

PIOTR STEC

O potrzebie zmian przepisów o ewaluacji jednostek naukowych 2022–2025*

On the Need for Changes to the Rules of Evaluation of Research
Units 2022–2025

Abstract

The Polish Research Evaluation Exercise was created in 2018 as a tool for an impartial research assessment and allocation of block grants to universities. However, the system has been distorted by arbitrary political interference and poorly designed metrics. This paper proposes legislative changes that will make the system workable and return it to its „factory conditions”: an impartial system based on calculable, unbiased, and hard-to-manipulate metrics.

KEYWORDS: research assessment,
higher education, science,
bibliometrics, reform

SŁOWA KLUCZOWE: ewaluacja
badań, szkolnictwo wyższe, nauka,
bibliometria, reforma

PIOTR STEC, doktor habilitowany nauk prawnych, profesor Uniwersytetu
Opolskiego, ORCID – 0000-0003-3797-1321, e-mail: piotr.stec@wp.pl

* Tekst dofinansowany 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 o całkowita wartość projektu 276 920,00 zł.

1 | Wstęp

Jednym ze stałych elementów pejzażu naukowego w Polsce są dokonywane zasadniczo co cztery^[1] lata oceny jakości badań^[2] prowadzonych przez jednostki naukowe. Niezależnie od stałej krytyki rozwiązań służących ocenie jakości badań należy stwierdzić, że od niej nie uciekniemy. Kraje, w których nauka uprawiana jest na wysokim poziomie, z reguły mają jakieś zewnętrzne systemy pozwalające ocenić jakość prowadzonych badań^[3]. Rozwiązania szczegółowe się różnią: Wielka Brytania ma system ekspercki, Włosi mieszany (niektóre dyscypliny oceniane są bibliometrycznie, inne – ekspercko)^[4], itd.

Przedmiotem niniejszego artykułu są propozycje zmian w systemie ewaluacji jednostek naukowych prowadzące do nadania mu racjonalnego kształtu umożliwiającego przeprowadzenie rzetelnej oceny za lata 2022-2025.

¹ Wyjątkiem był okres 2017-2021 obejmujący 5 lat ze względu na problemy związane z epidemią COVID-19.

² Grażyna Praweńska-Skrzypek, Justyna Maciąg, „Rozdział IV. Główne nurty krytyki ewaluacji polityki naukowej i innowacyjnej oraz sposoby jej doskonalenia na przykładzie wybranych krajów”, „Rozdział IV. Główne nurty krytyki ewaluacji polityki naukowej i innowacyjnej oraz sposoby jej doskonalenia na przykładzie wybranych krajów”, [w:] *Ewaluacja w procesie tworzenia i realizacji polityki naukowej i innowacyjnej*, red. Grażyna Praweńska-Skrzypek (Warszawa: Elipsa, 2017), 303-384.

³ Grzegorz Wierczyński, „Problemy oceny parametrycznej polskich czasopism naukowych z dziedziny nauk prawnych” *Państwo i Prawo*, nr 9 (2020): 3-22.

⁴ Gianfranco Rebori, Matteo Turri, „The UK and Italian Research Assessment Exercises Face to Face” *Research Policy*, nr 9 (2013): 1657-1666. <https://doi.org/10.1016/j.respol.2013.06.009>; Massimo Franceschet, Antonio Costantini, „The First Italian Research Assessment Exercise: A Bibliometric Perspective” *Journal of Informetrics*, nr 2 (2011): 275-291. <https://doi.org/10.1016/j.joi.2010.12.002>; Aldo Geuna, Matteo Piolatto, „Research Assessment in the UK and Italy: Costly and Difficult, but Probably Worth It (at Least for a While)” *Research Policy*, nr 1 (2016): 260-271. <https://doi.org/10.1016/j.respol.2015.09.004>; Katharine Barker, „The UK Research Assessment Exercise: The Evolution of a National Research Evaluation System” *Research Evaluation*, nr 1 (2007): 3-12. <https://doi.org/10.3152/095820207X190674>; Lewis Elton, „The UK Research Assessment Exercise: Unintended Consequences” *Higher Education Quarterly*, nr 3 (2000): 274-283. <https://doi.org/10.1111/1468-2273.00160>.

2 | Polski model ewaluacji jednostek naukowych

Zasadniczo sposób oceny jednostek naukowych pozostaje stały co najmniej od czasów minister Barbary Kudryckiej, odpowiedzialnej za wprowadzenie pierwszej po 1989 kompleksowej reformy systemu nauki i szkolnictwa wyższego, obejmującej także ocenę jakości badań opartą w przeważającej mierze na algorytmicznej ocenie osiągnięć jednostek^[5]. Reforma ministra Gowina^[6] poza kosmetyczną w istocie zmianą nazwy całego procesu z “parametryzacji” na „ewaluację” wprowadziła szereg nowości mających w założeniu uczynić system oceny bardziej efektywnym i przejrzystym. Po pierwsze, grupy wspólnej oceny zostały zastąpione przez dyscypliny. Po drugie, poszerzono zakres baz referencyjnych o SCOPUS oraz (w ograniczonym zakresie) ERIH+. Po trzecie, zrezygnowano z miar potencjału naukowego (nadane stopnie, stopnie posiadane przez kadre) i wprowadzono jako kryterium eksperckie wpływ badań na otoczenie. Zmiany te miały być w założeniu projakościowe i zachęcić jednostki naukowe, zwłaszcza w naukach społecznych, humanistycznych i teologicznych do wyjścia poza krajowe podwórko. Niestety, z przyczyn, o których niżej, mimo szlachetnych intencji całość doprowadziła do tych skutków, którym zamierzała zapobiec: naukom humanistycznym i społecznym nadal nie opłaca się umiędzynarodowienie, a miary, które miały wprowadzić nas w międzynarodowy obieg naukowy stały się celem samym w sobie. Nie znaczy to oczywiście, że cały system jest zły. Wręcz przeciwnie, jednym z jego skutków jest rosnąca liczba polskich czasopism na SCOPUS, co wiąże się ze zwiększoną widocznością polskiej nauki, choć niekoniecznie na poziomie, jakiego oczekiwali Minister Nauki czy opinia publiczna^[7].

⁵ *Oceny nauki: Tomaszowice, 16-18 listopada 2013, Debaty PAU, t. I, red. Szczepan Biliński* Kraków: Polska Akademia Umiejętności, 2014; Emanuel Kulczycki, Marcin Korzeń, Przemysław Korytkowski, „Toward an Excellence-Based Research Funding System: Evidence from Poland” *Journal of Informetrics*, nr 1 (2017): 282-298, <https://doi.org/10.1016/j.joi.2017.01.001>.

⁶ Ucieleśniona w postaci tzw. „Ustawy 2.0” albo „Konstytucji dla nauki” tj. ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce, tekst jednolity Dz. U. 1018, poz. 1668 t.j. Dz. U. z 2023 r. poz. 742, 1088 ze zn., dalej jako p.s.w.i.n.

⁷ Obecnie Scimago Journal Rank oparty na SCOPUS indeksuje 566 polskich czasopism, 770 hiszpańskich, 637 włoskich, 542 francuskie. Mieścimy się więc w europejskim peletonie, choć daleko nam do 1510 czasopism niemieckich w tym samym wykazie nawet, jeśli przyjąć, że chodzi o kraj z dwa razy większą niż Polska populacją. Dane wg. www.scimagojr.com [dostęp: 16.04.2024].

