

Kraków, 2018.08.15

dr hab. Władysław Marek Kolasa, prof. UP

Instytut Nauk o Informacji

Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie

Ul. Podchorążych 2

30-084 Kraków

Recenzja rozprawy doktorskiej mgra Łukasza Opalińskiego pt. *Starzenie się publikacji naukowych w języku polskim i angielskim w perspektywie zachowań warunkujących proces cytowania w naukach o Ziemi* napisanej pod kierunkiem dra hab. Remigiusza Sapy

Przedmiot i cele badań

Głównym przedmiotem rozprawy jest proces starzenia się publikacji naukowych. Tak określony przedmiot ma wprawdzie charakter interdyscyplinarny, lecz ze względów metodologicznych wpisuje się w tradycję badań naukoznawczych, które w Polsce zwyczajowo uprawiane są w ramach bibliologii i informatologii. Ze względu na rozmiar materiału Doktorant nie badał szczegółowo tytułowego procesu, lecz skupił swą uwagę na naukach o Ziemi, spośród których dokładniej przebadal kilkanaście dyscyplin zgrupowanych w sześciu korpusach. Rdzeń badań tworzyły cztery dyscypliny podstawowe: tj. geofizyka i geochemia, geografia fizyczna, geografia społeczno-ekonomiczna oraz geologia; kilka dyscyplin ulokowanych na pograniczu nauk o Ziemi i nauk technicznych (m.in. urbanistyka, geoinżynieria i geotechnika, inżynieria środowiska, górnictwo, geodezja) oraz dyscypliny usadowione na pograniczu nauk o Ziemi i nauk biologicznych (tj. ekologia, ochrona środowiska, botanika, mikrobiologia).

Zasadniczym celem pracy było „zidentyfikowanie bibliometrycznych właściwości starzenia się piśmiennictwa” z zakresu „nauk Ziemi oraz ich interpretacja w perspektywie wiedzy teoretycznej i empirycznej dotyczącej zachowań naukowców w zakresie cytowania źródeł” (s. 8). Aby właściwie usytuować przedmiot badań Doktorant zdefiniował siedem celów szczegółowych, które zawierały następujące pytania: 1 – w jaki sposób liczba cytowań prac anglo- i polskojęzycznych prac zależy od osiągniętego przez nie wieku, na jaki okres przypada moment kulminacji cytowań i jak długo trwa oraz w jakim tempie maleje; 2 – jak można wykryte prawidłowości zinterpretować w kategoriach naukoznawczych; 3 – czy polscy naukowcy z zakresu nauk o Ziemi bazują na najnowszych międzynarodowych publikacjach; 4 – jaka jest ilościowa skala wykorzystania piśmiennictwa polskojęzycznego w

komunikacji naukowej; 5 – czy istnieje związek pomiędzy cytowalnością a liczbą punktów przyznawaną przez MNiSW; 6 – czy w poszczególnych dyscyplinach występują różnice pod względem tempa starzenia się publikacji; 7 – czy wyniki podjętych badań można aplikować w praktyce (np. w bibliotekarstwie). Odpowiedź na postawione pytania wiodła poprzez empiryczną weryfikację ośmiu hipotez badawczych, które w większości były operacjonalizacją powyższych celów.

Badania empiryczne przeprowadzono na imponującej próbie badawczej, obejmującej pozycje cytowane przez polskojęzyczne periodyki z zakresu nauk o Ziemi w tzw. listy „B” MNiSW z 2015 roku. Próba ta liczyła: 24582 pozycji (w tym 15541 polskojęzycznych i 9041 anglojęzycznych), pochodzących z 1088 artykułów zawartych w 43 czasopismach.

