

Propozycja nowego modelu ewaluacji jednostek naukowych

Piotr Stec / Instytut Nauk Prawnych, Uniwersytet Opolski

e-mail: pstec@uni.opole.pl

ORCID: 0000-0003-3797-1321

Streszczenie

Dotychczasowy system ewaluacji jednostek naukowych pomimo wielu wad pozwala na osiągnięcie dwóch podstawowych celów: oceny jakości badań i ustalenia rankingu, według którego dzielone są środki na działalność naukową. Jego skuteczność zależy jednak od istnienia przejrzystego i wolnego od wpływów naukowych lobbies systemu narzędzi pomiarowych. W artykule zaproponowano system oceny jednostek naukowych pozwalający na stworzenie rankingu na poziomie krajowym oraz ocenę stopnia, w jakim wyniki polskich jednostek naukowych są zbliżone do tych osiągniętych przez zagraniczną konkurencję. System ten uzupełnia ocena ekspercka służąca wyłącznie jako narzędzie zarządzania jakością.

Słowa kluczowe: **Ocena badań naukowych, miary, zarządzanie jakością.**

A Proposal for a New Model for the Evaluation of Scientific Units

Abstract

The current system of research assessment in Poland is despite its numerous defaults a relatively reliable tool for performance-based funding and for ranking the institutions. Its workability depends largely on the quality of metrics used. These metrics should be transparent and designed without undue influence of academic lobbyists. In this paper a system for ranking the research quality at various institutions has been proposed. This system is supplemented by a set of metrics allowing an international comparison of universities in terms of research performance. Additionally a system for peer-review as a quality management tool has been proposed.

Keywords: **Research assessment, metrics, quality management.**

WPROWADZENIE

Ewaluacja wyników badań naukowych jest jednym z ważniejszych problemów nie tylko polityki naukowej, ale i prawa nauki i szkolnictwa wyższego. Dobrze skonstruowany system ewaluacji badań pozwala uczelniom na dokonanie oceny własnych dokonań, zapewnia rozliczalność wobec interesariuszy, pozwala sprawdzić, czy pieniądź publiczny jest dobrze wydany, a polityka naukowa państwa właściwie realizowana (Kulczycki, 2019). Do tego potrzebujemy miar pozwalających na rzetelną ocenę jednostek, a przy tym mało podatnych na manipulacje czy to ze strony organizatorów systemu (władzy publicznej), czy to jego uczestników. Problem ten, jak się wydawało, dotyczył głównie nauk humanistycznych i społecznych, dla których istotne znaczenie miały czasopisma krajowe, rozszerzył się jednak na pozostałe dziedziny, które również chętnie skorzystały z możliwości optymalizacji punktowej. Po doświadczeniu z kolejnymi parametryzacjami oraz niemal dwoma zakończonymi ewaluacjami wiemy już, że o ich, delikatnie mówiąc umiarkowanym, sukcesie zdecydował czynnik ludzki. To ludzie decydują o doborze miar (kryteriach oceny) i wagach przypisywanych poszczególnym osiągnięciom. Tu na nieuniknione błędy wynikające z niepełnego dostępu do informacji nałożyły się wyniki nacisków środowiskowych i próby ręcznego poprawienia systemu tam, gdzie wydawałoby się rzetelne miary zawiodły.

We wcześniejszych tekstach pokazano, jak wyniki ewaluacyjne i preconcepcje leżące u podstaw naszego systemu oceny jednostek naukowych zderzają się z rzeczywistością (Garayová et al., 2023; Stec, 2022, 2023; Wierczyński, 2020, 2022; Winczorek, 2022). Zaproponowano też na potrzeby kolejnej ewaluacji doraźne zmiany, pozwalające nadać jej sensowny kształt bez wprowadzania (Stec, 2024). W tym tekście proponuję nowy model ewaluacji, składający się z oceny ilościowej pozwalającej na ustalenie rangi jednostki w populacji polskich uczelni oraz jej pozycji w odniesieniu do światowej lub międzynarodowej konkurencji uzupełniony o ocenę ekspercką. Ta pierwsza stanowiłaby narzędzie pozwalające na efektywny podział środków, ta druga zaś – prawdziwe narzędzie zarządzania jakością.

1. MIĘDZYNARODOWA POZYCJA POLSKICH JEDNOSTEK NAUKOWYCH

Obowiązujące przepisy prawa wymagają, by przy ustalaniu progów dla każdego z trzech kryteriów ewaluacji obliczać progi na poszczególne kategorie z uwzględnieniem pozycji międzynarodowej Polski w międzynarodowych bibliograficznych bazach danych¹. Dyrektywę tę należy uznać za ze wszech miar słuszną, gdyż porównywanie wewnętrzne w obrębie jednego kraju prowadzi co najwyżej do grading on the curve i mało interesującej poznawczo konstatacji, że duża i stara uczelnia wygrywa zwykle z małą i prężną (Cieśliński & Stec, 2018). Oczywiście nasze naukowe zawody odbywają się w kategorii open, trudno jednak uznać sensowność kosztownej przecież procedury, której wyniki są znane z góry, chyba że pojawi się błąd w systemie, który wykorzystają mniejsi uczestnicy gry o punkty.

Dotychczasowe próby ustalenia pozycji międzynarodowej Polski na potrzeby ewaluacji należy uznać za niezadowolające. (Uchwała nr 1/2021 Komitetu Nauk Prawnych Polskiej Akademii Nauk z 15.02.2021 r. w sprawie wykazu i punktacji czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych, 2021; Wróblewski, 2013). Wbrew pierwszej i dość naturalnej intuicji uważam, że ustalenia takie są możliwe nawet w naukach humanistycznych i społecznych, a także w dyscyplinach posługujących się w znacznej mierze językami narodowymi. Niezależnie od widocznej dominacji „anglosfery” w SCOPUS czy WoS, w bazach tych są indeksowane także czasopisma wydawane w językach innych niż angielski, w tym polsko- lub dwujęzyczne czasopisma krajowe, trudno więc postawić zarzut zupełnej niereprezentatywności tych baz. Dodatkowo w obrocie dostępny jest otwarty odpowiednik tych baz w postaci OpenAlex AI, będący znacznie bardziej inkluzywnym, choć wciąż niedoskonałym narzędziem porównań.

Problemem jest natomiast ustalenie miar pozwalających na odniesienie polskich osiągnięć naukowych do ich odpowiedników w skali międzynarodowej albo choćby europejskiej. Obecnie, jak wiadomo z wypowiedzi członków KEN publikowanych na łamach Forum Akademickiego, podstawowym punktem odniesienia przy ustalaniu międzynarodowej pozycji naukowej Polski były dane ze SCIMAGO Country and Journal Rank (Skoczeń, 2025). Jest to wiarygodne źródło danych tworzone w oparciu o SCOPUS przez niezależny od Elsevierów ośrodek badawczy. Pytanie jednak, czy na podstawie tych danych można przeprowadzić ocenę zgodną z wytycznymi zawartymi w rozporządzeniu ewaluacyjnym? Jak można domniemywać z wypowiedzi członków KEN, podstawową miarą zastosowaną do wyliczeń były dane SCIMAGO dotyczące ich podstawowego rankingu państw. I tu zaczynają się problemy. Podstawowy ranking w zasadzie opiera się na dwóch wskaźnikach – liczbie publikacji i indeksie Hirscha. Ta pierwsza liczba powinna być związana z wielkością danego kraju i liczbą publikujących uczonych. Ta druga – choć jest to uznana, co nie znaczy, że niekontrowersyjna miara bibliometryczna – jest zupełnie niepowiązana z wynikami ewaluacji. Indeks Hirscha mierzy, ile razy była cytowana n-ta publikacja danego uczonego, a w naszym systemie jakość i prestiż czasopisma liczona zupełnie innymi wskaźnikami. Do oceny jakości konfitur używamy więc miar przeznaczonych do oceny jakości

¹ § 26 ust. 2 pkt 2 Rozporządzenia Ministra Nauki i Szkolnictwa Wyższego z dnia 22 lutego 2019 r. w sprawie ewaluacji jakości działalności naukowej, Dz. U z 2019 r., poz. 392.

destylatów, bo i jedno i drugie robi się z owoców. O zawodności tego systemu świadczy też fakt, że daje on zupełnie różne wyniki zależnie od tego, które z kryteriów wybierzemy i jak bardzo będziemy chcieli uzyskać konkretny wynik. Najprawdopodobniej dane ze SCIMAGO posortowano według defaultowego kryterium liczby dokumentów z danego kraju. To kryterium jest zależne m.in. od wielkości danego kraju, a zatem stosunkowo wysoka pozycja Polski nie dziwi. Problem ilustrują dane ze SCIMAGO Journal and Country Rank dla wybranych dziedzin i dyscyplin (klasyfikacja rankingu nie w pełni da się przełożyć na polską klasyfikację nauk). Wybrano arts & humanities i social sciences jako odpowiedniki dość szeroko rozumianych nauk humanistycznych i społecznych, chemię, medycynę i fizykę jako dyscypliny mocno bibliometryczne, w których powinniśmy mieć dobrą pozycję światową. Dodatkowo uwzględniłem nauki o zarządzaniu, ekonomię i ekonometrię oraz psychologię jako dyscypliny nauk społecznych, w których stosowanie miar bibliometrycznych jest normą, a także politologię jako dyscyplinę, w której toczy się bardzo ostry spór o sposób jej uprawiania w Polsce (Górecki, 2019). W tej Stichprobe wziąłem pod uwagę trzy wskaźniki: liczbę dokumentów w SCOPUS, liczbę cytowań na dokument i indeks Hirscha. Dane są za 2023 r., czyli najświeższe dostępne w chwili pisania tekstu. W próbie tej wziąłem pod uwagę dane dotyczące pozycji światowej Polski, czyli to, do czego odnoszą się przepisy ewaluacyjne, jak i kraje Unii Europejskiej jako zbioru obejmującego państwa wysokorozwinięte z podobnymi systemami nauki i szkolnictwa wyższego.