3 | Punkty krytyczne w procesie ewaluacji

3.1. Wykazy czasopism i wydawnictw

Pierwszym, i to dość oczywistym punktem krytycznym w ewaluacji jest konstrukcja wykazu czasopism i wydawnictw punktowanych. Zgodnie z przepisami rozporządzenia w sprawie wykazu czasopism i wykazu wydawnictw^[8] ustalanie punktacji tych ostatnich jest stosunkowo proste – eksperci decydują, czy wydawnictwo jest dobre, czy bardzo dobre.

Zgodnie z § 2 pkt 1 r.wyd do podstawowej grupy wydawców zaliczamy tych, którzy wydają publikacje recenzowane i zobowiązują się przestrzegać zasad etycznych Komitetu do spraw Etyki Publikacyjnej COPE. Wydawcy poziomu 2 to zgodnie z § 2 pkt 2 r.wyd muszą dodatkowo spełniać łącznie cztery kryteria: publikowanie monografii wnoszących istotny wkład w rozwój światowej nauki,^[9] prowadzą politykę wydawniczą zmierzającą do upowszechniania publikacji w skali światowej^[10], są uznawane za wiodące w opinii ekspertów tudzież w wykazach klasyfikujących wydawnictwa,^[11] stosują jednakowe zasady przyjęcia do druku książek niezależnie od wniesienia opłaty za publikację. Już na pierwszy rzut oka widać, że w ewidencji na poziomie 1 zmieści się praktycznie każde wydawnictwo mające pretensje do naukowości, podczas gdy poziom 2 jest zastrzeżony dla nielicznej grupy, wybieranej według nieznanymi bliżej kryteriów szczegółowych, bo uzasadnienia wpisu na jedną bądź drugą listę nie są publicznie dostępne. Nadto, przynajmniej w naukach prawnych, sporo międzynarodowych wydawnictw z poziomu 1 spełnia w mojej ocenie kryteria dla poziomu 2.

⁸ Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 12 lutego 2020 r. w sprawie sporządzania wykazów wydawnictw monografii naukowych oraz czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych, Dz. U. z 2018 r., poz. 2152, t. j. Dz. U. z 2020 r., poz. 349.

⁹ Złośliwy interpretator tego przepisu wskazałby zapewne, że każde wydawnictwo, które opublikowało osiągnięcie habilitacyjne w postaci monografii spełnia ten warunek, bo osiągnięcia takie muszą stanowić *ex lege* istotny wkład w rozwój dyscypliny.

¹⁰ Rozporządzenie nie zawiera żadnych wskazówek odnośnie do tego, jak badać istnienie i treść tych polityk. Wypada więc przyjąć, że publikowanie po angielsku, w otwartym dostępie lub sprzedaż via Amazon i podobne platformy spełnia ten warunek.

¹¹ Dziwi, że prawodawca niemający problemu ze wskazaniem bibliometrycznych baz referencyjnych dla czasopism nie był w stanie wskazać wykazów referencyjnych dla wydawnictw.

Jednocześnie analiza wykazu wydawców pokazuje trzy podstawowe grupy problemów związanych z jego tworzeniem.

Po pierwsze, w podstawowej grupie wydawców znajdują się znani międzynarodowi gracze oraz wydawnictwa krajowe o różnym poziomie naukowym. Po drugie, grupa wydawców z poziomu drugiego jest dość niejednolita, obejmuje multidyscyplinarnych wydawców oraz wydawnictwa o niszowych specjalnościach. W ten sposób szanse publikacyjne danej dyscypliny zależą częściowo od tego, ile wydawnictw niszowych znalazło się w wykazie na poziomie 2. Po trzecie wykaz nie uwzględnia niektórych znanych i renomowanych wydawnictw.

Nie lepiej jest z wykazem czasopism, choć tu sytuacja byłaby nieco mniej skomplikowana, gdyby potraktować serio założenia reformy ministra Gowina i oprzeć punktację czasopism naprawdę na wskaźnikach bibliometrycznych. Zgodnie z rozporządzeniem do wykazu wchodzi czasopisma ujęte w bazach SCOPUS i Web of Science oraz, dodatkowo, niektóre czasopisma indeksowane w ERIH+ oraz inne czasopisma o uznanej renomie włączone do wykazu przez ekspertów. System ten, na pozór obiektywny i łączący algorytmiczny przydział punktów z oceną ekspercką, traci swój obiektywizm już na początku procesu. Na pierwszym etapie czasopismom przypisuje się punktację zależną od centyla, w którym się znajdują przy uwzględnieniu odpowiedniego wskaźnika bibliometrycznego. Ponieważ punktacja ma odpowiadać znaczeniu czasopisma we wszystkich dyscyplinach, do których je przypisano, ostateczny wynik punktowy to średnia z cząstkowych ocen dyscyplinarnych^[12]. Różnice w ocenach cząstkowych mogą znacznie obniżyć ostateczny wynik, prowadząc do degradacji przyzwoitego czasopisma do parteru.

Przykładem problematyczności takiego procesu punktowania są czasopisma prawnicze i z zakresu innych nauk humanistycznych i społecznych, którym obok podstawowej (a czasem jedynej) dyscypliny dodano przynależność do nauk ścisłych lub przyrodniczych obniżając ich punktację. Kolejnym – arbitralnie dokonywane zmiany w punktacji czasopism^[13].

¹² Pomijam tu przypadki, w których dowartościowano czasopismo punktami *ex gratia* za szczególne w opinii ekspertów znaczenie dla dyscypliny.

¹³ „Uchwała nr 1/2019 Komitetu Nauk Prawnych Polskiej Akademii Nauk z 2.10.2019 r. w sprawie koniecznych i pilnych zmian dotyczących listy naukowych czasopism prawniczych” *Państwo i Prawo*, nr 11 (2019): 141-143; „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” *Państwo i Prawo*, nr 3 (2021): 137-138.

Ostatnie z nich wywołały tak potężny sprzeciw środowiska, że ich cofnięcie stało się nawet (spełnioną) obietnicą wyborczą Platformy Obywatelskiej^[14]. Trzeba zaznaczyć, że zmiany te, choć dokonywane w trudnym do akceptacji trybie zmierzały do wyrównania pozycji tych ośrodków, zwłaszcza w naukach humanistycznych i społecznych, które publikowały tradycyjnie w periodykach krajowych lub niżej punktowanych acz uznanych zagranicznych. Dlaczego ministerstwo zdecydowało się na taki „tryb doraźny”, zamiast po prostu zmienić przepisy i przeprowadzić aktualizację wykazu, pozostaje zagadką. Niezrozumiałe jest też, dlaczego KEN, przygotowując ostatnią propozycję wykazu czasopism, nie zdecydował się na aktualizację punktacji czasopism na SCOPUS i WoS: niektóre z nich mają już nowe wskaźniki, inne weszły do baz i powinny być oceniane na nowych zasadach. Dotyczy to m.in. kilkunastu polskich czasopism prawniczych, a także „Humanities and Social Sciences Communications”, notowanego w 97 percentylu SCOPUSa dla General Arts and Sciences, a figurujących w wykazie pod starym tytułem Palgrave Communications i z oszałamiającą liczbą 20 pkt. Nie jest to oczywiście przypadek jedyny – „Heliyon”, „Crime Science” i „Sage Open” znajdują się w podobnej sytuacji.

