

TERESA ŻMIJEWSKA-JĘDRZEJCZYK

Badania internetowe

Internet stanowi tkankę naszego życia. Gdyby technologię przetwarzania informacji uznać za dzisiejszy odpowiednik tego, czym była elektryczność w erze przemysłowej, to Internet, z uwagi na jego zdolność do przekazywania siły informacji na wszystkie sfery ludzkiej działalności, można by porównać zarówno do sieci energetycznej, jak i silnika elektrycznego. Ponadto tak jak nowe techniki produkcji i dystrybucji energii elektrycznej uczyniły z fabryk i dużych firm organizacyjne podstawy społeczeństw industrialnego, tak Internet stanowi techniczną bazę dla struktury organizacyjnej epoki informacji – sieci.

(Castells 2001: 11)

Wprowadzenie

Narodziny Internetu, który obecnie dla wielu ludzi stanowi codzienny sposób komunikowania się, związany jest z powstaniem w 1990 roku sieciowego systemu wymiany informacji WWW – World Wide Web, opracowanego przez angielskiego programistę z Europejskiego Centrum Badań Jądrowych (CERN) z siedzibą w Genewie (Castells 2001). W roku 1995 Internet stał się ogólnodostępną siecią (Castells 2001) zarówno dla biznesu, jak i dla zwykłych ludzi. Od tego momentu Internet bardzo się rozpowszechnił jako wolny, nie poddany cenzurze sposób komunikowania się, wymiany refleksji, dzielenia

się opiniami i wiedzą. W niniejszym artykule przedstawię ukształtowane na przestrzeni ostatnich kilku lat rodzaje badań internetowych, ich możliwości i ograniczenia oraz kwestie etyczne związane z ich prowadzeniem.

W rozwoju nowych metod badawczych warsztatu socjologa możemy wyróżnić dwa główne nurty związane z pomiarem i opisem aktywności w przestrzeni wirtualnej. Pierwszy z nich opiera się na technikach obserwacji zjawisk zachodzących w sieci WWW, dotyczących przede wszystkim zachowań użytkowników sieci, drugi zaś opiera się na tradycyjnych metodach badawczych (realizowanych *off-line* i/lub *on-line*) i dotyczy przede wszystkim postaw i opinii.

Badania zachowań użytkowników Internetu

W badaniach zachowań użytkowników Internetu narzędziem gromadzenia danych są systemy elektronicznej rejestracji cech ruchu internetowego. Badania te dzielą się na dwa podtypy: *site/server centric* (od strony serwera) i *user centric* (od strony użytkownika). Obie metody gwarantują znaczną precyzję pomiaru badanych cech, obie jednak mają swoje ograniczenia.

Badania site/server centric

W przypadku projektów *site/server centric* badanie obejmuje pełną populację badanych jednostek (użytkowników Internetu), a analizie poddane są wszystkie połączenia dokonane z witryną WWW. Badanie polega na niezależnym rejestrowaniu kontaktów użytkowników z badanymi serwisami, służy więc przede wszystkim do określenia widowni i sposobu korzystania z badanych witryn WWW.

Badanie *site/server centric* jest badaniem typu *opt-in*: badane są tylko te serwisy, które zgłoszą się do badania. Nie obejmuje zatem wszystkich zachowań użytkowników Internetu, lecz tylko zachowania w obrębie badanych witryn. Przykładem badań typu *site centric* jest Web Audience Measurement¹, gdzie przedmiotem analizy jest zapis z dziennika serwera stron WWW. W typowych badaniach *site/server centric* zbierane są następujące informacje: liczba odsłon, liczba użytkowników oraz użytkowników powracających, średnie długości wizyty, częstość wizyt, strony odsyłające, strony początkowe i końcowe, lokalizacja geograficzna użytkowników. Zbierane dane umożliwiają użycie następujących wskaźników:

¹ Web Audience Measurement (WAM) jest produktem TNS OBOP.

1. Liczbę pobrań stron serwisu w danej jednostce czasu (liczba odsłon – *page views*).
2. Liczbę użytkowników odwiedzających serwis (liczba unikalnych użytkowników – *unique users*). Przy czym wskaźnik ten tylko przybliża rzeczywistą liczbę użytkowników i ze względu na kasowalność *cookie* jest coraz rzadziej stosowany do wyznaczania zasięgu.
3. Liczbę i długość wizyt/sesji – *user sessions*. Sesja jest rozumiana jako seria następujących po sobie odsłon stron serwisu oglądanych przez pojedynczego użytkownika. Granice sesji wyznaczane są przez system na podstawie odstępów czasowych pomiędzy odsłonami. Przeważnie sesją jest seria odsłon oddalonych od siebie czasowo nie więcej niż 30 minut.
4. Średnia liczba sesji użytkownika w danym okresie (częstość wizyt/sesji – *frequency of access/user frequency*).
5. Średnia liczba odsłon w sesji dla danego użytkownika (intensywność wizyt/sesji – *sessions intensity*).

Wskaźniki te są podstawą do wyznaczania zestawień statystycznych obrazujących sposób korzystania z badanych serwisów WWW.

Coraz powszechniejsze stają się badania typu *ad-centric*. Badania te opierają się w całości na metodologii *site/server centric*. *Ad-centric* stosowane są do analizy kampanii reklamowych realizowanych w Internecie. Każda z reklam wyświetlonych użytkownikom Internetu zawiera odpowiednie skrypty umożliwiające pozyskanie danych typowych dla systemu *site/server centric* (odsłony, unikalni użytkownicy). Jednym z podstawowych ograniczeń tego typu badań jest możliwość uogólniania wyników na użytkowników Internetu. W badaniach *site/server centric* badaną populacją nie są, wbrew pozorom, internauci, a wyniki nie dotyczą użytkowników Internetu. Systemy *site/server centric* opierają się na znacznikach (*cookie*) przyporządkowanych przeglądarkom internetowym (np. Internet Explorer). Przyjęcie takiej metodologii pomiaru niesie poważne ograniczenia w ekstrapolacji wyników na użytkowników Internetu. Najważniejsze z nich to:

1. Kasowalność znaczników (*cookies*) – ta sama osoba kasująca *cookie* jest liczona kilka razy przez system jako unikalny użytkownik.
2. Korzystanie przez jedną osobę z wielu komputerów – ta sama osoba korzystając z wielu przeglądarek (np. w domu i pracy) posiada więcej niż jeden znacznik, liczona jest zatem wielokrotnie.
3. Korzystanie kilku osób z jednego komputera – osoby korzystające z jednego komputera mają jeden znacznik, liczone są zatem jako jeden użytkownik.

