

Magdalena Szpunar
magdalena_sz@wp.pl
WNSS
AGH
Kraków

Czy prowadzenie badań sondażowych drogą online ma sens? Elektroniczne narzędzia do badań online – casus eBadań

Badania prowadzone drogą online ciągle zyskują na znaczeniu. Coraz więcej osób decyduje się na prowadzenie badań przez Internet, nie tylko jakościowych, ale ilościowych. Niestety, publikowane w Sieci ankiety niejednokrotnie traktowane są jako gorsze edycje, tradycyjnych, papierowych wydań kwestionariuszy. Niejednokrotnie osoby decydujące się na prowadzenie badań sondażowych w Internecie, nie mają świadomości wad związanych z prowadzeniem badań tą drogą, ale także niewątpliwych zalet, jakie oferuje to medium. W związku z rosnącą liczbą badań sondażowych prowadzonych w interaktywnym środowisku, palącą potrzebą wydaje się być wypracowanie standardów prowadzenia badań w tej przestrzeni.

Niejednokrotnie poddaje się w wątpliwość sens prowadzenia badań w Sieci, zarzucając im niską jakość, nieprzestrzeganie standardów metodologicznych, niski stopień reprezentatywności, co w konsekwencji utrudnia ekstrapolację uzyskanych wyników na populację generalną. Zastrzeżenie to dotyczy szczególnie osób realizujących internetowe badania sondażowe do swoich prac licencjackich czy magisterskich. Narzędzia badawcze obfitują w błędy metodologiczne, np. nierozłączne zakresy odpowiedzi, niewyczerpywalne kafeterie, czy brak instrukcji do pytań. Autorzy ankiet online nierzadko nie dbają o formę i jakość publikowanych kwestionariuszy. Łatwość publikowania w Sieci jednopytaniowych sond, możliwa dzięki wielu narzędziom oferującym tę usługę, sprawia, iż wiele osób jest przekonanych, że w Internecie nie można prowadzić poprawnych metodologicznie, a tym bardziej reprezentatywnych badań. Ze względu na ograniczone ramy artykułu pominę kwestie związane z zaletami prowadzenia badań w Sieci, o czym pisałam niejednokrotnie [zob. Szpunar, 2007; Szpunar, 2007a; Szpunar, 2007b]. W niniejszym artykule przedstawione zostaną kwestie problematyczne związane z prowadzeniem badań sondażowych w Sieci, ze szczególnym naciskiem na kwestie reprezentatywności i możliwości uogólniania uzyskanych drogą online wyników.

Socjologowie podnoszą dwie sprawy związane z prowadzeniem badań w Sieci. Po pierwsze – zastanawiają się, jak przystosować tradycyjne narzędzia badawcze do zapośredniczonego przez medium środowiska, po wtóre – jak stworzyć unikalne narzędzia, przestrzegające standardów metodologicznych, pozwalające na prowadzenie sensowych badań w Sieci. Wychodząc naprzeciw temu wyzwaniu stworzyliśmy System Zarządzania Badaniami *eBadania.pl*.

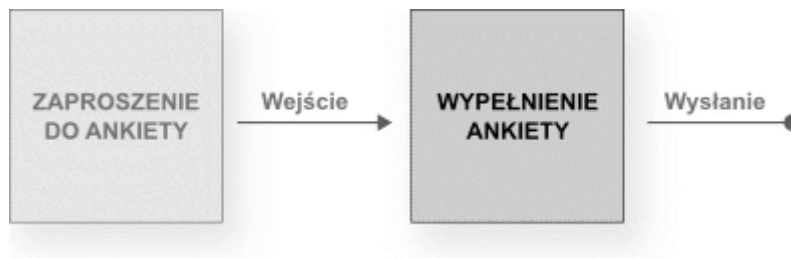
Utworzony pod koniec 2006 roku serwis eBadania bazuje na solidnych standardach badań socjologicznych. W odróżnieniu od serwisów oferujących przygotowanie prostej

sondy w trzech krokach staraliśmy się nie wmawiać użytkownikom, że stworzenie ankiety jest banalnie proste. Dla osób nieposiadających wiedzy z zakresu metodologii badań empirycznych stworzono rozbudowany system pomocy, umożliwiający postawienie pierwszych kroków w badaniach kwestionariuszowych i skonstruowanie poprawnej ankiety. Socjologom lub zaawansowanym badaczom narzędzie umożliwia przygotowanie kwestionariusza zgodnie z regułami metodologicznymi obowiązującymi przy standardowych ankietach [Szpunar, 2007a, s. 14].