Tabela 1 *Międzynarodowa pozycja Polski w wybranych dyscyplinach. Opracowanie własne na podstawie danych SCIMAGO Journal & Country Rank 2024*

Dyscyplina	Pozycja wg liczby dokumentów		Pozycja wg liczby cytowań na dokument		Pozycja wg indeksu Hirscha	
	Świat	UE	Świat	UE	Świat	UE
Chemia	15	5	97	18	23	9
Nauki medyczne	20	7	73	21	23	11
Fizyka z astronomią	19	8	84	21	19	8
Arts & Humanities	32	5	145	24	32	13
Social Sciences	16	6	125	20	35	14
Nauki o zarządzaniu	26	8	47	11	38	14
Ekonomia i finanse	22	6	53	11	49	15
Psychologia	24	9	82	23	31	13
Politologia i stosunki międzynarodowe	14	8	44	22	35	15

Przedstawiony tu wybór pokazuje nam dwie ważne, choć niepowiązane ze sobą rzeczy. Po pierwsze, dobór wskaźników decyduje o pozycji międzynarodowej Polski w danym kwartylu. Na przykład pod względem liczby cytowań i wskaźnika Hirscha polska chemia plasuje się międzynarodowo w górnych 25%, a według liczby cytowań w górnych 50%. Ta sama zależność widoczna jest przy dyscyplinach nauk społecznych, natomiast przy dziedzinach nauk humanistycznych i nauk społecznych liczba cytowań na dokument znacznie się obniża i przechodzi do przedziału 50–75%. W grupie krajów UE widać, że nauki ścisłe i przyrodnicze plasują się na wysokich pozycjach, a nauki społeczne i humanistyka tworzą peleton dochodzący do środka skali. Jedynie na marginesie, bo wymaga to dalszych badań, a wyniki dla tak małej próby mogą być niewiarygodne, analiza współczynnika korelacji Spearmana wykonana za pomocą programu Jamovi (Jamovi Core Team, 2025) pokazała, że nie ma statystycznie istotnych korelacji pomiędzy pozycjami dyscyplin w każdej z tych kategorii. To znaczy, że liczba dokumentów, liczba cytowań na dokument i indeks Hirscha nie są wzajemnie ze sobą powiązane.

Co nam to mówi o pozycji naukowej Polski – dość dużo, a wnioski są znacznie bardziej optymistyczne niż to, co możemy przeczytać w publicystycznych tekstach na ten temat. Najwyraźniej nie uprawiamy „nauki równoległej”, jak w odniesieniu do politologii twierdził Maciej Górecki (Górecki, 2019), nie ma też już chyba podziału na „aspirujące centrum” i „zadowolone peryferia” zdiagnozowanego przez K. Szczygileckiego (Szczygilecki, 2023). Zajmujemy nienajgorszą pozycję, zwłaszcza jeśli uwzględnić to, że jesteśmy jako sektor permanentnie niedofinansowani, a w przypadku nauk społecznych i humanistycznych przepisy o kategoryzacji jednostek zmniejszyły premię za publikowanie w kraju dopiero w ostatniej ewaluacji. Na 194 kraje ujęte w rankingu SCIMAGO mieścimy się w pierwszym kwartylu pod względem liczby publikacji i indeksu H. Pod względem średniej liczby cytowań balansujemy (z wyjątkiem humanistyki i nauk społecznych) w okolicach mediany.

Jeśli weźmiemy pod uwagę wyłącznie kraje UE, to dobrze wypadamy pod względem liczby dokumentów na SCOPUS (widać tu skutki reformy J. Gowina), gorzej jeśli chodzi o liczbę cytowań i indeks H. Należy jednak pamiętać, że w grupie krajów UE znajdują się kraje będące od lat w światowej czołówce. Jest to miejsce, jakiego można się spodziewać po kraju, który do międzynarodowej konkurencji o zasoby i prestiż przystąpił stosunkowo późno i intensywnie nadrabia wieloletnie zaniedbania.

Powyższe dane pokazują też, że argumenty przeciwko stosowaniu baz, takich jak SCOPUS czy Web of Science jako referencyjnych w humanistyce i naukach społecznych, chyba straciły rację bytu. W ciągu ostatniej dekady nawet w opornych naukach prawnych liczba krajowych czasopism na SCOPUS wzrosła kilkudziesięciokrotnie, a polscy autorzy zaczęli publikować za granicą. Oczywiście, nie zmienia to specyfiki tych dziedzin i konieczności publikowania także poza SCOPUS, jest to jednak czynnik do uwzględnienia przy ostatecznej ocenie, a nie argument przeciwko wyłączeniu baz referencyjnych z oceny dyscyplin humanistycznych i społecznych, a także teologii, biblistyki czy nauk o rodzinie. Troską organów władzy publicznej powinno być zapewnienie polskim wydawcom narzędzi pozwalających na wprowadzenie czasopism do wykazów. Przykładowo, program wsparcia czasopism mógłby obejmować usługi profesjonalistów pomagających wydawcom zrobić audyt czasopisma i przygotować wniosek o indeksację. Wynagrodzenie tych konsultantów powinno być częściowo zależne od odsetka skutecznych wniosków o indeksację.

Pomimo że ranking SCIMAGO daje wiarygodny obraz podstawowych praktyk publikacyjnych (przy zastosowaniu bardziej skomplikowanych wskaźników, np. w rankingu uczelni SCIMAGO wypadamy gorzej), w dalszym ciągu nie nadaje się on do ustalania progów na kategorie. Nawet gdybyśmy przyjęli, że podstawą kategoryzacji jest indeks Hirscha, to da się go zastosować tylko przy kryterium pierwszym. Tymczasem cała ewaluacja, wliczając w to także kryterium grantowe i wpływ na otoczenie, ma zgodnie z przepisami progi na kategorie ustalone w odniesieniu do pozycji międzynarodowej Polski wyliczanej zgodnie z danymi z bibliometrycznych baz danych. Przyjmując, że tego przepisu by nie było, ewaluacja ma sens wyłącznie w przypadku porównywania naszych jednostek – nie tylko ze sobą, ale i z osiągnięciami zagranicznych konkurentów.

Czy zatem znalezienie takich punktów odniesienia jest możliwe? Uważam, że tak, i to takich, które będą pozwalały na wolną od ingerencji władzy publicznej i stosunkowo mało podatną na manipulację ocenę oraz będą bazować na danych ze SCOPUS. Pierwsza z tych baz zawiera także informację o uzyskiwanych przez ośrodki naukowe prestiżowych grantach, jest więc punkt odniesienia w tym zakresie. Nadto SCOPUS pozwala na mierzenie wpływu na otoczenie. Rozumianego nieco inaczej, niż chce tego nasz ustawodawca, ale mierzalnego. SCOPUS jako główną miarę proponuje liczbę tekstów realizujących cele zrównoważonego rozwoju proponowane przez ONZ. Jest to wskaźnik niezależny od języka czy tematyki badań, bo cele te ujęte są bardzo szeroko. Takie wskaźniki jak odsetek publikacji wspólnych z innymi ośrodkami czy powstałych we współpracy międzynarodowej stosowane przez SCOPUS również nadają się do zastosowania. W tym miejscu od razu trzeba odeprzeć argument, który na pewno pojawi się w debacie po stronie nauk humanistycznych i niektórych społecznych, na przykład prawnych: „ale w naszych dyscyplinach większość publikacji jest jednoautorska”. Owszem, przez lata tak było, ale i to się zmienia. Nawet monografia wieloautorska, praca zbiorowa czy komentarz prawniczy od zawsze były wynikiem współpracy między uczonymi z różnych ośrodków. Poza tym miara ta, jak i pozostałe, ma nam posłużyć do porównania z zagranicą w ramach danej dyscypliny. Jeśli więc humanistyka szwedzka, włoska czy portugalska mają takie same praktyki publikacyjne, to i odsetek prac wspólnych powinien być porównywalny. Jeśli okaże się, że tak nie jest, może to znaczyć, że zostaliśmy po prostu w tyle.

2. LICZBA N, MINIMA EWALUACYJNE, MAKSYMALNA LICZBA PUBLIKACJI, PRZYPISANIE DO DYSCYPLINY

Problemy, których dotyczy ta sekcja, wydają się na pierwszy rzut oka mało interesujące, wręcz techniczne. Tymczasem to one decydują i będą decydować w przyszłości o kształcie ewaluacji i polityce kadrowej ewaluujących się jednostek.

