

LICZBOWA OCENA PRACOWNIKÓW NAUKOWYCH

Prof. dr hab. inż. Jarosław Świdorski
Instytut Technologii Elektronowej, Warszawa
swider@ite.waw.pl

Streszczenie

Czy i dlaczego należy oceniać pracowników naukowych? Czy można ich oceniać jedynie „jakościowo”, czy też można w tym przypadku zastosować również ocenę ilościową? Co zrobić, aby ocena nie zmieniła się w osądzanie i pozostała oceną „owoców”? Pytania te należą do fundamentalnych zagadnień etycznych w nauce. Pochodnymi tych zagadnień są dwie dalsze grupy pytań: jak i kto ma oceniać pracowników naukowych oraz – czemu ta ocena ma służyć? Odpowiedź na ostatnie z pytań determinuje podejście do tego postawionego na wstępie i tak koło się zamyka. Poniżej zostanie przeanalizowane modne ostatnio paradoksalne zestawienie osiągnięć (i ich twórców) wielkich z osiągnięciami ważnymi, a następnie zostaną opisane systemy oceny ilościowej stosowane w najwyższej i najniższej ocenianych placówkach naukowych. Wreszcie podjęta zostanie próba udowodnienia dwóch związanych ze sobą tez: że brak dokładnej i szybkiej oceny ważnych osiągnięć (zwłaszcza w naukach stosowanych) najczęściej prowadzi do poważnych, wymiernych strat, a ocena zbyt dobra jest równie krzywdząca, jak ocena niezasłużenie zła.

1. Czy można i czy warto oceniać naukowców?

Na przełomie tysiącleci modne było wyszukiwanie w sieciach informatycznych różnego typu mniej lub bardziej mądrych „rekordów”. Zainteresowałem się kiedyś przeglądem najczęściej występujących nazwisk. Okazało się, że wśród Polaków najpopularniejsze w sieci było nazwisko Jana Czochralskiego, który wtedy nieznacznie wyprzedzał Fryderyka Chopina, Karola Wojtyłę i Lecha Wałęsę. Następni spośród polskich ludzi nauki, Maria Skłodowska-Curie i Mikołaj Kopernik znajdowali się na początku drugiej dziesiątki najpopularniejszych Polaków. Czochralski mógłby mieć znacznie więcej „cytowań”, gdyby nie to, że w części źródeł, zwłaszcza amerykańskich, figurował jako Rosjanin, a w źródłach rosyjskich... jako Czech. On sam, który za największe swoje dokonanie uważał stop do wykorzystywanych w kolejnictwie łożysk (ten właśnie patent pozwalał mu żyć stosunkowo dostatnio), zapewne zdziwiłby się ogromnie, że tak wielką sławę kilkadziesiąt lat później przyniesie mu jego metoda monokryształizacji, do której za życia nie przywiązywał większej wagi. Wielkie, ale wówczas niezbyt ważne opracowanie mogło zostać docenione dopiero w kontekście rewolucji technicznej końca

XX w., do której samo walenie się przyczyniło! A twórca tego opracowania pozostał nawet we własnej ojczyźnie dalece niedoceniony. Opisaniami niedoceniania uczonych i ich osiągnięć można by wypełnić niejedną bibliotekę, choć złośliwcy twierdzą, że tak naprawdę osiągnięcia te były wielkie i ważne tylko dla ich twórców, ale z braku kompetentnych ocen dyskusja nad tym zjawiskiem jest zwykle całkowicie jałowa.

Podobno każdy z pracowników nauki, który ma za sobą przynajmniej dziesięcioletnie doświadczenie w ocenianiu naukowego dorobku, może wskazać co najmniej jeden przypadek odwrotny. Ze zrozumiałych względów nie przytoczę tu konkretnych nazwisk, ale też znam ze dwie dziesiątki przykładów, w których dobroduszni recenzenci (a i sam nie jestem bez winy) tak przecenili rozpatrywane osiągnięcia i ich twórców, że ci przez resztę życia usiłowali zastosować coś, co się zastosować nie dało i kończyli swą działalność w poczuciu głębokiej krzywdy i żalu do całego świata.