Rys. 1. Zrzut ekranu głównego Systemu Zarządzania Badaniami eBadania.pl

Serwis eBadania został przygotowany z myślą o badaczach zamierzających prowadzić badania sondażowe drogą online. Składa się on z dwóch zasadniczych części. Pierwszy z modułów to część przeznaczona dla respondenta, w której zapraszany jest on do udziału w ankiecie, ma możliwość jej wypełnienia ankiety oraz wysłania.



Rys. 2. Moduł respondenta

Drugi moduł systemu eBadania – ankierski – składa się z panelu, umożliwiającego edycję wszystkich typów pytań, tj. zamkniętych, otwartych, półotwartych, skali oraz rangi. Badacz ma możliwość ustalenia charakteru pytania (jednokrotnego lub wielokrotnego wyboru), ustalenia zakresu odpowiedzi (liczby minimalnej i maksymalnej dopuszczalnych odpowiedzi). Ma również możliwość „wymuszenia” na respondencie konieczności udzielenia odpowiedzi według ustalonego schematu, dzięki pojawiającym się alertom, co pozwala eliminować do zera braki odpowiedzi. *System Zarządzania Badaniami* eBadania umożliwia również dodawanie, usuwanie i zmianę kolejności pytań, dołączanie instrukcji do pytań, dodawanie, usuwanie i zmianę kolejności odpowiedzi, wprowadzanie reguł przejścia, które umożliwiają „przefiltrowanie” respondentów, którzy nie powinni odpowiadać na dane pytanie lub np. zadeklarowali brak wiedzy w danym temacie. Dużą zaletą eBadań jest szeroki wachlarz możliwości eksportu wyników do programów służących do statystycznej analizy danych (SPSS, Statistica) oraz do najpopularniejszego programu analityczno-kalkulacyjnego – MS Excel. Pozwala to eliminować żmudny proces wprowadzania danych, a co się z tym wiąże, częstych na tym etapie błędów, związanych z czynnikiem ludzkim. Warto odnotować, iż inne narzędzia służące do prowadzenia badań kwestionariuszowych w Internecie oferują głównie sumaryczne zestawienie danych, bez możliwości ich dalszej analizy statystycznej. Badacz ma również możliwość załączania na stronie eBadań raportów z przeprowadzonych przez siebie badań. *System Zarządzania Badaniami* generuje unikalny link do strony z ankietą. Badacz może go wklejać na innych stronach WWW promujących jego badanie lub załączać do maili zachęcając potencjalnych respondentów do wypełnienia ankiety [Szpunar, 2007a, s. 15].



Rys. 3. Moduł ankietera

Mimo, że Internet zwiększa możliwości eksploracji dla badacza, jednocześnie pojawia się problem jakości uzyskiwanych danych i możliwości generalizacji otrzymanych wyników. Wskazuje się, że populacja internautów, jak i osób niekorzystających z sieci znacznie się od siebie różni pod względem demograficznym, społecznym i psychologicznym. W USA użytkownikiem sieci jest częściej osoba rasy białej, młoda i posiadająca dzieci. W Polsce jest osobą młodą (do 35 lat), uczącą się lub studiującą, mieszkającą w dużej miejscowości. Nie tylko próba może mieć charakter stroniczy, gdyż kolejne problemy pojawiają się w związku z dokonywaniem samowyboru przez internautę, który wyraża zgodę lub nie na wypełnienie ankiety [Kraut i inni, 2004, s. 107-108]. Zauważa się również, że procenty odpowiedzi uzyskiwane online są niższe, niż w przypadku ankiety pocztowej, czy wywiadu telefonicznego, choć wyniki różnych badań ukazują znaczne rozpiętości i różnice w tym zakresie. Poza tym respondenci ciągle jeszcze preferują tradycyjną, papierową wersję narzędzia. Nawet osoby, które sprawnie posługują się nowymi technologiami, np. uczniowie czy studenci, mimo wszystko preferują tradycyjne – papierowe wersje ankiet [Schonlau, Asch, Du, 2003].