Przepisy ewaluacyjne zakładają konieczność istnienia pewnej ewaluacyjnej masy krytycznej pozwalającej na ocenę jakości badań, a co więcej na przyjęcie, że będzie ona wystarczająco duża, by stworzyć zespół, w którym da się kształcić i promować doktorantów. Wymóg istnienia masy krytycznej nie jest nowy. Co więcej, przewidywały go nieobowiązujące przepisy o tytułach i stopniach (Ustawa z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki, Dz. U. z 2017 r. poz. 1789), a także przepisy wprowadzające ustawę, tj. Prawo o szkolnictwie wyższym i nauce (Art. 175 ust. 1 ustawy z dnia 3 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce, Dz. U. z 2018, poz. 1669 ze zm).

Podobne minima istniały odnośnie prowadzenia kierunków studiów. Nasz ustawodawca jest więc w tym zakresie konsekwentny. Otwartym pozostaje pytanie, dlaczego akurat 12 a nie 11 czy 13 osób tworzy minimalną liczbę N. 12 osób to minimalna obsada dla kierunku studiów prowadzonych w danej dyscyplinie na I i II stopniu². Teoretycznie to całkiem rozsądne

² Przyjmując, że na 30 pkt ECTS w semestrze wypada około 10 h godzin zajęć z prowadzącym, studia I stopnia dają 1800 h dydaktycznych dla jednej grupy wykładowej oraz 1200 h dydaktycznych dla II studiów II stopnia – otrzymujemy 3000 h do zagospodarowania. Po odjęciu zajęć niezwiązanych bezpośrednio z dyscypliną (języki, zajęcia ogólnouczelniane, itd.), otrzymujemy pensum dla tuzina osób ze średnim obciążeniem 210 h rocznie.

podejście. Oczywiście eliminuje ono mniejsze zespoły badawcze czy pojedynczych pracowników, którzy albo są nieewaluowani albo przypisywani na siłę do jakiejś dyscypliny. Dla jednostek małych należałoby więc stworzyć osobną kategorię „M” (od „małe”), w której mieściłyby się takie właśnie niewielkie zbiorowości, jednak dyscypliny w tej kategorii nie miałyby praw do nadawania stopni naukowych (Stec, 2024, pp. 438–439).

Proponowane tu rozwiązanie zakłada wymóg istnienia masy krytycznej tam, gdzie potrzebne są zasoby ludzkie do kształcenia następnych pokoleń doktorantów. Będzie się to jednak wiązało z koniecznością zapewnienia uczelniom większej elastyczności zatrudnienia, tak by mogły one tworzyć swobodnie etaty badawcze. To problem na osobną analizę. Opowiadałbym się przy tym za wzmocnieniem tendencji do tworzenia zespołów z dużą masą krytyczną przez odejście od liczenia liczby N wymiarem etatu. Pracownicy badawczy i badawczo-dydaktyczni mają zróżnicowane obowiązki i ci ostatni nie są w stanie poświęcić tyle samo czasu na badania, co ci pierwsi. Możemy przyjąć, że o ile obowiązki organizacyjne pochłaniają około 20% czasu pracy, to badacz pozostałe 80% przeznacza na osiągnięcie efektów ewaluacyjnych, natomiast ktoś na etacie badawczo-dydaktycznym może poświęcić na prowadzenie badań 40% swojego czasu. Resztę zajmuje dydaktyka. Dlatego proponuję, by przeliczać etaty na potrzeby ustalania liczby N, uwzględniając różnicę między tymi dwoma kategoriami pracowników. Badacz byłby liczony z wagą 1, a pracownik badawczo-dydaktyczny z wagą 0,5. Odpowiednio do tego należałoby dostosować normy publikacyjne (2–3 prace w okresie ewaluacyjnym dla badacza, 1–2 dla pracownika badawczo-dydaktycznego). W połączeniu z elastycznie konstruowanym stanem kadrowym pozwoliłoby to m.in. uczelniom mającym w danej dyscyplinie mało studentów, a dużo dobrych pracowników, bardziej racjonalnie dysponować kadrami. Oczywiście łączy się z tym możliwość utraty uprawnień przez niektóre uczelnie. Konieczne byłoby więc odejście od fetyszyzacji pełni praw akademickich na rzecz promowania jakości badań. Znowu wypada zauważyć, że szczegółowa analiza wykracza poza ramy tego artykułu.

Do liczby N nie wliczaliby się w dalszym ciągu doktoranci i dydaktycy, ale ich prace mogłyby być uwzględniane w ewaluacji. Nie mogłyby jednak stanowić więcej niż 20% zgłoszonych do ewaluacji tekstów. O ile wyniki badań doktorantów w sposób oczywisty łączą się z prowadzeniem badań w jednostce, o tyle pewne wątpliwości mogłyby się pojawić odnośnie dydaktyków niemających przecież badań w obowiązkach służbowych. Należy przy tym pamiętać, że dydaktyk ma obowiązek rozwoju zawodowego, a prowadzenie badań i research-informed teaching są dopuszczalnymi formami realizacji tego obowiązku.

3. WYKAZY CZASOPISM I WYDAWNICTW. PUNKTACJA OSIĄGNIĘĆ

Przy tworzeniu wykazów czasopism i wydawnictw proponuję na pierwszym etapie reformy ewaluacji ograniczyć się do kryteriów formalnych i weryfikowalnych bez udziału czynnika eksperckiego. Doświadczenia ostatnich kilku lat z wykazem eksperckim modyfikowanym później dość swobodnie przez ministra (Kulczycki, 2021) skłaniają jednak do tego, by optować za niemal automatycznym systemem oceny.

Proponuję, jeśli chodzi o wykaz czasopism, oparcie się na bibliometrycznych i niebibliometrycznych bazach danych o uznanej reputacji: SCOPUS, Web of Science, listy fińską i norweską oraz ERIH+ (National Board of Scholarly Publishing, 2025), DOAJ i EBSCO. Dla nauk prawnych dodatkową bazą referencyjną byłby HeinOnline, a dla informatyków – wykaz konferencji CORE. Pozostawiając na boku szczegółowe wyliczenia i kryteria doboru baz referencyjnych w poniższej klasyfikacji posłużę się częściowo podziałem kwartyłowym SCIMAGO Journal Rank. O tyle ciekawym, że choć bazuje na danych ze SCOPUS, jest efektem pracy niezależnego niekomercyjnego podmiotu. Oczywiście wskaźniki te można dopasować także dla pozostałych baz. Proponowany poniżej system to jeden z wielu dostępnych. Jego wariant bazowy jest obecnie używany przez CERES (CERES Research School for International Development, 2025).

Kategorię A tworzyłyby czasopisma z najwyższego (Q1) kwartyła SCIMAGO, kategorię B – czasopisma z kwartyli 2–4 oraz nierangowane czasopisma ujęte w SCOPUS lub WoS. Kategorię C – czasopisma indeksowane przez listy fińską i norweską, być może także włoski wykaz ANVUR (Agenzia nazionale di valutazione del sistema universitario e della ricerca, 2023) oraz dla nauk prawnych przez HeinOnline. W kategorii D byłyby czasopisma indeksowane w ERIH+, DOAJ, EBSCO. W przypadku konferencji informatycznych najlepsze z nich byłyby w grupie A, zaś pozostałe w B. W tym systemie należałoby zrezygnować z tworzenia osobnych zasad dla czasopism polskich. Redakcja, która nie potrafi zakwalifikować swojego czasopisma do grupy C lub D, prawdopodobnie powinna zająć się czymś innym.

Podobny zabieg należałoby zastosować do wykazu wydawców, opierając się na już istniejących listach, takich jak listy fińska i norweska, ranking hiszpański (Consejo Superior de Investigaciones Científicas, 2025) czy stosowany przez University of Limerick (University of Limerick, 2024), a najbardziej zbliżony do proponowanego tutaj, system oparty na wyszukiwaniu wydawców uznawanych za prestiżowych przez większą liczbę rankingów. Żeby sporządzić taki wykaz należałoby znaleźć część wspólną zbioru, to jest wydawców, którzy we wszystkich bazach referencyjnych znaleźli się w czołówce. Tworzyliby oni wydawnictwa kategorii A. Pozostałe wydawnictwa zakwalifikowane na którejkolwiek z tych list do czołówki, tworzyłyby kategorię B. Reszta wydawnictw rangowanych byłaby w kategorii C. Kategorię D tworzyłyby wydawnictwa praktyczne albo o lokalnym znaczeniu. Ponieważ indeksacja w tych wykazach jest dla polskich czasopism nowością, proponowałbym tu wyjątkowo wprowadzić dodatkowe zasady kwalifikacji. Wydawnictwa krajowe, w których opublikowano w okresie 4 lat poprzedzających okres ewaluacyjny prace będące osiągnięciami habilitacyjnymi lub osiągnięciami uwzględnionymi w postępowaniu o nadanie tytułu profesora, znalazłyby się w grupie B. Pozostałe wydawnictwa naukowe natomiast w grupie C. Przyjęcie takiego sposobu

wyłaniania polskich wydawców grupy B uzasadnione jest tym, że osiągnięcia habilitacyjne muszą wywierać istotny wpływ na rozwój dyscypliny, a profesorskie cechować się wybitnością. O ile więc w danej liczbie (jakiej – do dyskusji) przypadków recenzenci uznali, że wydawnictwo prezentuje wysoki poziom naukowy i znajdują się tam prace wybitne lub istotnie wpływającą na rozwój dyscypliny, to należy to docenić.

