszystkich dyscyplinach. Przedziały ufności dla współczynników nachylenia pozwoliły określić średnie tempo wzrostu w ciągu roku.

Każda dyscyplina miała inny czas wzrostu odsetka kobiet naukowców o jeden punkt procentowy. Najszybszy wzrost zanotowano w przypadku ENVIR (1,24 roku), AGRI (1,37) i MED (1,41). Dziewięć dyscyplin potrzebowało nieco więcej czasu (1,64–2,39 roku) na taki sam wzrost, natomiast najdłużej rósł udział kobiet w wielkiej czwórce MATH, COMP, PHYS i ENG – od 3,03 do 3,69 roku, tabela 7).

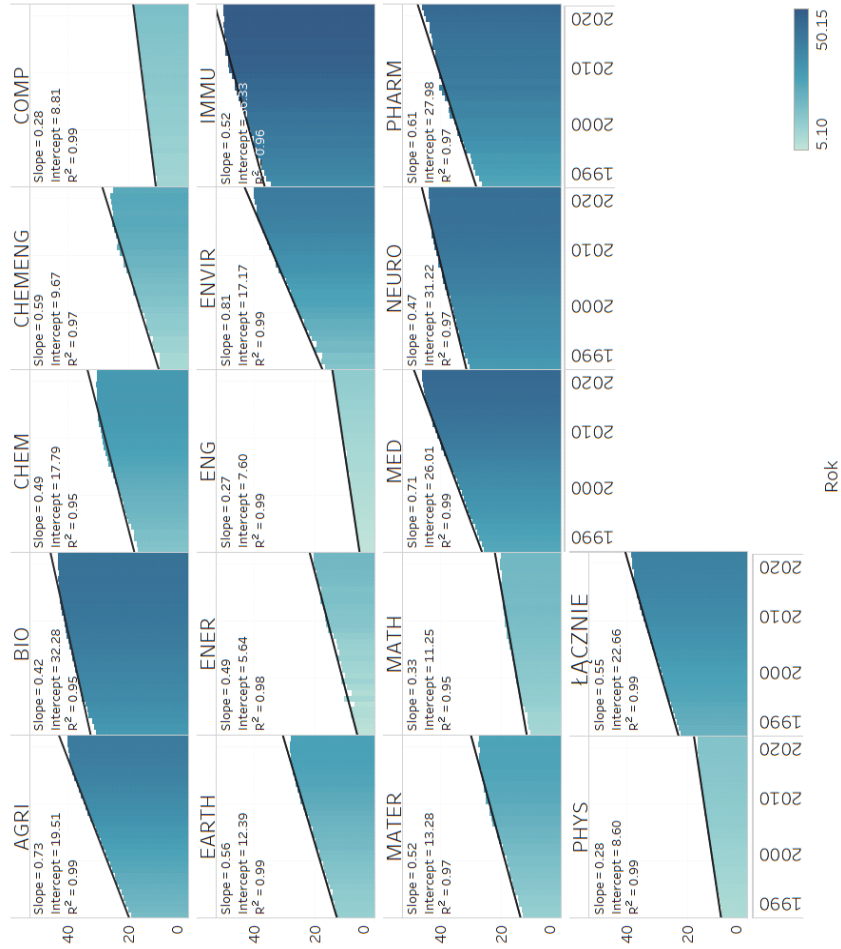
Hipotetycznie, przy stabilnych warunkach dostępu zawodowego do dyscyplin i przy obecnych trendach dotyczących uczestnictwa kobiet w nauce według dyscyplin wyznaczonych na podstawie ostatnich trzech dekad, z których żaden nie może być zagwarantowany w przyszłości, można oczekiwać, że równomierny rozkład mężczyzn i kobiet naukowców w ramach dyscypliny (czyli 50% kobiet i 50% mężczyzn) w czterech dyscyplinach zostanie osiągnięty za mniej więcej sto lat: po upływie 91 lat dla MATH (rok 2112), 113 lat dla COMP (rok 2134), 119 lat dla PHYS (rok 2140) i 134 lat dla ENG (rok 2155); we wszystkich pozostałych dyscyplinach taki równomierny rozkład naukowców może zostać osiągnięty między 2027 i 2028 rokiem (PHARM i CHEMENG) a 2081 rokiem (ENER). Jediną dyscypliną, w której rozkład ten został już osiągnięty, jest IMMU (zob. szczegóły w tabeli 7). Aby obliczyć datę jego osiągnięcia dla dowolnej dyscypliny, wykorzystaliśmy punkty procentowe brakujące do osiągnięcia poziomu 50% z tabeli 3 i brakującą liczbę lat pomnożyliśmy przez wyznaczony przez nas średni czas potrzebny do osiągnięcia zmiany o 1 p.p.