Problem pojawia się również w momencie, gdy chcemy prowadzić longitudinalne badania. Internauci często zmieniają adresy poczty elektronicznej, co uniemożliwia ponowne dotarcie do tego samego respondenta [tamże]. Nieaktualne mogą być nie tylko adresy kont pocztowych, ale także np. usługi, poprzez które wysłano zaproszenie do użytkowników. Internet jest medium dynamicznie się zmieniającym, a okres 2-3 lat to w Sieci niemal wieczność.

Nie jest jednak tak, że Internet nie daje możliwości doboru respondentów według interesujących nas zmiennych. Można tego dokonać – na przykład – przy pomocy okienka pop-up. Jest to graficzny element „wyskakujący” w osobnym oknie przeglądarki internetowej po wejściu użytkownika na dany serwis lub stronę WWW. W okienku tym możemy prosić respondenta o podanie interesujących nas zmiennych, tj. płci, wieku, miejsca zamieszkania, wykształcenia itp. W momencie, kiedy respondent jest tym, którego poszukujemy przechodzi do zasadniczej części badania, w której znajduje się kwestionariusz ankiety, jeśli zaś podane przez niego informacje nie odpowiadają poszukiwanym przez nas danym, okienko się zamyka i nie ma możliwości dalszego wypełniania ankiety [Szpunar, 2007, s. 88].

Analizując badania realizowane drogą internetową podkreśla się również brak kontroli tego, kto naprawdę odpowiada na pytania ankiety, co podważa wartość uzyskiwanych danych [Zmijewska-Jędrzejczyk, 2004, s. 245]. Z analogiczną sytuacją mamy do czynienia w ankiecie pocztowej, czy prasowej, nie jest to, więc wyłącznie specyfika badań internetowych. Ze względu na niską penetrację użytkowników Internetu, podważa się również reprezentatywność i możliwość ekstrapolacji uzyskanych wyników na populację generalną. Mamy tutaj do czynienia z tzw. błędem pokrycia (ang. *coverage error*), który oznacza, że nie wszystkie jednostki badanej populacji mają szansę dostania się do próby wylosowanej do badania – nie wszystkie osoby w Polsce mają dostęp do Internetu, a wręcz stanowią one mniejszość¹, stąd zbyt duża liczba jednostek populacji ogólnej nie ma szans dostać się do próby [Sobieszek, 2006, s. 369]. Według danych CBOS w marcu 2008 roku 48% gospodarstw domowych w Polsce posiadało dostęp do Internetu, a 57% było

wyposażonych w komputer [Wenzel, 2008, s. 2]. Pod względem liczby Internautów (14,1 mln) Polska zajmuje w Europie 9 miejsce. Wśród narodów europejskich najliczniejszą grupę w Sieci stanowią Niemcy – 53,2 mln internautów, Brytyjczycy – 40,4 mln, Francuzi 34,9 mln oraz Włosi 33,1 mln. Najwyższym poziomem penetracji Internetu mogą się poszczycić Norwegowie 88%, Holendrzy 87,8%, Szwedzi 77,3% oraz Portugalczycy 73,1% [<http://www.internetworldstats.com>]. Internauci stanowią populację o specyficznych cechach. Stanowią zbiorowość, w znaczący sposób różniącą się od populacji generalnej Polaków. Statystyczny polski internauta jest osobą młodą (do 24 lat), legitymuje się wykształceniem wyższym lub jest w trakcie pobierania nauki, mieszka w mieście powyżej 500 tysięcy mieszkańców, co potwierdzają liczne badania: *Diagnoza Społeczna 2005*, *Krakowski Eksperyment Internetowy*, czy badanie prowadzone przez autorkę w 2006 roku – *Społeczna przestrzeń Internetu*. Wśród polskich internautów odsetek osób zajmujących kierownicze stanowiska, posiadających własne firmy oraz pracujących umysłowo jest znacznie wyższy, niż w całej populacji Polaków – 34% internauci, 17% Polacy, a odsetek uczniów i studentów wśród internautów jest trzy razy większy, niż w przypadku populacji generalnej Polaków [SMG/KRC, 2003]. Należy jednak zwrócić uwagę, iż wraz ze wzrostem penetracji Internetu zbiorowość internautów będzie coraz bardziej przypominać populację generalną Polaków. W krajach o wysokim stopniu nasycenia tą technologią zbiorowość internautów w dużym stopniu odzwierciedla populację generalną, pozwalając – między innymi – na ekstrapolację wyników uzyskanych przy pomocy ankiet online na całe społeczeństwo.

