

MIROŚLAW KRUKOWSKI

<https://orcid.org/0000-0002-4089-6329>

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

Wydział Nauk o Ziemi i Gospodarki Przestrzennej

al. Kraśnicka 2D, 20-718 Lublin, Polska

miroslaw.krukowski@umcs.pl

Niepewność, której nie mierzymy. Problem jakości danych w geoinformacji

The Uncertainty We Do Not Measure: The Problem of Data Quality in Geoinformation

Abstract: This article examines the problem of uncertainty in GIS data that extends beyond the standard technical parameters of the “big five”. It is argued that traditional reports overlook semantic and conceptual uncertainty, which is crucial for poorly-defined geographical objects. Using examples of marsh and forest definition changes, it demonstrates how a lack of conceptual context leads to erroneous analytical conclusions. The paper proposes extending metadata to include ontologies, institutional context, and the opacity of GeoAI models. The goal is to shift towards a holistic fit-for-purpose approach, integrating measurement precision with a reflection on the meaning of geographic information.

Keywords: spatial data quality; uncertainty of geographic information; vagueness; metadata

Abstrakt: W artykule analizie poddano problem niepewności w danych GIS, wykraczający poza standardowe parametry techniczne tzw. wielkiej piątki. Wskazano, że tradycyjne raporty pomijają niepewność semantyczną i koncepcyjną, co jest kluczowe dla obiektów słabo zdefiniowanych. Na przykładach zmian definicji bagien i lasów wykazano, jak brak kontekstu pojęciowego prowadzi do błędnych wniosków analitycznych. Praca zawiera postulat rozszerzenia metadanych o ontologie i kontekst instytucjonalny oraz uwzględnienie nieprzejrzystości modeli GeoAI. Celem jest przejście ku holistycznemu podejściu adekwatności do celu, integrującemu precyzję pomiaru z refleksją nad znaczeniem informacji geograficznej.

Słowa kluczowe: jakość danych geograficznych; niepewność informacji; nieostrość; metadane

WSTĘP

Jakość danych geoinformacyjnych zazwyczaj jest oceniana poprzez mierzalne parametry, takie jak dokładność położenia, kompletność czy spójność logiczna, które podlegają raportowaniu w metadanych oraz procedurach walidacyjnych (Chrisman, 1991; Federal Geographic Data Committee [FGDC], 1998; ISO, 2023). W praktyce dokumentacyjnej powszechnie utrwalił się zestaw pięciu kategorii (tzw. wielka piątka): dokładność pozycyjna, dokładność tematyczna, spójność logiczna, kompletność oraz rodowód (*lineage*). Model ten, zaproponowany pierwotnie w standardzie SDTS, znajduje szerokie zastosowanie u producentów danych w celu zapewnienia interoperacyjności procesów transferu zbiorów (Bielecka, 2010; Guptill i Morrison, 1995). Jednocześnie analiza literatury z zakresu jakości i niepewności danych przestrzennych wykazuje, że pole badawcze ulega dynamicznej ewolucji, obejmując nie tylko metryki techniczne, lecz także problematykę propagacji niedoskonałości w złożonych strukturach analitycznych i procesach decyzyjnych (Bielecka i Burek, 2019; Heuvelink, 1998).

Zasadniczy problem wynika z faktu, że niepewność pojmowana w ujęciu holistycznym – jako niepewność dotycząca semantyki, adekwatności oraz interpretacji informacji (Burrough, 1996; Duckham i in., 2001) – bywa w tradycyjnych systemach raportowania niedoreprezentowana. Stan ten jest determinowany trudnością w formalizacji tych aspektów i sprowadzeniu ich do pojedynczych miar statystycznych. Fisher (2003) określił tę sytuację metaforą „dwóch okrętów mijających się nocą”, wskazując na izolowany rozwój teoretycznych studiów nad niepewnością względem procesów standaryzacji jakości danych. W konsekwencji współczesne systemy informacji geograficznej (GIS – Geographic Information Systems) mogą nie dostarczać wyczerpującej wiedzy o immanentnych ograniczeniach zbiorów, zwłaszcza w zakresie niepewności konceptualnej i semantycznej związanej z definicją i interpretacją klas obiektów (Comber i in., 2005; Salgé, 1995). Zagadnienie to ma kluczowe znaczenie metodologiczne oraz epistemologiczne, gdyż bezpośrednio warunkuje poziom zaufania użytkownika do generowanych wyników (Agumya i Hunter, 2002).

