

Warszawa, dnia 24 października 2015 roku

Prof. dr hab. Bogdan Stefanowicz  
Szkoła Wyższa Informatyki Stosowanej i Zarządzania  
pod auspicjami Polskiej Akademii Nauk

### Recenzja

rozprawy doktorskiej mgr Macieja Mozejewskiego

„Model pomiaru jakości informacji publicznej na potrzeby ponownego użycia”

#### **Charakterystyka rozprawy**

Zgodnie z tytułem rozprawy, Doktorant zajął się problematyką jakości informacji w kontekście wielokrotnego wykorzystania tych samych treści informacyjnych w różnych zastosowaniach.

Praca składa się ze wstępu, pięciu rozdziałów, zakończenia („podsumowania”), spisu rysunków i tabel, wykazu wykorzystanych źródeł literaturowych (bibliografii), trzech załączników w postaci tabel oraz streszczenia w języku angielskim. W przytoczonej bibliografii Autor zamieścił ok. 180 tytułów, w tym dwa tytuły swoich prac opublikowanych oraz szereg prac ze swojego środowiska naukowego. W spisie można znaleźć pewne publikacje z dość odległego okresu, w jakimś stopniu już nieaktualne lub mało interesujące. Odczytuję to jednak jako chęć podkreślenia przez Doktoranta faktu, że szereg problemów, którymi zajmują się obecnie specjaliści, wywodzi się z minionego okresu.

We wstępie Autor scharakteryzował poruszaną przez siebie problematykę oraz sformułował cel rozprawy (s. 7): „opracowanie autorskiej koncepcji modelu pomiaru jakości informacji publicznej na potrzeby ponownego użycia”. Przedstawił zarys wybranej ścieżki badawczej oraz w skrócie scharakteryzował treść poszczególnych rozdziałów.

W rozdziale 1 („Informacja publiczna i jej ponowne użycie”) Autor przedstawił główny obiekt swoich badań – informację publiczną, przytaczając literaturowe interpretacje tego pojęcia oraz wybrane cechy informacji. W tymże rozdziale sformułował kryteria oceny cech jakościowych *informacji publicznej do ponownego*

użycia (w recenzji będę stosował symbol *IP* na oznaczenie tego rodzaju informacji). Przedstawił obraz stanu wykorzystania *IP* w Polsce i w innych krajach.

Można odnieść wrażenie, że miejscami obraz ten jest nadmiernie rozwinięty jak na główny temat dysertacji (od s. 21 do s. 55; stanowi to ok. 15% całej rozprawy). Na przykład zostały przytoczone szczegóły dot. opisu ustawy o *IP* w Wielkiej Brytanii lub opis zasad prawnych w tym zakresie we Francji.

Można mieć nadzieję, że Doktorant wyjaśni swoje intencje pod tym względem podczas obrony swojej dysertacji na Radzie Wydziału.

W dalszej części rozdziału zostały scharakteryzowane korzyści wynikające z korzystania z *IP* oraz bariery i ograniczenia utrudniające upowszechnianie tego rodzaju praktyk.

Przedstawiony przez Autora materiał faktograficzny może stać się źródłem inspiracji do dalszych badań w tym zakresie, w szczególności w zakresie własności informacji wtórnego wykorzystania (*IP*).

W rozdziale 2 („Wybrane źródła informacji publicznej do ponownego użycia”) Doktorant przedstawił charakterystykę wybranych źródeł informacji, które – jego zdaniem – niosą bogate zasoby *IP*. W szczególności przedstawił *Biuletyn Informacji Publicznej* w Polsce oraz opisał portale *open data* w krajach wymienionych w rozdziale 1. Doktorant scharakteryzował główne cechy tych systemów, ich funkcje i gromadzone zasoby informacyjne oraz wskazał możliwości wielorakiego wykorzystania gromadzonych tam informacji.

Przedstawiony w rozdziale materiał stanowi podstawę do formułowania tezy w sprawie celowości budowania i udostępniania systemów *IP*. Można mieć jednak pewne wątpliwości odnośnie do celowości przytaczania szeregu dość szczegółowych opisów, jak na przykład ilustracje obrazujące baner promujący Europejski Open Data Portal (s. 84) lub wygląd strony portalu *open data* według [www.data.gouv.fr](http://www.data.gouv.fr) (s. 88).

Rozdział 3 („Determinanty jakości informacji publicznej”) został poświęcony informacji w ogólnym ujęciu ze szczególnym wyróżnieniem informacji publicznej. Doktorant analizuje czynniki mające wpływ na jakość *IP* (nazywa je determinantami jakości informacji) oraz dokonuje przeglądu technologicznych uwarunkowań procesów przekazywania informacji publicznej do ponownego użycia.