Niektórzy badacze wskazują również, iż osoby skłonne do udziału w badaniu online różnią się od tych, które nie podejmują takiej decyzji, np. aktywnie korzystają z Internetu, posiadają stałe łącze [Zajac, Batorski, 2007]. Warto więc badania prowadzone w Internecie traktować raczej w kategoriach eksploracyjnych niż opisowo-wyjaśniających.

Wśród ograniczeń badań realizowanych przez Internet należy wymienić [Mącik, 2005, s. 109]:

- częstą samoselekcję respondentów (próba składa się głównie z ochotników, którzy zechcieli wziąć udział w badaniu),
- możliwość wielokrotnego udziału w badaniu tych samych osób²,
- brak reprezentatywności populacji użytkowników Internetu dla populacji generalnej, nawet w najbardziej spenetrowanych przez Internet krajach świata (z reguły internauci są młodszy, lepiej wykształceni, bardziej zamożni od osób niekorzystających z Internetu),
- krótszy czas koncentracji podczas uczestniczenia w badaniu online – zmęczenie respondenta następuje po ok. 25-30 pytaniach,
- obniżający się w długim okresie realizacji badania odsetek wypełnień,
- brak możliwości zinterpretowania zachowania respondenta podczas badania ze względu na brak bezpośredniego kontaktu z ankierem (warto podkreślić, iż jest to wada typowa dla wszystkich nieosobistych form badań ankietowych).

Prowadząc badanie ankietowe drogą online, nie ma sensu wprowadzać pytań o wiedzę. Analogicznie bowiem, jak w przypadku ankiety pocztowej, nie wiemy, jaka część respondentów faktycznie posiadała tę wiedzę, a jaka sięgnęła do innych źródeł, by tę wiedzę zdobyć [por. Sztabiński, 1997, s. 120]. Ponadto w przypadku występowania pytań

trudnych, takich, z którymi respondenci nie mogą sobie intelektualnie poradzić, respondenci odczuwają znaczne lub co najmniej odczuwalne zaniżenie się ich samooceny [Rostocki, 1992, s. 42]. Z pewnością bezradność intelektualna wobec wymagań stawianych przez badacza jest dla respondenta przykra emocjonalnie [tamże, s. 47]. Odpowiadając na pytania w ankiecie pocztowej, jak i internetowej, respondenci mają możliwość poszukać odpowiedzi w książkach, zapytać małżonka lub przeszukać zasoby Internetu, w celu znalezienia trafnej odpowiedzi. Wysoki odsetek odpowiedzi na pytanie o wydarzenia rozgrywające się w 1989 roku na Placu Tiananmen, niekoniecznie będzie wskazywał na wysoki poziom świadomości historycznej, a raczej na umiejętności uzyskania tej informacji. Specyfika medium, jakim jest Internet, pozwala te informacje uzyskiwać szybko i efektywnie, nie odkrywając przed badaczem niewiedzy w danym zakresie. Trudno określić, jaki odsetek respondentów nie ma problemów z przyznaniem się do własnej niewiedzy, jaki zaś czuje się zażenowany, gdy nie zna odpowiedzi na pytanie i chce ją odnaleźć.