W artykule podjęto próbę analizy powyższej problematyki poprzez identyfikację luk w obecnych praktykach standaryzacyjnych (m.in. w kontekście normy ISO 19157-1:2023) oraz systematyzację pojęcia niepewności informacji przestrzennej. Przedstawiono propozycję rozszerzenia koncepcji jakości danych o nieostrość (*vagueness*) oraz inne, dotychczas marginalizowane aspekty niepewności, które w modelowaniu kartograficznym odgrywają rolę równie istotną, co błędy pomiarowe (Krukowski, 2021). Poprzez połączenie podejścia technicznego z analizą ontologiczną sformułowano postulaty dotyczące kierunków rozwoju metadanych, które uwzględniają kontekst semantyczny informacji geograficznej oraz specyficzne potrzeby ich odbiorców.

WSPÓŁCZESNE RAPORTOWANIE JAKOŚCI DANYCH GEOINFORMACYJNYCH

Obecne metodyki dokumentowania jakości danych przestrzennych są ściśle sformalizowane przez standardy międzynarodowe oraz krajowe, które opisują jakość za pomocą rygorystycznego zestawu mierzalnych elementów technicznych (Guptill i Morrison, 1995; Veregin, 1999). Zestawy te obejmują przede wszystkim pięć podstawowych parametrów, zdefiniowanych pierwotnie w raporcie Moelleringa z 1987 roku dla Międzynarodowej Asocjacji Kartograficznej, które stanowią fundament oceny przydatności zbiorów danych (Moellering, 1987). Należą do nich:

- dokładność położenia – określa stopień zgodności współrzędnych obiektów zapisanych w bazie danych z ich rzeczywistą lokalizacją w przyjętym układzie odniesienia, uwzględnia błędy pomiarowe oraz błędy transformacji;
- dokładność atrybutowa – rozumiana jako poprawność cech opisowych przypisanych do obiektów przestrzennych, obejmuje zarówno błędy klasyfikacji cech jakościowych, jak i błędy pomiaru cech ilościowych;
- kompletność – informuje o stopniu nadmiarowości (obecności obiektów niepotrzebnych) lub braków w zbiorze danych w stosunku do obiektów, które powinny się w nim znajdować zgodnie ze specyfikacją;
- spójność logiczna – weryfikuje wewnętrzną poprawność struktury zbioru, w tym relacje topologiczne między obiektami oraz zgodność z regułami zdefiniowanymi w modelu danych;
- aktualność czasowa – opisuje zbieżność stanu informacji w bazie danych z rzeczywistym stanem zjawisk w świecie fizycznym w określonym momencie lub przedziale czasu.

Fundamenty dla tych rozwiązań położył amerykański komitet FGDC (komitet odpowiedzialny za standaryzację federalnych danych geograficznych w Stanach Zjednoczonych) poprzez opracowanie *Content Standard for Digital Geospatial Metadata* w 1998 r., który wprowadził hierarchiczną strukturę opisu danych opartą na rodowodzie oraz precyzji pomiarowej (FGDC, 1998). Najważniejszą rolę odgrywa tu obecnie ISO (International Organization for Standardization – Międzynarodowa Organizacja Normalizacyjna, tworząca ogólnosiwiatowe normy techniczne) wraz z najnowszą normą ISO 19157-1:2023, która dąży do unifikacji modeli opisu jakości poprzez wprowadzenie ujednoczonych procedur walidacyjnych oraz raportowania wyników w formie cyfrowych metadanych (ISO, 2023). Ewolucja tych norm pokazuje wyraźne przesunięcie od prostych opisów tekstowych w stronę złożonych schematów XML (*Extensible Markup Language* – rozszerzalny język znaczników służący do strukturalnego zapisu danych) oraz JSON (*JavaScript Object Notation* – format wymiany danych), które mają ułatwiać automatyczną wymianę informa-

cji między systemami klasy GIS, wciąż jednak pozostają one silnie zakorzenione w podejściu ilościowym.

