

PRACE GEOGRAFICZNE, zeszyt 140

Instytut Geografii i Gospodarki Przestrzennej UJ

Kraków 2015, 9–23

doi: 10.4467/20833113PG.15.001.3530

OCENA JAKOŚCI INFORMACJI GEOGRAFICZNEJ (VOLUNTEERED GEOGRAPHIC INFORMATION – VGI) NA PRZYKŁADZIE MONITORINGU IMISJI ODORU W TARRAGONIE I SYSTEMU ORNITOLOGICZNEGO SABAP2

Kamil Jawęgiel

Rating criteria of the VGI quality on the examples of odour immission monitoring in Tarragona and the SABAP2 ornithological system

Abstract: Since 2007 there has been a discussion on the quality of data gained from volunteering. The information, which is gained from the volunteers, is based on the search for new experience and is usually considered as inaccurate and incorrect. This is why the evaluation of the VGI system in various forms is introduced. In the criteria analysis of the quality of VGI the Myszczyk's (2012) dimensions of proper functioning of the VGI, the Haklay's (2010) VGI classification levels and the Cooper's et al. (2012) set of the VGI quality characteristics were used, which formed the basis of the information types development and a plane for the comparative analyzes. Because of their utilitarian meaning, their goal is to raise the standards of the provided information and improvement of the developed information systems quality. Analysis of the quality of volunteering geographic information (VGI) was carried out on the example of monitoring odour immission in Tarragona and ornithological system SABAP2.

Keywords: VGI, Geographic Information, GIS, volunteering, odour

Zarys treści: Od 2007 r. trwa dyskusja na temat jakości danych uzyskanych z wolontariatu (VGI). Informacje, które uzyskuje się od ochotników, często wynikają z poszukiwania przez

nich nowych doświadczeń i są zwykle uważane za niedokładne i błędne. W pracy przedstawiono wprowadzane w różnych postaciach oceny VGI. W analizie kryteriów jakości VGI wykorzystano podstawowe wymiary sprawnie funkcjonującego systemu VGI wyróżnione przez Myszczyka (2012), klasyfikację poziomów zaangażowania wolontariatu VGI Haklaya (2010) oraz zestaw cech opisujących VGI według Coopera i in. (2011). Stwierdzono, że atrybuty te stanowiły podstawę rozwoju typów informacji i płaszczyznę dla analiz porównawczych. Ze względu na użyteczność tej klasyfikacji celem opracowania jest określenie standardów dostarczanej informacji, co może posłużyć poprawie jakości rozwiniętych już systemów VGI. Analizę jakości Informacji Geograficznej (VGI) przeprowadzono na przykładzie monitoringu emisji odorów w Tarragonie i systemu ornitologicznych SABAP2.

Słowa kluczowe: VGI, Informacja Geograficzna, GIS, wolontariat, odór

Wstęp

Ewolucja technologiczna pociąga za sobą zmiany sposobu pozyskiwania i wykorzystania danych geograficznych. Coraz powszechniejsze są interaktywne serwisy mapowe, jak GoogleMaps czy Microsoft Virtual Earth, które mogą być wykorzystywane przez każdego użytkownika z dostępem do internetu i dają możliwość rozpowszechniania autorskich map i informacji. Coraz częściej władze państwowe i samorządowe uzyskują informacje geograficzne od mieszkańców, na których podstawie dostrzegają ich potrzeby społeczne i rozwiązują lokalne problemy przestrzenne. Coraz więcej telefonów komórkowych czy aparatów cyfrowych wykorzystuje technologie GPS, aby zapewnić użytkownikom informacje o lokalizacji, a geotagging (geotagowanie) pozwala na znakowanie obiektów z informacją o ich położeniu. Wszystkie te zmiany przyczyniają się do rozwoju społecznościowej informacji geograficznej (VGI; Goodchild 2007). Dane przestrzenne są tworzone przez ochotników, którzy korzystają z narzędzi opisanych wyżej, w celu zebrania i rozpowszechniania swoich obserwacji i wiedzy geograficznej (Elwood 2008).

Popularyzacja Informacji Geograficznej (VGI) wspiera rozwój Systemów Informacji Geograficznej (*Geographic Information System* – GIS). Dane pozyskiwane przez wolontariuszy wpływają na rozwój badań interdyscyplinarnych, w których GIS znajduje nowe pole badawcze. W tym kontekście VGI może dać w przyszłości nowe odpowiedzi na pytania z różnych dziedzin nauki (Rodríguez-Pose 2011). Zaletą VGI jest także powszechny dostęp do danych, a poprawa ich jakości wpłynie na rozwój dalszych badań w różnych dziedzinach (Goodchild, Haining 2004).

Potencjał VGI dostrzegany jest przez firmy, przedsiębiorstwa i organizacje, które tworząc system społecznościowego pozyskiwania informacji, gromadzą i wykorzystują dane oraz informację geograficzną na własne potrzeby i dzielą się nią z klientami. Zdobywanie informacji w taki sposób jest stosunkowo tanie, biorąc pod uwagę ich ilość. Problemem staje się dokładność i wiarygodność pozyskanych informacji.



Zadawalające rezultaty w postaci danych o odpowiedniej jakości mogą przynieść tylko dobrze zaprojektowane, odpowiednio realizowane i na bieżąco rozwijane VGI (Jawgiel 2014).

Od 2007 r. naukowcy prowadzą dyskusję na temat jakości danych pozyskiwanych przez ochotników w ramach VGI. Jednym z przeciwników takiego sposobu zbierania informacji jest Budhathoki (Budhathoki i in. 2008), który twierdzi, że ważne elementy danych geograficznych, takie jak metadane czy normy opracowań, zostaną pominięte przez niedoświadczonych obserwatorów, dlatego postuluje usystematyzowanie i harmonizacja danych VGI. Elwood (2008) sugeruje natomiast badania relacji pomiędzy danymi twardymi i danymi uzyskanymi w ramach systemu VGI w celu określenia ich użyteczności. Coraz częściej również podważana jest trafność terminu *Volunteered Geographic Information* (Gorman 2012), ponieważ niejednokrotnie ochotnicy pobierają wynagrodzenie w ramach projektów. Dyskusje dotyczą także sposobu oceny wartości danych pozyskanych w ramach VGI oraz możliwości podniesienia ich jakości i dokładności. Informacje uzyskane od ochotników, ze względu na ich subiektywny charakter, często uznawane są za niedokładne, a nawet błędne. Dlatego wprowadzono różne sposoby oceny VGI.