Struktura pracy

Rozprawa mgr Łukasza Opalińskiego (nie licząc aneksów) liczy 426 stron, zawiera przeszło osiemdziesiąt zestawień (21 tabel, 61 histogramów) i dzieli się na 4 rozdziały. Dwa pierwsze rozdziały (I, II) mają charakter kontekstowy i służyły do sformułowania założeń tudzież operacjonalizacji hipotez; w rozdziale III ułożono założenia metodologiczne, zaś w IV – raport z badań, czyli weryfikację hipotez oraz interpretację wyników. Pracę zamyka ponadto solidne podsumowanie oraz aparat naukowy, w tym 10 obszernych aneksów.

Struktura pracy z naukowego punktu widzenia jest logiczna, gdyż odzwierciedla tok pracy badawczej, niemniej jest obarczona błędem proporcji. Rozdziały wstępne (I i II) są bowiem nieproporcjonalnie obszerne, gdyż zajmują prawie połowę dysertacji (łącznie liczą 204 strony, czyli 47% objętości pracy). Wnikliwa lektura ich zawartości (o czym piszę niżej) przekonuje, że z badawczego punktu widzenia niektóre podrozdziały były zbędne. Drugim mankamentem pracy jest jej objętość. Ponieważ tekst rozprawy był napisany niestandardową czcionką (11,2) z zachowaniem niestandardowych interlinii na jednej stronie mieściło się nie 1800 znaków (jak przewidują standardy), ale 3690 (czyli ponad 2 razy więcej). W efekcie rozprawa o objętości 542 stron (wraz z aneksami) w rzeczywistości liczyła ich przeszło 1100. Z naukowego punktu widzenia zbyt duża objętość (jako taka) błędem nie jest, lecz świadczy, że Doktorant ma problemy z syntetyzowaniem. Liczne partie tekstu są bowiem napisane zbyt rozwlekłe i wymagają od czytelnika dużej dyscypliny, aby powiązać wątki. Występują ponadto liczne powtórzenia oraz odwołania, co sprawia, że podczas lektury bardzo często zachodzi konieczność sięgania wstecz lub do załączników. W efekcie lektura tekstu była bardzo trudna.

Wbrew wyrażonej wyżej krytyce można jednak wskazać argumenty, które usprawiedliwiają Doktoranta. Obowiązująca ustawa *o stopniach naukowych i tytule naukowym...* (2003, art. 13) wymaga, aby kandydat poza oryginalnym rozwiązaniem problemu naukowego wykazał się „ogólną wiedzę teoretyczną”. I to zapewne legło u podstaw zbytniej rozbudowy dwu pierwszych rozdziałów.

Ocena merytoryczna

Rozprawę otwierają dwa rozdziały o charakterze wprowadzającym, które opracowano przy pomocy metod deskryptywnych. Rozdział I, zatytułowany *Zachowania uczonych warunkujące proces cytowania (citation behaviour)*, ma charakter kontekstowy i dzieli się na pięć części. W poszczególnych podrozdziałach mgr Opaliński usystematyzował wiedzę na temat cytowań (rozumianych jako akt lub proces) oraz zaprezentował teorie i badania z tym zakresie. Szczególnie wnikliwie omówił dwa dominujące paradygmaty: normatywny (tj. wyrosły na gruncie prac Roberta K. Mertona – rozdz. 1.2.1) oraz opozycyjny względem niego paradygmat konstruktywistyczny (zapoczątkowany przez Nigela Gilberta – 1.2.2). Ważną rolę odgrywał także podrozdział dotyczący badań nt. zachowań związanych z językiem cytowań (1.3) oraz krytyka omówionych teorii i badań (1.5). Rozdział zamyka krótki tekst nt. związku teorii z badaniami własnymi (1.6), gdzie Autor wskazał w jakim zakresie korzysta z zaprezentowanego dorobku.