W przypadku gdyby ustawodawca zdecydował się na taki wariant, proponuję odejście od skali nieliniowej na rzecz liniowej z punktacją w przedziale 1 do 4 zależnie od kategorii, do której przypisałibyśmy osiągnięcie.

4. KRYTERIA 2 I 3 EWALUACJI

Podobny zabieg należałoby zastosować do kryterium 2. ewaluacji, zaliczając do kategorii A granty ERC i granty indeksowane przez SCOPUS. W grupie B byłyby granty przyznawane przez narodowe agencje grantowe, NATO i pozostałych grantodawców organizujących międzynarodowe konkursy grantowe. W grupie C znalazłyby się prace zlecone i wyniki komercjalizacji, a w D – małe granty. Liczbę osiągnięć w tej kategorii ograniczyłbym do nie więcej niż 20% z N. Zważywszy na próg skuteczności aplikacji grantowych wypełnienie normy może tu być trudne i należałoby się zastanowić, czy w przypadku, gdy dana jednostka ma dobre osiągnięcia publikacyjne, powinna być karana za to, że osiąga je małym kosztem? O ile nie budzi wątpliwości dodatkowe premiowanie pozyskiwania grantów w algorytmie finansowania, o tyle w przypadku oceny jakości badań jest ono nieoczywistym składnikiem ze względu na to, że nawet w ramach jednej dyscypliny kosztowność badań może być różna. Prawnik – dogmatyk badający prawo obowiązujące metodami klasycznymi, będzie *sit venia verbo* tańszy w utrzymaniu niż jego koleżanka – kryminolożka prowadząca badania empiryczne czy komparatysta, którego badania wymagają kwereń zagranicznych. Przy takiej konstrukcji kryterium 2 można stosunkowo łatwo wyznaczyć przy pomocy SCOPUS-a liderów. Analiza wyników grantowych w niższych kategoriach byłaby trudniejsza, choć nie niemożliwa – dane o osiągnięciach zagranicznych z grupy B należałoby zebrać albo ręcznie, albo przygotowując odpowiednie oprogramowanie. Natomiast osiągnięcia z grup C i D miałyby tylko znaczenie lokalne. Ponownie przydział punktów w każdej grupie mieściłby się w przedziale 1 do 4.

Jeśli chodzi o wpływ na otoczenie, to żeby go ocenić w odniesieniu do jednostek zagranicznych, musielibyśmy odstąpić od obecnego systemu oceny eksperckiej i zastosować miary, które pozwolą nam na dokonanie porównania. I tu jednym z możliwych narzędzi jest SCOPUS z jego miarami wpływu, innym – narzędzia pozwalające wyszukiwać cytowania prac naukowych w dokumentach dotyczących polityk publicznych. Przy użyciu takich narzędzi możliwe byłoby stworzenie miary pozwalającej na zbudowanie sensownego systemu oceny wpływu na otoczenie. Zastrzegam tu, że narzędzia do badania wpływu przez analizę szarej literatury są dopiero w fazie wstępnej i niekoniecznie odzwierciedlają rzeczywisty stan współpracy nauki z otoczeniem. Dlatego w pierwszej fazie wdrażania proponowanego tu systemu być może należałoby się oprzeć tylko na danych ze SCOPUS-a. Tak skonstruowana miara obejmowałaby wspomniane wcześniej elementy – odsetek tekstów odnoszących się do celów zrównoważonego rozwoju (SDG), odsetek tekstów we współautorstwie krajowym, odsetek tekstów we współautorstwie międzynarodowym, odsetek tekstów pisanych wspólnie z praktykami i odsetek tekstów cytowanych w dokumentach dotyczących polityk publicznych, opisach patentowych czy szerzej mówiąc szarej literaturze. W każdej z tych grup przypisywano by 1 punkt za każdy decyl. Jeśli tekstów odnoszących się do SDG byłoby w dorobku jednostki do 10%, otrzymywałaby ona za nie 1 pkt, jeśli od 10% do 20% – 2 pkt, itd.

5. KATEGORYZACJA JEDNOSTEK: MODEL ALGORYTMICZNO-OPISOWY

Bez pełnego przejścia na system oceny eksperckiej decydujemy się na uproszczony, oparty na wyliczeniach, system kategoryzacji. To, że poszczególnym osiągnięciom nadajemy wagę, nie oznacza jednak, że jesteśmy skazani na pogoń za punktami czy posługiwanie się wyłącznie rangami nadanymi przez bezmyślnie mielący dane algorytm. Przeciwnie, dane liczbowe użyte w sposób odpowiedzialny, pozwalają na uzyskanie przejrzystego opisu działalności badawczej jednostek. Pozwala to zorientować się w profilu publikacyjnym i pozostałych osiągnięciach jednostek. Co zatem można skwantyfikować, a co przedstawić w sposób opisowy?

Jeśli chodzi o strukturę publikacji i grantów, opis dorobku jednostki byłby stosunkowo prosty. Wystarczy pokazać, jaki procent osiągnięć jednostki został zaliczony do danej kategorii. Czyli na przykład Uniwersytet Najlepszy w Dyscyplinie ma 40% publikacji z grupy A, 30% z grupy B i po 15% z grup C i D. Natomiast Uniwersytet Przeciętny ma 15% publikacji z grupy A, 30% z grupy B 40% z grupy C, 15% z grupy D. Idem dla pozostałych kryteriów.

Do czego zatem potrzebna by była ocena punktowa? Do analizy, jak sobie dana jednostka radzi w stosunku do innych polskich uczelni. Uzyskany przez nią średni wynik punktowy pozwoliłby umieścić ją w jednym z czterech kwartyli Q1 dla najlepszych, Q2 i Q3 dla przeciętnych i Q4 dla jednostek radzących sobie najsłabiej. Będzie ona miała także znaczenie przy wyliczaniu subwencji według nowego algorytmu.

Powyższa propozycja nie rozwiązuje oczywiście problemu podstawowego – porównania wyników polskich jednostek z ich zagranicznymi konkurentami. Proponowałbym oprzeć, tak jak dotychczas, porównanie międzynarodowe na danych z bazy SCOPUS wspomaganymi danymi ze SCIMAGO odnośnie rang czasopism liczonych według Scimago Journal Rank. Taki wybór

podyktowany jest kilkoma czynnikami. Po pierwsze, SCOPUS jest jedną z uznanych międzynarodowo baz danych używanych do analiz bibliometrycznych. Odwołując się do niej, unikamy zarzutu dobierania bazy danych tak, żeby pasowała do oczekiwanych wyników. Po drugie, SCOPUS jest inkluzywny w tym sensie, że obejmuje znacznie więcej czasopism z nauk humanistycznych i społecznych niż Web of Science. Po trzecie, ze względu na to, że czasopisma do bazy dobiera wydawca, zawartość bazy jest trudniejsza do zmanipulowania. W szczególności nie zdarzy nam się to, co jest jednym z grzechów głównych obecnego wykazu czasopism – skuteczny lobbing lokalnych środowisk naukowych na rzecz „dowartościowania” tego czy innego czasopisma. Dodatkowym zabezpieczeniem byłoby oparcie analizy na danych SCIMAGO i SJR. Wskaźnik ten, choć dotyczy czasopism indeksowanych przez SCOPUS, został opracowany niezależnie i ze względu na jego wieloparametrowość wydaje się trudniejszy do manipulacji. Odwołując się do danych ze SCOPUS, bierzemy pod uwagę wyłącznie czasopisma spełniające trochę wyższe standardy jakości niż cała nieindeksowana reszta. Oczywiście proponowana tu metoda zadziała z każdą bibliometryczną bazą danych i zastosowanie jej do danych z OpenAlex czy Web of Science nie będzie problemem.

Wreszcie wypada zauważyć, że odniesienie się do SCOPUS nie ogranicza w niczym nauk humanistycznych i społecznych. Jeśli prawdziwa dla nich jest teza, że ze względu na mnogość kanałów publikacyjnych SCOPUS pokazuje tylko wycinek rzeczywistości, to problem ten dotyczy wszystkich uczelni, w których uprawiana jest dana dyscyplina, a nie tylko Polski. Zatem, skoro wszystkich dotyczą te same ograniczenia, to wyniki będą porównywalne.

Oczywiście w niektórych dyscyplinach konieczne byłoby, na co zwraca uwagę G. Marzec (Marzec, 2025), dopasowanie dyscyplin do kategorii używanych przez SCIMAGO. Wiemy, że takie dopasowanie już jest, bo powstało na potrzeby ostatniej ewaluacji, jednak nie zostało nigdy upublicznione ani tym bardziej poddane środowiskowej krytyce. Alternatywnie, można by po prostu przyjąć jedną ze stosowanych międzynarodowo klasyfikacji nauk – to rozwiązanie rodziłoby jednak nowe problemy.

Proponowana tu miara, którą można by mierzyć pozycję danej jednostki w stosunku do osiągnięć międzynarodowych, jest bardzo prosta: stosunek liczby tekstów opublikowanych w czasopismach zaliczanych do Q1 w rankingu SCIMAGO do liczby pozostałych tekstów opublikowanych w czasopismach indeksowanych na SCOPUS lub, gdybyśmy chcieli zaostrzyć kryteria, do tekstów publikowanych w czasopismach z pozostałych kwartyli w rankingu SCIMAGO.