W tym miejscu warto i trzeba powrócić do samej istoty oceniania [1]. Co chcemy osiągnąć, podejmując trud oceny, recenzji czyjś dorobku lub oszacowania wartości konkretnego dzieła? Czy chcemy na tej podstawie wyrokować o czyichś losach, czy może tylko wybrać lub opracować dalszy plan działań? Dla chrześcijan jest w tym przypadku, moim zdaniem, zupełnie jednoznaczna wskazówka zawarta w Biblii. Nie wolno nam osądzać człowieka, ale możemy, a nawet musimy poznać dokładnie owoce jego działalności i to w tym właśnie celu, by pomóc jemu i tym, którzy są z nim związani we właściwym wyborze dalszego postępowania.

2. Czym jest ocena naukowców i ich dzieł oraz jak powinna być przeprowadzana

Każda ocena jest, trywializując nieco, rodzajem pomiaru [2]. Natomiast pomiar jest porównywaniem wielkości mierzonej ze wzorcem lub dwóch wielkości między sobą. Oceniając pracę badawczą, projekt, pomysł, artykuł itp., bardzo trudno jest znaleźć obiektywnie istniejący wzorzec. Pewną namiastką stosowania takiego wzorca są np. oceny publikacji dokonywane na podstawie wartości czasopisma (tzw. lista filadelfijska), w którym została opublikowana lub oceny na podstawie ilości i rodzaju cytowań; w przypadku patentu – oceny na podstawie jego zasięgu, w przypadku prac rozwojowych – zysku, jaki przyniosły w określonym czasie. Trzeba jednak stale pamiętać o słabościach takich ocen wynikających z nieprzewidywalności i oryginalności poszczególnych osiągnięć. Łatwo się o nich przekonać, „oceniając” tymi metodami wielkich twórców z przeszłości (np. Newton nie mieściłby się w pierwszej dwudziestce tak ocenianych współczesnych mu fizyków angielskich) lub niespodzianki znane nam z autopsji (rekordowe liczby cytowań publikacji zawierających poważne błędy prostowane „na wyścigi” przez licznych czytelników). Za to oceny dokonywane na podstawie porównywania dzieł ze wzorcami sprawdzają się dobrze w przypadku dużej liczby elementów ocenianych (im ich więcej, tym ocena jest wiarygodniejsza). W przypadku oceniania pojedynczych naukowców i pojedynczych dzieł znacznie lepiej sprawdza się metoda porównywania między sobą podobnych przypadków, metoda tworzenia rankingów ocenianych obiektów. Prawie zawsze można określić, że praca (projekt, artykuł) A jest lepsza, równa lub gorsza od pracy B. Czasem trzeba takie porównanie zawęzić, zwłaszcza dla prac odległych tematycznie i powiedzieć, że jedna z nich jest np. bardziej odkrywczą, bardziej pracochłonną, bardziej przejrzystą, bardziej ciekawą czy potrzebną.

Oddzielnym problemem jest sama technika oceniania. Jeśli decydujemy się na ocenę określoną wyżej jako porównywanie ze wzorcem, to musimy dążyć do maksymalnego

zmniejszenia subiektywnego zawsze „czynnika ludzkiego”. Ideałem jest w tym przypadku ocena komputerowa, w czasie której człowiek sprawdza tylko zgodność wprowadzonych danych z programami i stanem źródeł. Niebezpieczeństwo w tym przypadku polega na łatwym pominięciu sytuacji szczególnych niemieszczących się (lub źle mieszczących się) w stosowanym programie. Przy porównywaniu dzieł „nieprzeliczalnych” wszystko zależy od oceniającego. Należy tu rekomendować wszelkie mechanizmy pozwalające na uniknięcie nawet podświadomej stronniczości. Najprostsze pod tym względem i najstarsze jest stosowanie jednostronnej lub obustronnej anonimowości, gdy oceniający pozostaje znany tylko dla organizatora oceny lub gdy oceniający, pozostając nieznanym, sam też nie ma możliwości identyfikacji ocenianego. Słabą stroną tego systemu jest duże utrudnienie w zakresie możliwości żądania i udzielania wyjaśnień, a także, zwłaszcza w małych środowiskach – pewna fikcyjność takiej anonimowości. Znacznie skuteczniejszym i nowocześniejszym systemem oceniania jest ocenianie wieloosobowe, gdy np. dziesięciu recenzentów otrzymuje do oceny dziesięć dzieł (każdy wszystkie) i po zestawieniu dziesięciu rankingów wszyscy spotykają się i dyskutują jedynie przypadki sporne. W tym systemie łatwiej jest też o wprowadzenie oceny liczbowej, gdyż wyraźnie wzrasta dokładność przy ustalaniu miejsca ocenianego obiektu w rankingu. Warto zauważyć, że Europejska Karta Naukowca [1] przy omawianiu systemów oceny szczególnie podkreśla wagę niezależności oceniających komisji, w tym zatrudniania w nich naukowców z różnych krajów.