Jestem pod dużym wrażeniem przedstawionego rozdziału. Doktorant ze swobodą porusza się w ogromnej literaturze przedmiotu (w rozdziale cytował ok. 200 prac), potrafi identyfikować wątki, selekcjonować fakty i łączyć je całości. Posiada też rzadką umiejętność operowania na wysokim poziomie abstrakcji oraz umiejętnie operować aparatem pojęciowym innych nauk, nawet tak odległych od informatologii jak: fizyka kwantowa czy biologia (by przywołać cytowane w tekście analogie do teorii nieoznaczoności Heisenberga [s. 98] czy hipotezy *size-efficiency* [s. 52-53]). W trakcie pisania rozdziału Autor perfekcyjnie wyzyskał dostępną literaturę i przedstawił spójny obraz wiedzy na temat zachowań warunkujących proces cytowania. Była to – dodajmy – pierwsza w polskim piśmiennictwie próba systematycznego zgłębienia tematu. Fakt ów zasługuje na pochwałę, zaś recenzent zachęca, aby tę część pracy (po niezbędnych przeróbkach) opublikować. Jeśli jednak oceniamy rozdział w kontekście celów dysertacji, to niektóre jego części były zbędne. Uwaga ta dotyczy w szczególności podrozdziałów nt. analizy kontekstu i zawartości semantycznej (1.4.1), badań motywacji (1.4.2) tudzież krytyki badań (1.5) – gdyż zagadnienia te nie były przedmiotem badań własnych.

Zbliżony charakter ma rozdział II, zatytułowany *Starzenie się publikacji naukowych jako przedmiot analiz naukoznawczych*, który dzieli się na 10 części, w których mgr Opaliński zestawil wiedzę teoretyczną i praktyczną na temat głównego przedmiotu badań. Autor wyczerpująco omówil dwa dominujące kierunki refleksji, a więc nurt teoretyczny (2.2) i praktyczny (2.3) oraz kilka specjalistycznych zagadnień związanych z analizą starzenia się publikacji, m.in. symetrii między wynikami analiz synchronicznych i diachronicznych (2.4) i typologiami dokumentów (2.6) i in. W sensie konstrukcyjnym rozdział przypomina podręcznik zawierający systematyczny wykład nt. problematyki starzenia się publikacji. Autor miał jednak inne cele. Rozdział miał ulokować badania własne w szerszym kontekście, a w szczególności uzasadnić wybór tematyki (s. 17) oraz „rozbudzić w polskim środowisku informatologicznym, po ewentualnym opublikowaniu tej pracy, szerszą dyskusję na temat poruszanych tutaj zagadnień” (s. 102). Jeśli weźmiemy pod uwagę pierwszy cel, nietrudno o konkluzję, że został on osiągnięty w już pierwszych pięciu poddziałach (2.1-2.5), natomiast pozostałe fragmenty (2.6-2.10) – jakkolwiek zgodne merytorycznie z głównym tematem pracy – są zbędne, bowiem Autor nie badał ich empirycznie w ramach badań własnych. Gdy natomiast rozważymy cel drugi, cały analizowany rozdział przedstawia koherentną całość i po odpowiednich przeróbkach mógłby stać się osnową do osobnej publikacji.

Podobnie jak poprzednia część dysertacji, także rozdział II jest napisany kompetentnie i przekonująco oraz solidnie udokumentowany. Autor doskonale się orientuje w literaturze, umiejętnie ją wykorzystuje i potrafi wyciągać trafne wnioski. W toku dyskusji literatury mgr Opaliński cytował tu ok. 400 prac, głównie anglojęzycznych. Wnikliwa analiza przywołanego korpusu artykułów, rozpraw i monografii pozwala na stwierdzenie, że dostępna literatura została wyzyskana perfekcyjnie. Drobne przypadki pominięć dotyczą jedynie prac kontekstowych. Np. przy omawianiu badań starzenia się z zakresu humanistyki (s. 120-121) Autor ograniczył refleksję niemal wyłącznie do prac chińskich uczonych (Yanhui Song, Feng Ma i Siluo Yang) [„Scientometrics” 2015, vol. 102(1), pp. 365-388] podczas, gdy wcześniej prace z tego zakresu uwieźzone licznymi publikacjami prowadził m.in. recenzent [np. „Scientometrics” 2012, vol. 90(3), pp. 905-923]. W 2013 roku wydał on ponadto obszerną monografię [*Historiografia prasy polskiej (do 1918 roku) – naukometryczna analiza dyscypliny 1945-2009*], w której oś narracji zasadzała się na jednej z odmian wskaźnika starzenia się publikacji – *citing half-life*.