Tabela 6. Statystyki modelu regresji: trendy dotyczące odsetka kobiet naukowców według dyscyplin, 1990–2021

Dyscyplina	Współczynnik kierunkowy <i>a</i> (nachylenie)						Współczynnik kierunkowy <i>b</i> (przecięcie)				Miary jakości	
	Wartość współczynnika	Błąd standardowy	<i>t</i> -wartość	<i>p</i> -wartość	95% przedział ufności – dolny	95% przedział ufności – górny	Wartość współczynnika	Błąd standardowy	<i>t</i> -wartość	<i>p</i> -wartość	R ²	Błąd standardowy
AGRI	0,73	0,012	60,513	< 0,0001	0,704	0,753	19,51	0,217	89,842	< 0,0001	0,992	0,629
BIO	0,42	0,018	23,611	< 0,0001	0,381	0,454	32,28	0,319	101,175	< 0,0001	0,949	0,924
CHEM	0,49	0,021	22,715	< 0,0001	0,443	0,531	17,79	0,387	46,011	< 0,0001	0,945	1,120
CHEMENG	0,59	0,019	31,706	< 0,0001	0,549	0,624	9,67	0,334	28,983	< 0,0001	0,971	0,966
COMP	0,28	0,005	55,445	< 0,0001	0,272	0,292	8,81	0,092	96,007	< 0,0001	0,990	0,266
EARTH	0,56	0,012	47,946	< 0,0001	0,538	0,586	12,39	0,211	58,632	< 0,0001	0,987	0,612
ENER	0,49	0,014	34,941	< 0,0001	0,466	0,524	5,64	0,255	22,089	< 0,0001	0,976	0,739
ENG	0,27	0,004	67,841	< 0,0001	0,263	0,279	7,60	0,072	105,555	< 0,0001	0,994	0,209
ENVIR	0,81	0,015	53,321	< 0,0001	0,778	0,840	17,17	0,274	62,716	< 0,0001	0,990	0,793
IMMU	0,52	0,019	26,908	< 0,0001	0,481	0,560	36,33	0,349	104,162	< 0,0001	0,960	1,010
MATER	0,52	0,016	32,913	< 0,0001	0,483	0,547	13,28	0,282	47,001	< 0,0001	0,973	0,818
MATH	0,33	0,014	23,208	< 0,0001	0,302	0,361	11,25	0,258	43,670	< 0,0001	0,947	0,746
MED	0,71	0,012	57,674	< 0,0001	0,684	0,734	26,01	0,222	117,338	< 0,0001	0,991	0,642
NEURO	0,47	0,014	32,908	< 0,0001	0,436	0,494	31,2	0,255	122,402	< 0,0001	0,973	0,738
PHARM	0,61	0,020	30,966	< 0,0001	0,568	0,649	27,98	0,355	78,940	< 0,0001	0,970	1,026
PHYS	0,28	0,006	47,373	< 0,0001	0,269	0,294	8,60	0,107	80,276	< 0,0001	0,987	0,310
RAZEM	0,55	0,010	56,971	< 0,0001	0,535	0,574	22,66	0,176	129,059	< 0,0001	0,991	0,508

Tabela 7. Trendy dotyczące odsetka kobiet naukowców według dyscyplin (współczynnik nachylenia, przecięcia i tempo zmian), 1990–2021

Dyscyplina	Współczynnik kierunkowy <i>a</i> (nachylenie)	Współczynnik kierunkowy <i>b</i> (przecięcie)	Czas potrzebny do zmiany o 1 p.p. (w latach)	Czas potrzebny do osiągnięcia równości płci (kobiety 50%) w latach (oraz data)	Czas potrzebny do osiągnięcia równowagi płci (kobiety 40%) w latach (oraz data)
ENVIR	0,81	17,17	1,24	13,5 (2035)	0 (osiągnięto)
AGRI	0,73	19,51	1,37	16,1 (2038)	0 (osiągnięto)
MED	0,71	26,01	1,41	40,6 (2062)	0 (osiągnięto)
PHARM	0,61	27,98	1,64	5,4 (2027)	0 (osiągnięto)
CHEMENG	0,59	9,67	1,70	6,6 (2028)	25,5 (2047)
EARTH	0,56	12,39	1,78	39,4 (2061)	21,6 (2043)
IMMU	0,52	36,33	1,92	0 (osiągnięto)	0 (osiągnięto)
MATER	0,52	13,28	1,94	44,4 (2066)	25,0 (2046)
ENER	0,49	5,64	2,02	60,0 (2081)	39,8 (2061)
CHEM	0,49	17,79	2,05	40,6 (2062)	20,1 (2042)
NEURO	0,47	31,22	2,15	13,4 (2035)	0 (osiągnięto)
BIO	0,42	32,28	2,39	16,1 (2038)	0 (osiągnięto)
MATH	0,33	11,25	3,02	90,5 (2112)	60,3 (2081)
COMP	0,28	8,81	3,55	112,9 (2134)	77,4 (2099)
PHYS	0,28	8,60	3,55	118,5 (2140)	83,3 (2105)
ENG	0,27	7,60	3,69	133,5 (2155)	96,6 (2118)
RAZEM	0,55	22,66	1,82	–	–



Ryc. 11. Różne punkty wyjścia wzrostu poziomu uczestnictwa kobiet w nauce w czasie.
Trend dotyczący odsetka kobiet naukowców według dyscyplin, 1990–2021 ($N = 4\,314\,666$)

Zamiast rozkładu (50%/50%) dla wszystkich, możemy skupić się również na takim rozkładzie tylko dla najmłodszych pokoleń naukowców, przeliczając wyniki dla tej grupy wiekowej. Równomierność rozkładu w takim ujęciu została już osiągnięta w sześciu dyscyplinach (np. AGRI, BIO i MED, patrz ryc. 4) i niemal osiągnięta we wszystkich dyscyplinach łącznie. Ponadto zamiast równomiernego rozkładu możemy przyjąć również podejście alternatywne: równowagę, która odnosi się do obecności mężczyzn i kobiet w nauce i mieści się w przedziale od 40% do 60% (EC 2021: 20). Przeliczyliśmy wyniki dla równowagi między kobietami i mężczyznami dla wszystkich naukowców, ze znacznie krótszymi okresami, w których można ją osiągnąć: w siedmiu dyscyplinach równowaga ta została już osiągnięta (tabela 8). Analityka predykcyjna wykracza jednak poza nasz obszar zainteresowań.