3. Potrzeba i przeznaczenie oceny liczbowej naukowców

Większość ocen naukowców i ich dzieł przeprowadza się w systemie zero-jedynkowym. Nadaje się lub nie na dane stanowisko, zasługuje lub nie na dany stopień czy tytuł, można przyjąć lub nie daną publikację do druku itp. Ale coraz częściej potrzebne są też oceny liczbowe i rankingi. Dotyczy to przede wszystkim rozstrzygnięć finansowych. Komu przyznać grant, komu i o ile dać więcej środków na prowadzone już badania, ale także – którą publikację zakwalifikować do druku, gdy chętnych jest znacznie więcej niż miejsca w czasopiśmie itp. Z tego względu większość fundacji i agencji zajmujących się finansowaniem badań wprowadziło bądź wprowadza obecnie różne systemy oceny liczbowej naukowców i ich dzieł. W Polsce próbował to wprowadzić już Komitet Nauki i Techniki, a skutecznie rozwinął różne formy oceny, w tym oceny liczbowej, Komitet Badań Naukowych, którego działalność kontynuowana jest obecnie przez Ministerstwo Nauki i Szkolnictwa Wyższego. Pierwsze formy opierały się na podziale ocenianych obiektów na kategorie, do których to kategorii przywiązane były określone „widelki” przyznawanych (przeważnie w procentach potrzeb) środków. Najbardziej znanymi były (i w większości przypadków są nadal) systemy oceny instytucji naukowych, projektów badawczych (grantów) i różnych przedsięwzięć, w tym wydawnictw naukowych, czasopism itp. Obecnie we wszystkich tych przypadkach stosowane są systemy „podwójnej odpowiedzi”, tj. odpowiedzi na dwa prawie niezależne (i poprzedzone niezależną oceną) pytania: czy przyznać pomoc finansową (ocena zero-jedynkowa, ale podjęta przeważnie na podstawie oceny liczbowej) i w jakiej ją przyznać wysokości (na podstawie oceny liczbowej). Praktycznie wszystkie te oceny sprowadzają się do tworzenia rankingów opartych na liczbowych ocenach dokonywanych wg znanych ocenianych kryteriów. Stąd tylko krok do systemów ocen liczbowych używanych w instytucjach naukowych do oceny własnych pracowników stosowanej następnie do wewnętrznego podziału środków na badania. Proces ten w Polsce wciąż jest jeszcze w fazie formowania

się, choć trzeba przyznać, że w wielu dobrych placówkach badawczych taki lub podobny system był już z powodzeniem wykorzystywany zanim jeszcze jego powstanie zaczęły wymuszać czynniki zewnętrzne. Co sprawniejsze instytucje naukowe systemy oceny naukowców tworzyły i posługiwały się nim do prowadzenia właściwej polityki kadrowej oraz do wybierania i wspomagania finansowego najbardziej obiecujących kierunków badań.

4. Przykłady systemów oceny naukowców

Poniżej zostaną opisane [3] systemy oceny stosowane w dużym instytucie naukowym (ok. 300 pracowników) zajmującym we wszystkich dotychczasowych rankingach ogólnokrajowych miejsca w pierwszej trójce oraz system oceny stosowany w średnim instytucie (ok. 100 pracowników) lokującym się w takich rankingach bliżej końca niż środka.

W pierwszym przypadku stosowane są jednocześnie trzy powiązane ze sobą systemy. Najbardziej subiektywny służy do oceny wykonania rocznego etapu pracy naukowej. Kierownik pracy (najczęściej jednocześnie kierownik zakładu) przedstawia uzyskane wyniki w formie seminaryjnej. Po każdym seminarium kilkunastoosobowa komisja opracowuje na zamkniętym spotkaniu „znormalizowany” protokół odbioru. Po przeprowadzeniu wszystkich takich seminariów, co trwa (w końcu roku) ok. dwa tygodnie komisja ta w tajnym głosowaniu (z odrzuceniem najwyższej i najniższej oceny) ustala punktową wartość każdego z odbieranych etapów w czterech kategoriach: prezentacji, opracowania złożonego przez kierownika pracy w znormalizowanej formie pisemnej, osiągnięć przeliczalnych i znaczenia pracy. Uzyskany wynik punktowy wchodzi w skład oceny kierownika oraz pozwala na określenie liczby punktów, jakie dane kierownik może rozdzielić według własnego uznania swoim pracownikom w czasie ich oceniania w drugim systemie – systemie oceny indywidualnej.

