

## Laboratorium Badawcze Cyfrowych Nauk Społecznych

Rozwój technologiczny przesunął znaczną część ludzkiego życia w kierunku komputera lub/i ekranów smartfonów, zmieniając w zauważalny sposób niektóre z ludzkich zachowań i indukując nowe. Sposób, w jaki ludzie, komunikują się, współdziałają i poszukują informacji zmienił się radykalnie w tej cyfrowej transformacji. W roku 2019 około 4,2 miliarda ludzi korzystało z Internetu na co dzień. To 54% ludności świata i 87% ludności świata rozwiniętego. Intensywność korzystania z technologii cyfrowych szybko rośnie. Szacuje się, że w 1992 r. łączny ruch internetowy wynosił około 100 gigabajtów dziennie. Liczba ta wzrosła w 2017 r. o 17,5 miliona. Około 90 proc. danych w 2017 r. zostało wygenerowanych w ciągu ostatnich dwóch lat. W coraz większej liczbie dziedzin życia ludzie polegają na technologii cyfrowej. W roku 2017 z dysku Google Drive korzystało 800 milionów aktywnych miesięcznie użytkowników (czyli ponad 15% ludności świata), a każdego dnia przesyłano trzy miliardy obiektów i 1,2 miliarda zdjęć. W 2020 roku Google przetwarza średnio ponad 40 000 zapytań co sekundę, co przekłada się na ponad 3,5 miliarda wyszukiwań dziennie i 1,2 biliona wyszukiwań rocznie. Przeciętny użytkownik Internetu na całym świecie zadaje swojemu urządzeniu różne rodzaje pytań mniej więcej 2 razy dziennie. W USA przeciętny dorosły użytkownik sprawdza swój telefon 30 razy dziennie, podczas gdy w przypadku nastolatków liczba ta sięga 157 razy dziennie. W Wielkiej Brytanii osoby dorosłe spędzają ponad połowę czasu pracy angażując się w działalność medialną lub komunikacyjną, a w ciągu ostatniej dekady czas ten podwoił się. Ponad 99% respondentów w wieku od 16 do 54 lat posiada telefon komórkowy, a ponad 80% respondentów w wieku od 16 do 44 lat ma profil w mediach społecznościowych. Zachowania online, w tym interakcje społeczne, stały się regularną częścią codziennego życia miliardów ludzi na całym świecie.

Większość z tych zachowań jest rejestrowana i przechowywana jako drobnoziarniste dane opatrzone znacznikami czasu na poziomie indywidualnym (tzw. ślady cyfrowe). Są one rejestrowane u milionów lub miliardów użytkowników. Taki rząd wielkości danych pojawia się po raz pierwszy dla nauk społecznych i może być porównywany tylko z takimi przedsięwzięciami jak akceleratory cząstek czy badania genomiczne w naukach przyrodniczych. Dla niektórych badaczy nowo opracowane narzędzia do obserwacji aktywności w sieci, które obiecują przynieść podobne efekty transformacyjne dla nauk społecznych jak mikroskop elektronowy, teleskop, akcelerator cząstek i obrazowanie rezonansem magnetycznym w naukach przyrodniczych, są dowodem na to, że można by zainicjować nowe narzędzia i metodologie. Ten punkt widzenia zwykle nie jest podzielany przez socjologów zajmujących się badaniem opinii społecznej, naukowców z dziedziny metodologii badań lub szerzej - metodologii nauk społecznych. W tych dziedzinach naukowcy w dużej mierze ignorowali rewolucję cyfrowych danych. Nie bez powodu. Dane cyfrowe zostały uznane za czarną skrzynkę ograniczającą możliwości zagłębienia się w problem. A nawet więcej, bezkrytyczne analizy dużych zbiorów danych doprowadziły do kilku spektakularnych niepowodzeń tego podejścia.

Uważamy jednak, że cyfrowe ślady nie wyeliminują potrzeby bezpośredniego zadawania pytań ludziom w ankietach, ale znacznie zwiększą użyteczność badań. Klasyczne dane pozwolą na kalibrację danych cyfrowych, weryfikację krzyżową i sprawdzenie kluczowych założeń. Prawdopodobnie analizowane dane online, wspomagane przez klasyczne instrumenty, mogłyby pomóc naukowcom w rozwiązywaniu podstawowych kwestii z zakresu nauk społecznych. Jak wskazali Golder i Macy w *ciągu ostatniego stulecia, nie brakowało teorii społecznych, ale istniały poważne ograniczenia w dostępie do danych* (2014, s. 130). Bezprecedensowe możliwości związane z dużymi danymi gromadzonymi w Internecie na masową skalę i na poziomie mikro zarazem, mogą zmienić tę sytuację.

Projekt ma na celu rozwinięcie metod wykorzystania cyfrowych śladów wraz z danymi ankietowymi, a tym samym przyczyni się do zrozumienia obecnych społecznych problemów badawczych dotyczących zaufania, postaw wobec migracji, środowiska, rządów, a także błędnych informacji i polaryzacji postaw. W celu opracowania metod wykorzystania cyfrowych śladów proponujemy przeprowadzenie szeroko zakrojonych badań metodologicznych, które wykorzystają pięć źródeł danych: reprezentatywne dane ankietowe, cyfrowe dane online, dane ze specjalnie wydzielonego badania panelowego, które połączy dane ankietowe i online, badania eksperymentalne i dane jakościowe.