Innym ograniczeniem przyjętej w systemach *site/server centric* metodologii jest brak możliwości określenia cech demograficznych użytkowników serwisów WWW; jest to badanie wyłącznie zachowań. Pewnym rozwiązaniem

tego problemu było wprowadzenie zastosowanie hybrydowej metodologii łączącej dane *site/server centric* z danymi demograficznymi pochodzącymi z ankiet wypełnianych online (CAWI). Należy jednak pamiętać, że dane demograficzne uzyskane tą drogą dotyczą przeglądarek internetowych, a nie osób. Dlatego też tego typu badania posiadają wszystkie ograniczenia zwykłych systemów *site/server centric*.

Badania user centric

Drugim rodzajem badań zachowań użytkowników są badania typu *user centric*. Są one najbardziej zbliżone do klasycznych badań telemetrycznych: dobierany jest panel osób, których korzystanie z Internetu jest w sposób ciągły monitorowany za pomocą odpowiedniego oprogramowania instalowanego na komputerach panelistów. Podstawowe zalety badań *user centric*, to objęcie badaniem populacji rzeczywistych użytkowników Internetu (a nie przeglądarek jak ma to miejsce w *site/server centric*), dostarczenie obok zmiennych behawioralnych, informacji o zmiennych demograficznych oraz objęcie badaniem całego Internetu, zamiast kilku wybranych witryn.

Do podstawowych ograniczeń tego badania należy zaliczyć: problem z dobraniem dostatecznie licznej i reprezentatywnej grupy panelistów, oraz koszty utrzymania badania. Liczebność panelu determinuje szczegółowość pomiaru – w przypadku Internetu, medium składającego się z wielu tysięcy wydawców, liczebność ta powinna być znacząco wyższa od liczebności klasycznego panelu telemetrycznego. Od prawidłowego (najlepiej losowego) doboru panelistów do panelu zależy jakość pozyskiwanych danych, tj. możliwość ich uogólniania na populację użytkowników Internetu. Badania *user centric* najczęściej są wykorzystywane do badania profilu widowni poszczególnych witryn/stron WWW.

Obserwacja, której narzędziem są skrypty programów komputerowych, nie jest jednak w istocie techniką realizacji badań socjologicznych czy marketingowych. Jest to jedynie zapis, bierna obserwacja użytkowników sieci.

Badanie postaw i opinii użytkowników Internetu

Drugim głównym nurtem związanym z pomiarem i opisem aktywności w przestrzeni wirtualnej są techniki badawcze realizowane w sposób tradycyjny, tzw. off-line'owe (realizowane poza Internetem) oraz badania CAWI – *Computer Assisted Web Interviews* (wywiady internetowe wspomagane komputerowe), tzw. on-line'owe (realizowane przez Internet). Wypełniają one lukę

informacyjną dotyczącą sfery badawczej wykraczającej poza zakres identyfikacji i opisu cech ruchu internetowego.

Badania off-line'owe

Do badań tradycyjnych opisujących Internet i sposób jego użytkowania można zaliczyć badania *ad-hoc* (ilościowe i jakościowe) oraz trackingi (np. *Interbus*²). Wachlarz technik i tematów obejmujących badania *ad-hoc* jest analogiczny, jak dla innych narzędzi komunikacji i nazbyt szeroki, by go w tym miejscu omawiać. Natomiast głównym celem badań trackingowych jest monitorowanie: wskaźników dostępu do Internetu, sposobu, w jaki uzyskiwany jest dostęp, częstotliwości korzystania z Internetu, celów korzystania, znajomości i korzystania z najpopularniejszych stron WWW, ich funkcjonalności oraz ogólna charakterystyka społeczno-demograficzna użytkowników Internetu.

Badania on-line'owe

Znane w Polsce od kilku lat badania realizowane *on-line* budzą wciąż wiele kontrowersji. Najbardziej rozpowszechniona metoda CAWI ma zarówno zdecydowanych zwolenników, upatrujących w niej wolną od jakichkolwiek wpływów technikę surveyową, jak również zdecydowanych przeciwników, wedle których niska penetracja Internetu w Polsce wyklucza zasadność jej stosowania, w szczególności na próbach osób. Wskazują oni również brak możliwości kontroli tego, kto w rzeczywistości odpowiada na pytania ankiety, co może podważyć wartość uzyskiwanych odpowiedzi. Pośrednio jest również podważana zasadność stosowania tej metody na próbach celowych.

Nie sposób jednak zakwestionować zalet CAWI. Po pierwsze, krótki czas realizacji badania – skrócenie, w porównaniu z badaniami tradycyjnymi, fazy terenowej oraz fazy opracowania danych. W badaniach marketingowych, ale również w badaniach opinii publicznej krótki czas realizacji badania i przekazania wniosków z badania może mieć kluczowe znaczenie dla tzw. gorących tematów i wszędzie tam, gdzie wydłużony czas realizacji może negatywnie wpłynąć na homogeniczność zebranego materiału. Po drugie, zastosowanie CAWI eliminuje (lub w niektórych odmianach ogranicza do minimum) udział ankieterów, a więc i możliwość wystąpienia efektu ankieterskiego. Przy zastosowaniu CAWI mamy pewność standaryzacji, tzn. w ten sam sposób wyświetlana jest treść pytań, wszelkie instrukcje, ewentualna rotacja pytań. Nie jest również możliwe przejście do następnego pytania bez udzielenia odpo-

² Produkt TNS OBOP.

wiedzi na pytanie wcześniejsze (oczywiście o ile badacz nie dopuści takiej możliwości). Ograniczeniu błędów przy wpisywaniu odpowiedzi przez respondentów sprzyja kontrola zakresów i zgodności logicznej wpisywanych odpowiedzi oraz w przypadku reguł przejścia – automatycznego przechodzenia do właściwego pytania.

Niewątpliwą zaletą metody CAWI jest również niski koszt realizacji. Główna oszczędność wynika z wyeliminowania kosztów przejazdu, wynagrodzeń ankieterów i koordynatorów oraz kosztów związanych z drukiem materiałów i wprowadzeniem danych.

Ilościowe badania on-line'owe (głównie metoda CAWI) stanowią w istocie adaptację tradycyjnych metod gromadzenia informacji do nowego medium, jakim jest Internet. Najbardziej rozpowszechnione techniki badań ilościowych realizowanych przez Internet to: *e-mailing* i *web surveys*.