Punktem odniesienia byłyby średnie wyniki koszyka najlepszych uczelni w danej dyscyplinie. Koszyk ten można by skonstruować, biorąc pod uwagę wyniki najlepszych uczelni indeksowanych w SCIMAGO Institutions Ranking, po jednej z każdego kraju. Taki dobór zapewniłby właściwe proporcje pomiędzy krajami Globalnej Północy, zwłaszcza anglosaskimi, mającymi w powszechnym przekonaniu łatwiejszy dostęp do czasopism na SCOPUS a krajami Globalnego Południa. Alternatywnie można by było ograniczyć się do krajów Unii Europejskiej. Porównywalibyśmy wtedy system polski z systemami krajów z podobnego kręgu kulturowego, o podobnych standardach publikacyjnych i zasadach finansowania tudzież oceny wyników badań. Odpowiednio ten sam zabieg należałoby przeprowadzić w odniesieniu do liczby grantów (kryterium 2) oraz wpływu na otoczenie (kryterium 3).

W zależności od wyniku takiego porównania, uczelnia otrzymywałaby w danym kryterium rangę A, B, C lub D. Uczelnia, która w danym kryterium uzyskaby wynik co najmniej odpowiadający średniej międzynarodowej, otrzymywałaby rangę A, pozostałe – odpowiednio niżej przy progach równych 75% (B), 50% (C) i 25% (D) średniej. Dodatkowo tę miarę można uzupełnić o wskazanie, jaki odsetek pracowników uzyskał ten wynik, używając gwiazdek do pokazania stopnia aktywności ocenianego zespołu (**** jeśli powyżej 75% pracowników publikuje na SCOPUS, *** jeśli ponad 50% i mniej niż 75%, ** jeśli ponad 25% a mniej niż 50% uzyskało ten wynik oraz * jeśli od 0 do 25% pracowników publikowało teksty na SCOPUS).

Ostatecznie ocena opisowa w każdym kryterium składałaby się z dwóch elementów. Pierwszy to liczba, określająca w którym kwartyli dana jednostka mieści się w porównaniu z innymi jednostkami krajowymi w danej dyscyplinie. Drugi to oznaczenia literowe, wskazujące na odległość jednostki od średniej międzynarodowej. Dodatkowo gwiazdki pozwalałyby pokazać strukturę aktywności publikacyjnej zespołu. Przykładowo bardzo dobra uczelnia, mająca przyzwoite osiągnięcia we wszystkich kategoriach i intensywnie publikujący zespół, mogłaby być opisana następująco: 1A1B2B****, a uczelnia przeciętna, z niewielką liczbą osób publikujących na SCOPUS: 2A, 3C, 2B*.

6. ROLA EKSPERTÓW

Istotą proponowanego tu modelu jest ograniczenie subiektywnej oceny do minimum. Wynika to z tego, że dotychczasowe doświadczenia z tworzeniem wykazów czasopism i wydawców, ustalaniem progów na kategorie i oceną kryterium 3 ewaluacji są niezbyt dobre. Towarzysząca im arbitralność zmniejszyła zaufanie do oceny środowiskowej, dlatego uważam, że co najmniej przez jeden okres ewaluacyjny należałoby ograniczyć się do oceny opartej na twardych miarach. Nie oznacza to, że jestem przeciwnikiem oceny eksperckiej, wręcz przeciwnie, uważam, że docelowo powinniśmy połączyć oba modele – algorytmiczny i ekspercki – w harmonijną i efektywną całość. Na obecnym etapie widziałbym rolę ekspertów w dwóch momentach: po pierwsze przy ocenie, czy tekst w czasopiśmie spoza dyscypliny mieści się w jej szeroko zakreślonych ramach, po drugie przy tworzeniu ewentualnej czarnej listy czasopism, które pomimo znalezienia się w wykazie nie powinny być brane pod uwagę. Jak taki wykaz mógłby wyglądać, postaram się pokazać w dalszej części artykułu.

Ocenie eksperckiej powinny podlegać także ośrodki, których wyniki znacznie odbiegają od przeciętnej w dyscyplinie in plus lub in minus. Naturalnymi kandydatami do takiej oceny byłyby jednostki uzyskujące ocenę 1A lub 4D we wszystkich kryteriach ewaluacji. Zadaniem ekspertów w pierwszym przypadku byłaby kompleksowa ocena jednostki oraz wypracowanie planu badań i rozwoju. Zadania te byłyby realizowane w ramach grantu blokowego dla dyscypliny, finansowanego ze środków ministra właściwego do spraw nauki. Granty te zastąpiłyby dotychczasowe finansowanie IDUB-ów. Natomiast w przypadku najgorszych ośrodków, ocena dotyczyłaby całości badań i miałyby na celu zarówno zidentyfikowanie przyczyn niepowodzenia ewaluacyjnego, jak i zaproponowanie środków zaradczych. Dodatkowo eksperci musieliby ocenić merytorycznie dorobek jednostki, żeby sprawdzić, czy nie wpadła ona w „czarną dziurę” przez to, że badania były dobre i publikowane we właściwych miejscach, ich specyfika jednak spowodowała, że w sposób naturalny ogłaszano je w czasopismach niższej rangi.

Docelowo należałoby wprowadzić instytucjonalną ocenę profilu badawczego uczelni obejmującą wszystkie uprawiane tam dyscypliny i działalność szkoły doktorskiej. Podstawą do opracowania metodyki mógłby być holenderski SEP i francuska metodologia stosowana przez HCRES. Ewaluacja taka, służąca jedynie poprawie jakości i niepociągająca negatywnych konsekwencji dla uczelni, byłaby przeprowadzana raz na 6–7 lat. Szczegóły tego przedsięwzięcia wykraczają jednak poza ramy niniejszego artykułu. Wypada jedynie zauważyć, że dobrym punktem wyjścia do wprowadzenia takiej oceny byłyby nie tylko wyniki refleksji nad recenzjami profili badawczych w poprzedniej i kolejnych ewaluacjach, ale także wyniki ewaluacji szkół doktorskich. Niezależnie od intencji ustawodawcy ocena szkół doktorskich jest mocno zbliżona do modelu eksperckiej oceny 360 tego, co dzieje się w danej dyscyplinie. Połączenie jej z oceną badań na uczelni wydaje się więc być dobrym pomysłem.

7. MECHANIZMY ZABEZPIECZAJĄCE PRZED NIEUCZCIWYMI PRAKTYKAMI PUBLIKACYJNYMI

Jednym z problemów, zidentyfikowanych podczas ostatniej ewaluacji i w obecnie trwającym okresie ewaluacyjnym, jest zjawisko, które moglibyśmy nazwać „optymalizacją punktową”. Pod tym pojęciem rozumiem działania ocenianych jednostek zmierzające do osiągnięcia jak najlepszego wyniku jak najmniejszym wysiłkiem. Sama optymalizacja nie jest niczym nagannym, uczelnie i instytuty dysponują ograniczonymi zasobami i muszą je racjonalnie spożytkować. Problem pojawia się natomiast, jeśli na źle skonstruowaną punktację czasopism i wydawnictw nałożą się niefrasobliwość w korygowaniu wykazu i przydzielonych punktów oraz naturalny dla kierownictw ocenianych jednostek strach przed utratą praw akademickich i obciążeniem finansowania. Ewaluacja w obecnym wydaniu jest bowiem grą hazardową o wysokiej stawce. Grą, w której próg wygranej ustalany jest przez bankiera dopiero po odsłonięciu kart przez wszystkich graczy, a gracz nawet z dobrą pulą kart, do samego końca nie wie, czy to wystarczy. Bankier może więc swobodnie wskazać zwycięzców, a także sumę wypłacanych im nagród. I, jak mówił jeden z bohaterów „Wiecha”, „pisz pan apelację do ślepej kieszki”.

Nic więc dziwnego, że oceniane jednostki starały się uzyskać jak największą liczbę punktów jak najniższym kosztem. Najbardziej widoczne w tym wyścigu były *Energies* i *Sustainability*, co skłoniło dość licznych autorów do zarzucania ich redakcjom nierzetelnych praktyk wydawniczych, także dlatego, że pojawia się tam wyjątkowo dużo polskich tekstów (Brzoza-Brzezina i in., 2022). Oczywiście nie tylko polscy uczeni mają wątpliwości co do praktyk wydawniczych stosowanych przez dużych wydawców open access (Ángeles Oviedo-García, 2021b, 2021a; Christos Petrou, 2023; David Nicholas i in., 2023), a argument, że jeśli twoje dochody zależą od liczby opublikowanych tekstów, będziesz gotów przymknąć oko i puścić teksty nieco słabsze, może być zasadny. Jednakże to nie było rozstrzygające. Jak się wydaje, zadecydowało to, że niektóre czasopisma otrzymały już w pierwszej wersji wykazu czasopism więcej punktów od ekspertów, niż wynikałoby to z ich wskaźników bibliometrycznych³. Niektóre tak przecenione czasopisma mają teraz nadmiar autorów z Polski, choć nikt nie zarzuca im nierzetelnych praktyk wydawniczych (pars pro toto: *Muzeologia*⁴ i *Præhistorische Zeitschrift*⁵). Minister Nauki nie zdecydował się nawet na pozbawienie punktów *European Research Studies Journal*, choć ten został skreślony ze SCOPUS-a za nietrzymanie standardów (Kieraciński, 2001; Thalassinou, 2001; Wroński, 2001, 2021). Dodatkowo, raczej prawem kaduka, dowartościowywano czasopisma krajowe i niektóre zagraniczne poza wykazami. W naukach prawnych na przykład skorzystały z tego niektóre jednostki wydające dobrze punktowane czasopisma. Teksty w tych czasopismach własnych stanowiły w niektórych przypadkach powyżej 50% puli

³ Obecnie wskaźniki te są znacznie wyższe i nie można mówić o przeszacowaniu tych czasopism.