Na system oceny indywidualnej dokonywanej na początku każdego roku składa się ok. dwudziestu kryteriów. Każdy pracownik z wyższym wykształceniem pracujący w zakładach badawczych ma obowiązek wypełnić ankietę, wyliczając w formie znormalizowanej wszystkie swoje osiągnięcia w dziedzinach określonych zarządzeniem: awanse (stopnie, tytuł) naukowe, publikacje, wystąpienia konferencyjne, patenty i zgłoszenia patentowe, osiągnięcia związane z rozwojem kadry (w tym promotorstwo), uzyskiwanie projektów badawczych (i kierowanie nimi), organizację konferencji, nagrody i inne „zewnętrzne” wyrazy uznania, opracowania nowej metody, nowej technologii, nowego przyrządu, w tym modele, prototypy, wdrożenia, zamówienia zewnętrzne, sprzedaż licencji, produktów wytwarzanych laboratoryjnie i inne przychody pozyskane dla instytutu wreszcie – wypełniane przez zwierzchników punkty stanowiące wyraz ich subiektywnego uznania. Kilkuosobowa komisja ocenia wszystkie te osiągnięcia i sporządza rankingi wszystkich ocenianych pracowników w kilku kategoriach, w tym w kategorii osiągnięć czysto naukowych, w pracach rozwojowych, w zdobywaniu środków finansowych dla instytutu, a przede wszystkim w kategorii wszystkich osiągnięć łącznie. Ten ostatni ranking ma podstawowe znaczenie przy podziale środków finansowych między zakłady badawcze oraz w polityce personalnej. Naukowcy zajmujący w nim czołowe miejsca są odpowiednio nagradzani, a ok. dziesięciu zajmujących ostatnie miejsca jest proszonych o wyjaśnienie swojej sytuacji, a w przypadku powtarzania się tych lokat w kolejnych latach – muszą pożegnać się z instytutem. W ostatnich latach system ten ewaluuje w kierunku zmniejszania roli uznaniowości. Początkowo np. pewną pulą

punktów (dla wyrównywania nietypowych przypadków) dysponował dyrektor, istniały też dość szerokie „widełki” przypisane do każdego typu osiągnięcia, a decyzje podejmowała komisja w jawnej dyskusji. Dziś komisja ta nie dyskutuje nad większością osiągnięć, a co najwyżej ustala (bez świadomości kogo dana sprawa dotyczy) pewne precedensy. Niemal wszystkie dane o osiągnięciach są pobierane z odpowiednich zbiorów prowadzonych przez ogólnoinstytutowe komórki organizacyjne: dział dokumentacji, ochrony patentowej, księgowość, szkolenie, planowanie i in. Rola ocenianych pracowników sprowadza się do sprawdzenia, czy (i ewentualnie dlaczego) coś nie zostało pominięte. Należy jeszcze dodać, że wycena punktowa poszczególnych osiągnięć opiera się na wycenie używanej przez Ministerstwo, zmienianej jedynie w tych obszarach (np. pozyskiwanie grantów), których nagradzaniem zainteresowany jest tylko instytut. Wartość punktowa każdego osiągnięcia jest odnoszona do przychodów instytutu. Oznacza to np., że suma punktów rozdysponowana za osiągnięcia związane ze statutową działalnością naukową odpowiada sumie dotacji otrzymanych przez instytut na tę działalność, suma punktów rozdysponowanych za osiągnięcia związana z produkcją laboratoryjną – sumie przychodów ze sprzedaży wytworzonych wyrobów itd.

Ostatnim, trzecim systemem oceny stosowanym w omawianym instytucie jest system wymuszony ustawą i powołany w niej zarządzeniem ministra. Przy pomocy tego systemu komisja wybrana przez radę naukową ocenia pracowników naukowych i badawczo-technicznych raz na 5 lat (profesorów i docentów) lub raz na 3 lata (pozostałych). Ocenę taką przeprowadza się analizując umiejscowienie danego pracownika w trzech lub w pięciu ostatnich rankingach. A że w tej ocenie stosuje się tylko cztery określenia (wyróżniający się, dobry, zadawalający i niezadawalający) dyskusji (z udziałem bezpośredniego przełożonego) podlegają jedynie pojedyncze przypadki graniczne.