3.2. Kryterium II ewaluacji

Obecnie kryterium II ewaluacji pełni trojaką funkcję. Po pierwsze, stanowi miarę prestiżu: przychody z grantów prestiżowych (europejskich i międzynarodowych) mają większą wagę niż przychody z tytułu grantów krajowych. Po drugie, w pewnym sensie stanowią miarę jakości badań: uzyskanie grantu łączy się z dodatkową oceną propozycji badawczej przez zewnętrznych recenzentów. Po trzecie, pozyskiwanie środków zewnętrznych tak z grantów, jak i w ramach prac zleconych jest popieraną przez rząd metodą finansowania działalności badawczej, która coraz mniej będzie zależeć od subwencji. Wszystkie trzy cele co do zasady są nie tylko akceptowalne, ale można uznać, że stanowią mechanizm sterowania nauką – mocną zachętę do pozyskiwania środków zewnętrznych.

Trzem funkcjom pełnionym przez kryterium drugie towarzyszą też, zupełnie przypadkowo, trzy poważne mankamenty podważające jego sensowność w obecnym kształcie.

¹⁴ *Propozycje Programowe Koalicji Obywatelskiej*. [dostęp: 16.04.2024]. <https://100konkretow.pl/edukacja/>.

Po pierwsze, jako jedynne kryterium ewaluacyjne ma ono charakter otwarty, to znaczy, że nie jesteśmy w stanie nie tylko wyliczyć ani nawet zaprognozować z pewnym prawdopodobieństwem maksymalnych wartości jakie może ono osiągnąć. W przypadku kryterium I i III maksima takie są do wyliczenia i wynikają z ograniczeń co do liczby osiągnięć jakie mogą być zgłoszone do ewaluacji oraz maksymalnej punktacji za dane osiągnięcie. W przypadku kryterium II ewaluacji znane są tylko wagi punktowe przypisane poszczególnym kategoriom przychodów, nie znamy natomiast maksymalnej liczby osiągnięć jakie można zgłosić w tej kategorii. Tym samym przypadkowe osiągnięcia wykraczające znacznie poza to, czego można się spodziewać (wartości skrajne, *Outliers*) będą decydowały o progach punktowych dla dyscypliny. W dyscyplinie takiej jak prawo, gdzie przeciętny grant nie przekracza 300-400 tysięcy złotych, uzyskanie przez jakąś jednostkę grantu na milion euro podniesie poprzeczkę na kategorie o lata świetlne.

Po drugie, kryterium II w jego obecnym kształcie zachęca do niegospodarnego planowania wydatków i niepotrzebnego mnożenia kosztów. W przypadku dwóch jednostek mających takie same osiągnięcia w kryterium I i III, a różne w kryterium 2, jako „lepsza” badawczo zostanie oceniona ta, która wydała więcej pieniędzy, by osiągnąć ten sam efekt. Jeśli dodatkowo przyjmiemy, że w niektórych dyscyplinach mamy podział na badania teoretyczne i eksperymentalne (fizyk-teoretyk jest mniej kosztochłonny niż fizyk potrzebujący do swoich badań zderzacza hadronów, badania cywilisty są mniej kosztowne niż kryminologa czy kryminalistyka), obecna konstrukcja prowadzi do rezultatów niesprawiedliwych także w obrębie tej samej dyscypliny – liczba uzyskanych punktów może zależeć nie tyle od jakości, co od profilu badawczego jednostki.

Po trzecie, nacisk położony na uzyskiwanie funduszy ze środków zewnętrznych przekłada się w pewnym stopniu na funkcjonowanie grantodawców, takich jak Narodowe Centrum Nauki. Wprowadzenie jako jednego z alternatywnych kryteriów wymogu kierowania lub bycia wykonawcą grantu w postępowaniach awansowych i konsekwentne uwzględnianie przychodów z tego tytułu najpierw w przepisach o parametryzacji, a potem ewaluacji badań, w oczywisty sposób stanowią bodziec do starania się o znaczne środki zewnętrzne nawet tam, gdzie wystarczałyby mały grant, albo wprost – środki z subwencji. Oznacza to, że nawet gdyby finansowanie NCN utrzymać na poziomie uwzględniającym inflację i ogólny wzrost kosztów, rosnąca liczba wniosków grantowych i tak doprowadziłaby do znacznego obniżenia współczynnika sukcesu, być może nie do dzisiejszych 10%, a do 15%-20%, jednakże zmiana i tak byłaby znacząca.

3.3. Kryterium III

Kryterium III w jego obecnej postaci jest w moim przekonaniu niemal niereformowalne. Jedyne, co może uratować jego wiarygodność to kompetentni recenzenci i system odwoławczy pozwalający na odniesienie się merytoryczne zarówno do oceny jednostki, jak i zarzutów stawianych osiągnięciom. Jednakże i tu mamy do czynienia z punktami krytycznymi, bez uwzględnienia których kryterium to będzie bezwartościowym poznawczo, za to mającym ukrytą siłę sprawczą narzędziem pozwalającym dostosować wyniki ewaluacji do potrzeb tego, kto ją kontroluje.

Pierwotnie kryterium III badało swoisty zasięg geograficzny wpływu badań na społeczeństwo – najwyżej punktowano osiągnięcia międzynarodowe, niżej – krajowe i regionalne. Taka bardzo prosta konstrukcja została uznana za niewystarczającą po badaniu pilotażowym poprzedzającym ewaluację¹⁵. W konsekwencji miary te wzbogacono dodatkowo o element jakościowy, jakim jest znaczenie wpływu. Niestety ustawodawca nie wskazał nawet w niebędących źródłami prawa podręcznikach dla ekspertów, jakimi kryteriami należy się kierować przy ocenie tego jakościowego elementu. Mogło to prowadzić do rozbieżnych ocen tego samego osiągnięcia będącego wynikiem wspólnych badań kilku ośrodków. Nadto, ponieważ decydujące znaczenie, jeśli chodzi o punktację, miała ocena dokonywana przez osobę niereprezentującą danej dyscypliny, samo niezrozumienie rządzących nią zasad mogło wpływać na punktację. Ktoś przyzwyczajony, że we wniosku patentowym cytuje się znaną literaturę przedmiotu, automatycznie założy, że tak samo postąpią legislatorzy przy pisaniu uzasadnienia projektu ustawy, itd. itp. Nie można też wykluczyć pokusy „ręcznego” poprawienia pozycji punktowej dobrego ośrodka, który ze względu na źle ustawione wagi w kryterium 1 wypadł słabo przez podrasowanie jego wyniku w kryterium 3, tudzież obniżenia w ten sam sposób wyników konkurenta, któremu za dobrze poszło w kryterium 1.

¹⁵ Przemysław Korytkowski, Emanuel Kulczycki, *Raport z pilotażu procedur oceny wpływu społecznego* (2020). <https://doi.org/10.6084/M9.FIGSHARE.12402686>.

3.4. Zespoły małe i jednostki nowe

Reforma ministra Gowina, skupiając się na sprawach wielkich, pominęła kilka spraw małych z perspektywy całości ustawy, jednak newralgicznych z perspektywy stosujących ją jednostek. Przebudowując system tak, żeby było nowoczesnie, sprawiedliwie i światowo, zapomniano bowiem o dwóch kategoriach: nowych zespołach i jednostkach badawczych oraz o małych grupach uczonych niemogących stworzyć ewaluowanego zespołu w ramach dyscypliny.