E-mailing

E-mailing – to technika najwcześniejsza i w najmniejszym stopniu odzwierciedlająca zalety CAWI. W e-mailu lub do e-maila (w formie załącznika) dołączana jest treść ankiety. Następnie jest on rozsyłany do zdefiniowanej listy potencjalnych respondentów, którzy zapraszani są do udziału w badaniu. Po wyrażeniu zgody i udzieleniu odpowiedzi na pytania ankiety badani proszeni są o odesłanie jej na wskazany adres e-mailowy – zazwyczaj ten sam, z którego przyszło zaproszenie. Badani nie muszą być użytkownikami stron WWW, wystarczy, że korzystają z programu pocztowego. *E-mailing* wykorzystuje Internet jedynie jako kanał dystrybucji ankiety i w największym stopniu przypomina wykorzystywaną od wielu lat technikę ankiety pocztowej. Posiada jej zalety, jakimi są możliwość dotarcia do respondentów trudno dostępnych przy zastosowaniu innych metod oraz swoboda w udzielaniu odpowiedzi, tzn. respondent ma czas na sformułowanie odpowiedzi, zebranie potrzebnych informacji, może ankietę wypełnić w dogodnym dla siebie czasie, ma większe poczucie prywatności i na jego odpowiedzi nie wpływa ankietujący.

Ma również wady cechujące ankietę. Do najważniejszych zaliczyć należy: niepewność, czy ankietą trafiła do adresata, czy została wypełniona samodzielnie, kto właściwie odpowiada na pytania ankiety oraz czy została wypełniona prawidłowo. Wpływ na wartość zebranego materiału może mieć również fakt, że respondent przed rozpoczęciem wypełniania ankiety może ją obejrzeć i ustalić strategię jej wypełniania (Mangione 1995).

Stosowanie *e-mailingu* rodzi problem standaryzacji na poziomie przygotowania, dystrybucji ankiety i wprowadzania danych. Nie mamy także pewności, czy każdy element ankiety zostanie wyświetlony w ten sam sposób.

Może się zdarzyć, że ze względu na różne wersje oprogramowania uczestnik badania nie odczyta właściwe treści pytania – na przykład opisy skali przy pytaniu przesuną się. Nie mamy również możliwości kontrolowania kompletności i poprawności wypełnianej ankiety, a stosowanie wszelkiego rodzaju reguł przejścia może zwiększać ryzyko błędów w wypełnianiu ankiety.

Podsumowując, zastosowanie tej techniki jest godne polecenia, jeśli podejrzewamy, że znaczna część potencjalnych respondentów nie przegląda stron WWW, a jedynie obsługuje program do wymiany poczty.

Web surveys

Wraz z rozwojem technologicznym Internetu oraz upowszechnieniem korzystania ze stron WWW pojawiły się *web surveys* (ankiety edytowane na stronach WWW), które posiadają wcześniej wymienione zalety CAWI. *Web surveys* to wszelkiego rodzaju ankiety umieszczone w Internecie prezentowane respondentom, w zależności od sposobu ich pozyskania do badania, na przykład w formie zaproszenia rozsyłanego e-mailem z linkiem odsyłającym do ankiety czy *pop-up*. Poniżej omówię ten rodzaj badań CAWI pod kątem sposobu doboru respondentów.

Rozpocznę od sposobu najprostszego – od samodoboru. Ankieta internetowa lub zaproszenie (z linkiem) umieszczone jest na stronie WWW, w miejscu dostępnym dla każdego odwiedzającego. Przy zastosowaniu tej metody nie przewiduje się żadnych kryteriów doboru w stosunku do badanych. *Response rate* dla tej techniki jest łatwy do oszacowania, jeśli tylko dostępne są informacje o oglądalności (wielkości widowni) danej strony. Jednak wraz z upowszechnieniem Internetu technika ta ma coraz niższy *response rate* i obecnie szacuje się, że wynosi niespełna 1%. Techniki samodoboru są najłatwiejszym, ale jednocześnie obciążonym największym błędem na poziomie konstruowania próby sposobem pozyskiwania respondentów.

Wraz z popularyzacją umieszczania reklam na stronach WWW pojawiły się techniki rekrutowania *Site Intercept*³ – nazwę tę zaproponował MacElroy. Technika ta polega na zachęcaniu potencjalnych respondentów do wypełnienia ankiety poprzez stosowanie różnego rodzaju reklam bannerowych, tekstów reklamowych z linkem do ankiety lub innych elementów umieszczonych na stronie. Najbardziej rozpowszechnione są ankiety z użyciem bannerów (*banner surveys*). Idea tego typu badań polega na umieszczeniu reklamy bannero-

³ Nowe terminy angielskie nie mają swoich polskich odpowiedników. Zanim je znajda (jeżeli kiedykolwiek) i wejdą do słownika języka polskiego, muszą się zakorzenić, co jest naturalnym sposobem tworzenia terminów. Powszechność stosowania terminów anglojęzycznych odnoszących się do Internetu wynika również ze specyfiki tego medium – język angielski jest językiem Internetu.

wej na wybranej stronie (lub stronach) WWW, która zachęca odwiedzających tę stronę (strony) do kliknięcia na tę reklamę, a po kliknięciu uaktywnia się ankieta, tzn. wyświetlane jest nowe okno dialogowe lub podstrona z informacjami na temat badania, a następnie treść pytań. Dla tej techniki *response rate* jest relatywnie niski, rzędu kilku procent. Jedynie dla stron o wysokiej oglądalności nawet 2–3% może okazać się wystarczające. Ze względu na niski odsetek odpowiedzi oraz całkowity brak kontroli doboru respondenta ten sposób stosuje się coraz rzadziej.

Rozwinięciem poprzedniej pasywnej techniki rekrutowania jest *Interstitial Window* – termin również zaproponowany przez MacElroya. Od wcześniej omawianej techniki różni ją to, że uwagę potencjalnych uczestników badania, po wejściu na określoną stronę WWW, przykuwa *pop-up* lub automatycznie otwierająca się podstrona z zaproszeniem do udziału w badaniu i ankietą. Nie ma więc możliwości, by *pop-up*, czy podstrona nie zostały zauważone. Na etapie przygotowania tego typu ankiet najwięcej uwagi poświęca się temu, by zwrócić uwagę internautów i w krótkim komunikacie przekonać potencjalnych uczestników badania o ważności podejmowanego przez nich wysiłku. Pierwszy kontakt z *pop-up* lub wyświetlona podstrona ma zatem decydujący wpływ na zgodę na udział w badaniu. Przy użyciu tej techniki mogą być kontrolowane niektóre elementy doboru respondenta do badania. Zachęta do wypełnienia ankiety może być wyświetlana losowo wybranym internautom odwiedzającym daną stronę (co „n-temu” wyświetlany jest *pop-up*) na różnych etapach jego wędrówki. W zależności od wielkości widowni określany jest odsetek odwiedzających, którzy zostaną losowo wybrani do badania. Użytkownikom może zostać wyświetlony *pop-up* z zaproszeniem, gdy wchodzi na konkretny portal czy serwis, przy wychodzeniu z konkretnej podstrony, przy wychodzeniu z portalu czy serwisu, gdy obejrzą wskazaną zawartość (element serwisu, reklamę) na podstronie czy stronie.

