

Prof. dr hab. WŁODZIMIERZ BOLECKI

Instytut Badań Literackich PAN

Wiceprezes Fundacji na rzecz Nauki Polskiej

Członek korespondent PAU

OCENA JEDNOSTEK NAUKOWYCH W HUMANISTYCE (PARAMETRIZACJA A EWALUACJA)

Nazwa „nauki humanistyczne” w tytule mojego wystąpienia jest zbiorczym określeniem dwóch różnych dyscyplin: nauk społecznych i nauk humanistycznych, a w zasadzie nawet trzech, bo ze względu na kryteria oceny jednostek naukowych, wyodrębniono z nauk humanistycznych jeszcze nauki o sztuce. Na użytek naszej konferencji w takim znaczeniu będę się posługiwał sformułowaniem „humanistyka”.

Oceny badań w naukach humanistycznych zawsze sprawiały kłopot. Ilekroć w ciągu ostatnich lat stosowano identyczne kryteria oceny jednostek naukowych dla wszystkich dziedzin nauki, tylekroć były one kwestionowane przez humanistów. I słusznie, bo nie można porównywać ptaka z rybą. To znaczy można, ale trzeba zacząć od sporządzenia katalogu różnic. I tak się stało. Mimo że od parametryzacji z 2010 r. do oceny parametrycznej z 2013 upłynęły niespełna trzy lata, to zmianę kryteriów oceny instytucji badawczych w naukach humanistycznych trzeba uznać za rewolucyjną. Zmiana ta była skutkiem dwóch działań: publicznych dyskusji na temat różnic pomiędzy humanistyką a naukami ścisłymi i eksperymentalnymi (*hard sciences*) oraz wprowadzonej w ramach reformy sektora nauki nowej instytucji, tj. Komitetu Ewaluacji Jednostek Naukowych (KEJN). Przypomnę, że KEJN został podzielony na zespoły odpowiednie dla poszczególnych dziedzin naukowych. W rezultacie przedstawiciele nauk humanistycznych (w szerokim rozumie-

niu) przygotowali kryteria specyficzne dla prac prowadzonych w swoich dyscyplinach.

Ocena parametryczna

Najpierw słówko o różnicach (między *hard sciences* a humanistyką), które po raz pierwszy zostały uwzględnione w dokumentach Ministerstwa Nauki regulujących działalność KEJN-u. Należą do nich: inna niż w *hard sciences* funkcja języka w humanistyce (stąd kryterium języka właściwego dla konkretnej dyscypliny); specyfika badań nad kulturą narodową (nb. problem wspólny dla wszystkich krajów EU) oraz powstanie Narodowego Programu Rozwoju Humanistyki, finansującego badania dokumentacyjne, nieprowadzone w ramach *hard sciences* i niemożliwe do realizacji w ramach krótkoterminowych grantów NCN (leksykony, słowniki, katalogi, starodruki i inne zbiory specjalne, archiwa, opracowania filologiczne, komentowanie źródeł, komentarze naukowe i tłumaczenia, np. z języków starożytnych, tekstów filozoficznych, religijnych i in., etc.). Zróznicowanie kryteriów oceny dotyczy też gatunków wypowiedzi naukowej. W naukach ścisłych i eksperymentalnych podstawowym gatunkiem są artykuły, w naukach humanistycznych – monografie. Innym gatunkiem wypowiedzi naukowej, całkowicie odmiennie rozumianym w tych dwóch typach nauk, są recenzje. W naukach humanistycznych są to rozprawy o objętości od kilku do kilkudziesięciu stron. W naukach ścisłych i eksperymentalnych – często kilkudziesięciu opinii. Te pierwsze mają najczęściej wymowę ogólną i są często drukowane jako osobne artykuły o książkach czy nawet rozprawy o dużej wartości naukowej, te drugie dotyczą wyłącznie przedmiotu recenzji.

