

Zastosowanie danych administracyjnych dla lepszego kształtowania polityki w państwach EŚW

Policy Brief

Władze publiczne w krajach postsocjalistycznych gromadzą dużą ilość danych na temat obywateli i przedsiębiorstw, lecz dane administracyjne rzadko wykorzystywane są do badań naukowych lub dla lepszego kształtowania polityki. Niniejszy dokument analizuje przyczyny tej sytuacji i przedstawia zalecenia, mające umożliwić skuteczniejsze wykorzystanie danych administracyjnych do prowadzenia polityki opartej na dowodach.

Postsocjalistyczne kraje europejskie odziedziczyły po centralnie planowanych gospodarkach tradycję gromadzenia przez państwo bardzo dużej ilości danych. Dane administracyjne zbierane przez urzędy podatkowe, fundusze emerytalne i inne organy rządowe są często (choć nie bez wyjątku) dość wysokiej jakości. Dane administracyjne dotyczące osób fizycznych i firm mają swoje ograniczenia, ale pod wieloma względami są lepsze od danych ankietowych. Na przykład często są bardziej rzetelne, zwłaszcza w porównaniu do badań ankietowych, w których osoby proszone są o przypomnienie sobie czegoś z przeszłości albo opisanie różnych źródeł dochodów. Poza tym w przypadku danych administracyjnych w przystępnej cenie otrzymuje się duże próby badawcze. Oprócz tego, łącząc ze sobą bazy danych, można wygenerować taką ilość informacji na temat (zanonimizowanych) przedsiębiorstw lub osób fizycznych w ramach skomplikowanej struktury danych, której nigdy nie udałoby się uzyskać w badaniach ankietowych.

W Europie Zachodniej regularnie korzysta się z danych administracyjnych do oceny interwencji rządów. Stanowią one poza tym wspaniałą zasób badawczy dla społeczności naukowej. Kraje nordyckie mają długą tradycję udostępniania danych administracyjnych dla celów naukowych (zob. [Figlio et al 2015](#)), natomiast w innych państwach podobne inicjatywy powstały stosunkowo niedawno. Na przykład niemiecki rząd powołał komisję do zbadania tej kwestii w 1999 r., dwa lata później ustanowił [Niemieckie Forum Danych](#), a w 2003 r. otworzył dwa centra obsługi danych ([GML](#) i [IDSC](#)). Pod koniec 2011 r. Wielka Brytania powołała [zespół roboczy](#), mający zbadać najbardziej optymalne metody udostępniania danych administracyjnych dla celów badawczych. Wskutek jego rekomendacji w 2013 r. ustanowiono [Sieć Badawczą Danych Administracyjnych](#), której zadaniem jest opracowanie infrastruktury zapewniającej łatwy i etyczny dostęp do danych. Na szczęblu unijnym w 2003 r. przyjęto dyrektywę promującą ponowne wykorzystanie informacji sektora publicznego ([Dyrektywa 2003/98/WE](#)), która dziesięć lat później została zaktualizowana ([Dyrektywa 2013/37/UE](#)). [Centrum](#)

Kompetencyjne Badań Mikroekonomicznych Komisji Europejskiej również aktywnie promuje wykorzystanie danych administracyjnych przez rządy państw członkowskich UE.

Korzystanie z danych administracyjnych często wymaga zaawansowanych umiejętności z zakresu oczyszczania i analizy danych. W większości krajów Europy Środkowo-Wschodniej (EŚW) jest też wielu świetnych naukowców i statystyków, którzy mogliby sprostać temu wyzwaniu i wykorzystywać dane administracyjne do lepszego kształtowania polityki.

Czemu więc nie korzystamy z tych zasobów?

Jak na ironię jedną z przeszkód są ustawy dotyczące ochrony danych osobowych wprowadzone na początku lat 90., gdy w państwach EŚW na nowo wprowadzono instytucje demokratyczne. Choć badacze pracują na zanonimizowanych danych, nawet ich użycie może zostać skutecznie zablokowane przez sztywne przepisy dot. ochrony danych. W większości krajów EŚW są one bardzo restrykcyjne, istnieją też dodatkowe obostrzenia dotyczące informacji o pochodzeniu etnicznym, stanie zdrowia lub innych cechach wrażliwych. Podczas gdy w większości państw obywatele mają ustawowe prawo dostępu do informacji publicznej, dotyczy ono w większości danych zagregowanych, zaś jego egzekwowanie i instytucje za nie odpowiedzialne są dość słabe. Bez przejrzystych ram prawnych i procedur formalnych dostęp do indywidualnych danych administracyjnych dla celów naukowych jest często określany *ad hoc*, w nieformalnych procedurach.

Kolejną przeszkodą wspólną dla krajów EŚW jest brak zaufania między organizacjami akademickimi i rządowymi, jak też w ramach samych rządów. Na przykład na Węgrzech, po serii skandali medialnych dotyczących nadużywania danych osobowych przez urzędników publicznych w latach 90., znacznie spadło zaufanie opinii publicznej do sposobu, w jaki rząd wykorzystuje informacje o obywatelach. W sektorze publicznym pojawiły się zaś obawy dotyczące potencjalnych konsekwencji korzystania z danych zindywidualizowanych. Właściciele danych zazwyczaj boją się utracić pozycję monopolisty w zakresie publikowania (lub sprzedaży) danych albo tego, że użytkownicy zewnętrzni odkryją jakieś niedociągnięcia w ich jakości. Zainteresowane strony mogą też nie dostrzegać konieczności usprawnienia dostępu do danych, jeśli nie są świadome postępów w metodach statystycznych i potencjału korzystania z danych zindywidualizowanych.

