

18 kwietnia 2016 r.

Prof. dr hab. Mirosław Szreder
Kierownik Katedry Statystyki
Uniwersytetu Gdańskiego
ul. Armii Krajowej 101
81-824 Sopot

**Recenzja rozprawy doktorskiej mgr. Macieja Beręsewicza
pt. *Internet data sources for real estate market statistics*
napisanej pod kierunkiem prof. UEP dr hab. Elżbiety Gołaty, z udziałem
promotora pomocniczego dr. Marcina Szymkowiaka**

Powierzona mi do oceny rozprawa doktorska dotyczy ważnego i bardzo aktualnego zagadnienia, jakim jest ocena możliwości wykorzystania w badaniach statystycznych internetowych zbiorów danych oraz kombinacji tych zbiorów z wybranymi rejestrami administracyjnymi. Podjęty w dysertacji mgr. Macieja Beręsewicza temat otwartości statystyki na nowe zasoby informacji, w tym przede wszystkim na zwiększające się szybko ilości danych dostępnych w Internecie, uznaję za wart badań naukowych oraz szerokiej dyskusji w gronie daleko wykraczającym poza środowisko statystyków. Problematyka dużych zbiorów danych i ogromnych, nieznanych nigdy wcześniej możliwości ich gromadzenia i przetwarzania, a więc tego wszystkiego co powszechnie rozumie się pod pojęciem Big Data, zajmuje uwagę nie tylko statystyków, ale również specjalistów z zakresu: sztucznej inteligencji, informatyki i socjologii. Jednym z kluczowych wyzwań – jak słusznie zauważa Autor rozprawy – jest zdolność użytkowników tych zbiorów danych do oceny ich jakości, w tym reprezentatywności, przesądzającej o możliwości dokonywania na ich podstawie uogólnień na większe populacje.

Celem rozprawy doktorskiej mgr. Macieja Beręsewicza jest „ocena internetowych źródeł danych dla potrzeb analiz statystycznych rynku nieruchomości” (s. 5). Uważam, że jest to poprawnie sformułowany cel badawczy. Sądzę ponadto, że w tytule rozprawy lepiej było dodać słowo „ocena”, jako że obecne brzmienie tytułu: *Internet data sources for real estate market statistics* nie sugeruje w żaden sposób rodzaju zagadnień z zakresu internetowych zbiorów danych, którym praca jest poświęcona. Pozytywnie oceniam sformułowane na s. 5 rozprawy trzy cele szczegółowe, podporządkowane głównemu celowi badawczemu. Zastrzeżeń nie zgłaszam do hipotez badawczych, z których naczelną hipotezę, brzmiącą następująco: *Internetowe źródła danych umożliwiają estymację charakterystyk rynku nieruchomości z akceptowaną wielkością błędu*, uważam za ambitną i odważną. Weryfikacja tej hipotezy wymaga bowiem rozwiązania kilku problemów natury ogólniejszej, wykraczającej poza specyfikę rynku nieruchomości. W szczególności chodzi o określenie wielkości i