Prawidłowe funkcjonowanie systemu Informacji Geograficznej wymaga dobrej organizacji pracy na każdym etapie, od doboru ochotników przez sposób zbierania i udostępniania informacji aż do określenia funkcji systemu w zakresie przechowywania i przetwarzania zgromadzonych danych. Nie każdy z tych elementów jest uwzględniany i rozwijany przez twórców baz VGI (Sui i in. 2012). Ze względu na złożoną strukturę VGI, zróżnicowane cechy obserwowanych zjawisk oraz zakres i metody ich pomiarów, jakość tworzonej na zasadzie społecznościowej informacji geograficznej jest różna. Stąd potrzeba opracowania wskaźnika umożliwiającego ocenę systemu zbierania informacji, wiarygodności obserwacji i pomiarów oraz określającego słabe i mocne strony systemu. Do dokładnej analizy i wykazania braków w funkcjonowaniu VGI niezbędny jest wskaźnik łączący wiele aspektów, który umożliwi niezależną ocenę każdego systemu, grupy wolontariuszy czy sposobu prowadzenia obserwacji.

Celem artykułu jest opracowanie metody ilościowej oceny jakości informacji zbieranych w ramach VGI. Przedmiotem oceny są kategorie opisane we wcześniejszych opracowaniach określające w sposób jakościowy zakres struktury i funkcji oraz użyteczność dostarczanych danych. Kategorie te zostały wyrażone w ujednoczonej, punktowanej skali. Należą do nich własności jakości Haklaya (2010), Coopera i in. (2011) i Myszczuka (2012). Punktacja ta uwzględnia zarówno pozytywne, jak i negatywne właściwości systemu informacji oraz gromadzonych przez niego danych. Analiza jakości informacji w ramach VGI została przeprowadzona na przykładzie monitoringu imisji odoru w Tarragonie (Ramos, Pindado 2013) i systemu ornitologicznego SABAP2 (<http://sabap2.adu.org.za/> – 20.12.2014). Projekty będące



przedmiotem tego opracowania zostały tak dobrane, aby pokazać uniwersalność i interdyscyplinarność Informacji Geograficznej i potencjalną możliwość zastosowana opracowanej przez autora punktacji, w różnych dziedzinach nauki.

Charakterystyka ocenianych projektów wykorzystujących VGI

Społecznościowy system monitoringu emisji odorów w Tarragonie powstał w 2010 r. Miasto położone jest na obszarze zasobnym w złoża naturalne. Ich eksploatacja oraz dalsze przetwórstwo są przyczyną nieprzyjemnego zapachu. Działalność zakładów produkcyjnych otaczających miasto negatywnie wpływa na środowisko naturalne (głównie na powietrze i wodę), co wywołało lokalną debatę na temat ochrony i monitoringu środowiska oraz postrzegania zagrożenia przez społeczeństwo. Władze gminne we współpracy z Centrum Technologii Chemicznej (CTQC) uznały, że najbardziej uciążliwe w mieście są trzy zakłady (parki technologiczne), które poddano stałemu monitoringowi olfaktometrycznemu (Ramos, Pindado 2013).

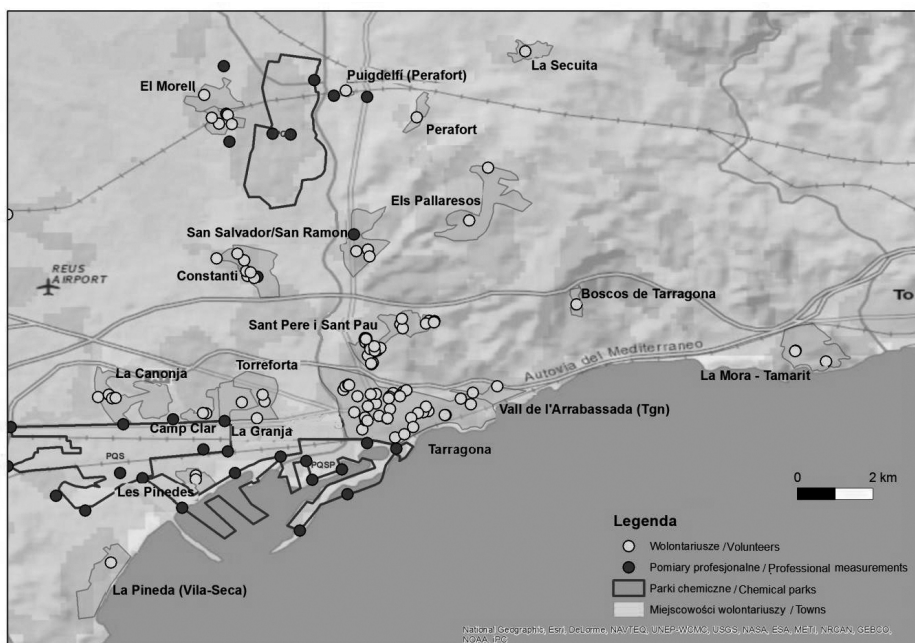
Obserwacje wokół wspomnianych parków technologicznych prowadzone są w ramach dwóch niezależnych systemów pomiarowych. Pierwszy z nich opiera się na badaniach profesjonalnych, prowadzonych przez specjalnie przeszkolone grupy strażaków, które wykonują pomiary dwa razy w ciągu dnia, w godzinach 8–11 i 15–18. Pomiary olfaktometrem terenowym Nasal Ranger wykonywane są w tych samych punktach pomiarowych, zlokalizowanych w łatwo dostępnych miejscach o dużej ekspozycji. Wokół każdego zakładu znajduje się 10–11 takich punktów. Badanie polega na odczytaniu z urządzenia intensywności odorów oraz pobraniu próbki powietrza do analizy fizykochemicznej w laboratorium (Ramos, Pindado 2013).