Praktycznie każda uczelnia i instytut badawczy zatrudnia pewną liczbę osób, których badania nie mieszczą się w głównym nurcie ewaluowanych dyscyplin. Na uniwersytetach są to na przykład prawnicy kanoniści, wykładający tę dyscyplinę teologom i prawnikom, którzy powinni znać nie tylko prawo świeckie, ale też rzymskie i kanoniczne, choćby po to, żeby umieć poruszać się w więcej niż jednym zbiorze reguł. Na uczelniach niebędących klasycznymi uniwersytetami, jak uczelnie medyczne, techniczne czy ekonomiczne, wykładane są matematyka, prawo, nauki o polityce, historia itd. Osoby prowadzące te zajęcia są co do zasady czynne naukowo. Zazwyczaj też nie można ich przypisać do dyscyplin ewaluowanych, trudno bowiem z badacza systemów partyjnych zrobić np. inżyniera mechanika o specjalności „trybologia polityczna”. Praktyka dotychczasowa poszła w takim kierunku, że tych mniejszościowych pracowników próbuje się jednak na siłę przypisać choć w części do dyscyplin ewaluowanych, względnie przenosić na etaty dydaktyczne. Ci, którzy zostają na etatach badawczo-dydaktycznych mają mniejsze finansowanie, bo ich pracy (nieewaluowanej przecież) algorytm subwencji nie widzi. Są więc skazani na wegetację, jeśli nie wprost na wymarcie.

Ustawodawca jak wiadomo jest racjonalny, choć nie zawsze rozsądny. Wypada więc przyjąć, że w obu wskazanych wyżej przypadkach kierował się on jakimś zamysłem politycznym. Pytanie: jakim? Jedyne chyba możliwe wyjaśnienie jest takie, że ustawodawca zamierzał wprowadzić system bodźców negatywnych do tworzenia nowych jednostek i zespołów badawczych. Takimi bodźcem jest ta okoliczność, że nowo tworzony zespół przez kilka lat będzie w praktyce pozbawiony środków nie tylko na rozwój, ale i funkcjonowanie. Pośrednio przepisy te utrudniały tworzenie w uczelniach akademickich kierunków lekarskich i lekarsko-dentystycznych, ze względu na wymóg posiadania co najmniej kategorii B+ w naukach medycznych lub naukach o zdrowiu. Utrzymywanie zespołów prowadzących działalność badawczo-dydaktyczną w dyscyplinach pomocniczych,

takich jak np. prawo na uczelniach medycznych, technicznych i ekonomicznych, również staje się nieopłacalne i uczelnie otrzymywały mocny bodziec do przenoszenia reprezentantów tych dyscyplin na etaty dydaktyczne bądź zastępowania ich zleceniobiorcami. Konsekwentne stosowanie tych bodźców negatywnych spowodowałoby petryfikację krajobrazu naukowego – zbudowanie nowego ośrodka konkurencyjnego wobec pozostałych łączyłoby się z trudnościami odstrasżającymi od podejmowania takich zadań. W przypadku nauk medycznych uczelnie musiałyby podejmować olbrzymie nakłady inwestycyjne nie mając pewności uzyskania pozwolenia na prowadzenie kierunku ani tym bardziej na zwrot z inwestycji pozwalający na utrzymanie z subwencji zespołu czysto badawczego. Utworzenie nowego instytutu badawczego oznacza wegetację do pierwszej ewaluacji, a jeśli instytut powstał pod koniec okresu ewaluacyjnego, mógł po prostu nie mieć szansy na osiągnięcie przyzwoitego wyniku. Trudno od jednostki zaczynającej działać w czwartym czy piątym roku ewaluacji, by miała takie wyniki, jak okrzepłe ośrodki. Ten przykład nie jest czysto teoretyczny – w opisanym tu sytuacji znalazł się nowoutworzony Międzynarodowy Instytut Mechanizmów i Maszyn Molekularnych (IMOL) PAN, który pomimo zatrudniania najwyższej klasy uczonych ma kategorię C, bo po prostu nie byli w stanie „wyprodukować” odpowiedniej liczby punktów w bardzo krótkim czasie^[16].

Takie odczytanie polityki państwa, jeśli chodzi o utrudnienie tworzenia, często na wyrost, nowych ośrodków, nie wytrzymuje jednak konfrontacji z późniejszymi działaniami ustawodawcy, który poluzował rygory do tworzenia kierunków lekarskich i lekarsko dentystycznych tak, że pozwolenie może uzyskać uczelnia niemająca kategorii A lub A+ lub B+ w dyscyplinie nauki medyczne lub naukach o zdrowiu. Co więcej praktyka udzielania zezwoleń na otwarcie takich kierunków była tak liberalna, że zaprzeczała restryktywnej polityce wynikającej z ustawy^[17]. Należy też zauważyć, że nowoutworzona Szkoła Główna Mikołaja Kopernika uzyskała ona w drodze fikcji prawnej kategorię B+ we wszystkich uprawianych w niej dyscyplinach.

¹⁶ Dziękuję w tym miejscu dyrektorce IMOL PAN, prof. Agnieszce Chacińskiej za udostępnienie danych ewaluacyjnych instytutu. Dane te zostaną wykorzystane niżej do pokazania, jak proponowane zmiany w ewaluacji nowych jednostek mogą wpłynąć na ich byt.

¹⁷ Piotr Wasilewski, Agnieszka Kalita, Małgorzata Szczęsna, „Kształcenie kadry medycznej – problemy w ochronie zdrowia” *Kontrola Państwowa*, nr 6 (2023): 708-711. <https://doi.org/10.53122/ISSN.0452-5027/2023.1.36>.

4 | Propozycje korekty systemu

Po zidentyfikowaniu punktów krytycznych obecnego systemu ewaluacji możemy przejść do propozycji zmian. Zaznaczam tu ponownie, że proponowane zmiany mają charakter doraźny. Jeśli ustawodawca zdecyduje się na utrzymanie obecnego systemu, będą one mogły być przystosowane do potrzeb ewaluacji 2026-2029. Jeśli nie, mam nadzieję, że przyczynią się one do nadania obecnej ewaluacji bardziej racjonalnego charakteru.

4.1. Wykaz wydawców i wykaz czasopism

Jak rozwiązać problem punktacji czasopism i wydawnictw bez naruszania zasad, na których oparta jest ewaluacja? Należy pamiętać, że z uwagi na to, że jesteśmy w połowie okresu ewaluacyjnego, jakkolwiek dalej idąca ingerencja naruszałaby zasadę zaufania obywatela do państwa a mówiąc mniej uczenie – byłaby po prostu nie *fair* wobec uczelni i uczonych, którzy dostosowują swoją działalność naukowo-badawczą do wytycznych rządowych.

Rozwiązanie tak postawionego problemu nie jest proste, zważywszy, że obecnie obowiązuje (trochę na zasadzie *hoc volo, sic iubeo*) nowy wykaz przygotowany według nieznanych kryteriów i wiążący na okres 2024-2025 oraz stary, przygotowany przez ministra Czarnka, obowiązujący w okresie 2022-2023. Zróżnicowanie wykazów wynika obecnie jedynie z treści komunikatu ministra, który to komunikat nie jest źródłem prawa powszechnie obowiązującego. Nie ma też w przepisach ustawy ani rozporządzenia ewaluacyjnego podstawy prawnej dla takiego podziału. Należy przy tym wskazać, że przepisy te nie rozstrzygają wyraźnie, czy zmiany w punktacji czasopism lub wydawnictw obowiązują jedynie na przyszłość, czy w całym okresie ewaluacyjnym. Dotychczas stosowana wykładnia prowadziła do wniosku, że najnowszy obowiązujący wykaz dotyczy całego okresu ewaluacyjnego, wypierając jako nowszy wszystkie starsze wykazy.