4. Podsumowanie, dyskusja i wnioski

Przeanalizowaliśmy zmieniającą się demografię globalnej kadry naukowej w ciągu ostatnich trzech dekad (zdefiniowanej w tym badaniu jako naukowcy należący do obszaru STEMM, pochodzący z 38 krajów OECD, posiadający status nieokazjonalny w nauce, czyli posiadający co najmniej trzy artykuły indeksowane w bazie Scopus), ze szczególnym uwzględnieniem zmieniającego się uczestnictwa w nauce młodych kobiet i mężczyzn.

Nasze badanie cechowała duża skala (4,3 mln naukowców); pokoleniowość (naukowcy zostali przydzieleni do 10 grup wieku akademickiego, z zasadniczym podziałem na młodą kohortę, doświadczenie akademickie 10 lub mniej lat, i starszą kohortę, 31–50 lat); oraz podejście zarówno przekrojowe (rok 2021), jak i wzdłużne, czyli dynamiczne (obejmujące lata 1990–2021 i rok 2000 vs. rok 2021).

Aby kompleksowo przeanalizować cztery wymiary (płeć, wiek, dyscyplina i czas), zastosowaliśmy dwa podejścia: w tym, co nazwaliśmy podejściem poziomym, skupiliśmy się na rozkładzie naukowców pod względem płci w tych samych grupach wieku akademickiego w różnych dyscyplinach; a w tym, co nazwaliśmy podejściem pionowym, zajęliśmy się koncentracją kobiet i mężczyzn naukowców osobno w grupach wieku akademickiego i w ramach dyscyplin.

Naszym zasadniczym wyborem metodologicznym było użycie indywidualnych naukowców (z ich atrybutami), a nie indywidualnych publikacji (z ich cechami) jako jednostki analizy. Użyliśmy mikrodanych na poziomie indywidualnym wyekstrahowanych z surowych danych z bazy Scopus, ponieważ nasze badania w dużym stopniu opierały się na identyfikatorach autorów, a baza Scopus zapewnia dane bibliometryczne z precyzją na poziomie 98,1% i przywołaniem na poziomie 94,4% (Baas et al., 2020).

Nasze badanie ma charakter ilościowy i eksploracyjny: zadajemy pytania „co”, na tym etapie nie pytając „dlaczego”. Z tego powodu prezentowane wyniki można uzupełnić o kolejne badania ilościowe na małą skalę (oparte na globalnych i krajowych danych

ankietowych) oraz badania jakościowe oparte na metodologii wywiadów pogłębionych i grup fokusowych (jak sugeruje Fox 2020 w badaniach kobiet i stopni akademickich).

Nie dysponujemy wiedzą na temat podobnego badania polegającego na mapowaniu rozlokowania młodych mężczyzn i kobiet naukowców w różnych dyscyplinach w kontekście młodszych i starszych grup wieku akademickiego (pod względem doświadczenia akademickiego czy zawodowego). Choć raporty statystyczne dostarczające danych na temat mężczyzn i kobiet w nauce są niezwykle przydatne, nie wydają się stanowić części globalnej konwersacji naukowej na temat obecności kobiet w nauce.

Nasze badanie nie testuje różnych hipotez dotyczących nierówności między kobietami i mężczyznami w nauce, ponieważ ma charakter eksploracyjny; jednak nasze ustalenia ogólnie wspierają podstawy teoretyczne omówione we Wprowadzeniu. Poziom odejścia kobiet z nauki jest wysoki (teoria „przeciekającego rurociągu”), wyraźnie istnieją dyscypliny, które – z różnych powodów – nie są przyjazne dla kobiet (hipoteza „chłodnego klimatu”) i w których struktura pokoleniowa kadry naukowej nie ulega zmianie (hipoteza „autoselekcji”, patrz dyscypliny zmatematyzowanej wielkiej czwórki i jej stabilny rozkład wieku i płci w czasie).

Kadra naukowa zmienia się pod względem struktury płci i wieku z różną intensywnością w różnych dyscyplinach. Zmiany te mają charakter ciągły i globalny. Wśród 16 dyscyplin STEMM większość jest obecnie zdominowana liczbowo przez mężczyzn, ale w niektórych już teraz dominują kobiety, a procesy zmian są szybkie w niektórych i powolne w innych dyscyplinach.

Nieco zaskakuje, nawet w kontekście pandemii COVID-19, fundamentalna rola badań medycznych dla globalnej kadry naukowej, zwłaszcza w przypadku kobiet naukowców: prawie połowa wszystkich naukowców w obszarze STEMM (45,98%) została zdefiniowana w naszej metodologii jako kadra zajmująca się badaniami medycznymi (medycyna jako dyscyplina dominująca, oparta na wszystkich cytowanych odwołaniach bibliograficznych w publikacjach z całego życia każdego naukowca). Koncentracja kobiet naukowców jest wyraźna w poszczególnych dyscyplinach: ponad połowa (55,02%) z nich jest zlokalizowana w MED, a co siódma (15,91%) w BIO. Tym samym około 70% (70,93%) wszystkich kobiet naukowców na świecie, we wszystkich dyscyplinach z obszaru nauk STEMM, skupia się właśnie w tych dwóch dyscyplinach.