Przed określeniem wielkości próby analizowana jest oglądalność danej strony. Wybierając miejsce na stronie, które zostanie użyte do wyświetlania zaproszenia, również analizowana może być jego widownia i popularność w porównaniu z innymi miejscami na stronie. Także dzień tygodnia i pora dnia są kontrolowane przy wyświetlaniu zaproszeń. Profil odwiedzających dane miejsce rano lub wieczorem oraz w dzień powszedni lub w weekend może się znacznie różnić. Wymienione powyżej czynniki mogą w znacznym stopniu kształtować strukturę próby wypełniających ankietę oraz wpłynąć na liczbę zebranych ankiet. Określając liczbę ankiet, jaka ma zostać zebrana, można się posłużyć prostym zabiegiem uwzględniającym dane o widowni i przewidywanych odsetkach zwrotu: liczba zebranych ankiet równa się liczbie odwiedzających stronę pomnożonej przez planowany odsetek, jaki obejmujemy badaniem, następnie przez prognozowany odsetek, jaki rozpocznie wypełnianie

ankietę oraz przez prognozowany odsetek, jaki zakończy wypełnianie ankiety. Uwzględnienie różnych odsetków dla rozpoczętych i ukończonych ankiet ma znaczenie przede wszystkim wtedy, gdy badani mają do wypełnienia długą, kilkunastominutową ankietę. Przy zastosowaniu tej bardziej agresywnej techniki rekrutowania *response-rate* może się zwiększyć nawet dwukrotnie w porównaniu z wcześniej opisanymi. Niestety, zdaniem internautów jest to bardzo irytujący sposób zachęcania ich do udziału w badaniu.

Kolejnym, znacznie bardziej czasochłonnym i wymagającym zaangażowania sposobem pozyskiwania respondentów do badań CAWI, jest prekrutacja *on-line*. Wykorzystuje ona najczęściej powyżej opisane techniki rekrutacji: *pop-up surveys*, *banner surveys*. Respondenci, którzy wyrażą zgodę na udział w badaniach, proszeni są na początku o wypełnienie ankiety rekrutacyjnej, która w głównej mierze składa się z pytań metryczkowych (np. płeć, wiek, wykształcenie, miejsce zamieszkania, stan cywilny). Po zrekrutowaniu założonej liczby potencjalnych uczestników badania, rozsyłane są zaproszenia do udziału w badaniu zasadniczym. W zależności od charakteru badania i jego specyfiki, do badania rekrutowanych jest 2-, 3- lub nawet 4-krotnie więcej respondentów, niż zakładamy objąć badaniem. Określając wielkość „zapasu” brane są pod uwagę m.in.: dostępność respondentów w planowanym terminie badania zasadniczego, tematyka, poziom skomplikowania ankiety, jej długość.

Ważnym aspektem przy zastosowaniu tej techniki jest utrzymanie kontaktu z już zrekrutowanymi do badania respondentami. Brak takiego kontaktu lub zbyt duże odstępstwa między kolejnymi badaniami mogą spowodować utratę części potencjalnych uczestników badania.

Niepodważalną zaletą takiej metody rekrutacji jest kontrola struktury respondentów biorących udział w badaniu. Prekrutacja do badań CAWI może się odbywać również poza siecią, przy wykorzystaniu tradycyjnych metod. Doświadczenia wskazują jednak, że trudniej jest zachęcić potencjalnych respondentów do udziału w badaniu internetowym tradycyjnymi metodami. Prekrutacja *off-line* jest najbardziej czasochłonną i kosztochłonną metodą pozyskiwania respondentów, jednak w niektórych projektach warto ponieść wysokie koszty rekrutacji, aby zapewnić kontrolę parametrów doboru uczestników do badania.

Rozwinięciem technik prekrutacyjnych, zarówno *on-line*, jak i *off-line* jest budowanie paneli o charakterze syndykatowym, tzn. panel jest utrzymywany i wykorzystywany przez wielu zlecniodawców. Realizacja prób dobieranych z takich paneli charakteryzuje się relatywnie wysokimi wskaźnikami *response rate*. Rzadziej stosowanym i bardzo kosztownym rozwiązaniem jest budowanie paneli na potrzeby jednego klienta.

Na zakończenie wspomnieć należy o istniejących bazach danych e-maili potencjalnych respondentów. Jest to jedna z bezpiecznych metod realizacji badań CAWI, pod warunkiem że posiadamy aktualną, sprawdzoną i kompletną bazę danych osób, które wyraziły chęć udziału w badaniach i przekazały swój adres mailowy. Przy zastosowaniu tej metody możliwe jest również osiągnięcie relatywnie wysokich wskaźników *response rate*.

Jakościowe badania *on-line* są również w dużej mierze próbą aplikacji tradycyjnych badań jakościowych do środowiska Internetu. Najpopularniejszą techniką jakościowych badań realizowanych przez Internet jest FGI (*Focus Group Interview* – sesje dyskusyjne w gronie od 6 do 10, a czasem nawet 12 osób dobranych w sposób celowy. Sesja jest nadzorowana i kierowana przez specjalnie wyszkolonego moderatora, który kieruje nią zgodnie z ustalonym scenariuszem. W trakcie sesji możliwe jest stosowanie różnych technik badawczych). Jego odmiana internetowa nosi nazwę *On-line Focus Group Interview*. Rekrutacja do tego typu badań przeprowadzana jest w sposób analogiczny, jak w przypadku wyżej opisanych technik ilościowych. Uczestnicy najczęściej dobierani są *on-line* w trakcie „łapanek” na stronach WWW (*pop-up surveys, banners*), w czasie chatowania lub z paneli syndykatowych.