Spośród czterech kryteriów tzw. kompleksowej oceny jednostek naukowych (osiągnięcia naukowe i twórcze; potencjał naukowy; materialne efekty badań naukowych oraz „konkurs piękności”, czyli dziesięć największych osiągnięć wskazywanych przez instytucje naukowe), kryterium trzecie (materialne efekty wyników badań) w praktyce nie ma żadnego zastosowania w humanistyce. Nie jest to wyłącznie kwestia niemożności komercjalizacji wyników badań (taka możliwość istnieje, ale nie jest wykorzystywana, poza tym ma zasięg ograniczony), ale przede wszystkim sprawa niskiego poziomu ochrony własności intelektualnej w humanistyce. Mówiąc krótko: w humanistyce nie ma ochrony patentowej czy tzw. wdrożeń, nie mówiąc o pomocy w zakresie komercjalizacji – np. inkubatorów przedsiębiorczości, *seed money*, firm uniwersyteckich typu *spin off* czy *spin out*, *business angels* itd.

Różnice obejmują także brak wskaźników bibliometrycznych (zwłaszcza w badaniach humanistycznych *sensu stricto*) lub nikłe, a nawet deformujące

skutki ich stosowania. W konsekwencji publikacje w naukach ścisłych i eksperymentalnych (artykuły) i w części nauk społecznych (psychologia, ekonomia) są indeksowane, a w humanistyce (monografie) nie są indeksowane – w związku z tym nie widać ich w bazach danych typu Web of Science. Chodzi też o bardzo ograniczone – a często niemożliwe – stosowanie takich wskaźników, jak IF czy CI.

Te wszystkie różnice dały jednak efekt zdumiewająco pozytywny. Wszystkie bowiem (a nawet znacznie więcej) zostały uwzględnione w kryteriach oceny jednostek naukowych przez KEJN. Jest to zmiana przełomowa, gdyż tak szczegółowych, a zarazem specyficznych dla humanistyki kryteriów nigdy dotąd nie uwzględniano. Nb. pojawiające się co pewien czas w środowisku humanistycznym zarzuty, że jacyś „oni” (w domyśle urzędnicy ministerstwa) znów coś nam narzucili, świadczy o nieznanym zasad reformy nauki. Kryteria oceny (KEJN) stworzyli bowiem przedstawiciele środowiska naukowego, wybrani przez nas samych specjalnie w tym celu.

Podsumowując tę część mojej wypowiedzi, zaczynam od pytania: co otrzymaliśmy w wyniku tej wielokryterialnej, bardzo szczegółowej i bardzo rzetelnie przeprowadzonej oceny parametrycznej jednostek naukowych?

Odpowiedź jest prosta – otrzymaliśmy podział instytucji badawczych na cztery grupy (A+, A, B, C). Fakt ich zróżnicowania, a zwłaszcza wyodrębnienia ok. 4% jednostek najlepszych (w porównaniu ze spłaszczającą oceną w 2010 r.) jest bez wątpienia sukcesem KEJN-u. Ale zadajmy kolejne pytanie: czemu ten podział ma służyć i jaką wiedzę otrzymaliśmy o badaniach prowadzonych w polskich jednostkach naukowych (w tym humanistycznych)?

Otóż otrzymaliśmy informacje statystyczne w postaci siedmiu tabel¹, zawierających podział na grupy i liczby punktów przyznanych poszczególnym instytucjom – liczonych ogółem oraz w każdym kryterium osobno itd. Czemu zatem miała służyć ta parametryzacja? Odpowiedź znamy – jej celem było stworzenie rankingu czterech grup instytucji naukowych w formie podziału ich na jednostki lepsze i słabsze (według punktów). Jest to oczywiście wiedza potrzebna (a w Polsce nawet konieczna), ale o bardzo wąskim zakresie zastosowania. Wiedza ta bowiem ma pomóc MNiSW w realizacji tylko jednego celu (nie znalazłem innego wyjaśnienia), to znaczy – zmianie zasad finansowania instytucji naukowych. I tak, kategoria A+ dostanie dodatkowe pieniądze (w sumie niewielkie), kategoria A zachowa finansowanie bez zmian, kategorie B i C otrzymają mniejszą dotację statutową (a jeśli się nie zreformują, to może grozić im nawet likwidacja). Żadnego innego efektu oceny parametrycznej chyba nie przewidziano.

¹ Zob. *Forum Akademickie* 10, 2013.