Tradycyjna wizja wybranych dyscyplin STEMM jako silnie zdominowanych przez mężczyzn została zweryfikowana w skali globalnej: obecność kobiet w COMP, ENG, MATH i PHYS jest niezwykle ograniczona (poniżej 20% w 2021 r.). W większości dyscyplin w 2021 r. udział kobiet w każdej kolejnej młodszej kohorcie był wyższy (i był zwykle najwyższy dla najmłodszej kohorty: naukowców z 5-letnim lub krótszym doświadczeniem akademickim); natomiast w przypadku COMP, ENG, MATH i PHYS zasada ta nie obowiązuje, a różnice między kohortami są marginalne (ryc. 3).

Nasza analiza trendów w ramach poszczególnych dyscyplin w latach 1990–2021 wykazała, że obecność kobiet naukowców w globalnej nauce rośnie we wszystkich dyscyplinach, choć z różnymi punktami wyjścia w 1990 r. i z różną intensywnością, zgodnie z wcześniejszymi badaniami dotyczącymi problematyki „kobiet w nauce”. W przypadku najmniej nasilonych tendencji wzrost odsetka kobiet naukowców o jeden punkt procentowy średnio zajął 3,03 roku w przypadku MATH, 3,55 w przypadku COMP i PHYS oraz 3,69 roku w przypadku ENG.

Hipotetycznie można zatem oczekiwać, że równomierny rozkład mężczyzn i kobiet w ramach dyscypliny (50% kobiet naukowców, 50% mężczyzn naukowców) w tych czterech dyscyplinach zostanie osiągnięty za mniej więcej sto lat: w przypadku MATH w roku 2112, w przypadku COMP w roku 2134, w przypadku PHYS w roku 2140, a w przypadku ENG w roku 2155; we wszystkich pozostałych dyscyplinach równomierny rozkład może zostać osiągnięty między rokiem 2027 i 2028 (PHARM i CHEMENG) i 2081 (ENER). Natomiast w przypadku mniej restrykcyjnego podejścia równowaga (definiowana jako 40% kobiet naukowców, 60% mężczyzn naukowców) została już osiągnięta w siedmiu dyscyplinach (szczegóły w tabeli 7).

Jednak z perspektywy zdezagregowanej ze względu na wiek, w 6 z 16 dyscyplin odnotowano już więcej najmłodszych kobiet niż mężczyzn naukowców (IMMU, PHARM, NEURO, MED, AGRI, BIO), a dyscypliną najbardziej otwartą na kobiety jest IMMU (59,04%). Co ciekawe, 8 na 10 kobiet naukowców z obszaru STEMM na świecie pracuje właśnie w tych sześciu dyscyplinach (82,90%). We wszystkich dyscyplinach z obszaru STEMM łącznie większość kobiet obecnie zaangażowanych w publikowanie artykułów stanowią kobiety młode (z 10-letnim lub krótszym doświadczeniem akademickim).

Co najciekawsze, we wszystkich dyscyplinach z obszaru STEMM stwierdzono wyższą koncentrację młodych kobiet niż młodych mężczyzn i wyższą koncentrację starszych mężczyzn niż starszych kobiet. W każdej dyscyplinie udział młodych kobiet naukowców wśród wszystkich kobiet naukowców w danej dyscyplinie był wyższy niż udział młodych mężczyzn naukowców wśród wszystkich mężczyzn naukowców (korzystając z ujęcia pionowego). W każdej dyscyplinie udział starszych mężczyzn wśród wszystkich mężczyzn naukowców w danej dyscyplinie był znacznie wyższy niż udział starszych kobiet naukowców wśród wszystkich kobiet naukowców. Prawidłowości są jednoznaczne: we wszystkich dyscyplinach kobiety naukowcy są zazwyczaj młodsze, a mężczyźni – starsi.

Przechodząc od standardowych danych do danych zdezagregowanych ze względu na płeć dla poszczególnych grup wieku akademickiego, zaczynamy rozumieć, co oznacza globalna izolacja kobiet naukowców w takich dyscyplinach jak MATH, PHYS czy ENG. W tych dyscyplinach w 2021 roku udział starszych kobiet naukowców wynosił około 10%

lub mniej (czyli różnica ilościowa między kobietami i mężczyznami była dziesięciokrotna lub większa, np. ENG, MATH i PHYS: odpowiednio 6,31%, 11,09% i 9,21%). W starszych pokoleniach kobiety były samotnymi jednostkami wśród swoich kolegów mężczyzn w podobnym wieku. Liczby pokazują więcej niż odsetki (tabela 5): na przykład, w grupie wieku akademickiego 36–40 lat globalnie 84 kobiety pracowały obok 1466 mężczyzn w ENG i 396 kobiet pracowało obok 3726 mężczyzn w PHYS.

Ważny jest jednak kontekst zmieniających się czasów: dla tych samych trzech dyscyplin ENG, MATH i PHYS izolacja młodych kobiet naukowców znacznie się zmniejszyła, z 10-krotnej różnicy w przypadku starszych kohort do 5-krotnej różnicy w przypadku kohort młodych (tj. do 16,47% w przypadku ENG, 22,04% w przypadku MATH i 20,23% w przypadku PHYS). W tych trzech zdominowanych przez mężczyzn dyscyplinach w 2021 roku kobiety w młodych kohortach były co najmniej dwukrotnie bardziej reprezentowane niż kobiety w starszych kohortach.