Po zrekrutowaniu respondentom rozsyłany jest URL (adres strony WWW) odpowiednio zabezpieczony przed innymi użytkownikami (zazwyczaj jest to system logowania). Respondenci otrzymują instrukcje, w jaki sposób mają się zalogować i rozpocząć udział w badaniu. Po zalogowaniu respondenci witani są informacją wyjaśniającą zasady uczestnictwa w badaniu. Po ich zaakceptowaniu każdy z respondentów proszony jest o wpisanie imienia, nazwiska i danych adresowych umożliwiających przesłanie gratyfikacji po zakończeniu dyskusji. Następnym krokiem jest wybranie przez każdego uczestnika pseudonimu (*nick*), którym się będzie posługiwał w czasie badania. Po wejściu zaproszonych do grupy, prowadzący badanie otrzymuje informację, czy wszyscy są zalogowani. Moderator prosi każdego o przedstawienie się pozostałym uczestnikom dyskusji, gdy wszyscy zaproszeni zakończą proces logowania i rozpoczną dyskusję. Również informacje o wylogowaniu któregoś z uczestników grupy są natychmiast sygnalizowane moderatorowi. Cała dyskusja jest widziana przez każdego z uczestników, ale dodatkowo prowadzący może kontaktować się z wybranym respondentem bez wiedzy innych uczestników.

Ekran widziany przez respondenta zazwyczaj podzielony jest na trzy sekcje. W górnej części umieszczone są dane identyfikacyjne oraz linki do stron WWW, które mają zostać wykorzystane w trakcie badania do przedstawiania materiałów poglądowych. Dolna część ekranu przeznaczona jest do edycji, gdzie respondenci wpisują swoje odpowiedzi i komentarze. Centralną część zajmuje historia dyskusji. Moderator widzi tę samą zawartość ekranu, a do-

datkowo widzi okienko, w którym może otrzymywać wiadomości od autora, czy zleceniodawcy ze wskazówkami. Możliwy jest zatem jednoczesny kontakt z respondentami i z autorem badania. Autor ma również możliwość wpływania na przebieg dyskusji wnosząc poprawki, uzupełniając o dodatkowe pytania wyjaśniające dotychczasowe odpowiedzi badanych. Współpraca między moderatorem a autorem badania w czasie trwania badania może prowadzić do daleko idących zmian w scenariuszu dyskusji. Jest to ogromna zaleta w porównaniu z badaniami tego typu prowadzonymi metodami tradycyjnymi, ale i wyzwanie dla moderatora.

Oprócz okien dialogowych moderator zazwyczaj w prawym górnym rogu ma otwarty zwykły edytor tekstu ze scenariuszem badawczym. Możliwość szybkiego skopiowania z edytora poszczególnych pytań czy instrukcji znacznie usprawnia przebieg dyskusji. Jednak rola scenariusza badawczego jest ograniczona. Przy badaniach jakościowych prowadzonych *on-line* niezbędne jest dostosowanie przez moderatora pytań do języka uczestników, tak by dyskusja przybrała możliwie naturalny charakter.

Grupy fokusowe trwają około godzinę, co umożliwia zadanie około 40 pytań. Kilkanaście minut po ich zakończeniu transkrypcje z ich przebiegu są gotowe do analiz i mogą być przekazywane autorowi. Ze względu na zastosowaną technikę badawczą w krótkim czasie przygotowywane są również transkrypcje w podziale na poszczególnych respondentów. Grupy fokusowe przypominają działanie zwykłych chatów, a udział w nich nie wymaga szczególnego oprogramowania. Powszechnie używane przeglądarki umożliwiają udział w badaniu.

Podsumowując, niewątpliwymi zaletami grup dyskusyjnych prowadzonych *on-line* są:

1. Niezależność geograficzna – możemy zaprosić do udziału w badaniu osoby zamieszkujące odległe miejsca. Jest to szczególnie cenne przy realizacji badań z trudno dostępnymi kategoriami respondentów.
2. Możliwość wypowiedzenia się przez każdego z uczestników – możemy uniknąć w ten sposób problemu „dominującego uczestnika” grupy i dać szansę wypowiedzenia się każdemu. Możemy również, bez wiedzy innych uczestników, dopytać o szczegóły wybranych respondentów.
3. Anonimowość – każdy uczestnik badania jest w pełni anonimowy, co sprzyja udzielaniu szczerych, niczym nieskrępowanych odpowiedzi. Jest to bardzo istotne, zwłaszcza gdy prowadzimy badania na tematy drażliwe.
4. Oszczędność kosztów i czasu na etapie realizacji i przygotowania transkrypcji – nie ma konieczności wynajmowania sal, zapewniania obsługi oraz opłacania transkrybentów. Jest jednak pewne zastrzeżenie: badacz musi być wyposażony w niezawodny sprzęt komputerowy.

5. Autor (zleceniodawca) badania może aktywnie uczestniczyć w badaniu. Może towarzyszyć moderatorowi osobiście w trakcie trwania sesji albo obserwować i komentować ją *on-line* z dowolnego miejsca, w którym może się połączyć z Internetem.
6. Testowanie całych stron albo reklam. Dzięki możliwości umieszczenia linków do stron WWW każdy uczestnik grup fokusowych może na swoim ekranie obejrzeć przygotowaną dla niego stronę czy reklamę i natychmiast opisać własne wrażenia.
Z kolei do wad badań jakościowych realizowanych *on-line* zaliczyć należy:
 7. Brak możliwości obserwacji zachowań niewerbalnych. Jest to najpoważniejszy zarzut wobec badań jakościowych prowadzonych *on-line*. W tradycyjnych badaniach przyjęło się, że obserwacja niewerbalna stanowi istotny element na poziomie analizy i przygotowywania rekomendacji.
 8. Problem z utrzymaniem dynamiki grupy i stymulowaniem interakcji między uczestnikami. Kierowanie przez moderatora dynamiką grupy może być utrudnione przez różne czynniki, przede wszystkim sam fakt jedynie werbalnego sposobu komunikowania się. Napisanie własnej opinii zabiera czas i wymaga lepszego uporządkowania niż spontaniczne wygłoszenie opinii, która jest mniej „zobowiązująca”. Potrzebny jest również czas na odczytanie wiadomości od innych uczestników. Uwagę uczestników grupy może skupiać w większym stopniu konstruowanie odpowiedzi na postawione przez moderatora pytanie i tym samym ograniczać ich skłonność do zabierania głosu w dyskusji. Interakcje między uczestnikami ulegają osłabieniu i w większym stopniu koncentrują się na relacji: uczestnik–moderator. Obniżenie poziomu interakcji między uczestnikami grupy może stanowić duże ograniczenie dla FGI prowadzonych *on-line*. Podstawą tradycyjnego FGI jest założenie, że opinie tworzą się lub przynajmniej są wzmacniane w wyniku interakcji między członkami grupy (Morgan 1996) – jest to pochodna założenia o społecznej genezie opinii. Powstaje zatem wątpliwość, czy podczas FGI realizowanych *on-line* zebrane opinie są wynikiem interakcji grupowej czy też jedynie sumą indywidualnych wypowiedzi. Otwartą kwestią pozostaje zatem, czy FGI realizowany *on-line* zachowuje istotę tradycyjnego FGI.
 9. Utrzymanie pełnej koncentracji zarówno przez moderatora, jak i uczestników. Łącząc się z różnych miejsc uczestnicy mogą być rozpraszeni przez inne osoby. Tymczasem dla pełnego udziału konieczne jest ciągłe śledzenie korespondencji i aktywny w niej udział każdego uczestnika.
10. Sprawy techniczne – warunkiem powodzenia badania jest sprawność w posługiwaniu się komputerem każdej ze stron uczestniczącej w bada-