Kolejne dwa rozdziały (III, IV) mają charakter badawczy. W rozdziale III (*Metodologia badań*) mgr Łukasz Opaliński zaprezentował strategię badawczą oraz metody i



narzędzia zastosowane do badań własnych. W kolejnych 10 podrozdziałach zaprezentował: metodę (3.2), przedmiot i zakres badań (3.3.-3.4), hipotezy (3.8), cztery rozdziały szczegółowe (3.5-3.9) oraz podsumowanie (3.10). Przedmiot badań (24582 jednostek, o których wspomniano na wstępie) został opisany bardzo ogólnie. W analizowanym fragmencie Doktorant koncentrował się raczej na uzasadnieniu jego doboru, podczas gdy dokładniej został on zademonstrowany dopiero w następnym rozdziale (4.1). Jakkolwiek podziałam większość przytoczonych argumentów, wszelako trudno się zgodzić z twierdzeniem jakoby „kryteria selekcji czasopism, przedstawione w Komunikacie MNiSW (2015) zapewniały należyty charakter i poziom merytoryczny, redakcyjny i językowy wyselekcjonowanych czasopism” (s. 227). Wyliczone w kryteria mają bowiem wyłącznie charakter formalny i w żaden sposób nie mogą gwarantować poziomu merytorycznego. Na poziom merytoryczny wpływa przede wszystkim właściwy sposób recenzowania, co bynajmniej nie jest tożsame z deklaracjami na temat recenzowania, składanych przez redakcje w kolejnych ocenach (gdyż nie weryfikował tego żaden audyt).

W dalszych fragmentach Autor wiele miejsca poświęca zastosowanej metodologii (ilościowej analizie synchronicznej – 3.2) oraz szczegółowo referuje hipotezy badawcze (3.8). Odpowiednio do siedmiu celów ogólnych, które zawarto w wstępie (i powtórono w 3.1) Doktorant sformułował osiem hipotez. Założył przeto, że: wiek najczęściej cytowanych publikacji angielskojęzycznych będzie wyższy niż wiek najczęściej cytowanych publikacji polskojęzycznych (hipoteza 1); dla prac angielskojęzycznych obserwowalny będzie wyższy poziom cytowalności (hipoteza 2); skala wykorzystania materiałów polskojęzycznych jest wysoka i tworzy podstawy, aby uznać, że publikowanie w języku polskim jest zasadne (hipoteza 3); publikacje angielskojęzyczne będą starzeć się wolniej w porównaniu do prac polskojęzycznych (hipoteza 4); stadium dojrzenia cytowalności, poprzedzające osiągnięcie punktu cytowalności maksymalnej, będzie trwać dłużej w przypadku publikacji angielskojęzycznych (hipoteza 5); zjawisko starzenia się literatury w poszczególnych dyscyplinach z zakresu nauk o Ziemi przebiega w różnym tempie (hipoteza 6); istnieją różnice w bibliometrycznych charakterystykach starzenia się różnych typów dokumentów [chodzi o cztery typy wydawnicze: artykuł, wydawnictwo zwarte, materiał konferencyjny, inne] (hipoteza 7) oraz że punktacja przydzielone polskim czasopismom w latach 1999-2015 przez MNiSW odzwierciedla naukową rangę publikowanych w nich artykułów, przy założeniu, że ranga ta może zostać zmierzona za pomocą poziomu cytowalności tych artykułów (hipoteza 8). Przyjęte hipotezy w świetle przytoczonej literatury wydają się słuszne, zaś ich dobór służył do osiągnięcia wyspecyfikowanych we wstępie celów. Warto

jednocześnie podkreślić, że powiązanie celów z hipotezami w dysertacji nie tworzyło prostej macierzy, lecz złożoną strukturę (s. 9-10), by nie rzec specjalnej ontologii.