Zmiana struktury uczestnictwa kobiet i mężczyzn w nauce przebiegała stopniowo, ale jej wzorzec był wyraźny: we wszystkich dyscyplinach z obszaru STEMM, zarówno tych silnie zdominowanych przez mężczyzn, jak i tych najbardziej zbliżonych do równego rozkładu, młodsze pokolenia zasadniczo zawsze charakteryzowały się większą liczbą kobiet naukowców i ich wyższym odsetkiem niż starsze pokolenia. Kobiety naukowcy są silniej obecne, przechodząc od grupy starszych naukowców do grupy młodych naukowców. Z perspektywy longitudinalnej (dynamicznej), dla wszystkich dyscyplin, udział naukowców w najmłodszej grupie wieku akademickiego w 2000 r. był wyższy niż w 2021 r. zarówno dla mężczyzn, jak i dla kobiet. Zmniejszyła się baza młodych naukowców, zarówno mężczyzn, jak i kobiet, a zwiększyła się baza starszych naukowców, zarówno mężczyzn, jak i kobiet.

Szersze podsumowanie naszych badań jest następujące: nie istnieją inne źródła danych niż źródła bibliometryczne, które pozwoliłyby na oszacowanie globalnego (a nie tylko krajowego) rozkładu populacji naukowców pod względem płci, dyscypliny i grupy wiekowej, zarówno w ujęciu przekrojowym, jak i dynamicznym.

Zmiana punktu ciężkości z publikacji na poszczególnych naukowców otwiera fundamentalnie nowe perspektywy przed badaniami karier naukowych, choć pojawiają się nowe ograniczenia. Nowa wiedza ma swoją cenę, którą należy poddać analizie. Heurystycznie użyteczne jest traktowanie globalnych zbiorów danych bibliometrycznych jako „ustrukturyzowanych” (w przeciwieństwie do „nieustrukturyzowanych” i „półustrukturyzowanych”) Big Data, o dużych rozmiarach i złożoności, w przypadku których wymagane są nowe techniki algorytmiczne w celu wydobycia użytecznych informacji (Holmes, 2017). W naszym przypadku oznaczało to przykładowo ustalenie dominującej dyscypliny dla wszystkich naukowców w oparciu o 1,43 mld cytowanych pozycji bibliograficznych. Analiza naszych podstawowych zmiennych obejmowała ponad 4 mld

danych o różnym charakterze, a każdego z ponad 80 mln naukowców charakteryzowało kilkadziesiąt podstawowych cech (zob. Kwiek, Szymula, 2023).

Większość ograniczeń bibliometrycznych zbiorów danych jest analizowana od lat (język angielski i koncentracja na obszarze STEMM, skrzywienie anglosaskie, ujmowanie niemal wyłącznie artykułów i pomijanie książek itp.; zob. Sugimoto & Larivière, 2018, s. 38–44, na temat „kulturowej stronniczości źródeł danych”). Jednak nasze wykorzystanie bibliometrycznego zbioru danych do określenia indywidualnych właściwości globalnej kadry naukowej wymaga krótkiego omówienia nowych ograniczeń:

- (1) Ustalenie płci: zastosowano podejście binarne z różnym pokryciem dla różnych krajów, ponieważ algorytmy stosowane w bazie Scopus (i inne narzędzia do określania płci, takie jak np. genderize.io czy Gender Guesser, zob. Halevi, 2019, s. 566; Mihaljević & Santamaría, 2020: 1477–1478) działają znacznie lepiej w przypadku wybranych krajów; wszystkie obserwacje bez rozpoznania płci zostały usunięte z naszej analizy.
- (2) Ustalenie dyscypliny: wykorzystano komercyjną klasyfikację czasopism akademickich jako przybliżenie bogactwa krajowych dyscyplin. Wykorzystano historię wszystkich publikacji indeksowanych w bazie Scopus każdego naukowca do określenia pojedynczego atrybutu, czyli jego dyscypliny (użyto pojedynczej dominującej wartości, potencjalnie tłumiącej przechodzenie między dyscyplinami w czasie).
- (3) Ustalenie kraju afiliacji: pojedyncza dominująca wartość, potencjalnie tłumiąca indywidualne historie migracyjne w okresie całego życia.
- (4) Ustalenie nieokazjonalnego statusu w nauce: próg trzech artykułów indeksowanych w bazie Scopus jako warunek wejścia do populacji był arbitralny, co rodzi niedocenywanie roli naukowców na bardzo wczesnych etapach kariery akademickiej; wyższy próg zmniejszyłby tę populację, a niższy ją zwiększył.
- (5) Ustalenie wieku akademickiego: chociaż korelacja między wiekiem biologicznym a wiekiem akademickim w dyscyplinach z obszaru STEMM jest duża (i wyższa niż 0,9, jak wykazaliśmy dla próby 20 000 polskich naukowców ze stopniem doktora; Kwiek & Roszka, 2022b), pierwsze publikacje w indywidualnych historiach publikacyjnych mogą pojawiać się w różnych momentach życia akademickiego w różnych dyscyplinach; ponadto wzorce publikowania wyraźnie zmieniają się w czasie (naukowcy rozpoczynają publikowanie wcześniej w swojej karierze dzisiaj niż kilkadziesiąt lat temu).

Kolejny wniosek to wyraźne różnice między badaniami przeprowadzanymi na poziomie krajowym, zwłaszcza gdy dane bibliometryczne są połączone z danymi administracyjnymi i biograficznymi, a globalnym badaniem kadry naukowej. Krótko mówiąc, badania krajowe mogą wykorzystywać komercyjne i niekomercyjne zbiory danych dostępne tylko dla kilku krajów (np. USA, Norwegia, Polska i Włochy: patrz Savage &

Olejniczak, 2021; Abramo et al., 2022; Abramo et al., 2016), które mogą zawierać informacje biograficzne bezpośrednio niedostępne na poziomie globalnym, takie jak płeć, data urodzenia, daty uzyskania doktoratu i innych stopni i tytułów naukowych, krajowe klasyfikacje dyscyplin czy pełna historia zatrudnienia.