W podrozdziale 3.1 przedstawia poglądy innych autorów w sprawie terminu *informacja*, zaznaczając, że jest to pojęcie bogate w swojej treści i trudne do jednoznacznego zdefiniowania. Doktorant poddaje krytyce niektóre poglądy cytowanych autorów oraz uzupełnia ich poglądy własnymi propozycjami. W szczególności na s. 104 zwraca uwagę na konieczność uwzględniania źródła informacji w jego wymiarze wiarygodności, rzetelności i rangi w odbiorze społecznym informacji. Zwraca też uwagę na czynniki sprzyjające poprawnemu odczytowi treści informacyjnych oraz czynniki utrudniające taką analizę. Wskazuje na konieczność zwrócenia uwagi na kryteria odbioru informacji (Doktorant określa te czynniki jako wyróżnione cechy odbieranych treści – s. 106). Formułuje także własne wnioski w sprawie czynników negatywnych (zniekształcających informację) oraz czynników podnoszących jej czytelność (na przykład oprawa graficzna – s. 107). Wprowadził też pewne własne określenia, jak na przykład *sprawcza ocena informacji*, istotne ze względu na tezy rozprawy.

Istotne w kontekście rozprawy są w tym rozdziale spostrzeżenia Doktoranta w sprawie informacji publicznej, a zwłaszcza jej roli w społeczeństwie – jej roli jako „siły łagodnej” w oddziaływaniu na społeczeństwo, jako składnika wiedzy zbiorowej czy czynnika integracyjnego.

W podrozdziale 3.2 Doktorant przedstawił szereg faktów na temat sposobów przekazywania (komunikowania) informacji *IP*. W szczególności dokonał przeglądu standardów przekazywania informacji w serwisach internetowych z wyróżnieniem serwisów administracji publicznej. Przedstawił w zarysie stosowane formaty danych oraz odpowiednie standardy. Sformułował przy tym interesujący wniosek w sprawie swobodnego dostępu do informacji publicznej (s. 121, 123), a mianowicie konieczność zapewnienie standardu znaczeniowego i syntaktycznego występujących pojęć, czyli wspólnego słownika metadanych i wspólnego systemu danych referencyjnych. Wytknął luki w rozwiązaniu oferowanym przez GUS w Repozytorium Danych Statystycznych (s. 122).

Miejscami, niestety, myśli Doktoranta są niejednoznaczne i polemiczne. Na przykład na s. 108 pisz on: „Informacja publiczna do ponownego użycia to informacja już zebrana w zbiory uporządkowane w różnym stopniu na poziomie datalogicznym oraz infologicznym”. Otóż należy zgodzić się z tą opinią w sprawie możliwości

porządkowania informacji na poziomie datalogicznym jako zbioru danych. Ale czy jakiegos porządkowania *a priori* można dokonać na poziomie infologicznym – na poziomie subiektywnym?

W podrozdziale 3.2 zbędne wydają się szczegóły i ilustracje odnoszące się do prezentowanych standardów – standardu *RDF*, *TriG*. W rozprawie mogłyby one znaleźć się w odpowiednich załącznikach. Podobnie w załącznikach mogłyby zostać ~~się~~ umieszczone obszerne dane podane w niektórych tablicach, jak na przykład tablica 11 „Zestawienie cech jakości informacji proponowanych przez różnych autorów”.

W tymże podrozdziale Doktorant zwraca uwagę na pewne koncepcje pomiaru jakości *IP*, formułowane przez innych autorów, poddając je krytyce. Upoważnia go to do formułowania tezy w sprawie celowości dalszego poszukiwania rozwiązania problemu tego pomiaru. Podkreśla istnienie istotnych problemów związanych z tym pomiarem. W szczególności zauważa, że samo pojęcie *informacja* jest pojęciem nieostrym; podobnie pojęcie *jakość*. Podkreśla, że ważne jest zdefiniowanie obu pojęć oraz określenia kryteriów jakościowych informacji w ogóle, w szczególności kryteriów odnoszących się ~~do~~ *IP*. Specyfikacji takich mierników dokonuje w końcowym podrozdziale

Rozdział kończy się następującą konkluzją (s. 157): „(...) atrybuty użyteczności stanowią miarę jakości, a przypisanie im wartości kwantyfikujących umożliwia ocenianie poziomu jakości informacji publicznej do ponownego życia”. Stanowi to podstawę do rozwinięcie autorskich koncepcji tego pomiaru w następnym rozdziale.