Lektura rozdziału III poświadcza, że mgr Opaliński doskonale orientuje się w metodologii naukowej i potrafi ją twórczo wykorzystać do rozwiązywania problemów naukowych. Potrafi również doskonale wykorzystywać prace poprzedników, czego przykładem są twórczo zinterpretowane dane zaprezentowane w tabeli 5 (s. 249). W rozdziale niczego nie brakuje, zaś jedynym zauważalnym mankamentem są fragmenty zbędne lub nadmiernie rozbudowane. Przykładem jest obszerny podrozdział 3.5, w którym Doktorant nazbyt dokładnie uzasadnia cały proces decyzyjny w projektowaniu badań własnych. Poza prezentacją przyjętej taktyki badawczej znalazły się w rozdziale ponadto liczne aluzje do tematów, których autor nie bada (np. 3.5.3) lub które w toku dyskusji w sposób uzasadniony odrzuca (3.6). Niezależnie od wskazanych mankamentów – tę część rozprawy oceniam wysoko.

W rozdziale IV (*Analiza bibliografii załącznikowych artykułów z polskojęzycznych czasopism z zakresu nauk o Ziemi*) Doktorant zaprezentował i zinterpretował wyniki badań. Tok narracji rozpoczyna wprowadzenie, gdzie znajdujemy dokładną charakterystykę statystyczną badanego zbioru (4.1), w tym interesujące zestawienie wskaźników *citing half-life* dla badanego korpusu (tab. 9, s. 289); dalej następuje blok zawierający wyniki weryfikacji hipotez badawczych. Wyniki testowania nie doprowadziły jednak do jednoznacznych rozstrzygnięć. W pełni pozytywny wynik udało się uzyskać dla trzech hipotez (3, 4, 7), zaś w przypadku czterech był on negatywny (1, 2, 5, 8). Udało się m.in. potwierdzić słuszność hipotezy 3, zakładającej, że skala wykorzystania materiałów polskojęzycznych jest wysoka i tworzy podstawy, aby uznać, że publikowanie w języku polskim jest zasadne. Testy dowiodły, że w badanym korpusie prace takie stanowią przeszło 63%. Pozytywny wynik dało także testowanie hipotezy 4 zakładającej, że publikacje angielskojęzyczne będą starzeć się wolniej w porównaniu do prac polskojęzycznych. Dzięki analizie indeksów *citing half-life* prawidłowość tą udało się potwierdzić dla 3 typów wydawnictw (artykuły, książki, materiały konferencyjne), choć – warto podkreślić – różnice były subtelne. Potwierdzono także empirycznie hipotezę 7, zgodnie z którą poszczególne typy wydawnicze dokumentów starzeją się inaczej. Pozostałe hipotezy nie znalazły jednoznacznego potwierdzenia (np. hipoteza 6) lub nie potwierdzały się zupełnie. Do interesującego wniosku doprowadziło testowanie hipotezy 8, która zakładała zależność pomiędzy rangą czasopisma (mierzoną liczbą cytowań), a punktacją przydzieloną przez MNiSW. Specjalny test udowadniał, że pomiędzy tymi wskaźnikami nie występuje żadna korelacja, bowiem współczynnik Pearsona był bliski zera

($r=0,0139$). Mimo trudności jakie stwarzał materiał i negatywny wynik większości hipotez Doktorant wnikliwie przeanalizował rezultaty i szczegółowo je zinterpretował. Jestem skłonny zgodzić się z większością przedstawionych kierunków interpretacji, wszelako sędzę, że fragmentami Doktorant zbyt ufa własnemu materiałowi badawczemu. Chodzi m.in. o te fragmenty, gdzie przedstawia analizy porównawcze trendów, a następnie operuje na wartościach liczbowych ich charakterystyk (4.2.4 i dalsze). Uważam, że nie ma matematycznych podstaw, aby uznać, że aproksymacje, a w ślad za tym korelacje generowane przez MS Word / Excel mogły służyć celom innych niż pogładowe.