Należy przy tym zauważyć, że kontrowersje wokół nieoczekiwanych zmian punktacji dotyczyły niemal wyłącznie czasopism nieznajdujących się w SCOPUS, Web of Science czy ESCI. Mówiąc wprost – głównie czasopism polskich i nielicznych innych wprowadzonych do wykazu w trybie „eksperckim.” Czasopisma znajdujące się w bibliometrycznych bazach danych nie były w większości „dowartościowywane” przez ministra. Te zaś, które dowartościowano, są nieliczne i przeliczenie ich punktacji zgodnie z regułami gry nie powinno zaburzyć systemu oceny. Jeśli zaś chodzi

o czasopisma polskie, tu sprawa jest znacznie bardziej skomplikowana, bo w ich przypadku rozporządzenie przewidywało pierwotnie najniższą liczbę punktów z możliwością podniesienia przez ekspertów o próg lub dwa. Stąd też żadne polskie czasopismo niemające wyliczonych wskaźników bibliometrycznych nie powinno mieć więcej niż 70 pkt. Niektóre mają także obecnie 100 lub 140 pkt, a w wykazie ministra Czarnka – nawet 200. Wprowadzenie zmian w tej części wykazu znacznie zaburzy już zaburzoną nieoczekiwanymi zmianami strukturę punktową publikacji w poszczególnych dyscyplinach, głównie społecznych, humanistycznych i teologicznych. Dla tych dyscyplin publikacje krajowe stanowią zazwyczaj trzon przedstawianego do oceny dorobku, każda taka zmiana będzie więc dla nich istotna.

Jeśli chodzi o czasopisma uwzględniane na SCOPUS/WoS, to jak się wydaje wystarczy uproszczenie mechanizmu tworzenia wykazu. Ponieważ mówimy tu o rozwiązaniu tymczasowym, musi on być prosty do wprowadzenia, przejrzysty i niebudzący wątpliwości co do sposobu wykonania obliczeń.

Głównymi bazami indeksującymi czasopisma są SCOPUS oraz Web of Science (WoS) i Emerging Sources Citation Index będący wraz z WoS produktem Clarivate. Do tego dochodzą tworzące osobną grupę czasopisma krajowe objęte programem wsparcia, czasopisma nieobjęte programem, ale uznane przez ekspertów za wartościowe i nieliczne czasopisma międzynarodowe uznane przez anonimowych ekspertów za godne umieszczenia w wykazie. Czasopisma z tej grupy w większości są indeksowane przez ERIH+, ale nie jest to warunkiem koniecznym indeksacji. Zatem zgodnie z logiką reformy Gowina mamy dwa zbiory czasopism o różnym poziomie prestiżu i umiędzynarodowienia. Grupę liderów prestiżu tworzą WoS + ESCI i SCOPUS, grupę pretendentów: pozostałe czasopisma. W przypadku liderów przybliżoną miarą jakości są wskaźniki bibliometryczne. W przypadku reszty – decyzja ekspercka o umieszczeniu w wykazie. Hierarchię w obu tych grupach powinna odzwierciedlać także punktacja przyznawana czasopismom w sześciostopniowej nieliniowej skali 20-40-70-100-140-200. Należy przy tym utrzymać zasadę, że czasopisma z grupy pretendentów nie powinny mieć więcej niż 70 pkt.

Dla baz bibliometrycznych grupujących „liderów prestiżu” najprostszym sposobem przypisania czasopism do konkretnych punktowanych grup byłby podział kwartyłowy. Obie bazy stosują go w podobny sposób, biorąc za podstawę Impact Factor dla WoS i CiteScore dla SCOPUS. Przypisanie do poszczególnych kwartyli odbywa się w odniesieniu do danej dyscypliny. W przypadku czasopism na SCOPUS obok Cite Score występują także inne wskaźniki, jak SNIP (normalizujący punktację wobec dyscypliny) oraz

SJR (Scimago Journal Rank) biorący pod uwagę nie tylko same cytowania, ale także inne wskaźniki, jak np. prestiż cytującego czasopisma. Wybór wskaźnika jest oczywiście kwestią do dyskusji.

Jak wspomniano wyżej, zaliczenie danego czasopisma do danego kwartyła odbywa się w ramach dyscypliny. Katalogi dyscyplin stosowane przez Elsevierów, Clarivate i Ministra Nauki się nie pokrywają, konieczne więc jest jakaś postać dostosowania wyników do systemu krajowego. Jak pokazano wyżej, zastosowany w dotychczas obowiązujących przepisach system ważenia i uśredniania wskaźników uzupełniony o ocenę ekspercką nie zdał egzaminu. Należy więc go zastąpić inną miarą, spełniającą przy tym warunek *one size fits all* – jednej punktacji czasopisma dla wszystkich dyscyplin, do których zostało ono przypisane. Takim wskaźnikiem mógłby być najlepszy kwartył dla CiteScore, SJR czy Impact Factor^[18]. Jak wiadomo, podział kwartyłowy następuje w różnych dyscyplinach. Możliwe jest więc, że czasopismo znajdzie się w kilku z nich lub w kategorii „multidisciplinary”. Najlepszy kwartył wybiera się, wskazując dyscyplinę lub dyscypliny, w których czasopismo przydzielono do najwyższego kwartyła. Przykładowo, czasopismo w Q2 w naukach o zarządzaniu, Q3 w naukach prawnych i Q4 w inżynierii chemicznej zostanie przypisane do drugiego kwartyła z odpowiednią liczbą punktów. Można tu przyjąć, że czasopismo, dla którego najlepszy kwartył to Q1 ma 200 pkt, Q2 – 140, Q3 100 a Q4 – 70.

Oczywiście przypisanie do danego kwartyła w Web of Science i SCOPUS nie muszą się pokrywać. Powstaje zatem pytanie, co zrobić, jeśli najlepszy kwartył dla danego czasopisma to w jednej bazie Q1 a w drugiej Q3? Nie jest to nieprawdopodobne – wskaźniki wyliczane są dla dwóch różnych zbiorów danych. Uśrednianie wyników nic tu nie da. Możliwe są więc dwa inne rozwiązania: przypisanie czasopisma do wyższego z dwóch konkurujących wskaźników (najlepszy z najlepszych kwartyli albo mówiąc prościej – wynik ustalamy jak w karcianej grze w wojnę) albo stosowanie Web of Science do dyscyplin STEM a SCOPUS do pozostałych. Mimo pewnej trudności technicznej pierwszy z tych systemów wydaje się lepszy, bo rozwiązuje problem, co zrobić, jeśli czasopismo z nauk społecznych jest notowane w WoS, ale nie w SCOPUS? Wybór jednej z tych dwóch metod powinien być poprzedzony konsultacjami z przedstawicielami zainteresowanych

¹⁸ Vladimir Pislyakov, „On Some Properties of Medians, Percentiles, Baselines, and Thresholds in Empirical Bibliometric Analysis” *Journal of Informetrics*, nr 4 (2022): 101322, <https://doi.org/10.1016/j.joi.2022.101322>.

dyscyplin, na przykład z komitetami nauk PAN jako ciałami mającymi mocną legitymację demokratyczną.