W 1995 r. Salgé postulował już rozszerzenie tego instrumentarium o parametr dokładności semantycznej, który miał być miernikiem poprawności kategoryzacji obiektów względem ich rzeczywistego znaczenia (Salgé, 1995). Niemniej koncepcje te nie zyskały powszechnej akceptacji w głównym nurcie standaryzacji i rzadko znajdują odzwierciedlenie w krajowych profilach metadanych, co wynika w znacznej mierze z wysokiego stopnia złożoności kwantyfikacji znaczenia oraz trudności w formalizacji uniwersalnych ontologii (Devillers i Jeansoulin, 2006). Instytucje normalizacyjne skoncentrowały swoje działania głównie na certyfikacji procesów produkcyjnych, co faworyzuje obiektywne parametry techniczne kosztem głębokiej ewaluacji semantycznej oraz utrwała model zorientowany wyłącznie na producenta informacji. Współczesne metadane pełnią więc rolę certyfikatu zgodności technologicznej i tarczy prawnej dla dostawców danych, często marginalizując realne potrzeby interpretacyjne użytkownika końcowego oraz generując lukę informacyjną, która uniemożliwia rzetelną ocenę przydatności danych do realizacji specyficznych celów analitycznych (Comber i in., 2006). Istniejące ramy raportowania skupiają się na tym, jak precyzyjnie wykonano pomiar, zamiast informować o tym, jak wiernie model oddaje naturę reprezentowanych zjawisk, co staje się szczególnie problematyczne przy integracji zasobów pochodzących z różnych instytucji.

Dominacja parametrów technicznych powoduje, że standardowe raporty jakości ignorują niepewność wynikającą z niedoskonałości wiedzy oraz nieostrości pojęć geograficznych, co zostanie szczegółowo omówiono w kolejnej części pracy. Język norm technicznych operuje w koncepcji ostrych map binarnych i klasyfikacji rozłącznych, gdzie każdy obiekt jest uznawany albo za w pełni obecny, albo za całkowicie nieobecny. Pomimo wprowadzenia w normie ISO 19157-1:2023 większej swobody w definiowaniu elementów jakości opisowej, raportowanie wciąż opiera się na rutynowym wypełnianiu wymogów formalnych, które nie tłumaczy użytkownikowi praktycznych implikacji prezentowanych wartości (Vukalić i in., 2024). Brak informacji o kontekście powstania danych oraz o koncepcjach leżących u podstaw modelowania sprawia, że użytkownik dysponuje zbiorem o wysokiej jakości geometrycznej, który może być jednak nieadekwatny pojęciowo do realizowanego zadania (Chrisman, 1991; Comber i in., 2004). Ostatecznie obecny stan standaryzacji promuje ułudę pełnej kontroli nad jakością poprzez statystyczne wskaźniki dokładności, natomiast w rzeczywistości pomija on fundamentalne wyzwania epistemologiczne związane z komunikacją wiedzy o niepewnej i złożonej przestrzeni geograficznej.

ISTOTA NIEPEWNOŚCI W DANYCH GEOINFORMACYJNYCH

Przed przystąpieniem do analizy typologii niepewności warto odróżnić dwa porządki pojęciowe, które w praktyce GIS często są ze sobą utożsamiane, czyli jakość danych (*data quality*) oraz niepewność informacji (*information uncertainty*). Jakość danych opisuje cechy produktu danych raportowane w metadanych, takie jak: dokładność, kompletność, spójność czy rodowód, a także odpowiada na pytanie o to, jak wykonano i jak technicznie opisano dane. Natomiast niepewność informacji odnosi się do stopnia uzasadnienia i wiarygodności twierdzeń, jakie formułujemy na podstawie danych, a więc do wątpliwości dotyczących znaczenia, adekwatności i konsekwencji wnioskowania. W tym sensie raportowana jakość jest tylko częścią niepewności, co oznacza, że może ją ograniczać, ale nie wyczerpuje jej źródeł, zwłaszcza gdy problem wynika z nieostrych definicji i konceptualizacji obiektów (Krukowski, 2021).