⁴ *Muzeologia a Kulturne Deditstvo* jest słowackim czasopismem wychodzącym od 2013 r. W tym czasie autorzy ze słowacką afiliacją opublikowali w nim 120 tekstów, a polscy – 67. Jednakże według danych SCOPUS w obecnym okresie ewaluacyjnym polscy autorzy opublikowali 37 tekstów, a słowaccy zaledwie 9.

⁵ Jest to czasopismo z przeszło stuletnią historią, którego rzetelność nie budzi wątpliwości. Według danych ze SCOPUS w latach 2010–2025o publikowano w nim 111 tekstów z niemiecką afiliacją i 102 teksty z afiliacją polską. Natomiast od początku istnienia czasopisma, czyli od 1909 r. do roku 2009 opublikowano w nim 273 teksty z niemiecką afiliacją i tylko 15 tekstów polskich. Ten i poprzedni przypadek pokazują, jak sterowanie przez państwo punktami i preferencjami dotyczącymi publikacji „międzynarodowych” rozumianych jako zagraniczne, wpływa na decyzje publikacyjne uczonych.

zgłoszonych do ewaluacji artykułów. Dotyczy to zarówno dużych uniwersytetów, jak i małych szkół prywatnych⁶. Urban legend – o jednej z uczelni technicznych, która uzyskała kategorię A, publikując niemal wyłącznie w wysoko punktowanym, a przy tym mającym niski stopień odrzuceń zgłaszanych tekstów czasopiśmie technicznym, które przypadkiem zaliczono także do nauk ekonomicznych – łatwo zweryfikować. Nazw tych jednostek nie podaję celowo. Nie chodzi bowiem o „naming and shaming”, lecz o pokazanie pewnego zjawiska.

W samym publikowaniu u siebie nie ma oczywiście nic złego, w naukach prawnych jest to w pewnym sensie norma światowa. Problem pojawia się, jeśli publikacje domowe lub w czasopiśmie o niskim progu wejścia stanowią zbyt duży odsetek ewaluowanych publikacji. Taka optymalizacja skłania do poszukiwania łatwych rozwiązań przez oceniane jednostki. Dlatego proponuję wprowadzenie do systemu bezpiecznika – ograniczenia liczby tekstów z jednego czasopisma do powiedzmy 15% z 3N. Zmusi to uczonych do większego zróżnicowania portfela publikacji. Oczywiście nie jest to mechanizm bezbłędny. Kontrola nad trzema dobrze ocenianymi publikacjami pozwoli zapełnić 45% slotów, jednak mechanizm ten wydaje się stosunkowo bezpieczny, bo nie naraża organów władzy publicznej na zarzut ograniczania konkurencji przez arbitralne zakazy publikowania w konkretnych czasopismach.

Publikowanie poza dyscypliną jest innym środkiem polepszenia własnej pozycji ewaluacyjnej. Jak wiadomo, różne dyscypliny mają różne praktyki publikacyjne, co przekłada się także na wskaźniki bibliometryczne czasopism. Dlatego jednym z ryzyk ewaluacyjnych jest zgłaszanie do oceny tekstów tylko teoretycznie związanych z dyscypliną albo publikowanie poza własnym ekosystemem czasopism. Pierwotnie w ewaluacji próbowano temu zaradzić, wprowadzając ograniczenie kwotowe, co zostało słusznie oprotowane przez środowiska naukowe. Obecnie to eksperci wykluczają niepasujące do dyscypliny teksty, jeśli pojawiły się w czasopiśmie do niej niezaliczanym. I ten model, jak się wydaje, należałoby utrzymać. Dodatkowo można by go wzbogacić o system tworzenia czarnej listy czasopism. Niekoniecznie „drapieżnych” sensu proprio, bo te prawdopodobnie w ogóle nie znalazłyby się wśród ocenianych, ale takich, których standardy wydawnicze, choć formalnie poprawne, budzą wątpliwości środowiskowe. Mechanizm oceny powinien być dwutorowy. Przede wszystkim do wszystkich osób uprawnionych do wybierania członków komitetów nauk PAN⁷ należałoby rozesłać wykaz wszystkich czasopism i wydawnictw uwzględnionych w ewaluacji w danej dyscyplinie. Adresaci ankiety mogliby wskazać nie więcej niż 5 czasopism lub wydawnictw, które ich zdaniem powinny być usunięte z wykazu w kolejnej ewaluacji. W oparciu o tę ankietę należałoby wyłonić czasopisma i wydawców budzących wątpliwość przynajmniej połowy respondentów, a eksperci poddali by je analizie pod kątem spełniania kryteriów naukowej rzetelności. W przypadku oceny negatywnej wyniki analizy należałoby przedstawić wydawcy, dając mu możliwość ustosunkowania się do nich, a potem KEN podejmowałby uchwałę w przedmiocie umieszczenia czasopisma lub wydawcy na czarnej liście. Listę zatwierdzałby i publikował minister właściwy do spraw nauki.

8. WPŁYW EWALUACJI NA FINANSOWANIE I PRAWA AKADEMICKIE JEDNOSTKI

Proponowany tu system wymagałby odłączenia ewaluacji od finansowania algorytmicznego. Ocena opisowa nie przekłada się bowiem wprost na wartość nadającą się do umieszczenia w algorytmie. Proponuję zatem, by algorytm w części badawczej skonstruować tak, by uwzględniał tylko krajową pozycję naukową jednostki. Utrzymując dotychczasowy system przypisania osiągnięciom w każdym kryterium wag, bralibyśmy pod uwagę ocenę końcową (OK) daną wzorem $OK = W1 \cdot K1 + W2 \cdot K2 + W3 \cdot K3$, gdzie:

Kn = suma punktów uzyskanych w danym kryterium;

Wn = waga przypisana osiągnięciom w n-tym kryterium.

Następnie należałoby wyznaczyć dla tak ustalonego zbioru kwartyle i przyjąć, że jeśli osiągnięcia jednostki mieszczą się w 2. i 3. kwartylu, pracownikowi przeliczeniowemu w algorytmie przypisujemy wagę 1 i odpowiednio wagi 1,1 i 0,9 dla jednostek w pierwszym i czwartym kwartylu. W ten sposób jednostki, których wyniki oscylują wokół mediany, dostaną subwencję w podstawowej wysokości, te trochę lepsze 10% więcej, a te z ogona – 10% mniej. Liczba pracowników przeliczeniowych odpowiadałaby liczbie N. Nie ma potrzeby większego różnicowania, ponieważ wszystkie te jednostki prowadzą działalność naukową zgodnie z polityką państwa, a najlepsze mają możliwość uzyskania dodatkowego wyróżnienia w ramach oceny eksperckiej.

⁶ Korzystam z danych udostępnionych na potrzeby grantu przez Ministerstwo (wówczas) Nauki i Szkolnictwa Wyższego. Analiza struktury ewaluowanych publikacji w ewaluacji 2017-21 będzie przedmiotem osobnego tekstu. W publikowaniu u siebie nie ma nic nagannego, a wykorzystanie możliwości optymalizacyjnych jakie dały nieoczekiwane podwyżki punktów należy traktować jak optymalizację podatkową. Pamiętajmy też o elemencie loteryjnym związanym z punktacją – przykładowo, Białostockie Studia Prawnicze w ten sposób skoczyły z 20 na 140 punktów, a Opolskie Studia Administracyjno-Prawne z 40 na 140, by spaść z powrotem na 40. Przewidzenie tych zdarzeń było niemalże niemożliwe.

⁷ Zdaje sobie sprawę, że kryterium to jest arbitralne, jednak elektorzy mają wystarczający dorobek publikacyjny, jak również znajomość własnej dyscypliny, żeby dokonać rzetelnej oceny.

Ocena opisowa nie pozwala także na proste powiązanie wyniku ewaluacji z prawami akademickimi. W tym zakresie mam dwie propozycje *de lege ferenda*. Proponuję odejść od przyznania uczelniom powiązanej z wysoką kategorią autonomii dydaktycznej. Niepotrzebnie różnicuje to uczelnie, a ponadto w praktyce pojawia się problem, gdyż negatywna ocena przez PKA kierunku utworzonego bez zgody ministra powoduje automatycznie konieczność jego likwidacji, co wynika wprost z art. 246 prawa o szkolnictwie wyższym i nauce. Tymczasem kierunki utworzone za zgodą ministra mogą być prowadzone dalej mimo negatywnej oceny do czasu likwidacji w drodze decyzji administracyjnej. Proponuję zatem, by ujednolicić system i powrócić do jednolitych reguł tworzenia kierunków. Uczelnia uzyskiwałaby pozytywną opinię PKA i o ile minister nie zgłosi sprzeciwu – prawa do prowadzenia kierunku. Negatywna opinia wymagałaby rozstrzygnięcia sprawy przez ministra w drodze decyzji administracyjnej. Alternatywnie można by dać wszystkim ewaluowanym uczelniom pełną swobodę dydaktyczną, ale wtedy proponowane studia byłyby opiniowane przez PKA, która mogłaby zaproponować poprawki w efektach uczenia się i siatkach godzin.