Ewaluacja

W roku 2007 brałem udział w ewaluacji uniwersytetu w Uppsali. Zaproszono do niej w charakterze ekspertów kilkuset uczonych z całego świata. Ekspertów podzielono na panele, przyporządkowane kompetencjami do odpowiednich wydziałów i jednostek uniwersyteckich. Ewaluacja trwała pięć miesięcy (nie licząc wielomiesięcznych przygotowań), przy czym poszczególne panele ekspertów pracowały na miejscu przez tydzień. Każdy członek panelu otrzymał dostęp do konta *on line*, na którym mógł znaleźć najważniejsze publikacje pracowników z ocenianych jednostek uniwersyteckich. Do tego doszły różne informacje ogólne, dotyczące charakteru prowadzonych prac, dostępne *on line* lub przekazane podczas spotkań ekspertów z władzami uniwersytetu. Eksperti spotykali się z każdym zespołem, prowadzili wielogodzinne rozmowy z pracownikami, z kierownikami zespołów i doktorantami. Potem spotykali się we własnym gronie, omawiali wyniki rozmów i sporządzali krótkie raporty – najpierw robocze, a następnie końcowe. Celem pracy ponad dwudziestu paneli ekspertów było zidentyfikowanie najmocniejszych obszarów badań prowadzonych przez uniwersytet i wskazanie możliwości prowadzenia nowych badań, które w przyszłości mogłyby odnieść sukces w perspektywie krajowej i międzynarodowej.

Wynikiem tej ewaluacji była książka „Quality and Renewal. An Overall Evaluation of Research at Uppsala University”, licząca blisko sześćset stron, zawierająca informacje o procedurze ewaluacji i jej wynikach. Dane statystyczne, w postaci diagramów, tabel, liczb, zamieszczono jako narzędzia pomocnicze, służące precyzyjnej informacji, a nie jako wynik oceny jakościowej.

Wynikiem tak przeprowadzonej ewaluacji była więc ocena jakości i wartości badań prowadzonych na uniwersytecie. Czyli odpowiedzi na podstawowe pytania: jaki jest poziom tych badań, czy warto je kontynuować – czy wzmocnić je finansowo lub personalnie? A może ograniczyć lub nawet zlikwidować?

Ocena parametryczna w Polsce przyniosła odpowiedź na pytanie, kogo finansować, ale nie – dlaczego. Sposób finansowania (który ma być efektem tej oceny) także nie jest zindywidualizowany, lecz został również z góry uśredniony w grupach.

Z kolei ewaluacja przeprowadzona na uniwersytecie w Uppsali przyniosła odpowiedź na pytanie, jaka jest jakość badań prowadzonych na poszczególnych wydziałach, instytutach czy nawet w zespołach. W jakich obszarach te badania są mocne lub słabsze, jaka jest perspektywa ich rozwoju, możliwości liczących się osiągnąć, to znaczy, czy konkretne badania warto wzmocniać, stabilizować czy wygaszać? I jaka jest ich wartość w skali kraju,

a jaka w skali międzynarodowej? Zastosowano następujące kryteria oceny badań: najwyższa jakość w skali światowej; wysoki standard międzynarodowy; międzynarodowa rozpoznawalność; standard akceptowalny.

Z obu przedstawionych tu typów ocen jednostek naukowych wyłania się odmienne rozumienie „instytucji naukowej” (i uprawiania nauki). W ocenie parametrycznej instytucja naukowa to zbiór aktywności prowadzących do wypracowania wyższej lub niższej liczby punktów. Pracodawca ma otrzymać informację, jak prowadzone w jego instytucji aktywności wypadają na tle aktywności podobnych instytucji naukowych i jak wpłyną na dotację statutową.

W perspektywie ewaluacji instytucja naukowa to zbiór indywidualnych badaczy lub zespołów, których działalność prowadzi do wyników bardziej lub mniej wartościowych w skali krajowej i międzynarodowej. A pracodawca musi znać odpowiedź na pytania, kogo, co, dlaczego i po co ma finansować.

W wyniku oceny parametrycznej nie otrzymaliśmy informacji, jaką wartość to, co robi się w poszczególnych instytucjach naukowych w Polsce (nawet w obrębie jednostek należących do Grupy Wspólnej Oceny), czy warto tę pracę kontynuować, podtrzymywać, czy ograniczyć lub wygasić itd.