W naszych dwóch ostatnich badaniach wzdłużnych (w wąskim sensie) zmieniających się klas produktywności 2326 profesorów tytularnych w ciągu 20–40 lat ich kariery (Kwiek & Roszka, 2023) oraz wpływu wczesnych i późnych, a także szybkich i powolnych awansów na produktywność na próbie 16 000 naukowców z obszaru STEMM (Kwiek & Roszka, 2022d), nasz zbiór danych obejmujący około miliona polskich publikacji indeksowanych w bazie Scopus z ostatnich 50 lat został wzbogacony o pełne dane biograficzne i administracyjne z rejestru 100 000 polskich naukowców (pochodzącego z OPI PIB).

W badaniach globalnych, w przeciwieństwie do badań prowadzonych na poziomie krajowym, wiek biologiczny musi być badany poprzez przybliżenie w postaci wieku akademickiego czy zawodowego, płeć musi być określana za pomocą progów prawdopodobieństwa, stopnie akademickie muszą być używane poprzez przybliżenie w postaci długości kariery od czasu pierwszej publikacji, a krajowe rankingi prestiżu akademickiego muszą być zastępowane przez przybliżenia w postaci rankingów globalnych typu Leiden Ranking czy ARWU.

Wszystkich naukowców zarejestrowanych w danym kraju trzeba zastąpić w badaniach globalnych naukowcami publikującymi, czyli posiadającymi publikacje indeksowane w bazie Scopus (czy Web of Science). Realnych naukowców posiadających krajowe numery identyfikacyjne dostępne w krajowych bazach danych trzeba zastępować identyfikatorami autorów zamieszczonymi w bazie Scopus, a niemal doskonale dane administracyjne i biograficzne trzeba zastępować danymi domyślnymi, wywnioskowanymi czy przybliżonymi. Niemniej jednak globalne badania eksploracyjne, tymczasowo mapujące teren i testujące najlepsze narzędzia i metodologie, są interesujące w swojej ogólności, zanim pojawią się kolejne, bardziej zaawansowane analizy.

Implikacje naszych badań są wielorakie. W aspekcie naukowym podejmujemy pierwszą próbę zdefiniowania i opisanego społeczności naukowej w skali globalnej za pomocą indywidualnych atrybutów, które do tej pory nie były wykorzystywane na dużą skalę. Odzworowanie zmieniającego się w czasie rozkładu populacji naukowców ze względu na płeć i wiek, jak również wgląd w dzisiejszą globalną kadrę naukową w jednym, wybranym roku, otwiera badania profesji naukowej (i profesji akademickiej) na dalsze, bardziej szczegółowe pytania. Kadra naukowa jest często omawiana w dwóch kontekstach polityki publicznej: starzenia się i towarzyszących mu problemów dla systemów szkolnictwa wyższego i innowacji oraz poszerzania dostępu do kariery naukowej dla młodych badaczy. Nasze metodologiczne podejście i szerokie ustalenia mogą być przy-

datne w badaniach złożonej kwestii wchodzenia do profesji akademickiej i odchodzenia z niej, wraz z towarzyszącymi pytaniami o zmieniającą się produktywność w cyklach życia naukowców czy o starzenie się i zmieniające się wzorce publikowania i współpracy w nauce etc. (szczególnie w sektorze akademickim).

Nasze badania mogą być użyteczne dla ustawodawców i decydentów, administratorów uczelni i dużych organizacji grantowych, pokazując, gdzie kadra naukowa koncentruje swoje wysiłki badawcze, jak duże jej segmenty są zaangażowane w badania w poszczególnych dyscyplinach oraz gdzie dyscyplinarnie ulokowani są młodzi naukowcy, mężczyźni i kobiety. Nasze analizy dotyczące znaczących różnic obecności między mężczyznami i kobietami w różnych dyscyplinach z obszaru STEMM (a zwłaszcza między ENG, COMP, MATH i PHYS a pozostałymi dyscyplinami) mogą dostarczyć nowych podstaw empirycznych przydatnych w dyskusjach na temat obecności kobiet w nauce i związanych z tym ograniczeń społecznych i instytucjonalnych w ramach poszczególnych dyscyplin. Z pewnością możemy dzisiaj wiedzieć radykalnie więcej o globalnej nauce i globalnych naukowcach (zob. monografię Kwiek 2022), w tym o młodej kadrze akademickiej, niż jeszcze kilka lat temu, co pozwala nam na szersze kontekstualizowanie polskich wyzwań w polityce naukowej.

Podziękowania

Marek Kwiek wyraża wdzięczność za uwagi gospodarzy i słuchaczy trzech zaproszonych seminariów poświęconych zagadnieniu globalnych analiz naukowców: (1) seminarium badawcze RISIS (Research Infrastructure for Science and Innovation Policy Studies), Uniwersytet w Lugano, Szwajcaria (11 maja 2022); (2) seminarium METRICS (Meta-Research Innovation Center), Uniwersytet Stanforda (23 lutego 2023); oraz (3) seminarium CGHE (Center for Global Higher Education), Uniwersytet w Oksfordzie (4 kwietnia 2023). Ponadto dziękuje za możliwość prezentacji i dyskusji w ramach konferencji „Doskonałość naukowa nie ma płci” zorganizowanej przez Akademię Młodych Uczonych PAN oraz UAM w Poznaniu (10 lutego 2023). Dziękujemy również za pomoc International Center for the Studies of Research (ICSR) Lab i Kristy James, Senior Data Scientist. Chcemy również podziękować dr. Wojciechowi Roszce za owocne dyskusje i wsparcie merytoryczne. Ponadto jesteśmy wdzięczni za wsparcie udzielone w ramach grantu MEiN NdS/529032/2021/2021.