Jeśli chodzi o czasopisma umieszczone w wykazie przez ekspertów lub finansowane z programu wsparcia czasopism system przypisywania im punktów z konieczności musi być bardziej skomplikowany. Obecnie ani łatwy do zmanipulowania Google Scholar^[19], ani OpenAlex^[20] nie dają w pełni wiarygodnych rezultatów, co oznacza, że nie mogą być jeszcze zastosowane do oceny bibliometrycznej czasopism na potrzeby ewaluacji^[21].

Proponuję, by o punktacji czasopism z tej grupy decydowały fakt indeksacji w innych niż wiodące bazach danych, struktura afiliacji autorów i czas funkcjonowania czasopisma.

W ten sposób 20 pkt otrzymywałyby wszystkie polskie czasopisma indeksowane w ERIH+ oraz te znajdujące się obecnie w wykazie i nie spełniające pozostałych kryteriów. 40 pkt. byłoby zarezerwowane dla czasopism indeksowanych w ERIH+ oraz EBSCO, HeinOnline/Index to Foreign Legal Periodicals (czasopisma prawnicze), Publication Forum, Listę Norweską lub DOAJ, w których nie więcej niż 30% tekstów pochodzi z jednego ośrodka. 70 pkt przysługiwałoby czasopismom spełniającym warunki do otrzymania 40 pkt i wydawanego co najmniej przez 10 lat.

W przypadku, gdy czasopismo polskie jest indeksowane przez WoS/SCOPUS/ESCI i jednocześnie jest objęte programem wsparcia czasopism, przyznawałoby mu się punktację wyliczoną według korzystniejszego algorytmu, czyli powiązaną z najwyższym kwartylem. Tak wyliczona punktacja powinna obowiązywać za cały okres ewaluacji 2022-2025.

¹⁹ Hazem Ibrahim et al., *Google Scholar is manipulatable*, 2024, <https://doi.org/10.48550/ARXIV.2402.04607>; Jan Winczorek, „Wykaz czasopism punktowanych MNiSW w dyscyplinie nauki prawne a bibliometria. Wyniki badania empirycznego” *Państwo i Prawo*, nr 5 (2022): 16-51.

²⁰ Lin Zhang i in., „Missing Institutions in OpenAlex: Possible Reasons, Implications, and Solutions”, *Scientometrics*, 5 lutego 2024. <https://doi.org/10.1007/s11192-023-04923-y>.

²¹ Odmienne: Grzegorz Wierczyński, „Wykaz czasopism naukowych z dziedziny nauk prawnych z perspektywy danych o cytowaniach dostępnych w bazie Google Scholar” *Państwo i Prawo*, nr 5 (2022): 3-15.

4.2. Wykaz wydawnictw

Stworzenie rozsądnego wykazu wydawnictw, zwłaszcza w naukach humanistycznych, społecznych i teologicznych jest jeszcze trudniejsze niż w przypadku czasopism. Tu bibliometria zawodzi o tyle, że nie dysponujemy stosunkowo kompletną bazą danych pozwalającą na wykonanie analiz bibliometrycznych. Warunku kompletności nie spełnia też OpenAlex AI, nieindeksująca np. czołowych niemieckich wydawców książek prawniczych, takich jak Mohr-Siebeck czy Duncker & Humblot (stan na 2 marca 2024). Pewnym rozwiązaniem byłoby odwołanie się do tzw. listy fińskiej Publications Forum^[22] oraz tzw. Listy Norweskiej (Norway Register for Scientific Journals, Series and Publishers^[23]). Obie te listy indeksują wydawców, przyznając im 0 (wydawnictwo nie-naukowe), 1, 2 lub 3 (tylko Lista Norweska) punkty. Najlepsi wydawcy (górne 20%) plasują się na poziomie 2 i 3. Polskie wydawnictwa w tych wykazach wypadają słabo – jest ich niewiele i większość to pozycje na poziomie 0.

Utrzymując obecny, dwustopniowy system publikacji, należałoby przyjąć weryfikowalne kryteria zaliczenia wydawcy do któregoś z poziomów. W przypadku podstawowego poziomu 1 mamy w tej samej grupie liczących się międzynarodowych wydawców i niewielkie krajowe i zagraniczne wydawnictwa drukujące praktycznie wszystko, od prac dyplomowych po habilitacje. Przy utrzymaniu dwustopniowej klasyfikacji proponowałbym zastosowanie podobnego kryterium jak przy czasopismach – zasięgu oddziaływania wydawnictwa, uznając za międzynarodowe te, które publikują autorów z co najmniej 10 krajów, a przy tym publikacje autorów z żadnego z tych krajów nie przekraczają 30% puli. Pomocniczo można przyjąć, że kumulatywna indeksacja na poziomie 2 – 3 (lista fińska) i na poziomie 2 (lista norweska) pozwalają na zaliczenie wydawcy do krajowego poziomu 2 (liczący się wydawcy międzynarodowi) bez weryfikacji pozostałych kryteriów.

Proponowane rozwiązanie jest oczywiście tymczasowe i tylko na obecny okres ewaluacyjny. Nie zmieni bowiem tego, że na poziomie 1 pojawiają się wydawnictwa o bardzo różnym profilu i jakości. Jeżeli ustawodawca

²² Publications Forum <https://julkaisufoorumi.fi/en/evaluations/classification-criteria>. [dostęp: 16.04.2024].

²³ Norway Register for Scientific Journals, Series and Publishers. https://kanalregister.hkdir.no/publiseringsskanaler/Forside.action?request_locale=en [dostęp: 16.04.2024].

zdecyduje się na utrzymanie algorytmicznego sposobu oceny, należałoby te wydawnictwa podzielić na 3 lub 4 grupy, być może z podziałem na wydawnictwa lokalne (publikują autorów z 1 – 3 ośrodków), krajowe (z rozkładem autorów odpowiadającym rozmieszczeniu geograficznemu ośrodków akademickich) i międzynarodowe (publikujące autorów z 5-10 krajów i publikujące teksty z więcej niż 10 ośrodków). Oczywiście jak wszystkie wskaźniki także i ten jest podatny na manipulacje.

4.3. Kryterium II ewaluacji

Jakie działania należałoby zaproponować, żeby uniknąć przynajmniej części wskazanych wcześniej problemów związanych z kryterium II? Jak się wydaje, mamy do wyboru kilka rozwiązań, nadających się do zastosowania łącznie lub osobno.

Pierwsze z nich to wskazanie górnego limitu punktów jakie można uzyskać w kryterium 2. Otwartym pytaniem jest, jak to kryterium wyliczyć. Przy zachowaniu dotychczasowych zasad liczenia punktów można pokusić się np. o sprawdzenie, w jakim przedziale mieściły się wyniki uzyskiwane w poszczególnych dyscyplinach po odcięciu wartości odstających i np. uśrednienie wartości maksymalnych. Takie podejście byłoby jednak żmudne rachunkowo i nie prowadziłyby do zadowalających rezultatów. W każdej ewaluacji i w każdej dyscyplinie te wartości byłyby inne, a więc kryterium dalej pozostawałoby otwarte.

Drugie możliwe rozwiązanie polegałoby na wskazaniu arbitralnie maksymalnej liczby punktów do uzyskania przez jednostkę, co wymagałoby jednak kalibrowania tych maksimów do nieprzewidywalnego rozkładu przychodów w poszczególnych dyscyplinach.