Drugi system pomiarowy funkcjonuje dzięki grupie 100 ochotników działających w ramach VGI. Oceniają oni intensywność odorów na podstawie własnych zmysłów i obserwacji. Badania te wykonywane są nieregularnie, bo tylko w momencie wyczućcia przez wolontariusza nieprzyjemnego zapachu pochodzącego z zakładu. Informacje o tym fakcie przesyłane są niezwłocznie do sieci monitoringu, gromadzącego dane zbierane zarówno przez profesjonalistów (strażaków), jak i przez ochotników. Wolontariusze oceniają intensywność zapachu w skali od 1 do 5 i notują wszelkie spostrzeżenia, które ułatwiają identyfikację procesu przez chemików. Wykorzystanie społecznościowego sposobu zbierania informacji poszerza zakres badań przestrzenne i ilościowo w stopniu, który nie byłby możliwy do osiągnięcia bez poniesienia znacznych kosztów na stworzenie profesjonalnej sieci pomiarowej. Obserwatorzy dokonują pomiarów stale w tym samym miejscu, czyli w najbliższej okolicy własnego domu, co nie wymaga od wolontariuszy każdorazowego określania położenia i chroni przed popełnieniem związanego z tym błędu (ryc. 1). Projekt Drugiego



Południowoafrykańskiego Atlasu Ptaków (SABAP2) jest realizowany od 2007 r. Celem projektu jest stworzenie kompleksowego zbioru miejsc występowania ptaków, głównie zagrożonych gatunków. Ornitolodzy i ochotnicy (amatorzy) gromadzą szczegółowe dane na temat miejsc bytowania ptaków. Za pomocą specjalnej platformy internetowej bądź drogą pocztową przesyłają je do operatorów systemów, którzy gromadzą dane i weryfikują zgłoszenia. Obszar kraju został podzielony na siatkę kwadratów, z których każdy ma wymiary 5' x 5' (ok. 9 km x 9 km), które rozmiarem odpowiadają potencjalnym wielkościom siedlisk ptaków i minimalnej dokładności określenia położenia obserwatora. Jako minimalną dokładność czasową raportowania informacji przyjęto okres 5 dni, który odpowiada w przybliżeniu maksymalnemu tempu migracji gatunków (Cooper i in. 2010).

Największym ograniczeniem metody jest poprawne zaklasyfikowanie obserwowanych ptaków do właściwej grupy systematycznej. Umiejętność ta jest bowiem uzależniona od wiedzy i doświadczenia obserwatora. Przekazywane informacje często są niepełne i mimo dołączonych zdjęć, nagrań odgłosów ptaków i innej dokumentacji niemożliwe staje się na tej podstawie prawidłowe określenie gatunku. Dlatego



Ryc. 1. Lokalizacja punktów pomiarowych (na podstawie mapy topograficznej ESRI)

Fig. 1. Measurement points (based on the ESRI topographic map)



obserwatorzy oznaczają najczęściej kategorie systematyczne, takie jak rząd, rodzina, rodzaj, a rzadziej gatunek. Dodatkowym elementem utrudniającym jednoznaczną ocenę są fałszywe zgłoszenia preparowane przez inwestorów w celu umożliwienia realizacji korzystnych dla siebie inwestycji, bez względu na negatywne ich oddziaływanie na środowisko (Cooper i in. 2011).

Prekursorem programu SABAP2 był projekt SABAP1, realizowany w latach 1986–1997 wyłącznie przez zespół profesjonalistów, którzy w podobny sposób jak opisany wyżej dokonywali obserwacji ptaków. Obecnie SABAP1 jest empirycznym zbiorem testowym dla badań realizowanych w ramach SABAP2. SABAP2 ze względu na zróżnicowane umiejętności obserwatorów jedynie w około 66% odpowiadał danym uzyskanym przez wolontariuszy, co świadczy o niewystarczających umiejętnościach części obserwatorów (Cooper i in. 2011; <http://sabap2.adu.org.za/> – 20.12.2014).

Metody badań

W ocenie jakości systemu Informacji Geograficznej i wyborze podstawowych atrybutów wykorzystano wymiary systemu VGI Myszczyka (2012), klasyfikację poziomów wolontariatu geograficznego Haklaya (2010) oraz zbiór cech opisujących VGI Coopera i in. (2011), stanowiących podstawę rozwoju tego typu informacji i płaszczyznę analiz porównawczych.

Według Myszczyka (2012) do opracowania systemu VGI niezbędne są cztery powiązane z sobą elementy: technologia, internet, społeczność, dane przestrzenne (tab. 1). Jedynie odpowiednie wykorzystanie tych elementów może zwiększyć ilość i jakość pozyskiwanych danych, a co za tym idzie, podnieść wartość każdego projektu VGI. Dodatkowymi czynnikami umożliwiającymi nadanie projektowi aplikacyjnego charakteru jest odpowiednie wykorzystanie GIS oraz przeprowadzenie obserwacji na wskazanym poziomie dokładności (Sui i in. 2012). Wykorzystanie internetu ma na celu automatyzację i usystematyzowanie zbierania danych (ryc. 2). Jedynie odpowiednia integracja wszystkich tych elementów może przynieść rezultat w postaci dobrze i skutecznie funkcjonującego VGI. Haklay (2010) wyróżnia cztery poziomy zaawansowania wolontariatu, z których każdy zależny jest od stopnia zaangażowania obserwatorów i ich wpływu na badania. Są to: *crowdsourcing*, *distributed intelligence*, *participatory science*, *extreme citizen science*. Wszystkie one zostały uwzględnione w przeprowadzonej analizie jakości Informacji Geograficznej (VGI).