Warto podkreślić, że we wszystkich opisanych wyżej systemach obowiązuje pełna dostępność wyników i uzasadnień dla ocenianych oraz ich przełożonych. Na początku stosowania pierwszych dwóch systemów, w połowie lat 90., co roku rozpatrywano odwołania ok. 20% ocenianych. Obecnie liczba reklamacji spadła poniżej 5%.

W drugim, mniejszym i znacznie słabszym z wyżej wymienionych instytutów stosuje się jedynie system oceny przez komisję rady naukowej analogiczny do trzeciego z używanych w dużym instytucie. Każdy pracownik naukowy lub badawczo-techniczny wypełnia ankietę, w której wymienia wg własnego rozeznania swoje najważniejsze osiągnięcia. Potem na tej ankiecie swój komentarz umieszcza jego przełożony. Komisja wybrana przez radę naukową w obecności tego przełożonego dokonuje („przedyskutowuje”) oceny analogicznej do ocen stosowanych w poprzednio omawianym instytucie w systemie trzecim. Uzyskane tak oceny odpowiadają z dobrym przybliżeniem obiegowym opiniom o danym pracowniku (sytuacja dość typowa dla małych i bardzo małych środowisk).

5. Próba podsumowania i wnioski

Popularna anegdota głosi, że Niels Bohr, w stosunku do pracowników, których wyniki plasowały się znacznie poniżej oczekiwań używał krótkiego zwrotu: „Napracował się pan” i nigdy nie stosował surowszych określeń. Nasze krajowe doświadczenia z ostatnich kilkunastu lat wskazują, że jedną z podstawowych słabości polskiej nauki, być może najpoważniejszą, jest słabość krytyki. Jeśli krytyka nie ma przekształcać się w osądzenie – musi być oparta na możliwie ścisłej, najlepiej liczbowej ocenie faktów (biblijnych „owoców”). Dzisiejsze naukoznawstwo dysponuje już w tej dziedzinie

dostatecznie wypróbowanymi, obiektywnymi metodami, a dość liczne doświadczenia krajowe potwierdzają możliwość ich stosowania w naszych warunkach. Można zaryzykować twierdzenie, że każda ocena ilościowa bardziej odpowiada chrześcijańskiemu poczuciu sprawiedliwości, niż ocena uznaniowa oparta na obiegowych opiniach. Również Biblia podpowiada nam, że każda ocena chwalaćca powinna też wskazać na niedoskonałości, a każda krytyczna – na możliwość poprawy. Z przykładów życia codziennego znamy złe skutki „przechwalenia” (częste marnowanie dobrych, ale nie aż tak, jak to przedstawiono, rezultatów) i krzywdy wynikłe z niedoceny. Typowo polską wadą w tym zakresie jest poszukiwanie osiągnięć wielkich (które i tak przeważnie wymusza sytuacja) a lekceważenie osiągnięć ważnych [4]. Sięgając do przedstawionego na wstępie przykładu, trzeba też przytoczyć opinię, że gdyby nie było Czochralskiego, to i tak ktoś musiałby wynaleźć tę metodę i stać się wielkim, ale naprawdę ważnym wydarzeniem było dopiero jej praktyczne zastosowanie do krzemu. I temu (tj. wychwytywaniu osiągnięć ważnych) także sprzyja powszechne stosowanie ocen ilościowych [5]. A że to czasem boli? Skutecznym ograniczeniem jest stare, pochodzące z przekładania prawd biblijnych na język pojęć współczesnych, powiedzenie: „Nie żałuj słów krytyki tak długo, jak długo sprawia ci to przykrość. Przestań natychmiast, gdy zaczniesz odczuwać przyjemność!”

Bibliografia

- [1] *Europejska Karta Naukowca*, *Dziennik Urzędowy Unii Europejskiej* z dn. 22.03.2005 poz.2005/251/WE.
- [2] Samsonowicz, H., Sławiński, J., Szczucki, L., (rada wydawnicza), *Krytyka i krytycyzm w nauce*, Warszawa 1998.
- [3] Świdorski, J., „Sposoby oceny działalności naukowej i ich praktyczne zastosowanie w elektronice”, *Zagadnienia Naukoznawstwa XXXVIII* (2002) nr 1–2, s. 59–63.
- [4] Samsonowicz, H., Sławiński, J., Szczucki, L., (rada wydawnicza), *Lepsze w nauce*, Wrocław 2000.
- [5] Świdorski, J., „Trudności występujące w procesie oceniania działalności badawczej w obszarze elektroniki”, *Elektronika*, w przygotowaniu.