Bibliografia

- Abramo G., Aksnes D.W., D'Angelo, C.A. (2021). Gender differences in research performance within and between countries: Italy vs. Norway. *Journal of Informetrics*, 15(2), 101144.
- Abramo G., D'Angelo C.A., Murgia G. (2016). The combined effects of age and seniority on research performance of full professors. *Science and Public Policy*, 43(3), 301–319.

- Arimoto A., Cummings W.K., Huang F., Shin J.C. (2015). *The changing academic profession in Japan*. Dordrecht: Springer.
- Baas J., Schotten M., Plume A., Côté G., Karimi R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377–386. 10.1162/qss_a_00019
- Boekhout H., van der Weijden I., Waltman L. (2021). Gender differences in scientific careers: A large-scale bibliometric analysis. Preprint: <https://arxiv.org/abs/2106.12624>
- Boothby C., Milojevic S., Larivière V., Radicchi F., Sugimoto C. (2022). Consistent churn of early career researchers: An analysis of turnover and replacement in the scientific workforce. Preprint: <https://doi.org/10.31219/osf.io/hdny6>
- Börner K. (2010). *Atlas of science: Visualizing what we know*. MIT Press.
- Britton D.M. (2017). Beyond the Chilly Climate: The Salience of Gender in Women's Academic Careers. *Gender & Society*, 31(1), 5–27.
- Calikoglu A., Jones G.A., Kim Y. (2023). *Internationalization and the Academic Profession. Comparative Perspectives*. Dordrecht: Springer.
- Elsevier (2020). *The researcher journey through a gender lens*. Elsevier.
- Fox M.F. (2020). Gender, science, and academic rank: Key issues and approaches. *Quantitative Science Studies*, 1(3), 1001–1006.
- Fox M.F., Mohapatra S. (2007). Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *Journal of Higher Education*, 78(5), 542–571.
- Fox M.F., Nikivincze I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. *Higher Education*, 81, 1237–1255.
- Halevi G. (2019). Bibliometric studies on gender disparities in science. [W:] W. Glänzel, H.F. Moed, U. Schmoch, M. Thelwall (red.), *Springer handbook of science and technology indicators* (s. 563–580). Springer.
- Hall R., Sandler B.R. (1982). *The Classroom Climate: A Chilly One for Women?* Association of American Colleges, Washington, D.C.
- Huang J., Gates A.J., Sinatra R., 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.
- Hyde J.S., Fennema E., Ryan M., Frost L.A., Hopp C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, 14(3), 299–324.
- Ioannidis J.P.A., Boyack K.W., Klavans R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLOS One*, 9(7): e101698.
- Kanter R.M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, 82(5), 965–990.
- King M.M., Bergstrom C.T., Correll S.J., Jacquet J., West J.D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, 3.
- Kwiek M. (2016). The European research elite: A cross-national study of highly productive academics across 11 European systems. *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. (2019). *Changing European academics. A comparative study of social stratification, work patterns and research productivity*. London and New York: Routledge.

- Kwiek M. (2020). Internationalists and locals: International research collaboration in a resource-poor system. *Scientometrics*, 124, 57–105.
- Kwiek M. (2022). *Globalna nauka, globalni naukowcy*. Warszawa: PWN.
- Kwiek M., Roszka W. (2021a). Gender disparities in international research collaboration: A large-scale bibliometric study of 25,000 university professors. *Journal of Economic Surveys*, 35(5), 1344–1388.
- Kwiek M., Roszka W. (2021b). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, 15(3), 1–38.
- Kwiek M., Roszka W. (2022a). Are female scientists less inclined to publish alone? The gender solo research gap. *Scientometrics*, 127, 1697–1735.
- Kwiek M., Roszka W. (2022b). Academic vs. biological age in research on academic careers: A large-scale study with implications for scientifically developing systems. *Scientometrics*, 127, 3543–3575.
- Kwiek M., Roszka W. (2022c). The Young and the Old, the Fast and the Slow: Productivity Classes and Rank Advancement of 16,000 STEMM University Professors. <https://arxiv.org/abs/2211.06319>.
- Kwiek M., Roszka W. (2023). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*. Online first <https://doi.org/10.1007/s10734-023-01022-y>.
- Kwiek M., Szymula Ł. (2023). Young Male and Female Scientists: A Quantitative Exploratory Study of the Changing Demographics of the Global Scientific Workforce. ArXiv, preprint. <https://arxiv.org/abs/2301.06196>.
- Larivière V., Ni C., Gingras Y., Cronin B., Sugimoto C.R. (2013). Global gender disparities in science. *Nature*, 504, 211–213.
- Maranto C.L., Griffin A.E. (2011). The antecedents of a 'chilly climate' for women faculty in higher education. *Human Relations*, 64(2), 139–159.
- Menard S. (2002). *Longitudinal research*. Sage.
- Mihaljević H., Santamaría L. (2020). Authorship in top-ranked mathematical and physical journals: Role of gender on self-perceptions and bibliographic evidence. *Quantitative Science Studies*, 1(4), 1468–1492.
- Morris L.K., Daniel L.G. (2008). Perceptions of a Chilly Climate: Differences in Traditional and Non-traditional Majors for Women. *Research into Higher Education* 49, 256–273.
- Morrison A.M., White R.P., Van Velsor E. (1987). *Breaking the glass ceiling: Can women reach the top of America's largest corporations?* Addison-Wesley.
- Nielsen M.W., Andersen J.P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences*, 118(7), e2012208118.
- Nygaard L.P., Piro, F., Aksnes, D. (2022). Gendering excellence through research productivity indicators. *Gender and Education*, 34(6), 690–704.
- Randolph C., Constantinople A., Gray J. (1988) The Chilly Climate: Fact or Artifact? *The Journal of Higher Education*, 59: 5, 527-55.
- Robinson-Garcia N., Costas R., Sugimoto C.R., Larivière V., Nane G.F. (2020). Task specialization across research careers. *eLife*, 9, e60586. 10.7554/eLife.60586
- Rowland D.T. (2014). *Demographic methods and concepts*. Oxford University Press.
- Ruspini E. (1999) Longitudinal Research and the Analysis of Social Change, [w:] Ruspini E. (Ed.) 'Longitudinal Analysis: A Bridge between Quantitative and Qualitative Social Research', Special Issue of *Quality and Quantity* 33(3), July-August.