Trzecie rozwiązanie polegałoby na zrezygnowaniu z uwzględniania w kryterium 2 z przeliczania przychodów na punkty i przypisaniu poszczególnym kategoriom grantów wag punktowych. Np. granty międzynarodowe dostawałyby 200 pkt, bycie wykonawcą w takim grantcie dawałoby 140 pkt, grant krajowy dawałby 100 pkt, itp. itd, przy czym liczbę zgłaszanych osiągnięć należałoby ograniczyć np. do 20% z N. Jest to bardzo szczerze określony margines, trudno bowiem założyć, że średnio w okresie ewaluacyjnym 20% pracowników zdobędzie liczące się granty. Ta możliwość wydaje się być rozsądnym kompromisem pomiędzy realizacją celów polityki publicznej a racjonalnym ustaleniem przedziału, w którym powinny mieścić się osiągnięcia dla tej kategorii.

4.4. Kryterium III ewaluacji

Jak się wydaje, do obrony byłoby twierdzenie, że ostatnia ewaluacja dość mocno podważyła zaufanie do *peer review*, choćby ze względu na, delikatnie mówiąc, przypadkowość wyników i niezbyt fortunną konstrukcję zasad oceny wpływu. Wypada więc zapostulować sparametryzowanie i tego kryterium tak, by wpływ społeczny był oceniany algorytmicznie przez przypisanie poszczególnych rodzajów osiągnięć do określonych grup i przyznanie osiągnięciom w każdej grupie odpowiedniej liczby punktów. Jeśli chodzi o grupy i poszczególne typy osiągnięć można się tu posłużyć charakterystykami stosowanymi przez francuskiego ewaluatora, HCERES^[24]. Przykładowo, w SH2, obejmującym normy społeczne (choć nieobejmującym ze względu na sprzeciw środowiskowy prawa), możemy wyodrębnić następujące rodzaje działań wpływających na otoczenie zewnętrzne^[25]. Elementy dowodzące uznania dla kompetencji zespołu: działalność ewaluacyjną i recenzencką przedstawicieli ocenianej jednostki, działalność ekspercką i opiniodawczą, organizację konferencji, przyjmowanie post-doców i stażystów. Dalej działalność dydaktyczną obejmującą tworzenie podręczników, materiałów e-learningowych itd, popularyzację nauki, w tym prace popularnonaukowe, udział w debatach publicznych, czy produkcje telewizyjne i filmowe. Oczywiście nie nadaje się on do skopiowania wprost, brakuje w nim np. działań w ramach tzw. *citizen science*, promocji nauki w mediach społecznościowych czy niezwykle ważnych w niektórych dyscyplinach prac przeznaczonych dla praktyków. Być może należałoby tu dodać jeszcze osiągnięcia studenckiego ruchu naukowego, w tym nagrody za najlepsze prace magisterskie i dyplomowe. Katalog uważam za otwarty i do dyskusji. Natomiast warto przejść z obecnie obowiązujących przepisów podział na uczelnie duże i małe, pozwalając tym pierwszym przedstawić do oceny 10, tym drugim 5 najważniejszych osiągnięć w tej grupie.

²⁴ „Guides des-produits de la recherche et activites de recherche”. <https://www.hceres.fr/fr/guides-des-produits-de-la-recherche-et-activites-de-recherche>. [dostęp: 16.04.2024].

²⁵ Francuskie procedury ewaluacyjne nie przewidują kategorii „wpływ na otoczenie” w naszym rozumieniu. Wskazane dalej osiągnięcia są rozproszone po kilku kategoriach. Analizowanie tej systematyki nie jest jednak przydatne ze względu na potrzeby tego tekstu.

4.5. Zespoły małe i jednostki nowe

Jak sądzę, dla przypadków takich jak nowo utworzone jednostki badawcze można zaproponować kilka możliwych rozwiązań. Najprostsze z nich to przyjęcie w drodze fikcji prawnej, że jednostka utworzona w trakcie okresu ewaluacyjnego do momentu ogłoszenia wyników najbliższej ewaluacji ma określoną kategorię badawczą (B lub B+), co pozwala na jej rozwój i zmierzenie się z kolejną oceną jakości badań. W tym wariantcie fikcji prawnej uzyskania kategorii B+ nie powinno chyba, choć może to być sporne, towarzyszyć prawo nadawania stopni doktora i doktora habilitowanego. Trudno bowiem wymagać by nowy, jeszcze nie okrzepły zespół praktycznie z dnia na dzień miał masę krytyczną potrzebną do kształcenia doktorantów. Przeciwno wyrażonemu tu pogładowi przemawia jednak dotychczasowa praktyka związana tak ze Szkołą Główną Mikołaja Kopernika, jak i z wyodrębnianiem nowych dyscyplin w obecnym okresie ewaluacyjnym, do czego nawiążę niżej.

Drugie możliwe rozwiązanie byłoby zbliżone do przewidzianego w art. 226a p.s.w.i.n., który to przepis odnosi się do uzyskiwania praw akademickich w przypadku powołania w okresie ewaluacyjnym nowej dyscypliny. Przesłankami ustawowymi uzyskania uprawnień do nadawania stopni doktora i doktora habilitowanego są:

- zatrudnienie w pełnym wymiarze czasu pracy co najmniej 12 pracowników mających dorobek w zakresie dyscyplin pokrewnych do nowoutworzonej
- posiadanie przez jednostkę (uczelnę lub instytut) kategorii A+, A lub B+ w co najmniej jednej dyscyplinie pokrewnej do nowoutworzonej,

przy czym o tym, jakie dyscypliny lub jaka dyscyplina jest pokrewna do nowoutworzonej, decyduje Rada Doskonałości Naukowej.

Oczywiście model ten nie może być zastosowany wprost do nowoutworzonych uczelni i instytutów, a nawet do nowych dyscyplin tworzonych w już istniejących jednostkach. Będzie on wymagał pewnej modyfikacji.

Bez większych zmian można zastosować do jednostek nowych pierwszą przesłankę, to jest wymóg zatrudnienia tuzina osób mających dorobek w danej dyscyplinie. Modyfikacji natomiast wymagałaby przesłanka druga. Proponowałbym tu przyjąć, że ocenie podlegałby dorobek publikacyjny osób wchodzących w skład takiego nowego zespołu, liczony za ostatni okres ewaluacyjny, niezależnie od tego, przy jakiej uczelni byłby on afiliowany.

Dodatkowo do tego dorobku można by wliczyć, obok publikacji, uzyskane przez członków tego zespołu granty, przy czym za każdy grant przysługiwałoby ryczałtowo 100 pkt^[26]. Uzyskanie przez zespół wyniku punktowego przekraczającego próg na kategorię B zapewniałoby finansowanie na tym poziomie. Uzyskanie wyniku przekraczającego próg na B+ zapewniałoby finansowanie na poziomie jednostek z tą kategorią. Jednostka natomiast do momentu uzyskania pierwszej kategorii badawczej nie miałaby jednak pozostałych uprawnień wiążących się z uzyskaniem kategorii B+. A lub A+. Na potrzeby statystyczne, rankingowe i ze względu na konieczność zapewnienia jawności danych, takie uczelnie tworzyłyby osobną kategorię N(owe), a wyniki kategoryzacji byłyby dostępne w systemie RADO.N.

Odnosząc propozycję do podanego wyżej przykładu IMOL PAN (mającego, jak pamiętamy, kategorię C) możemy wyliczyć, że dorobek publikacyjny zespołu (bez grantów i uwzględnienia niższej punktacji w latach 2017 i 2018) liczony według proponowanych przeze mnie kryteriów wynosiłoby 500,53 pkt., to jest 1,9 progu na kategorię B+ i 1,38 progu na kategorię A. Przy uwzględnieniu modyfikacji punktacji za lata 2017-2018 i przyjęciu, że w tym okresie pracownicy publikowali za 30 pkt., dorobek instytutu wyniósłoby 482,52 pkt., co stanowi odpowiednio 1,87 progu na B+ i 1,33 progu na kategorię A w dyscyplinie nauki biologiczne. Jak zatem widać, przyjęcie złożonej tu propozycji *de lege ferenda* pozwoliłoby dobrym zespołom badawczym na uzyskanie należytego finansowania do czasu pierwszej pełnej ewaluacji jakości ich badań.