niu. Uczestnicy oprócz sprawnego pisania muszą posiadać umiejętność poruszania się po Internecie. Z kolei moderator może być postawiony w sytuacji, w której musi sprawnie prowadzić kilka dyskusji jednocześnie, np. prowadzi dyskusję z uczestnikami i odpowiada na pytania autora. W takich przypadkach często podejmowana jest decyzja o prowadzeniu grup fokusowych przez dwóch moderatorów. Niestety takie rozwiązanie generuje dodatkowe koszty.

Oprócz najpopularniejszych badań fokusowych czasami realizowane są *Connected groups*. Są to tradycyjnie realizowane FGI z tą różnicą, że każdy uczestnik jest podłączony do Internetu, co umożliwi przeglądanie stron WWW w czasie trwania sesji. Przy użyciu tej techniki najczęściej testowane są nowe strony, serwisy już istniejących stron pod kątem ich funkcjonalności i atrakcyjności. Uczestnicy mogą również śledzić to, co dzieje się na ekranie ich komputerów za pośrednictwem projektora. Moderator może w ten sposób stymulować uczestników do wspólnych porównań, ocen, dyskusji.

Równie rzadko stosowane są *E-mail groups*. Technika ta polega na prowadzeniu dyskusji z określoną grupą uczestników na zadany temat w wyznaczonym terminie. Narzędziem komunikacji uczestników badania jest poczta elektroniczna. Za pośrednictwem Internetu realizowane są również indywidualne wywiady pogłębione – *E-mail In-Depth Interviews*. Moderator prowadzi dyskusję korespondencyjnie – z wykorzystaniem poczty elektronicznej – tylko z jednym respondentem. Obie metody – *E-mail groups* i *E-mail In-Depth Interviews* stosuje się, gdy respondentami są osoby trudno dostępne, które nie mogą uczestniczyć w innych formach badania *on-line* lub nie przeglądają stron WWW, nie korzystają z przeglądarek stron WWW.

Problemy deontologiczne badań internetowych

Z racji dynamicznego rozprzestrzeniania się badań internetowych, możliwości ich realizacji praktycznie przez każdego, zrodziła się potrzeba ustanowienia pewnego minimum reguł przeprowadzania badań internetowych. W założeniu zestaw takich reguł gwarantowałby realizację tego rodzaju badań na właściwym poziomie z zachowaniem prawideł metodologicznych oraz z poszanowaniem praw użytkowników Internetu – z jednej strony – i odbiorców wyników – z drugiej. Jedną z powszechnie uznanych organizacji, która postawiła sobie taki cel, jest ESOMAR (Stowarzyszenie Badania Rynku i Opinii, które ustanawia i promuje standardy zawodowe i etyczne osób działających w branży badawczej. Zrzesza zarówno badaczy rynku i opinii, naukowców, jak również użytkowników badań). Najogólniejszy postulat dotyczący

realizacji badań internetowych mówi o konieczności przestrzegania Kodeksu ESOMAR przygotowanego dla badań realizowanych metodami standardowymi (*off-line*) oraz respektowania postanowień innych branżowych instytucji. W badaniach internetowych szczególnej troski wymagają dane osobowe, które powinny być chronione zgodnie z prawodawstwem obowiązującym w danym kraju. W przedstawionym przez ESOMAR przewodniku prowadzenia badań marketingowych oraz opinii przy użyciu Internetu (*Conducting Marketing and Opinion Research Using the Internet. ESOMAR Guideline*) szczególną uwagę zwrócono na ochronę praw użytkowników Internetu oraz odbiorców wyników.

Ochrona praw użytkowników Internetu w procesie badawczym polega przede wszystkim na:

- zapewnieniu dobrowolności udziału w badaniach. Badacz powinien dołożyć wszelkich starań, by nie naruszać prywatności użytkowników Internetu. Respondent przed wyrażeniem zgody na udział w badaniu powinien zostać poinformowany o celu badania, jego charakterze (przebiegu) i sposobie opracowywania czy raportowania zebranych danych. Zgoda na gromadzenie danych osobowych musi być potwierdzona przez każdego respondenta. Jeżeli respondent został wylosowany z operatu, powinna zostać mu przedstawiona informacja na temat jego źródła i charakteru. Natomiast jeśli badacz chciałby ponownie zaprosić respondenta do udziału w badaniu, musi uzyskać jego zgodę na powtórny kontakt;
- zapewnieniu uczestnikowi badania anonimowości. Informacje na ten temat powinny być jasno przedstawione w łatwo dostępnym dla uczestnika miejscu. Jeżeli respondent wyrazi zgodę na identyfikację, należy zagwarantować mu, że zebrane informacje posłużą jedynie do realizacji celów badawczych i nie zostaną przekazane osobom trzecim;
- zagwarantowaniu uczestnikowi badania prawa do wycofania się z niego na każdym etapie realizacji. Konsekwencją takiej decyzji musi być wykluczenie danych respondenta ze zbioru;
- zapewnieniu uczestnikowi możliwości kontaktu z badaczem w przypadku jakichkolwiek wątpliwości. Kanały kontaktu należy dostosować do charakteru badania i możliwości badanych (np. adres e-mailowy, bezpłatna linia telefoniczna). Powszechnym rozwiązaniem jest przygotowanie dokumentu z odpowiedziami na najczęściej zadawane przez uczestników pytania (FAQ – *Frequently Asking Question*);
- uzyskaniu zgody rodziców/opiekunów w przypadku realizacji badania z udziałem dzieci lub niepełnoletniej młodzieży. Konieczne jest również uwzględnienie wszelkich innych ograniczeń prawnych wynikających z zaangażowania niepełnoletnich w badanie;

- absolutnym zakazie rozsyłania ankiet czy nawet zaproszeń do udziału w badaniu użytkownikom Internetu, o których wiemy, że odmówili udziału w badaniu.