Według Haklaya (2010) na najbardziej podstawowym poziomie zaangażowanie wolontariuszy jest niewielkie. Ochotnicy są zaangażowani jedynie w zbieranie danych, a nie ich analizę. To działanie określane jest jako *crowdsourcing*. Obserwowane przez wolontariuszy zjawiska i procesy są zazwyczaj opisywane dychotomicznie (zero-jedynkowo), dostarczając głównie informacje „gdzie?” i „kiedy?” Projekty



Tab. 1. Elementy systemów VGI wg Myszczyka (2012)

Table 1. Elements of the VGI systems by Myszczyk (2012)

Element / Component	Opis / Description
Technologia Technology	Urządzenia służące do przeglądania zasobów internetu, takie jak: komputery, smartphony, tablety itp.
Internet Internet	Globalne medium, protokół umożliwiający szybką wymianę informacji
Spółeczność Society	Ogół użytkowników, sensorów dostarczających bogatej wiedzy o powierzchni Ziemi i jej właściwościach (Goodchild 2007)
Dane przestrzenne Spatial data	Informacja, w tym również multimedialna, którą można ulokować w przestrzeni geograficznej

Źródło: Myszczyk (2012).

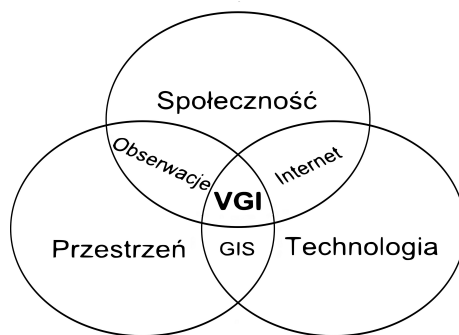
Source: Myszczyk (2012).

takie są bardzo często krytykowane, gdyż umiejętności i wiedza ochotników nie są efektywnie wykorzystane (Howe 2006).

Drugi stopień zaangażowania określany jest jako *distributed intelligence*. Na tym poziomie wykorzystywane już są bardziej umiejętności i wiedza uczestników. Polega on na przeszkoleniu wolontariuszy i ukierunkowaniu ich obserwacji na określone, konkretne informacje. Na tym poziomie prowadzi się najwięcej projektów ze względu na połączenie stosunkowo dokładnej informacji z małymi nakładami finansowymi (Haklay 2010).

Kolejny poziom zwany *participatory science* zakłada większe zaangażowanie naukowców. Wolontariusze uczestniczą w tworzeniu projektu, wspólnie z naukowcami opracowują metody pomiaru. Na tym etapie naukowcy i ochotnicy ściśle współpracują z sobą zarówno przy zbieraniu danych, jak i przy analizie wyników. Ten rodzaj współpracy występuje najczęściej w projektach odnoszących się do ochrony środowiska (Haklay 2010).

Ostatni poziom, *extreme citizen science*, jest zintegrowaną działalnością, w którą obok znaczącego udziału badaczy do rozwiązywania problemów naukowych włączone są osoby o ponadprzeciętnej znajomości warunków lokalnych. Wolontariusze uczest-



Ryc. 2. Elementy składowe VGI (na podstawie: Myszczyk 2012; Sui i in. 2012)

Fig. 2. Components of the VGI (on the basis of: Myszczyk 2012; Sui et al. 2012)



niczący w pracach na niższych poziomach rozwoju VGI mogą być również włączeni do analizy, ale preferowane jest bezpieczeństwo wyników i wiarygodność pomiarów. Zaangażowana społecznie nauka, w której nie braliby udziału naukowcy nadzorujący cały cykl obserwacji, nie pozwoliłaby osiągnąć założonego wcześniej celu (Haklay 2010). Różne poziomy zaangażowania wolontariuszy mogą być wykorzystywane w całej gamie zaangażowanych społecznie działań naukowych.

Zwykle największa liczba ochotników ogranicza swoje działania do zbierania informacji. Większość uczestników projektu może realizować jego zadania na najniższym poziomie, podczas gdy część z nich, wraz ze specjalistami, będzie na wyższych poziomach wspierać technicznie osoby z poziomów niższych. Dla bardzo zaawansowanych użytkowników powinien być też możliwy swobodny przepływ informacji pomiędzy poziomami, pozwalający na współpracę i analizy wyników badań. Znaczącą rolę ma również oddolna inicjatywa wolontariuszy, wskazująca problemy i mogąca wskazać nowe kierunki i metody badań (Haklay 2010).

Systematyzację cech opisujących jakość VGI oraz propozycję wyzwań stojących przed użytkownikami systemów przedstawił Cooper i in. (2011) (tab. 2). Zaproponowane przez Haklaya (2010), Coopera i in. (2011) oraz Myszczyka (2012) klasyfikacje, prezentujące między innymi, normy i zasady oceny jakości społecznościowej informacji geograficznej (VGI), mają na celu podniesienie standardów pozyskiwanych danych i jakości tworzonych systemów informacji.

Na podstawie przedstawionych wyżej form oceny wolontariatu geograficznego przygotowano autorski sposób oceny jego jakości. Punktacja mówiąca o użyteczności i sposobie jego funkcjonowania została opracowana w taki sposób, aby uwzględniła możliwie jak największą liczbę zmiennych (cech, właściwości) mających istotne znaczenie dla funkcjonowania VGI. W przeprowadzonej analizie uwzględniono: klasyfikację Haklaya (2010), cechy opisujące jakość systemów VGI według Coopera i in. (2011) oraz „automatyzację systemu” wytypowaną na podstawie autorskiego doświadczenia. Automatyzacja to sposób i szybkość pozyskiwania danych, możliwości reakcji operatorów na zgłoszenia wolontariuszy, a także łatwości w transferze informacji pomiędzy wszystkimi elementami systemu. Zmienne: dokładność tematyczna i dokładność semantyczna oraz liniowość danych i logiczna spójność zostały połączone w dwie grupy. Ostatecznie do dalszych analiz wybrano następujące cechy (właściwości systemów VGI): stopnie wg Haklaya (2010), dokładność geograficzną, dokładność czasową, dokładność tematyczną i semantyczną, kompletność danych, logiczną spójność i pochodzenie danych oraz automatyzację całego systemu. Podane wyżej zmienne pozwolą określić standardy oceny jakości danych przestrzennych. Wyselekcjonowano jedynie najistotniejsze elementy mające największe znaczenie przy tworzeniu VGI, a część z nichłączono w jedną grupę. Wszystkie cechy zostały opisane za pomocą czterostopniowej skali, stanowiącej odpowiednik skali Haklaya, która została zbudowana z wykorzystaniem