Pewną rolę może tu wreszcie odgrywać stosunkowo niska skuteczność zarządzania: rządy rzadko zamawiają stosowne (tzn. rzetelne z naukowego punktu widzenia) badania dotyczące skutków prowadzonej przez nie polityki. Ograniczone zainteresowanie mediów lub używanie zbytnich uproszczeń w dyskursie publicznym na temat odpowiedzialności rządzących może dodatkowo zmniejszyć zapotrzebowanie na zaawansowane analizy.

Sytuacja zdaje się jednak rozwijać w dobrym kierunku

Niektóre kraje EŚW zaczynają odpowiadać na wewnętrzne wyzwania i unijne dyrektywy. Nowej motywacji dostarczyła Dyrekcja Generalna ds. Zatrudnienia Komisji Europejskiej, która niedawno wprowadziła zasadę zobowiązującą oferentów ubiegających się o dostęp do Europejskich Funduszy Strukturalnych do opracowania co najmniej jednej ewaluacji w oparciu o mikrodane i metody kontryfakcyjne.

Inicjatywy dotyczące otwartych danych nabrały rozpędu w [Chorwacji](#), [Czechach](#), na [Słowacji](#) i w [Słowenii](#). Choć skupiają się one przede wszystkim na danych zagregowanych, mogą stanowić ważny pierwszy krok w tworzeniu środowiska użytkowników. W Polsce ustanowiono Ministerstwo Cyfryzacji, które [promuje](#) powtórne wykorzystanie danych administracyjnych. Oprócz tego regularnie korzysta się z [danych z systemu ubezpieczeń społecznych](#) do śledzenia kariery absolwentów uczelni wyższych na polskim rynku pracy oraz z danych dotyczących zarejestrowanych bezrobotnych – do oceny aktywnych polityk rynku pracy. Na czele stawki póki co plasują się Węgry, które w 2007 r. przyjęły Ustawę 101 usuwającą główną przeszkodę istniejącą w poprzednim prawie, umożliwiając właścicielom danych przetwarzanie danych osobowych w celu ich anonimizacji. Dzięki temu osiągnięto znaczące postępy w łączeniu baz danych administracyjnych i opracowaniu dostępnego dla naukowców [banku danych](#) w ramach [Instytutu Ekonomii Węgierskiej Akademii Nauk](#).

Naukowcy mogą wzmocnić te pozytywne trendy na wiele sposobów

Mogą zwiększyć publiczną świadomość istnienia i wagi badań opartych o dane administracyjne, prezentując ich wyniki w atrakcyjnej, zrozumiałej i dostępnej formie. Działania te będą skuteczniejsze, jeśli nawiążą oni kontakty z dziennikarzami i zaczną współpracować ze sobą nawzajem w zakresie zarządzania relacjami medialnymi. Naukowcy powinni też bardziej aktywnie reagować na badania ewaluacyjne polityk powstające na zamówienie rządu: chwając każdy przypadek sięgania po oceny zewnętrzne, jednocześnie zwracając uwagę na potencjalne niedociągnięcia wynikające z zastosowania nieodpowiedniej metody oceny. Wspólne wysiłki na rzecz wprowadzenia zajęć z metod oceny do programu szkoleń dla administracji publicznej również mogą być pomocne. Powyższe kroki miałyby pozytywny wpływ na budowanie zaufania i zrozumienia wśród zainteresowanych podmiotów, przygotowując grunt pod inicjowanie zmian w prawie. W międzyczasie zaś społeczność badawcza powinna zaangażować inne podmioty w monitorowanie przyjęcia nowego [unijnego rozporządzenia o ochronie danych osobowych](#), pilnując, aby badania naukowe i ewaluacja prowadzonej polityki zostały uwzględnione jako uprawnione powody przetwarzania, archiwizowania i analizowania danych wrażliwych.

* * *

Powyższe wnioski były efektem dyskusji przeprowadzonych w ramach warsztatu, który odbył się w Budapeszcie 21-22 września 2017 r. Uczestniczyli w nim przedstawiciele państw wyszehradzkich, Albanii, Bośni i Hercegowiny, Chorwacji, Rumunii, Serbii i Słowenii. Warsztat został przeprowadzony przez [Węgierskie Towarzystwo Ekonomistów](#), przy wsparciu [Funduszu Wyszehradzkiego](#) i zorganizowany we współpracy z [Instytutem Budapeszteńskim](#), [Środkowoeuropejskim Instytutem Badań nad Pracą – CELSI](#), [Centrum Badań Ekonomicznych i Kształcenia – Instytutem Ekonomicznym, CERGE-EI](#) oraz [IBS](#).

autor: Ágota Scharle

tłumaczenie: Aleksandra Szkudłapska

Kontakt do Instytutu Budapeszteńskiego

Ágota Scharle, agota.scharle@budapestinstitute.eu

Kontakt do Środkowoeuropejskiego Instytutu Badań nad Pracą (CELSI)

Martin Kahanec, martin.kahanec@celsi.sk

Kontakt do Centrum Badań Ekonomicznych i Kształcenia – Instytutu Ekonomicznego (CERGE-EI)

Martin Srholec, martin.srholec@cerge-ei.cz

Kontakt do Instytutu Badań Strukturalnych (IBS)

Iga Magda, iga.magda@ibs.org.pl