Ochrona praw odbiorców wyników badań sprowadza się przede wszystkim do obowiązku:

- przedstawienia metodologii doboru respondenta do badania;
- opisanie charakterystyki próby do badania, jego metodologii, odsetka uzyskanych odpowiedzi (*response rate*) oraz sposobu jego wyliczania;
- poinformowania o ograniczeniach w ekstrapolacji wyników badania będących skutkiem wybranej metodologii lub innych czynników wynikających z realizacji badania.

Jak widać, jest to rzeczywiście krótki zestaw zasad, który większość badaczy realizujących badania internetowe stosuje mniej lub bardziej świadomie.

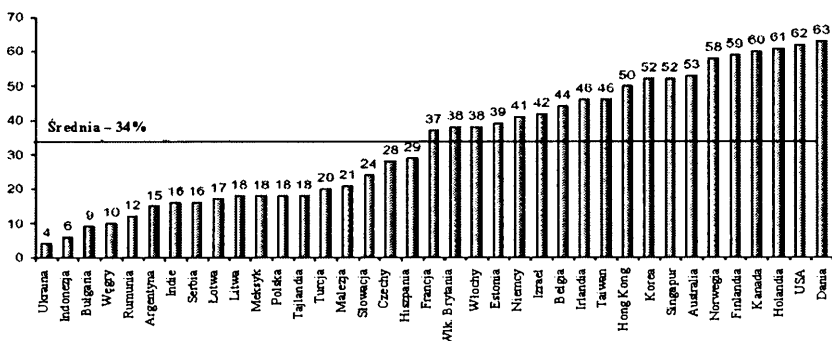
Możliwości i ograniczenia w stosowaniu badań CAWI

Realizując badania na próbach firm (B2B – *Business to Business*) ograniczenia w dostępie do Internetu zmniejszają się bardzo dynamicznie z roku na rok. Z tego powodu badania realizowane *on-line* mogą być w wielu wypadkach bardzo wartościową alternatywą dla badań realizowanych metodami tradycyjnymi. Dotyczy to w szczególności badań realizowanych na próbach pracowników trudno dostępnych, do których zaliczamy kadrę zarządzającą i wysokiej klasy specjalistów. Należy jednak mieć na uwadze, że odsetek firm posiadających dostęp do Internetu jest zróżnicowany i zależy przede wszystkim od wielkości firmy i charakteru działalności. Przykładowo, firmy produkcyjne, z branż przetwórstwa spożywczego czy usług komunalnych znacznie rzadziej wykorzystują w swojej działalności Internet niż np. firmy z branży doradztwa prawnego.

Zupełnie inaczej sprawa wygląda w przypadku badań realizowanych na próbach osób czy gospodarstw domowych. W niemal każdym kraju występuje ograniczenie w stosowaniu metod CAWI ze względu na możliwość dostępu do Internetu oraz możliwość korzystania z niego. Można się spodziewać, że na przestrzeni najbliższych kilku lat sytuacja nie ulegnie zmianie i będą istniały pewne grupy społeczne, dla których wskaźnik dostępu do Internetu będzie znacznie niższy niż dla innych.

Poniżej przedstawię wybrane dane z międzynarodowego badania kwestionariuszowego, przeprowadzonego w 37 krajach, na reprezentatywnych próbach ogólnokrajowych. Ogółem zrealizowano ponad 42 tys. wywiadów na przestrzeni marca i kwietnia 2002 roku (zob. wykres 1).

Wykres 1. Odsetek korzystających z Internetu na przestrzeni ostatniego miesiąca



Źródło: Global eCommerce Report 2002 (GeR, TNS 2002).

Na podstawie wyników tych badań, we wszystkich krajach objętych badaniem, około 34% populacji korzystało z Internetu na przestrzeni ostatniego miesiąca. W czołówce krajów o najwyższych wskaźnikach znalazły się: Dania, USA, Holandia, Kanada, Finlandia i Norwegia. W tych krajach odsetek korzystania z Internetu w ostatnim miesiącu wyniósł około 60%. Polska – ze wskaźnikiem 18% – znalazła się w grupie krajów o niskim poziomie penetracji Internetu, do której należą również: Litwa, Łotwa, Meksyk.

Choć z roku na rok obserwujemy systematyczny wzrost liczby osób, które korzystają z Internetu w Polsce, to jednak w porównaniu ze średnią dla wszystkich badanych krajów oraz w relacji do krajów zachodnioeuropejskich stale jest to dość niski wynik. Oznacza on m.in., że aż cztery osoby na pięć nie ma szans na znalezienie się w próbie ogólnopolskiej, co w oczywisty sposób wyklucza w najbliższych latach realizację przez Internet badań na ogólnopolskich reprezentatywnych próbach. Również wniosków z badań prowadzonych na internautach nie można uogólniać na populację mieszkańców Polski.

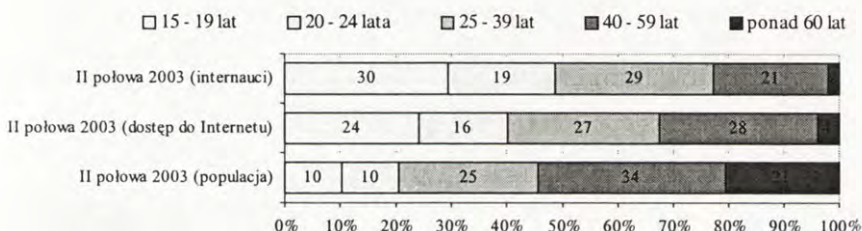
Przy analizowaniu zachowań użytkowników Internetu oraz przy konstruowaniu prób najczęściej wykorzystywana jest definicja internauty przyjęta za Internet Advertising Bureau. Za internautę uważa się osobę korzystającą z Internetu raz w miesiącu lub częściej. Biorąc pod uwagę internautów, mamy pewność, że wybrana do badania osoba ma nie tylko możliwość dostępu do Internetu, ale również korzysta z niego. Z badań wynika, że nadal niewielki odsetek badanych deklarujących możliwość dostępu do Internetu, nie korzysta z niego.

Powszechnie wiadomo, że zmienne społeczno-demograficzne w znacznym stopniu różnicują dostęp i korzystanie z Internetu. Dla zobrazowania wielko-

ści tych różnic przedstawię porównanie struktury osób posiadających dostęp do Internetu, korzystających z niego i odnośę do struktury populacji. Poniżej przedstawione dane pochodzą z wyników ogólnopolskich badań kwestionariuszowych, realizowanych cyklicznie od 2000 roku, metodą wywiadu bezpośredniego, wśród Polaków wieku 15 lat i więcej. Badanie realizowane jest na reprezentatywnej próbie gospodarstw domowych.