Tab. 2. Własności jakości i wyzwania dla VGI wg Coopera i in. (2011)

Table 2. Properties of quality and challenges for VGI by Cooper et al. (2011)

Poziom / Level	Cecha / Feature	Opis / Description
Własności jakości Quality properties	Dokładność lokalizacyjna	Określa, jak bardzo otrzymane dane są oddalone od rzeczywistego miejsca pomiaru oraz gęstość siatki pomiarowej, czyli liczbę wolontariuszy przypadającą na jednostkę powierzchni.
	Dokładność tematyczna	Określa, jak blisko atrybuty zbliżone są do rzeczywistości w postaci jakościowej bądź ilościowej. Wyznaczana jest za pomocą istniejących skal pomiarowych: ilorazowej, interwałowej, porządkowej i nominalnej.
	Dokładność semantyczna	Określa różnicę pomiędzy skalą interpretowaną przez ludzkie zmysły a skalą stosowaną przy pomiarach z użyciem sprzętu pomiarowego.
	Dokładność czasowa	Określa, jak bardzo zjawisko zmienia się w czasie i ile czasu upłynęło pomiędzy obserwacjami a jego wystąpieniem w rzeczywistości.
	Kompletność	Określa, jak zestaw danych odzwierciedla całość analizowanych zjawisk w aspekcie czasowym i przestrzennym, czy część danych została pominięta, czy zbadana kompleksowo.
	Logiczna spójność	Określa liczbę braków danych i ich logicznych błędów, takich jak: niespójne dane, niewłaściwe atrybuty, niewłaściwe relacje między cechami, brakujące ograniczniki. Wymagana jest logiczna spójność całości danych.
	Pochodzenie danych	Znajomość pochodzenia danych, w tym na temat wolontariuszy, na każdym etapie ich zaawansowania. Poznanie „cyklu życia” danych (zbierania, odbioru, przetwarzania) w celu odpowiedniego łączenia ich w wynikach badań.
Zadania VGI (wyzwania) VGI challenges	Uzależnienie od celu i kontekstu	Jakość danych powinna być uzależniona od kontekstu ich zapotrzebowania, a nie zbierana ponad miarę. Działanie takie zniechęca wolontariuszy do obserwacji.
	Nienormatywność	Obserwatorzy powinni oceniać w takich samych standardach i jednostkach mimo rozbieżności w metodach pomiarowych, zdolnościach poznawczych i osobniczych czy jakości sprzętu pomiarowego.
	Anonimowe wpisy	Są bardzo trudne do skontrolowania, co może powodować stronniczość bądź trudność w porównaniu zbiorów danych.
	Stronniczość	Oprócz klasycznych błędów badania nie powinny być przeprowadzane w sposób stronniczy, mający na celu promowanie lub dyskredytację pewnej opcji politycznej, religijnej bądź społecznej. Manipulowanie danymi może pochodzić zarówno od obserwatorów, jak i od twórców systemu.
	Aspekty jakościowe	Dane jakościowe powinny być tak opisywane, aby były zrozumiałe w różnych językach, bez używania slangów i wyrazów niezrozumiałych.

Źródło: opracowanie własne na podstawie: Cooper i in. (2011).

Source: own study based on Cooper et al. (2011).



sześciu najważniejszych i mierzalnych właściwości systemu, wybranych z połączenia metody Haklaya i Coopera. Każda z cech i sposób jej oceniania zostały opisane i odpowiednio zinterpretowane (tab. 3). Stopnie skali zostały uporządkowane rosnąco. Na najwyższym poziomie znajdują się funkcje systemu mające najkorzystniejszy wpływ na VGI, a na najniższym poziomie opisano sposoby obserwacji o mniejszej wartości użytkowej (4-3-2-1).

Wyniki

Na podstawie opisanej wyżej metody przedstawiono punktację dla opisywanych wcześniej systemów (Tarragona, SABAP2), a wyniki oceny jakości zostały zestawiono w tabeli (tab. 4). Oba systemy charakteryzują się niską kompletnością danych, której przyczyną jest nieregularne pokrycie terenu siatką pomiarów oraz nieregularność obserwacji. W obu przypadkach większość ochotników należała do grupy *distributed intelligence* (Haklay 2010), gdzie zdolności poznawcze osób nie są kompleksowo wykorzystane. Ze względu na dominującą liczbę ochotników zaklasyfikowanych do tej grupy uznano ją jako wyznacznik w punktacji. Należy pamiętać jednak, że granice pomiędzy poziomami mogą się zacierać, tzn. wolontariusze mogą jednocześnie wchodzić w skład kilku grup, wtedy należy wybrać tę dominującą. Obserwatorzy nie uczestniczą w tworzeniu projektu i jego realizacji. Cechy te otrzymały najmniejszą liczbę punktów, a co za tym idzie, powinny być rozwijane w obu przypadkach w pierwszej kolejności. Sumarycznie system Tarragona otrzymał 16/28 pkt, a system SABAP2 13/28 pkt, co może świadczyć o zbliżonym, średnio zaawansowanym poziomie rozwoju obu systemów. Niemniej oba z nich wykazują opisane wyżej niedoskonałości w funkcjonowaniu. Wyniki te mogą ukierunkować twórców i zarządzających systemami VGI do szczegółowej analizy ich struktury oraz określenia kierunków ich rozwoju pod kątem jakości pozyskiwanych danych.

System Tarragona charakteryzuje się wysoką dokładnością obserwacji i spójnością danych. Zaletą systemu jest również duża dokładność tematyczna i semantyczna oparta na wartościach progowych wyrażonych w jednostkach odoru na metr sześcienny ($\text{ou}\cdot\text{m}^{-3}$). Punkty pomiarowe są stałe, a obserwacje dokładne w czasie. Pozyskiwane dane są ujednolicone za pomocą spójnej skali pomiarowej. System SABAP2 cechuje wysoki stopień automatyzacji zgłoszeń dzięki specjalnej platformie internetowej, co stanowi główną zaletę systemu. Dzięki prowadzonej dokumentacji obserwacje są spójne logicznie i dokładne tematycznie, co wzmacnia również katalog empiryczny SABAP1, zawierający odpowiednią specyfikację prowadzenia obserwacji przez wolontariuszy. Największą wadą systemu jest niewielka dokładność geograficzna i czasowa, ze względu jednak na dużą zmienność obserwowanego zjawiska (duże i częste migracje plectwa) nie wpływa to znacząco na realizację projektu.