Puentując: mimo podziału jednostek humanistycznych na cztery grupy jakościowe, tak samo, jak przed oceną parametryczną, nie wiemy, jaka jest rzeczywista wartość prowadzonych w nich badań i dlaczego. A w rezultacie, jaki jest poziom humanistyki w Polsce (rzecz jasna, każdy uczony ma swoje zdanie na ten temat). Dlatego kończąc, chciałbym postawić dwie tezy i sformułować jedną propozycję.

Teza pierwsza: na ocenie parametrycznej (mimo jej bezsprzecznych osiągnięć i świetnej pracy KEJN-u) nie można poprzestać. Powinna stać się ona punktem wyjścia do pytań o wartość, jakość i celowość badań w poszczególnych instytucjach i w różnych obszarach humanistyki. Teza druga będzie powtórzeniem tego, co razem z dwojgiem kolegów (prof. prof. Ewą Dahlig-Turek, Przemysławem Urbańczykiem) napisaliśmy w 2012 r.: „Pierwszym krokiem [do reformy humanistyki w Polsce] powinna być dziś profesjonalnie przygotowana ewaluacja stanu i rozpoznanie potrzeb instytucji badawczych w humanistyce, a także poszczególnych dyscyplin, a w jej rezultacie – wskazanie najważniejszych celów, do jakich należy w tej chwili dążyć i koniecznych warunków ich realizacji”².

A moja propozycja jest następująca. Najwyżej ocenione przez KEJN instytucje naukowe powinny mieć możliwość przeprowadzenia ewaluacji

² „Po pierwszym konkursie NPRH”, *Forum Akademickie* 4, 2012.

jakościowej swoich badań (w modelu, który wyżej opisałem). Ale taka ewaluacja jest przedsięwzięciem bardzo kosztownym (szczególnie dla uniwersytetów; w mniejszej skali dla konkretnych wydziałów czy instytutów PAN) i bardzo trudnym do sfinansowania przez poszczególne jednostki naukowe. Dlatego wynikiem oceny parametrycznej powinien być dla najwyżej ocenionych jednostek (najlepszych, nie wszystkich!) hojny bonus, przyznany przez MNiSW, w postaci sfinansowania kosztów takiej (międzynarodowej!) ewaluacji. Ta jakościowa ewaluacja powinna zacząć się jak najszybciej – tym bardziej, że musi być procesem kilkuletnim.

Prof. dr hab. JANUSZ GIL
Uniwersytet Zielonogórski
Prorektor ds. nauki i współpracy z zagranicą

UWAGI NA TEMAT OCENY JEDNOSTEK NAUKOWYCH

Dnia 30 września 2013 r. MNiSW opublikowało dane dotyczące oceny parametrycznej jednostek naukowych, dokonanej przez KEJN. Zalety i wady tej oceny przedstawił ostatnio J. M. Brzeziński w artykule „Ocena parametryczna – co dobre, co zmienić”¹. W wyniku przeprowadzonej oceny parametrycznej KEJN przydzielił wydziałom uczelni wyższych kategorie naukowe A+, A, B i C, które odzwierciedlają (a przynajmniej powinny) ich poziom naukowy. Podział na kategorie został oparty na metodzie porównywania parami z zastosowaniem tzw. relacji przewyższania, przy użyciu naukowych metod analizy wielokryterialnej². Najpierw utworzono Grupy Wspólnej Oceny (GWO), które składały się z możliwie jednorodnych naukowo wydziałów. Proponowane składy GWO były wstępnie konsultowane środowiskowo i można uznać, że wszystkie wydziały zaakceptowały przynależność do swych GWO. Następnie w każdej GWO utworzono jednostki referencyjne, modelujące różnice między kategoriami. W każdej GWO porównywano parami wszystkie jednostki między sobą wraz z jednostkami referencyjnymi.

Wydziały oceniono według czterech odrębnych kryteriów, które zostały opatrzone różnymi wagami, aby zróżnicować odmienne aktywności naukowe oraz wyeksponować specyfikę danego środowiska naukowego: I. Osiągnięcia naukowe i twórcze (waga 60–65%); II. Potencjał naukowy (waga 10%);

¹ *Forum Akademickie* 11, 2013.

² Szczegółowe omówienie w artykule S. Kistryna i O. Hryniewiczza „Zrozumieć metodę porównań parami”, *Forum Akademickie* 7–8, 2012.