Jeśli zaś chodzi o uprawnienia do nadawania stopnia doktora, to przy założeniu, że uczelnie publikujące w aprobowanych przez państwo czasopismach i wydawnictwach prowadzą działalność naukową, trudno ograniczać ich prawa do nadawania stopni ze względu na to, że mają taki czy inny profil publikacyjny. Proponuję natomiast zmodyfikować zasady w ten sposób, że prawo do nadawania stopni przysługiwałoby tylko tym uczelniom, które zatrudniają co najmniej 6 osób z aktualnym dorobkiem na stanowisku profesora lub profesora uczelni (w podstawowym miejscu pracy). Przy czym aktualny dorobek musiałby obejmować odpowiednią liczbę prac kategorii A lub B opublikowanych w poprzednim okresie ewaluacyjnym. Liczba publikacji zależałaby od dyscypliny, a reguła ta pozwalałaby na sprawne funkcjonowanie uczelni także wtedy, gdyby ustawodawca zdecydował się na zlikwidowanie habilitacji i tytułu profesora.

9. W KIERUNKU OTWARTEGO SYSTEMU EWALUACJI

Jak wskazałem wyżej, o sensowności procesu ewaluacyjnego decyduje kompletność zbioru lub zbiorów danych, na których operujemy, dokonując oceny. Jest to prawdopodobnie najpoważniejszy problem, z którym musimy się zmierzyć, zwłaszcza przy ocenie badań w zakresie nauk humanistycznych i społecznych. W dziedzinach innych niż nauki humanistyczne i część nauk społecznych (wyjątkami są np. nauki ekonomiczne, psychologia czy, choć może to być wątpliwe, nauki o polityce i administracji) publikowanie w czasopismach indeksowanych przez SCOPUS lub Web of Science jest zasadniczo standardem i nie ulega wątpliwości, że „everybody who is anybody” publikuje właśnie tam. Nadto inne niż czasopisma kanały publikacyjne są w tych dyscyplinach znacznie mniej popularne. Możemy więc przyjąć, że w przypadku tych dyscyplin obie wspomniane wyżej bazy dają pełny obraz rzeczywistości. Natomiast dla nauk humanistycznych i pozostałych nauk społecznych, teologii, bibliistyki, itd. otrzymujemy tylko wycinek szerszej całości. Wycinek oczywiście istotny, bo zawarty w uznanej bazie danych, jednak dający niepełny obraz, zwłaszcza że dysponenti obu głównych baz bibliometrycznych kierują się przy ich uzupełnianiu własnymi interpretacjami, regułami, a ich rozstrzygnięcia nie podlegają kontroli instancyjnej. Uznanie środowiskowe oczywiście równoważy te braki, jednak istotą ewaluacji powinno być uzyskanie pełnego obrazu tego, jak nauka polska (lub dla współczesnych purystów: nauka w Polsce) wygląda na tle zagranicznej konkurencji.

Będąc sceptycznym wobec wykorzystywania do tego celu Google Scholar, choćby ze względu na indeksowanie przezeń wątpliwej jakości źródeł, za punkt odniesienia proponuję przyjąć OpenAlex jako bazę tworzoną z myślą o rozwoju otwartej alternatywy dla SCOPUS i Web of Science. Baza ta jest niezależna od konkretnej grupy wydawniczej czy dostawcy usług zarządzania nauką, co czyni ją mniej podatną na konflikty interesów. Oczywiście nie przeszkadza to w równoległym przygotowaniu podobnych, testowych rozwiązań z użyciem Google Scholar czy Dimensions. Dotychczasowe próby wydają się pokazywać, że za pomocą Scholar można uzyskać podobne wyniki jak przy ocenie eksperckiej. OpenAlex ma natomiast wbudowane narzędzia bibliometryczne ułatwiające stworzenie wskaźników do celów ewaluacyjnych. Posiada przy tym dość szeroką bazę źródeł, obejmującą, co ważne w naukach humanistycznych czy społecznych, także książki i ich wydawców.

Jak się wydaje, OpenAlex mógłby być użyty do stworzenia wskaźników bibliometrycznych pozwalających na dość wiarygodne odtworzenie obrazu nauki polskiej na tle światowym. Nie wdając się w szczegóły dotyczące tworzenia poszczególnych miar, można przyjąć, że pierwotną powinna być cytowalność mierzona ilorazem liczby cytowań danego czasopisma lub wydawcy do liczby opublikowanych tekstów. Miary te można oczywiście mnożyć, próbując dostosować je do średniego wyniku dla dyscypliny czy liczebności uczonych publikujących w danym języku lub danej specjalności, zasada pozostaje jednak taka sama: czasopismom i wydawcom wyliczamy wskaźniki cytowalności i przypisujemy do poszczególnych kwartyli. W tym przypadku o punktach przypisanych poszczególnym tekstom decydować będzie to, w którym kwartyli jest dane czasopismo lub wydawca. Pozostałe elementy oceny pozostają bez zmian. Narzędzia tego da się też użyć do stworzenia zbliżonych do proponowanych tutaj miar wpływu na otoczenie.

Ocena testowa powinna być prowadzona w kolejnym okresie ewaluacyjnym z zachowaniem pełnej jawności i dostępności danych: najlepiej, by interesariusze wewnętrzni i zewnętrzni widzieli wyniki i mieli dostęp do procedur ich uzyskiwania. Uważam, że pierwszą ewaluację, tę za lata 2026–2029, należałoby przeprowadzić według proponowanej tu metody z wykorzystaniem SCOPUS-a i pozostałych baz, a równolegle powinno się prowadzić prace nad stworzeniem narzędzi do oceny jakości badań przy użyciu systemów otwartych. Na koniec tego okresu ewaluacyjnego trzeba by pokazać i pozwolić przetestować

zainteresowanym działające narzędzie bibliometryczne. W ewaluacji 2030–33 ocena powinna być prowadzona równolegle na starych i nowych zasadach, przy czym na bieżąco należałoby analizować ewentualne rozbieżności w wynikach. Dopiero ewaluacja 2034–2037 mogłaby się odbyć wyłącznie na nowych zasadach. Oczywiście nic nie stoi na przeszkodzie, żeby pełną ocenę ekspercką – peer review powiązany z ewaluacją szkół doktorskich – przeprowadzić już we wcześniejszym okresie, począwszy od roku 2030.

10. WNIOSKI

Proponowany tu system nie jest idealny. Pozwala jednak na ocenę, jak bardzo uczelnie krajowe różnią się między sobą i jak bardzo zbliżają się do lub oddalają od międzynarodowej średniej. Jak się wydaje, metodami ilościowymi nie można osiągnąć więcej przy założeniu, że reguły gry muszą być dostosowane do wszystkich dziedzin i dyscyplin. Dlatego ewaluacja robiona metodami ilościowymi musi być uzupełniana o ocenę jakościową, ekspercką, dokonywaną na dwóch poziomach: uczelni (dyscypliny) i poszczególnych pracowników. Ta pierwsza powinna się skupiać na właściwej organizacji pracy jednostki, dbałości o rozwój pracowników i strategicznych kierunkach badawczych. Obok tego oceniana powinna być praktyka publikacyjna z uwzględnieniem specyfiki dyscypliny i zgodności publikacji z profilem jednostki. Do przemyślenia jest połączenie jej z ewaluacją szkół doktorskich. Ocena indywidualna powinna być prowadzona tak na etapie ocen pracowniczych, jak i przy procedurach awansowych i nie powinna zasadniczo opierać się na wskaźnikach bibliometrycznych. Przypuszczalnie będzie wymagało to zmiany podejścia i uelastycznienia procedur awansowych tak, by odzwierciedlały one zasady Europejskiego Kodeksu Naukowca.

Od strony praktycznej pierwszym krokiem powinno być wprowadzenie w latach 2026–2029 ewaluacji opartej na proponowanym tu modelu. Równolegle należałoby przygotować i przetestować na historycznych danych ewaluacyjnych oraz na danych z nowego okresu ewaluacyjnego ewaluację z wykorzystaniem narzędzia OpenAlex lub równoważnego i opracować jej wariant pozwalający na ocenę jakości badań w oderwaniu od zawartości komercyjnych baz danych obejmujących konkretne ekosystemy wydawnicze. Do przygotowania zaś modelu ocen eksperckich należałoby wykorzystać doświadczenia nie tylko nowej ewaluacji i badania jakościowego najlepszych i najgorszych jednostek, ale także recenzje jednostek aspirujących do kategorii A+ z ewaluacji 2016–2021 oraz doświadczenia ewaluacji szkół doktorskich. Dobrą praktyką byłoby włączenie w ten proces jak najszerszej grupy interesariuszy oraz pełna jawność danych, pozwalająca niezależnym zespołom na wypracowanie własnych rozwiązań, będących alternatywą do proponowanego tu modelu lub jego uzupełnieniem.