Tab. 3. Czterostopniowa skala jakości VGI
Table 3. Four-grade quality scale of the VGI

Cecha / Feature	Skala 4 / Scale 4	Skala 3 / Scale 3	Skala 2 / Scale 2	Skala 1 / Scale 1
Stopień zaangażowania wolontariuszy / Levels of participation and engagement (Haklay 2010)	<i>Extreme citizen science</i> : wolontariuszami są jedynie naukowcy, których działania nie ukierunkowane jest na rozwiązanie konkretnego problemu naukowego.	<i>Participatory science</i> : wolontariusze uczestniczą w tworzeniu projektu, wspólnie z naukowcami opracowują metody i rozdzielczość pomiaru.	<i>Distributed Intelligence</i> : wolontariusze obserwacji dokonują po przebyciu szkoleniu.	<i>Crowdsourcing</i> : wolontariusze nie zaangażowani w analizę, marnotrawione zdolności poznawcze.
Dokładność geograficzna Positional accuracy (Cooper i in. 2011, Myszczyk 2012)	Obserwacje dokonywane w różnych miejscach z dokładnym zaznaczeniem miejsca pomiaru przy pomocy technologii mobilnych.	Obserwacje w starych, strategicznych punktach tworzących regularną siatkę pomiarową.	Obserwacje w różnych miejscach pomiarowych, podawane z niewielką dokładnością, pozwalająca na zadowalające określenie lokalizacji zjawiska.	Obserwacje o czasie nieokreślonym, bądź o bardzo niewielkiej dokładności, uniemożliwiające dokładną interpretację zjawiska.
Dokładność czasowa Currency (Cooper i in. 2011)	Stale obserwacje zmienne w czasie, obserwowane w momencie wystąpienia zjawiska.	Obserwacje w stałym wyznaczonym czasie i dużej okładności (przy zjawiskach mało zmiennych w czasie).	Obserwacje sporadyczne i losowe, określone w czasie, ale z niewielką dokładnością (przy zjawiskach wysoko zmiennych w czasie)	Obserwacje o czasie nieokreślonym, uniemożliwiające dokładną interpretację zjawiska.
Dokładność tematyczna i semantyczna Attribute accuracy (Cooper i in. 2011)	Obserwacje dokonywane z dokładnością stosowanego sprzętu pomiarowego za pomocą zamontowanych na stałe urządzeń pomiarowych (skala Ilorzawa), np. poziom wodowoskazu.	Obserwacje oparte na charakterystycznych wartościach progowych wyznaczonych na podstawie specyfikacji realizowanego projektu, pozwalające porównywać obserwacje z twardymi danymi pomiarowymi (skala interwałowa), np. siła wiatru wg skali Beauforta.	Obserwacje określające siłę zjawiska, odróżniające jego intensywność, ale nie odnoszące się do skali stosowanych w nauce (skala porządkowa), np. temperatura: ciepło/zimno.	Obserwacje w pełni subiektywne, bez punktów odniesienia, na zasadzie zero-jedynkowej, zjawisko występuje lub nie występuje (skala nominalna), np. zerujące dzikie zwierzęta w danym miejscu.
Kompletność danych Completeness (Cooper i in. 2011)	Pomiar ciągły (niedyskretny), obserwacje nieustające, niewymagające dużego zaangażowania, ale zmysłowo, automatycznie wykrywalne przez wolontariuszy, np. opad deszczu, burza.	Obserwacje w wyznaczonych, strategicznych momentach uzależnionych od kontekstu realizowanego projektu, np. obserwacja wiatru podczas opadu gradu.	Obserwacje o stałych, wyznaczonych porach/miejscach tworzących regularną siatkę pomiarową, np. pomiary ciśnienia o godzinie 6.00.	Obserwacje wybiórcze, w momentach dogodnych dla wolontariuszy, niegwarantujących ciągłości i użyteczności.
Logiczna spójność i pochodzenie danych Logical consistency and lineage (Cooper i in. 2011)	Pełna znajomość pochodzenia danych, bezpośredni udział zespołu badawczego w każdym etapie związany z rejestracją obserwatorów i dokładnością obserwacji, która umożliwia głębsze ich poznanie. Możliwe wielokrotne dane analizy. Dane są spójne, a ich błędy i braki nie występują bądź są w odpowiedni, zunifikowany sposób zniwelowane.	Dobra znajomość pochodzenia i przetwarzania danych, pozwalająca na szeroką interpretację. Wolontariusze są zarejestrowani, ściśle określony jest sposób odbioru i przetwarzania danych, które są rozumiałe dla całego zespołu badaczy. Relacje pomiędzy atrybutami są poprawnie określone, a występujące błędy są możliwe do usunięcia.	Podstawowa wiedza na temat danych, cząstkowe informacje na temat wolontariuszy i sposobów transferu danych, pozwala na podstawowe analizy czasowo-geograficzne, wolontariusze anonimowi. Występujące błędy są liczne, ale po odpowiedniej analizie umożliwiają dalsze badania.	Niezajomość genety danych, otrzymywanie danych już obrabowanych bez wiedzy na temat obserwatorów, sposobów transformacji, charakterystyki, metadanych. Liczne i niezidentyfikowane błędy danych, uniemożliwiające analizę.
Automatyzacja / Automation	Obserwacje z użyciem technologii mobilnych, dostarczających informacji na bieżąco do administratora.	Obserwacje z użyciem platformy internetowej, przesyłane z opóźnieniem, ale niewymagające specjalnego zaangażowania obserwatora.	Obserwacje gromadzone i przesyłane komputerowo po pewnym czasie (np. na koniec miesiąca, tygodnia).	Obserwacje dostarczane analogowo (w formie papierowej), wymagające późniejszego wprowadzenia danych do komputera.