Warto podkreślić, iż nie dysponujemy żadną bazą danych zawierającą dokładną liczbę internautów, pozwalającą na przygotowanie operatu losowania, który mógłby stanowić podstawę do wyselekcjonowania próby badawczej. Internet z założenia jest anonimowy i dlatego nie istnieje – i prawdopodobnie nigdy nie będzie istnieć – pełna baza danych użytkowników sieci Internet. Nawet Telekomunikacja Polska jako główny dostawca Internetu dysponuje jedynie danymi osób dzierżawiących dane łącze, nie posiada ona natomiast danych dotyczących wszystkich osób korzystających z danego łącza, np. w jednym gospodarstwie domowym czy w jednej firmie. Konkludując, wykonanie badania reprezentatywnego realizowanego w Internecie jest na obecnym etapie niezwykle trudne. Aby zminimalizować niski stopień reprezentatywności badanej próby, można stosować ważenie obserwacji. Należy jednak pamiętać, iż w takim przypadku powinniśmy dysponować odpowiednio dużą próbą. Możemy również zastosować mechanizm losowania. Po wejściu na daną stronę WWW ankietę wyświetlana jest, co n-tej osobie, jednakże i ten mechanizm ma wiele ułomności i w minimalnym stopniu przyczynia się do zwiększenia stopnia reprezentatywności badanej zbiorowości. Problem ten rozwiązuje się sam w sytuacji, gdy zamierzamy prowadzić badanie celowe. W takiej sytuacji wystarczy wprowadzenie interesujących nas zmiennych metryczkowych na początku badania. Jeśli nas respondent nie odpowiada interesującej nas kategorii, okno z ankietą zamyka się, uniemożliwiając osobie badanej dalszy udział w ankiecie.

Badanie Internetu, mimo, iż jest to fenomen istniejący kilka lat rodzi określone problemy. Badanie tego medium rodzi te same problemy, kiedy stosuje się techniki ankiet samo zwrotnych, tj. brak ich reprezentatywności. Podstawowe zastrzeżenie dotyczące badań realizowanych drogą internetową (online) dotyczy ich reprezentatywności. Poddaje się w wątpliwość hipotezę, iż ludzie uczestniczący w badaniach sondażowych mogą stanowić odzwierciedlenie jakiejś populacji. E. Babbie zauważa, iż jeszcze do niedawna, ten sam rodzaj krytyki dotyczył sondaży telefonicznych, których rzetelności, nikt dzisiaj nie poddaje w wątpliwość. H. Taylor i G. Terhanian próbowali w roku 1998 oszacować wyniki wyborów na gubernatora i do senatu USA. W tym celu przeprowadzili cztery sondaże przez Internet [Babbie, 2004, s. 299]. Uzyskane przez nich wyniki mogą zaskakiwać – trafnie wytypowano 21 z 22 zwycięzców (95%), a wyniki sondaży online odbiegały od faktycznych

średnio o 6,8% (dla porównania sondaże telefoniczne odbiegały od faktycznych wyników średnio o 6,2%). Możemy, więc wnioskować, iż badania prowadzone w Internecie mogą stanowić odzwierciedlenie opinii i poglądów panujących w danym społeczeństwie, odpowiednio przygotowane i przeprowadzone mogą być reprezentatywne.

Badania Cronka i Westa wykazały również brak różnic pomiędzy wynikami badania realizowanego drogą online, jak i offline. Respondenci mieli do wyboru możliwość wypełnienia ankiety w domu bądź przy pomocy Internetu. Badacze nie stwierdzili istotnych różnic pomiędzy wynikami uzyskanymi w obu grupach [Cronk, West, 2002]. Różnic pomiędzy grupami realizującymi badanie online, jak i offline nie wykazano w badaniach dotyczących związków samokontroli z relacjami romantycznymi, atrakcyjnością fizyczną i seksualną, zaufaniem do siebie, czy zachowaniami i postawami seksualnymi. Wyniki badań przez Internet, jak i przeprowadzanych w laboratorium, są porównywalne [Rowiński, 2007, s. 302-303].

Na postawione w tytule artykułu pytanie należy zatem odpowiedzieć twierdząco. Prowadzenie badań sondażowych przez Internet ma sens, należy jednak podchodzić do nich z taką samą starannością, jak do badań prowadzonych przy pomocy tradycyjnych narzędzi badawczych. Przygotowywanie badań jest procesem złożonym, wymagającym gruntownej wiedzy z zakresu metodologii, jak i metod badawczych. Tylko solidne podstawy metodologiczne gwarantują sukces badawczy. Już w fazie konceptualizacji, a później operacjonalizacji należy odpowiedzieć sobie na zasadnicze pytania badawcze – między innymi – na jakie pytanie chcę uzyskać odpowiedź i jaka metoda, czy technika badawcza będzie optymalna do moich badań.