Rozdział 4 („Model pomiaru jakości informacji publicznej do ponownego użycia”) zawiera opis propozycji autorskiej modelu pomiaru jakości *IP* (modelu oznaczonego przez Doktoranta akronimem ODQM). W kolejnych podrozdziałach Autor przedstawił założenia w sprawie budowy modelu, zaprezentował jego strukturę, wprowadził definicje kluczowych pojęć wstępujących w modelu – definicje klas i cech jakościowych informacji *IP* oraz mierniki tej jakości.

Interesująca jest koncepcja Doktoranta w sprawie protokołu wejścia, który w modelu ODQM służy jako mechanizm do oceny poziomu jakości napływających informacji wejściowych. Protokół został zaprojektowany w postaci zestaw pięciu grup kryteriów jakościowych, odpowiadających pięciu opisanym wcześniej klasom jakości informacji. W każdym zestawie Autor wyróżnia pewien podzbiór kryteriów kluczowych,

czyli cech (nazywa je *krytycznymi*), które w procesie oceny jakości informacji odgrywają zasadniczą rolę. Cechy te traktuje jako *warunki konieczne* (nazywa je *wymaganiami niezbędnymi*), które muszą być spełnione, aby rozpatrywany zestaw informacji mógł zostać zakwalifikowany jako informacje o wystarczającym poziomie jakości.

Interesująca jest propozycja Doktoranta w sprawie infograficznej prezentacji uzyskanej wartości wskaźnika jakości: sprzyja to znacznemu zwiększeniu czytelności wyniku.

Doktorant uzupełnia opis modelu swoją oceną ze względu na jego zalety, jak i ograniczenia. Do zalet zalicza możliwość automatyzacji procesu analizy jakości informacji na podstawie modelu przy uniezależnieniu od platformy informatycznej oraz wprowadzenie wspomnianych infografik.

Do słabych stron obecnej wersji rozwiązania Doktorant zalicza brak uwzględnienia w modelu rangi (poziomu doświadczeń) ekspertów, których opinia służy za podstawę budowania mierników. Ale zaznacza, że model pozwala na usunięcie tego ograniczenia.

W rozdziale 5 („Zastosowanie modelu pomiaru w odniesieniu do portali *open data* wybranych krajów”) Doktorant przedstawił wyniki przeprowadzonych przez siebie badań empirycznych na bazie modelu ODQM do oceny informacji *IP* wybranych portali: polskiego portalu *open data* – *danepubliczne.gov.pl* oraz portali *open data* wybranych ośmiu innych krajów. Doktorant zademonstrował wyniki uzyskane według zaproponowanej przez siebie wcześniej (w rozdziale 4) procedury.

W „Podsumowaniu” (ostatni rozdział dysertacji) Doktorant dokonał podsumowania swojego dorobku naukowego i poddał generalnej ocenie swoją pracę. Podkreślił, że sformułowany we wstępie cel rozprawy („opracowanie autorskiej koncepcji modelu pomiaru jakości informacji publicznej na potrzeby ponownego użycia” – s. 7) został osiągnięty. Oznacza to, że Doktorant zbudował zarówno koncepcyjny model (ODQM), jak i zbudował niezbędne modelowe rozwiązanie informatyczne. Zweryfikował też jego użyteczność praktyczną na podstawie przeprowadzonych przez siebie obliczeń opartych na danych rzeczywistych, opisujących wybrane przez siebie portale *open data*, stanowiących źródło informacji o charakterze *IP*. Przeprowadzony test wykazał spójność modelu i jego praktyczną użyteczność.

Doktorant sformułował też kierunki dalszych badań w tym zakresie i cele uzupełnienia swojego rozwiązania.

### **Ocena rozprawy i dorobku naukowego Doktoranta**

Doktorant dokonał trafnego wyboru obszaru badawczego ze względu na jego wymiar poznawczy i znaczenie praktyczne. Postawił przy tym przed sobą ambitne zadanie: opracować instrument pomiarowy jakości informacji *IP*. Zadanie to jest o tyle trudne, że – jak pokazał w rozprawie – specjaliści podkreślają występowanie tutaj wielu problemów, wynikających z ontologicznej natury informacji, nieostrości pojęcia jakości i wynikających stąd trudności w zakresie wyspecyfikowania odpowiednich kryteriów jakościowych.

Doktorant „ominał” te trudności (zapewne coś jeszcze pozostało?), wprowadzając pojęcie protokołu wejścia jako swoistego filtra informacji ze względu na jej jakość, uzupełnionego względnie prostą procedurą postępowania. Jest to rozwiązanie oryginalne i eleganckie, ponieważ jest proste.