Zasadniczym ryzykiem jest niestabilność polityki publicznej. Ewaluacje, akredytacje, a wcześniej ocena parametryczna prowadzone są w czteroletnich odstępach. Wpływają nie tylko na strategię uczelni, ale także, a właściwie przede wszystkim, na kariery pojedynczych uczonych. Dlatego też potrzebny jest szeroki konsensus, że taki ewolucyjny model oceny jakości badań jest tym, czego potrzebujemy. Podejście „dłużej klastora niż przeora”, normalne w świecie akademickim czy organizacjach gospodarczych długiego trwania, jest trudniejsze do akceptacji dla polityków. Niemniej jednak, przy wszystkich jego wadach, kolejne ekipy rządzące, niezależnie od ich agendy politycznej, były w stanie kontynuować i modyfikować (nie zawsze ulepszać) system wprowadzony przez minister Barbarę Kudrykę. A to daje nadzieję, tym bardziej, że zgłoszona tu propozycja nie prowadzi do zerwania z długoletnią i mocno zinternalizowaną już praktyką, a tylko modyfikuje ją, przenosząc akcent z dystrybuowania zasobów i prestiżu na zarządzanie jakością.

FINANSOWANIE

Tekst powstał w ramach projektu „Projakościowa ewaluacja nauk prawnych” dofinansowanego ze środków budżetu państwa w ramach programu Ministra Edukacji i Nauki pod nazwą „Nauka dla Społeczeństwa”, nr projektu (nr NdS/529966/2021/2021); kwota dofinansowania 276 920,00 zł, całkowita wartość projektu 276 920,00 zł. <http://www.ewaluacjaprawa.com> (odsłona 25.05.2025).

BIBLIOGRAFIA

Agenzia nazionale di valutazione del sistema universitario e della ricerca. (2023). <https://www.anvur.it/it/ricerca/riviste/elenchi-di-riviste-classificate>

Oviedo-García, M.Á. (2021a). Erratum: Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI), *Research Evaluation*, 30(3), 420. <https://doi.org/10.1093/reseval/rvab030>

- Oviedo-García, M.Á. (2021b). Journal citation reports and the definition of a predatory journal: The case of the Multi-disciplinary Digital Publishing Institute (MDPI). *Research Evaluation*, 30(3), 405–419. <https://doi.org/10.1093/reseval/rvab020>
- Brzoza-Brzezina, M., Kolasa, M., & Makarski, K. (2022). PPP: Punktacja, Parametryzacja, Patologia. *Gazeta SGH – Życie Uczelni*. <https://buwlog.uw.edu.pl/polskie-publikacje-w-czasopismach-mdpi-czy-to-na-pewno-sukces/>
- CERES Research School for International Development. (2025). *Publishers*. <https://ceresresearchschool.nl/valuation-system/publishers/>
- Petrou, C. (2023, September 18). Guest post—Reputation and publication volume at MDPI and Frontiers. *The Scholarly Kitchen*. <https://scholarlykitchen.sspnet.org/2023/09/18/guest-post-reputation-and-publication-volume-at-mdpi-and-frontiers-the-1b-question/>
- Cieśliński, J. L., & Stec, P. (2018). Jak wyodrębnić uczelnie badawcze nie degradując reszty? [How to single-out research universities without degrading the rest]. *PAUza Akademicka*, 2, 385.
- Consejo Superior de Investigaciones Científicas. (2025). *Scholarly Publisher Indicators*. <https://spi.csic.es/>
- Nicholas, D., Herman, E., Abrizah, A., Rodríguez Bravo, B., Boukacem-Zeghmouri, C., Watkinson, A., Świgoń, M., Xu, J., Jamali, H. R., & Tenopir, C. (2023). Never mind predatory publishers... what about 'grey' publishers? *El Profesional de la Información*. <https://doi.org/10.3145/epi.2023.sep.09>
- Garayová, L., Nagy, Z., Radvan, M., Staničić, F., Stec, P., Szilágyi, J. E., & Tubić, B. (2023). The evaluation of scholarly activities in the field of law from the viewpoint of the Central European region. *Publicationes Universitatis Miskolcensis. Sectio Juridica et Politica*, 41(1), 21–26. <https://doi.org/10.32978/sjp.2023.002>
- Górecki, M. (2019). Nauka równoległa. Jak wyleczyć polską akademię z jej wstydlivej choroby? *Forum Akademickie*, 10. <https://prenumeruj.forumakademickie.pl/fa/2019/10/kronika-wydarzen/nauka-rownolegla/>
- Wierczyński, G. (2020). Problemy oceny parametrycznej polskich czasopism naukowych z dziedziny nauk prawnych. *Państwo i Prawo*, 9, 3–22.
- Wierczyński, G. (2022). Wykaz czasopism naukowych z dziedziny nauk prawnych z perspektywy danych o cytowaniach dostępnych w bazie Google Scholar. *Państwo i Prawo*, 5, 3–15.
- Jamovi Core Team. (2025). *Jamovi* (Version 2.6) [Software]. The Jamovi Project. <https://www.jamovi.org/>
- Winczorek, J. (2022). Wykaz czasopism punktowanych MNiSW w dyscyplinie nauki prawne a bibliometria. Wyniki badania empirycznego. *Państwo i Prawo*, 5, 16–51.
- Kieraciński, P. (2001). Czy na pewno przejrzyste? *Forum Akademickie*, 5. <https://miesiecznik.forumakademickie.pl/czasopisma/fa-5-2021/powinieniem-oczekiwac-gratulacji%e2%80%a9/>
- Kulczycki, E. (2019). *Procedury ewaluacji czasopism, współczynniki wpływu i listy czasopism*. Uniwersytet Adama Mickiewicza.
- Kulczycki, E. (2021). Polityczna ingerencja w Wykaz czasopism naukowych. Dlaczego wykazy muszą być w rękach naukowców. *Nauka*, 39–77. <https://doi.org/10.24425/nauka.2021.136315>
- Marzec, G. (2025). Szachy czy warcaby. W co gramy? *Forum Akademickie*, 3, 2025. <https://forumakademickie.pl/szachy-czy-warcaby-w-co-gramy/>
- National Board of Scholarly Publishing. (2025). *ERIH+*. <https://kanalregister.hkdir.no/publiseringskanaler/erihplus/>
- National Board of Scholarly Publishing. (2025). *Norwegian register for scientific journals, series and publishers*. <https://kanalregister.hkdir.no/en/informasjonsartikler/the-national-board-of-scholarly-publishing>

- Skoczeń, B. (2025). O wartości polskiego modelu ewaluacji. *Forum Akademickie*, 4. <https://forumakademickie.pl/o-wartosci-polskiego-modelu-ewaluacji/>
- Stec, P. (2022). Best of the best of the best? How international are top international law journals really? *Teka Komisji Prawniczej PAN Oddział w Lublinie*, 14(1), 429–440. <https://doi.org/10.32084/tekapr.2021.14.1-35>
- Stec, P. (2023). Bibliometric analysis of top-ranked European law schools' research outputs: An east-west comparison. *Krytyka Prawa*, 15(1), 15–33. <https://doi.org/10.7206/kp.2080-1084.569>
- Stec, P. (2024). O potrzebie zmian przepisów o ewaluacji jednostek naukowych 2022-2025. *Prawo i Wiąż*, 51(4), 421–441. <https://doi.org/10.36128/PRIW.VI51.917>
- Szczygielski, K. (2023). Dlaczego minister Czarnek robił to, co robił. I co dalej z polską nauką? *Polityka*. <https://www.polityka.pl/tygodnikpolityka/nauka/2236212,1,dlaczego-minister-czarnek-robil-to-co-robil-i-co-dalej-z-polska-nauka.read>
- Thalassinios, E. (2001). Powinienem oczekiwać gratulacji. *Forum Akademickie*, 5. <https://miesiecznik.forumakademickie.pl/czasopisma/fa-5-2021/powinienem-oczekiwac-gratulacji%e2%80%a9/>
- Uchwała nr 1/2021 Komitetu Nauk Prawnych Polskiej Akademii Nauk z 15.02.2021 r. w sprawie wykazu i punktacji czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych. (2021). *Państwo i Prawo*, 3, 137–138.
- University of Limerick. (2024). *University of Limerick Publication List*. <https://www.ul.ie/media/46017/download?inline=>
- Wroński, M. (2001). „Polskie” numery nie zapewniają prestiżu. *Forum Akademickie*, 5. <https://miesiecznik.forumakademickie.pl/czasopisma/fa-5-2021/powinienem-oczekiwac-gratulacji%e2%80%a9/>
- Wroński, M. (2021). Jak zapłacić sloty? *Forum Akademickie*, 2. <https://miesiecznik.forumakademickie.pl/czasopisma/fa-2-2021/jak-zapelnic-sloty%e2%80%a9/>
- Wróblewski, A. K. (2013). Pozycja nauki polskiej w międzynarodowych rankingach. *Nauka*, 4. <https://journals.pan.pl/dlibra/publication/106336/edition/92105/content>