Niepewność informacji geograficznej można zdefiniować jako wątpliwość co do prawdziwości lub adekwatności informacji o świecie (Krukowski, 2021). Ma ona wiele przejawów i przyczyn, a klasycznie dzieli się ją na dwa główne rodzaje w zależności od charakteru obiektu geograficznego (Fisher, 1999), czyli na niepewność dla obiektów dobrze zdefiniowanych oraz niepewność dla obiektów słabo zdefiniowanych.

Obiekty dobrze zdefiniowane (*well-defined*) to takie, których granice oraz kryteria klasyfikacji są względnie jednoznaczne, jak np. budynek, działka ewidencyjna czy odcinek drogi w określonej klasie technicznej. W ich przypadku dominują źródła niepewności o charakterze klasycznym, obejmujące błędy pomiaru, georeferencji i próbkowania, a więc odchylenia dające się opisywać probabilistycznie oraz poddawać estymacji w ramach teorii błędów. Tak rozumiana niepewność przekłada się przede wszystkim na dokładność pozycyjną i atrybutową oraz na ocenę stabilności wyników analiz poprzez propagację niepewności. W tej klasie problemów metadane jakości, jeśli są poprawnie prowadzone, faktycznie umożliwiają użytkownikowi przejście od danych do wiarygodności wnioskowania (Fisher, 2000).

Inaczej jest w przypadku obiektów słabo zdefiniowanych (*poor-defined*), czyli takich, których granice są rozmyte, a kryteria wydzielenia mają charakter kontekstowy, gradacyjny albo zależą od przyjętego schematu pojęciowego. Przykładami mogą być mokradła, tereny zabudowane, strefy przejściowe ekosystemów czy niektóre jednostki geologiczne. Dla tej klasy obiektów kluczowe stają się dwa rodzaje niepewności, czyli nieostrość oraz niejednoznaczność. Nieostrość dotyczy sytuacji, w której nie da się wskazać ostrej granicy przynależności, ponieważ obiekt zanika w otoczeniu, a jego zasięg ma charakter strefowy. W konsekwencji przynależność do klasy jest stopniowalna. Niejednoznaczność natomiast oznacza,

że ta sama sytuacja może zostać poprawnie opisana na więcej niż jeden sposób, ponieważ istnieją konkurencyjne lub równoległe interpretacje, takie jak różne definicje klas czy różne kryteria instytucjonalne.

Dobrym przykładem geograficznym jest rozróżnienie miasto–wieś, które staje się problematyczne w strefach suburbanizacji. Płynne zmiany gęstości zabudowy i struktury użytkowania terenu sprawiają, że wyznaczenie jednej ostrej linii demarkacyjnej jest aktem arbitralnym. Jak wskazują Smith i Mark (2003) w swojej analizie ontologii form terenu, nawet obiekty o pozornie oczywistej fizyczności, takie jak góry, rzadko posiadają granice typu *bona fide* wynikające z realnych nieciągłości w świecie. Większość obiektów geograficznych to konstrukty typu *fiat*, których granice są wyznaczane umownie, co czyni je nieostrymi, niezależnie od precyzji pomiaru. W konsekwencji przynależność do klasy obiektów pozostaje stopniowalna, a zasięg obiektu ma charakter rozmyty i strefowy (Krukowski, 2025). Dodatkowym czynnikiem pogłębiającym nieostrość i niejednoznaczność obiektów geograficznych jest zmienność czasowa i przestrzenna, która podważa założenie o ich stabilności ontologicznej. W wielu przypadkach granice obiektów nie są jedynie rozmyte, lecz także dynamiczne, ulegają bowiem okresowym lub długoterminowym przekształceniom. Dotyczy to m.in. sezonowych zmian zasięgu jezior i mokradeł, fluktuacji rozmieszczenia zwierząt związanych z migracjami, a także przekształceń stref klimatycznych i ekosystemów w warunkach zmian klimatu. Zmienność ta utrudnia jednoznaczne definiowanie i delimitację obiektów, ponieważ nawet precyzyjnie sformułowane kryteria klasyfikacyjne mogą być spełnione tylko w określonych momentach czasu lub w sposób nieciągły. W konsekwencji niepewność informacji nie wynika wyłącznie z nieostrości pojęć, lecz również z czasowej nietrwałości samych obiektów, co dodatkowo komplikuje proces wnioskowania na podstawie danych przestrzennych.