Przypisy

¹ W marcu 2007 dostęp do Internetu deklarowało 40% badanych – dane według CBOS: *Internet i komputery: wyposażenie gospodarstw domowych, sposoby i cele korzystania*.

² Tę niedogodność próbuje się niwelować stosując tzw. cookies oraz poprzez użycie unikalnych haseł lub dynamicznie generowanego adresu URL dla każdego respondenta. Jednakże obie metody mają niewielką skuteczność, cookie może być bowiem usunięte, a adresy IP są często przydzielane dynamicznie, co powoduje, że ten sam komputer podczas kolejnych sesji internetowych otrzymuje adresy różne.

Bibliografia

Babbie E.: *Badania społeczne w praktyce*. PWN, Warszawa 2004

Cronk B., West J.: *Personality research on the Internet: A comparison with traditional methods*. *Behaviour Research Methods*. „Instruments and Computers” 2002, nr 34 [podaję za:] T. Rowiński: *Internet w badaniach psychologicznych...*

Kraut R. E. i inni: *Psychological Research Online: Opportunities and Challenges*. „American Psychologist” 2004

SMG/KRC Net Track 2Q 2003, <http://www.tezmedium.pl/profile.xml>

- Mącik R.: *Wykorzystanie Internetu w badaniach marketingowych*. UMCS, Lublin 2005
- Rostocki W.: *Kłamstwo obronne w ujęciu M. Ossowskiej a etyczne problemy wywiadu kwestionariuszowego w socjologii*. [W:] *Analizy i próby technik badawczych w socjologii*. Tom IX. Z. Gostkowski (red.). Warszawa 1992
- Rowiński T.: *Internet w badaniach psychologicznych*. [W:] *Oblicza Internetu. Architektura komunikacyjna Sieci*. M. Sokołowski (red.). Państwowa Wyższa Szkoła Zawodowa, Elbląg 2007
- Schonlau M., Asch B., Du C.: *Web surveys as part of a mixed-mode strategy for populations that cannot be contacted by e-mail*. „Social Science Computer Review” 2003 [podaję za:]
- J. Zając, D. Batorski: *Jak skłonić do udziału w badaniach internetowych: Zwiększanie realizacji próby*. „Psychologia Społeczna” (w druku)
- Sobieszek K.: *Problem błędu braku odpowiedzi w badaniach Internetowych – rozważania teoretyczne*. [W:] *Re: Internet – społeczne aspekty medium. Polskie konteksty i interpretacje*. Ł. Jonak, P. Mazurek, M. Olcoń i inni (red.) WAIp, Warszawa 2006
- Szpunar M.: *Konstruowanie narzędzi do badań online na przykładzie serwisu eBadania.pl*. „e-mentor” 2007, nr 4, Warszawa
- Szpunar M.: *Realizowanie badań drogą on-line na przykładzie Systemu Zarządzania Badaniami eBadania.pl*. [W:] A. Szewczyk (red.): *Problemy społeczeństwa informacyjnego*. US, Szczecin 2007
- Szpunar M.: *Badania Internetu vs. badania w Internecie, czyli jak badać nowe medium – podstawowe problemy metodologiczne*. „Studia Medioznawcze” UW, Warszawa 2007, nr 2
- Sztabiński F.: *Ankieta pocztowa i wywiad kwestionariuszowy*. IFiS PAN, Warszawa 1997
- Wenzel M.: *Polacy w Sieci*. CBOS, BS/58/2008
- Zając J., Batorski D.: *Jak skłonić do udziału w badaniach internetowych: Zwiększanie realizacji próby*. „Psychologia Społeczna” 2007, nr 3-4, tom 2, s. 234-248
- Żmijewska-Jędrzejczyk T.: *Badania Internetowe*. [W:] *Nowe metody, nowe podejścia badawcze w naukach społecznych*. P. Sztabiński, F. Sztabiński, Z. Sawiński (red.). IFiS PAN, Warszawa 2004