Źródło: opracowanie własne na podstawie: Haklay (2010), Cooper i in. (2011), Myszczyk (2012).
Source: own study based on Haklay (2010), Cooper et al. (2011), Myszczyk (2012).

Tab. 4. Punktacja oceny jakości informacji wolontariatu geograficznego na przykładzie monitoringu odoru w Tarragonie i systemu ornitologicznego SABAP2

Table 4. Rating of the VGI quality on the example of odour monitoring in Tarragona and ornithological SABAP2 systems

Cecha / Feature	System Tarragona	Punkty Points	System SABAP2	Punkty Points
Stopień zaangażowania wolontariuszy Levels of participation and engagement (Haklay 2010)	<i>Distributed intelligence:</i> wolontariusze wprowadzeni w podstawy projektu.	2	<i>Distributed intelligence:</i> wolontariusze wtajemniczeni w podstawy projektu.	2
Dokładność geograficzna Positional accuracy (Cooper i in. 2011, Myszczyk 2012)	Obserwacje w stałych lokalizacjach, w pobliżu domu obserwatora.	3	Obserwacje w miejscach przypadkowych.	2
Dokładność czasowa Currency (Cooper i in. 2011)	Obserwacje nieregularne, uzależnione od miejsca pobytu, zjawisko wysoko zmienne w czasie.	2	Obserwacje niedokładnie wskazujące czas, określające miejsce bytowania ptaka, określenie jego siedliska.	1
Dokładność tematyczna i semantyczna Attribute accuracy (Cooper i in. 2011)	Obserwacje porównywalne ze stosowanym systemem. Stosowane progi wyczuwalności i rozpoznawalności zapachu.	3	Obserwacje w pełni subiektywne, na podstawie opisu wyglądu gatunku. Na podstawie fotografii i nagrań.	2
Kompletność danych Completeness (Cooper i in. 2011)	Obserwacje wykonywane jedynie w momencie, gdy obserwator jest w pobliżu domu, na zewnątrz. Nie uwzględniają wartości zerowych.	1	Obserwacje wykonywane jedynie w momencie, gdy obserwator znajduje się w miejscu bytowania ptaka.	1
Logiczna spójność i pochodzenie danych Logical consistency and lineage (Cooper i in. 2011)	Dobra znajomość danych, wolontariusze zarejestrowani, znajomość pochodzenia i określony sposób przetwarzania, nieokreślone powiązania między obserwacjami.	3	Niewielka znajomość danych, głównie ze względu na anonimowość wolontariuszy i niedokładność czasową i geograficzną.	2
Automatyzacja Automation	Obserwacje zbierane i przesyłane mailowo na koniec miesiąca.	2	Obserwacje dostarczane w dowolnym czasie różnymi metodami, głównie za pomocą specjalnej platformy internetowej.	3
SUMA / TOTAL		16		13

Punktacja zawarta jest w przedziale [7; 28] / rating comprised within the range [7; 28].

Źródło: opracowanie własne na podstawie: Haklay (2010), Cooper i in. (2011), Myszczyk (2012).

Source: own study based on Haklay (2010), Cooper et al. (2011), Myszczyk (2012).



Dyskusja

Jakość informacji geograficznej uzależniona jest od struktury systemu VGI, charakteru obserwowanych zjawisk oraz sposobu ich pomiarów i obserwacji. Zastosowane kryteria uznano za zasadne ze względu na ich kompleksowość oraz użycie różnych innych, wykorzystywanych już wcześniej skal. W zakresie konstrukcji punktacji oceny jakości VGI wskazują one potencjalne niedociągnięcia w funkcjonowaniu analizowanych VGI przez możliwość ilościowego porównania z innymi, podobnymi systemami.

Poziomy zaawansowania wolontariatu zaproponowane przez Haklaya (2010), uzależnione są w dużym stopniu od zaangażowania obserwatorów. Może ono wpływać na pozostałe zmienne opisujące VGI, dlatego powinno być traktowane jako czynnik nadrzędny, determinujący pozostałe z nich. W obu przypadkach wyniki były zbliżone i zostały sklasyfikowane na jednym poziomie (*distributed intelligence*), ale dopiero uzupełnienie i rozwinięcie metody Haklaya pozwoliło wskazać dokładne różnice pomiędzy SABAP2 i Tarragoną. Wyniki pokazały, że połączenie kilku zmiennych pozwala na określenie różnic pomiędzy systemami.

Przeprowadzone w opracowaniu analizy zostały oparte na skombinowanych klasyfikacjach Haklaya (2010), Coopera i in. (2011) i Myszczyka (2012). Podejście to cechuje się wielowymiarowością, czyli zwraca uwagę na różne cechy wolontariatu geograficznego. Podobnie jak Haklay autor wskazał wątpliwości co do słuszności klasyfikacji cechy w jednej tylko kategorii. W wielu przypadkach granice przedziału cech mogą być nieostre, jednoznaczna ich klasyfikacja może być utrudniona, a właściwości te mogą przenikać między poziomami, na przykład różne możliwości transferu danych stosowane w jednym systemie (automatyzacja). Dlatego ważne jest, aby w analizach jakości wolontariatu uwzględniać te atrybuty, które mają największe znaczenie bądź występują najliczniej.

Porównując zaproponowaną metodę oceny VGI z klasyfikacją Coopera i in. (2011), należy stwierdzić, że opracowanie zostało rozszerzone o podejście ilościowe. Każda z cech może mieć przypisaną wartość liczbową, pozwalającą na jednoznaczne określenie jej stopnia natężenia. Cooper opisuje jedynie właściwości systemów, nie interpretując ich wagi i znaczenia w funkcjonowaniu VGI. Zbiór cech mógł być odniesiony do każdej działalności, ale nie pozwalał ich porównywać i oceniać, co może stanowić barierę w rozwoju VGI. Podejście autorskie, zaproponowane tutaj, jest wzbogaconą formą, posiada wszystkie zalety opracowania Coopera i in. (2011) i również może odnieść się do każdego systemu. Należy podkreślić, że w celu udoskonalenia i implementacji rozwiązań w kolejnych systemach VGI zaproponowane podejście może ułatwić wyszukiwanie optymalnych środków dla rozwoju systemów oraz określić hierarchię podejmowanych działań opartą między innymi na zmiennych z najniższą punktacją. Wszystkie wzięte pod uwagę cechy mają znaczenie w rozwoju działalności VGI i konieczne jest ich rozwijanie niezależnie od siebie w celu



podniesienia jakości zbieranych informacji. Bez odpowiedniego przystosowania tych cech nie jest możliwe skuteczne funkcjonowanie systemu VGI.