Bardzo ważny w strukturze pracy jest [nienumerowany] rozdział *Podsumowanie i wnioski końcowe* (s. 406-426), gdyż dopiero tam Autor ustosunkował się do siedmiu celów badawczych wskazanych we wstępie. Ponieważ cele i hipotezy tworzyły następującą siatkę zależności: cel 1 (hipotezy 1, 2, 3, 4, 5 i 7) i dalej odpowiednio: 2 (4, 7), 3 (5), 4 (3, 7), 5 (8), 6 (6) odpowiedź na wszystkie postawione pytania nie mogła być jednoznaczna. Mimo to Doktorant dokonał wnikliwej syntezy i odniósł się do wszystkich postawionych na wstępie pytań. Jakkolwiek z niektórymi interpretacjami można dyskutować (np. uzasadnienia tempa starzenia piśmiennictwa w poszczególnych dyscyplinach s. 418-419), wszelko większość sformułowanych wniosków i hipotez wydaje się być słuszna. Osobiście sędzę, że obraz tego problemu byłby bardziej klarowny, gdyby Doktorant wykorzystał inną bazę empiryczną. Podejrzewam, że lepszy wynik dałaby analiza porównawcza dorobku kilku (najlepiej losowo dobranych) grup naukowców ściśle związanych z określonymi dyscyplinami aniżeli analiza korpusu czasopism formalnie przypisanych do dyscyplin.

Konkluzje

Magister Opaliński podjął bardzo ambitny temat i wyznaczył sobie równie ambitne cele. Jakkolwiek uzyskane wyniki nie doprowadziły do przełomowych stwierdzeń (przynajmniej w sferze teorii), Jego rozprawa z naddatkiem spełnia wszystkie wymagania sprawiane pracom doktorskim. Zawiera oryginalne rozwiązanie problemu naukowego, jest poprawna metodologicznie, została napisana dobrym językiem i solidnie udokumentowana. Autor swobodnie operuje słowem, potrafi perfekcyjnie wykorzystać literaturę przedmiotu tudzież stosować rozmaite metody i techniki badawcze (w tym statystyczną analizę danych).

Dodam, że w trakcie pisania recenzji uznałem, że pominię ocenę formalną dysertacji, gdyż pod tym względem jest ona napisana niemal wzorcowo. Dwa zarzuty, który można sformułować padły już na wstępie i odnoszą się do jej struktury pracy (zbyt rozbudowana część teoretyczna oraz rozwlekły styl pisania, który skutkował bardzo dużą objętością).

Pragnę jednak z całą mocą podkreślić, że nie umniejszają one naukowych walorów pracy. Recenzowana dysertacja nie jest wprawdzie łatwa w lekturze, wszelako została napisana w pełnej zgodzie z metodologią naukową.

Stwierdzam, że przedstawiona do oceny rozprawa doktorska mgr Łukasza Opalińskiego spełnia wszystkie warunki określone w art. 13 *Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki*. Jest rozprawą opartą na samodzielnie przeprowadzonych badaniach empirycznych, stanowi oryginalne rozwiązanie problemu naukowego oraz potwierdza ogólną wiedzę teoretyczną kandydata w zakresie bibliologii i informatologii tudzież umiejętność samodzielnego prowadzenia pracy naukowej. Wnioskuje przeto o dopuszczenie Doktoranta do kolejnych etapów przewodu doktorskiego.