Płeć respondentów w niewielkim stopniu różnicuje posiadanie dostępu do Internetu. Nieco tylko częściej dostęp do Internetu posiadają mężczyźni. Różnica ta jest w większym stopniu widoczna w strukturze internautów. W porównaniu z rozkładem w populacji, odsetek mężczyzn będących internautami jest wyższy o 8 punktów procentowych. Natomiast wiek i wykształcenie respondentów silnie różnicują dostęp do Internetu. Dostęp najczęściej deklarują osoby młode oraz najlepiej wykształcone. Różnice w strukturze ulegają zastrzeniu wśród internautów, co ilustruje wykres 2.

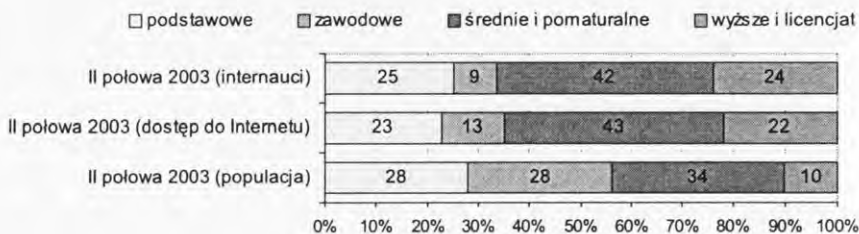
Wykres 2. Struktura wieku w populacji oraz wśród posiadających dostęp do Internetu i wśród internautów (II połowa 2003 roku)



Źródło: TNS OBOP, Interbus.

W dwóch kategoriach wykształcenia – średnim i pomaturalnym oraz wyższym – najostrzej widać różnice w strukturze dostępu do Internetu oraz w strukturze internautów w porównaniu z populacją. Odsetek osób z wykształceniem średnim i pomaturalnym posiadający dostęp do Internetu lub korzystających z Internetu przynajmniej raz w miesiącu jest wyższy od tego w populacji o blisko 10 punktów procentowych. Ponaddwukrotnie wyższy odsetek posiadających dostęp do Internetu i będących internautami występuje w grupie osób z wykształceniem wyższym – włączając licencjat (zob. wykres 3).

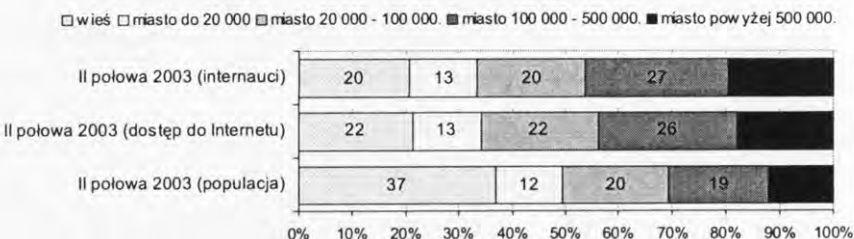
Wykres 3. Struktura wykształcenia w populacji oraz wśród posiadających dostęp do Internetu i wśród internautów (II połowa 2003 roku)



Źródło: TNS OBOP, Interbus.

Kategoria miejsca zamieszkania silnie różnicuje posiadanie dostępu do Internetu oraz korzystanie z niego (zob. wykres 4). Osoby zamieszkujące miasta o wielkości 100–500 000 mają najwyższy udział w strukturze osób posiadających dostęp do Internetu (26%) i w grupie internautów (27%). W populacji osoby zamieszkujące miasta o tej wielkości stanowią 19%. Jak można się spodziewać, na wsi ta relacja jest odwrotna. W populacji osoby zamieszkujące wsie stanowią 37%, w grupie internautów zaś zaledwie 20%.

Wykres 4. Struktura miejsca zamieszkania w populacji oraz wśród posiadających dostęp do Internetu i wśród internautów (II połowa 2003 roku)



Źródło: TNS OBOP, Interbus.

Podsumowując, różnice w strukturze populacji i osób posiadających dostęp do Internetu nie uprawniają do prowadzenia badań reprezentatywnych dla populacji mieszkańców Polski. Jednak osoby posiadające dostęp do Internetu są atrakcyjną grupą docelową na rynku badań konsumenckich. Również dynamika wzrostu dostępu do Internetu skłania do baczego śledzenia rozwoju badań internetowych i samego Internetu. Ważnym elementem przy realizacji badań *on-line* jest również charakterystyka korzystania z Internetu. Powszechnie wiadomo, że jest wielu internautów korzystających ze wszelkich pojawiających się nowinek i grupa osób wykorzystująca Internet jedynie w zakresie

niezbędnym do sprawnego komunikowania się. Zatem projektując badanie należy wziąć pod uwagę poziom zaawansowania potencjalnych uczestników badania.

Podsumowanie. Perspektywy rozwoju badań internetowych

Obszar zastosowań technik CAWI w badaniach socjologicznych determinowany jest stopniem rozwoju Internetu, ponieważ pomiarem objąć można wyłącznie próby internautów lub też podpróby z tej populacji. Biorąc pod uwagę niski stopień penetracji sieci w społeczeństwie polskim, stwierdzić należy, iż techniki CAWI są skutecznym narzędziem analitycznym w przypadku wąskich, specjalistycznych grup respondentów lub segmentów rynku charakteryzujących się wysokim współczynnikiem korzystania z Internetu. Niekwestionowaną zaletą technik CAWI jest możliwość dotarcia do grup trudno osiągalnych tradycyjnymi metodami badawczymi. W przypadku badań B2B zaliczamy do nich kadrę zarządzającą, wysoko wykwalifikowanych specjalistów, wyższych urzędników państwowych, zaś w badaniach na próbach osób są to przede wszystkim młodzi, uczący się na wyższych uczelniach lub już z wyższym wykształceniem, mieszkańcy dużych miast.

Na zakończenie dodam, że na przestrzeni ostatnich kilku lat w badaniach internetowych przeszliśmy od huraoptyzmu dotyczącego nieograniczonych możliwości, jakie otwiera nowy kanał komunikacyjny – Internet – do wyważonej konkluzji, że badania internetowe mają przed sobą przyszłość i ogromny potencjał, ale na dzień dzisiejszy wiele ograniczeń. W żadnym razie nie wyprą tradycyjnych metod komunikowania się z respondentami. Należy się spodziewać, że badania internetowe rozwiną się jako metoda komplementarna w stosunku do tradycyjnych.