Podsumowanie

VGI nie jest doskonałym źródłem informacji przestrzennej, tylko dobrze zaprojektowanym i funkcjonującym systemem dającym zadowalające wyniki. Rozwijanie wszystkich właściwości systemów VGI jest konieczne, aby pozyskiwane dane odznaczały się odpowiednią jakością. Powszechnie stosowane systemy VGI wciąż nie są doskonałe i stale powinny być rozwijane w wielu kierunkach. Odpowiednie ich rozpoznanie jest podstawą tworzenia danych o wyższej jakości i wartości nie tylko poznawczej, ale i praktycznej.

Za najlepszy sposób oceny wolontariatu uznaje się porównanie danych z innym zbiorem testowym, na przykład z pomiarami z użyciem profesjonalnego sprzętu, bądź danych empirycznych. Taka metoda weryfikacji zgłoszeń obserwatorów nadaje im bardziej profesjonalny charakter. W sytuacji gdy zbiór testowy nie jest dostępny, należy wykorzystać syntetyczne metody oceny danych, podobne do wykorzystanych w niniejszym opracowaniu.

Zaproponowany sposób oceny wskazuje słabe i mocne strony VGI i umożliwia ich syntetyczne porównywanie. Podejście ilościowe dopasowane jest do poziomu istotności statystycznej danej cechy w badaniach oraz jej wpływu na wyniki obserwacji. Pozwala na hierarchizację i integrację działania w rozwijaniu systemów i przeciwdziałaniu trudnościom w realizacji projektów z udziałem VGI. Stosowanie ocen w dalszej perspektywie potencjalnie może nie tylko podnieść jakość pozyskiwanych danych, ale także wpłynąć na sposób funkcjonowania systemu. Charakter ilościowy ujednoliconej punktacji daje szerokie możliwości nie tylko jako podstawy analizy SWOT, ale również innych metod statystycznych.

Przeprowadzone badania potwierdzają trudności w jednoznacznej ocenie jakości danych dostarczanych przez VGI. Oba analizowane systemy w wielu aspektach nie mogły zostać jednoznacznie ocenione. Zaproponowana punktacja nie daje możliwości oceny wszystkich potencjalnie istniejących kombinacji systemów VGI, ale dokładnie nakreśla ich właściwości, które mogą podlegać dalszej analizie jakościowej i ilościowej. Liczba systemów w miarę rozwoju i popularyzacji VGI będzie rosła. Zapewne będą dodawane nowe, bardziej zaawansowane funkcje systemów. Dlatego też sposób oceny VGI musi być wciąż rozwijany.

Literatura

Budhathoki N.R., Bruce B., Nedovic-Budic Z., 2008, *Reconceptualizing the role of the user of spatial data infrastructure*, GeoJournal, 72, 149–60.



- Cooper A.K., Coetzee S., Kaczmarek I., Kourie D., Iwaniak A., Kubik T., 2011, *Challenges for quality in volunteered geographic information*, AfricaGEO 2011, Cape Town, South Africa.
- Cooper A.K., Coetzee S., Kourie D., 2010, *Perceptions of virtual globes, volunteered geographical information and spatial data infrastructures*, *Geomatica*, 64, 1, 73–88.
- Elwood S., 2008, *Volunteered geographic information: Future research direction motivated by critical, participatory, and feminist GIS*, *GeoJournal*, 3–4, 173–183.
- Flanagin A.J., Metzger M.J., 2008, *The credibility of volunteered geographic information*, *GeoJournal*, 72, 137–148.
- Goodchild M.F., 2007, *Citizens as sensors: The world of volunteered geography*, *GeoJournal*, 69, 211–221.
- Goodchild M.F., Haining R.P., 2004, *GIS and spatial data analysis: Converging perspectives*, *Papers in Regional Science*, 83, 363–385.
- Gorman S., 2012, *Why VGI is the Wrong Acronym*, Fortius One.
- Haklay M., 2010, *How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets*, *Environment and Planning, B: Planning and Design*, 37, 4, 682–703.
- Howe J., 2006, *The rise of the crowdsourcing*, *Wired Magazine*, 14, 1–4.
- Jawgiel K., 2014, *Geowizualizacja imisji odoru według danych profesjonalnych i VGI na przykładzie Tarraony (Hiszpania) w oprogramowaniu GIS*, praca magisterska, UAM Poznań.
- Myszczuk M., 2012, *Geowizualizacja zmienności czasowo-przestrzennej aktywności VGI na przykładzie panoramio.com*, praca magisterska, UAM, Poznań.
- Ramos P., Pindado P., 2013, *Monitoratge d'olors als polígons químics de Tarragona – control intern*, Tarragona.
- Ramos P., Pindado P., 2013, *Projecte de participació ciutadana- sistema de monitorització d'olors a Tarragona*, Tarragona.
- Rodríguez-Pose A., 2011, *Economists as geographers and geographers as something else: On the changing conception of distance in geography and economics*, *Journal of Economic Geography*, 11, 2, 347–356.
- Sui D.Z., 2008, *The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS*, *Computers, Environment and Urban Systems*, 32, 1–5.
- Sui D.Z., Elwood S., Goodchild M.F., 2012, *Volunteered Geographic Information, Public Participation, and Crowdsourced Production of Geographic Knowledge*, Springer, Berlin.
- <http://sabap2.adu.org.za/> (20.12.2014).

Kamil Jawgiel

Uniwersytet im. Adama Mickiewicza

Instytut Geografii Fizycznej i Kształtowania Środowiska Przyrodniczego

ul. Dąbkielowa 27, 61-680 Poznań

e-mail: jawka@amu.edu.pl