- Salganik M.J. (2018). *Bit by bit. Social research in a digital age*. Princeton University Press.
- Savage W.E., Olejniczak A.J. (2021). Do senior faculty members produce fewer research publications than their younger colleagues? Evidence from Ph.D. granting institutions in the United States. *Scientometrics*, 126, 4659–4686.
- Selwyn N. (2019). *What is digital sociology?* Polity Press.
- Sexton K.W., Hocking K.M., Wise E., Osgood M.J., Cheung-Flynn J., Komalavilas P., ... Brophy C.M. (2012). Women in academic surgery: The pipeline is busted. *Journal of Surgical Education*, 69(1), 84–90.
- Shaw A.K., Stanton D.E. (2012). Leaks in the pipeline: separating demographic inertia from ongoing gender differences in academia. *Proceedings. Biological Sciences/The Royal Society*, 279(1743), 3736–41.
- Sheltzer J.M., Smith J.C. (2014). Elite male faculty in the life sciences employ fewer women. *Proceedings of the National Academy of Sciences of the United States of America*, 111(28), 10107–12.
- Sugimoto C., Larivière, V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press.
- Tang J. (1997). The glass ceiling in science and engineering. *The Journal of Socio-Economics*, 26(4), 383–406.
- Wachter K.W. (2014). *Essential demographic methods*. Harvard University Press.
- Way S.F., Morgan A.C., Clauset A., Larremore D.B. (2017). The misleading narrative of the canonical faculty productivity trajectory. *Proceedings of the National Academy of Sciences*, 114(44), E9216–E9223. 10.1073/pnas.1702121114
- Whitt E.J., Nora A., Edison M., Terenzini P.T., Pascarella E.T. (1999). Women's perceptions of a "chilly climate" and cognitive outcomes in college: Additional evidence. *Journal of College Student Development*, 40, 163–177.
- Wolfinger N.H., Mason M.A., Goulden M. (2008). Problems in the pipeline: Gender, marriage, and fertility in the ivory tower. *Journal of Higher Education*, 79(4), 388–405.
- Zhang S., Wapman K.H., Larremore D.B., Clauset A. (2022). Labor advantages drive the greater productivity of faculty at elite universities. *Science Advances*, 8(46).

Profesja akademicka w ujęciu globalnym: co Big Data mówią nam o udziale kobiet w nauce?

W prezentowanym studium analizujemy 4,3 mln nieokazjonalnie publikujących naukowców z 38 krajów OECD (w tym z Polski) z lat 1990–2021. Interesuje nas wzdłużny (longitudinalny) rozkład populacji młodych kobiet naukowców – do 10 lat po pierwszej publikacji naukowej – w czasie w 16 dyscyplinach STEMM (ściślych, technicznych, inżynieryjnych, matematycznych – oraz medycznych). Analizujemy zmieniającą się kadrę naukową z pomocą 5-letnich grup wiekowych w ramach każdej dyscypliny. Sprawdzamy przydatność globalnych źródeł surowych danych bibliometrycznych (ujmowanych jako ustrukturyzowane Big Data) w analizie kadry naukowej w czterech wymiarach: płci, wieku, dyscypliny i wpływu czasu. Tradycyjne zagregowane dane dotyczące naukowców ogółem skrywają zróżnicowany obraz zmieniającej się dynamiki uczestnictwa mężczyzn i kobiet nauki w ramach dyscyplin i wieku. Analizujemy ograniczenia bibliometrycznych

zbiorów danych, a badania globalne porównujemy z polskimi badaniami na poziomie krajowym. Wskazujemy wybory metodologiczne i ich implikacje, a także krótko omawiamy nowe możliwości globalnego badania aktywności publikacyjnej młodych kobiet (i mężczyzn) naukowców.

Słowa kluczowe: globalna kadra naukowa, kobiety w nauce, wiek akademicki, OECD, dane bibliometryczne, badanie wzdłużne

Global scientific workforce: What Big data say on women in science

The purpose of this article is to determine what can we learn about the changing demographics of scientific personnel around the world and over time using available global bibliometric data sources. We wanted to see how useful global data could be for analyzing the scientific workforce. We tested how the demographic transformation of the global scientific profession can be measured using new data sources, thus transcending the traditional approach in which national statistics from national statistical offices are aggregated to a higher level, as in the case of the scientific workforce databases produced by the OECD, UNESCO and the European Union (Eurostat).

Key words: global scientific workforce, women in science, academic age, OECD, bibliometric data, longitudinal study