Ten sam mechanizm można zastosować w przypadku małych zespołów – dyscyplin, w których jest od 1 do 11 pracowników. Również i tu wynik kategoryzacji decydowałby jedynie o wysokości finansowania w algorytmie subwencji bez pozostałych praw i przywilejów. Mógłby być oczywiście wykorzystywany w rankingach i porównaniach jakości badań, jednak wymagałoby to stworzenia osobnej kategorii jednostek M(ałe)^[27].

²⁶ Oczywiście wagi można by tu modyfikować np. w zależności od tego, czy to jest grant krajowy czy międzynarodowy itd.

²⁷ Podobną kategorię wyodrębniano w czasach oceny parametrycznej prowadzonej przez Komitet Badań Naukowych. Przy czym w tej kategorii znajdowały się jednostki mające nie więcej niż 15 pracowników samodzielnych na etacie (5 w przypadku instytutów). Wyodrębnienie tej kategorii motywowano wtedy mniejszą stabilnością badawczą. Informacja na zarchiwizowanej stronie KBN. https://web.archive.org/web/20110110195332if_/http://kbn.icm.edu.pl/finauki98/system/ost/ost_ho2.html. [odsłona 16.04.2024].

5 | Wnioski

Podsumowując dotychczasowe rozważania, należy stwierdzić, że możliwe, a przy tym konieczne jest wprowadzenie zmian w przepisach ewaluacyjnych ukierunkowanych na usunięcie najpoważniejszych wad systemu. Do zmian tych należy wprowadzenie nowego algorytmu przypisywania punktów czasopismom, stworzenie mającej własny algorytm przypisywania punktów listy czasopism spoza dwóch podstawowych referencyjnych baz danych, wprowadzenie maksimów punktowych w kryterium 2 oraz uproszczenie zasad oceny w kryterium 3 przez częściową przynajmniej algorytmizację obliczania jej wyniku.

Dodatkowo należy wprowadzić nowe zasady ewaluowania i finansowania jednostek utworzonych w trakcie trwania okresu ewaluacyjnego oraz zespołów małych, liczących mniej niż 12 osób. Pozwoli to tym pierwszym na finansowanie badań na poziomie umożliwiającym realne przygotowanie się do ewaluacji, tym drugim na zapewnienie środków z subwencji na przetrwanie i rozwój. Ograniczy to także przypisywanie uczonych z takich małych zespołów niejako na siłę do dyscyplin ewaluowanych.

Proponowane rozwiązania dotyczą obecnego modelu ewaluacji. Pilnym zadaniem stojącym przed społecznością akademicką jest wypracowanie nowych zasad, dostosowanych do specyfiki dziedzin i dyscyplin naukowych, które powinny być znane przed rozpoczęciem nowego okresu ewaluacyjnego.

Bibliografia

- Barker Katharine, „The UK Research Assessment Exercise: The Evolution of a National Research Evaluation System” *Research Evaluation*, nr 1 (2007): 3-12. <https://doi.org/10.3152/095820207X190674>.
- Elton Lewis, „The UK Research Assessment Exercise: Unintended Consequences” *Higher Education Quarterly*, nr 3 (2000): 274-283. <https://doi.org/10.1111/1468-2273.00160>.
- Franceschet Massimo, Antonio Costantini, „The First Italian Research Assessment Exercise: A Bibliometric Perspective” *Journal of Informetrics*, nr 2 (2011): 275-291. <https://doi.org/10.1016/j.joi.2010.12.002>.
- Geuna Aldo, Matteo Piolatto, „Research Assessment in the UK and Italy: Costly and Difficult, but Probably Worth It (at Least for a While)” *Research Policy*, nr 1 (2016): 260-271. <https://doi.org/10.1016/j.respol.2015.09.004>.

- Hazem Ibrahim Fengyuan Liu, Yasir Zaki, Talal Rahwan, *Google Scholar is manipulatable*, 2024. <https://doi.org/10.48550/ARXIV.2402.04607>.
- Korytkowski Przemysław, Emanuel Kulczycki, *Raport z pilotażu procedur oceny wpływu społecznego*, 2020. <https://doi.org/10.6084/M9.FIGSHARE.12402686>.
- Kulczycki Emanuel, Marcin Korzeń, Przemysław Korytkowski, „Toward an Excellence-Based Research Funding System: Evidence from Poland” *Journal of Informetrics*, nr 1 (2017): 282-298. <https://doi.org/10.1016/j.joi.2017.01.001>.
- Oceny nauki: Tomaszowice, 16-18 listopada 2013*, Debaty PAU, t. I, red. Szczepan Biliński Kraków: Polska Akademia Umiejętności, 2014.
- Pislyakov Vladimir, „On Some Properties of Medians, Percentiles, Baselines, and Thresholds in Empirical Bibliometric Analysis” *Journal of Informetrics*, nr 4 (2022): 101322. <https://doi.org/10.1016/j.joi.2022.101322>.
- Prawelska-Skrzypek Grażyna, Justyna Maciąg, „Rozdział IV. Główne nurty krytyki ewaluacji polityki naukowej i innowacyjnej oraz sposoby jej doskonalenia na przykładzie wybranych krajów”, [w:] *Ewaluacja w procesie tworzenia i realizacji polityki naukowej i innowacyjnej*, red. Grażyna Prawelska-Skrzypek. 303-384. Warszawa: Elipsa, 2017.
- Propozycje Programowe Koalicji Obywatelskiej*. <https://100konkretow.pl/edukacja/>.
- Rebora Gianfranco, Matteo Turri, „The UK and Italian Research Assessment Exercises Face to Face” *Research Policy*, nr 9 (2013): 1657-1666. <https://doi.org/10.1016/j.respol.2013.06.009>.
- Wasilewski Piotr, Agnieszka Kalita, Małgorzata Szczęsna, „Kształcenie kadry medycznej - problemy w ochronie zdrowia” *Kontrola Państwa*, nr 6 (2023): 71-85. <https://doi.org/10.53122/ISSN.0452-5027/2023.1.36>.
- Wierczyński Grzegorz, „Problemy oceny parametrycznej polskich czasopism naukowych z dziedziny nauk prawnych” *Państwo i Prawo*, nr 9 (2020): 3-22.
- Wierczyński Grzegorz, „Wykaz czasopism naukowych z dziedziny nauk prawnych z perspektywy danych o cytowaniach dostępnych w bazie Google Scholar” *Państwo i Prawo*, nr 5 (2022): 3-15.
- Winczorek Jan, „Wykaz czasopism punktowanych MNiSW w dyscyplinie nauki prawne a bibliometria. Wyniki badania empirycznego” *Państwo i Prawo*, nr 5 (2022): 16-51.
- Zhang Lin, Zhe Cao, Yuanyuan Shang, Gunnar Sivertsen, Ying Huang, „Missing Institutions in OpenAlex: Possible Reasons, Implications, and Solutions” *Scientometrics*, 5 lutego 2024. <https://doi.org/10.1007/s11192-023-04923-y>.