Jego rozprawa wnosi istotny wkład naukowy zarówno do teorii, jak i praktyki w obszarze nauk ekonomicznych. Na uwagę zasługują następujące fakty:

- Doktorant wzbogacił literaturową interpretację pojęcia *informacja publiczna ponownego użycia* o własne propozycje ze wskazaniem szczególnych cech i funkcji tego rodzaju informacji w komunikacji publicznej.
- Zoperacjonalizował wybrany problem badawczy w sposób pozwalający na podjęcie próby znalezienia jego rozwiązania – nadał mu charakter problemu mierzalnego.
- Sformułował i zbudował model konceptualny w postaci zestawu pięciu klas jakości informacji *IP* i opisał każdy z nich za pomocą listy cech jakościowych, którym nadał rangę kryteriów pomiarowych. Wyzaczył przy tym kryteria kluczowe (nazwał je kryteriami krytycznymi), których spełnienie decyduje o uznaniu informacji jako spełniającej oczekiwania użytkownika pod względem jej jakości.
- Opracował koncepcję protokołu wejścia, stanowiącego filtr informacji ze względu na jej jakości.
- Zdefiniował mierniki jakości informacji *IP*.

- Na tej podstawie opracował model ODQM. Przeprowadzony przez Doktoranta eksperyment obliczeniowy dowiódł jego poprawności obliczeniowej oraz możliwości jego praktycznego użytkowania.

Należy też odnotować rozwinięty i dojrzały warsztat badawczy, jakim posłużył się Doktorant: od wnikliwej i twórczej analizy literatury specjalistycznej począwszy, na opracowaniu autorskiego rozwiązania w zakresie jakości informacji i przeprowadzeniu eksperymentu badawczo-obliczeniowego skończywszy.

### **Uwagi i propozycje**

Z obowiązku recenzenta pragnę zwrócić uwagę na kilka kwestii o charakterze ogólnym, częściowo polemicznym:

- We wstępie brakuje założeń, na bazie których Autor sformułował tezy w sprawie istnienia szans na pomiar jakości informacji publicznej i – w konsekwencji – możliwości opracowania stosownego modelu postępowania. W obecnej wersji można jedynie domniemywać, że Autor takie założenia (podstawy) jest w stanie wykazać – na przykład podczas obrony swoich tez przed Radą Wydziału.
- Stosowany przez Autora zwrot „informacja publiczna do ponownego użycia” mogłby zostać zastąpiony w tekście jakimś skrótem lub symbolem – to mogłoby ułatwić czytelnikowi wyłowienie go tekście.
- Zbędne są niektóre szczegóły, jak na przykład ilustracje na rys.17 lub listingi na s. 130 i 132. W mojej opinii, nie wnoszą one treści związanych z tezami rozprawy.
- Przytoczony wykaz źródeł bibliograficznych jest interesujący. Z punktu formalnego Autor nie zadbał o jego uporządkowanie alfabetyczne.
- W pracy zaznaczyło się tempu przygotowania tekstu rozprawy. Autor nie ustrzegł się pewnych niezręczności stylistycznych oraz niedopatrzeń formalnych w zakresie interpunkcji i „literówek”.
- Wskazane byłoby wyraźniejsze wyartykułowanie autorskich myśli i propozycji – na przykład przez spisanie cech i funkcji *IP* w formie ich enumeratywnego wyliczenia z numeracją.
- Na s. 156 pojawiły się pewne drobne błędy i nieścisłości w przytoczonych tam wzorach i w tekście, utrudniające poprawny odczyt ich treści. Warto przy okazji

zapropnować ponumerowanie przytaczanych wzorów – podobnie jak numerowanie tablic. Ułatwia to czytelnikowi odnajdywanie się w tekście.

- Do słabszych stron rozprawy należy zaliczyć nieprecyzyjność formułowanych założeń oraz połączenie założeń z wnioskami w sprawie wykorzystania modelu.

### **Wnioski**

Rozprawa doktorska mgr Macieja Możejewskiego stanowi istotny wkład teoretyczny i praktyczny w rozwój wiedzy w zakresie informacji w ogóle, a w szczególności w odniesieniu <sup>do</sup> jakości informacji oraz w odniesieniu do informacji publicznej o charakterze wielokrotnego użycia. Dorobek naukowy Doktoranta, zaprezentowany w dysertacji, podkreśliłem w poprzednim punkcie. Teraz pragnę zaznaczyć, że Doktorant wykazał umiejętność krytycznego odbioru formułowanych w literaturze opinii i koncepcji oraz umiejętność formułowania własnych propozycji.

Stwierdzenia te upoważniają mnie do sformułowania końcowego wniosku: recenzowana rozprawa doktorska mgra Macieja Możejewskiego spełnia wymagania stawiane rozprawom doktorskim w Ustawie o stopniach naukowych i tytułach naukowych. W związku z tym wnioskuję o dopuszczenie jej Autora do publicznej obrony.

Bogdan Stefanowicz