W obrębie niejednoznaczności wyróżnia się dwa uściślenia, które pomagają rozdzielić niezgodność oraz nieokreśloność. Niezgodność zachodzi wtedy, gdy obiekt jest względnie ostry, ale w zależności od przyjętego systemu klasyfikacyjnego może zostać przypisany do różnych klas, przy czym każda z tych decyzji jest spójna w ramach własnej definicji. Przykładem może być klasyfikacja form terenu, gdzie ten sam obiekt może zostać nazwany górą lub wzgórzem w zależności od tego, czy przyjęty próg wysokości względnej wynosi np. 200 czy 300 metrów. Nieokreśloność (niekiedy określana jako niespecyficzność) dotyczy zaś samej definicji klasy, ponieważ jeśli jest ona ogólna, elastyczna lub mocno kontekstowa, to przypisanie obiektu pozostaje interpretacyjne. Przykładowo kiedy termin teren podmokły nie precyzuje dokładnego stopnia uwilgotnienia ani czasu jego utrzymywania, decyzja o zaklasyfikowaniu danego obszaru staje się subiektywna. W konsekwencji ta sama informacja przestrzenna może być formalnie poprawna,

a jednak epistemicznie niepewna, gdyż problem leży nie w pomiarze, tylko w pojęciach i decyzjach klasyfikacyjnych (Krukowski, 2021).

1. Formalizacja

W GIScience wypracowano szereg formalnych ujęć niepewności, które nie redukują jej do błędu pomiaru. Kluczowe znaczenie ma tu rozróżnienie między niedokładnością a nieostrością, gdzie pierwsza dotyczy odchylenia danych od referencji (błąd), a druga niejednoznacznych kryteriów klasyfikacji oraz przypadków granicznych wynikających z języka i konceptualizacji świata. W konsekwencji modelowanie nieostrości jest próbą zoperacjonalizowania tego, co w opisie geograficznym jest z natury stopniowalne i nieostre.

Najbardziej klasycznym narzędziem formalizacji nieostrości są zbiory rozmyte (*fuzzy sets*), w których przynależność do klasy ma charakter stopniowalny i jest opisywana funkcją przynależności (Zadeh, 1965). Zamiast ostrej granicy należy / nie należy, otrzymujemy kontinuum stopni bardziej–mniej, co pozwala modelować obszary przejściowe bez udawania, że są jednoznaczne. Komplementarne podejście stanowią zbiory przybliżone (*rough sets*), w których obiekt opisuje się przez przybliżenie dolne, obejmujące elementy pewne, oraz przybliżenie górne, obejmujące elementy możliwe, dzięki czemu da się wprost zapisać strefę niepewności przynależności (Pawlak, 1982). Tę logikę można traktować jako formalne ujęcie sytuacji, w której granica klasy jest niepewna, ale nie musi być probabilistyczna.

Równolegle rozwijają się ujęcia filozoficzne porządkujące źródło nieostrości. Kluczowe jest tu rozróżnienie między nieostrością *de re* (dotyczącą fizycznej natury rzeczy) a nieostrością *de dicto* (wynikającą z niedoskonałości języka i definicji). Nieostrość *de re* obserwujemy tam, gdzie same obiekty fizyczne pozbawione są wyraźnych krawędzi. Klasycznym przykładem jest linia brzegowa, która ze względu na dynamikę pływów i erozji stanowi obiekt o nieostrości fizycznej, a jej reprezentacja jako ostrej kreski w bazie danych jest jedynie kognitywnym uproszczeniem. Z kolei nieostrość *de dicto* pojawia się w procesach mapowania, gdy próbujemy narzucić nazwy (etykiety) na ciągłą rzeczywistość (Krukowski, 2025). Z tym wiąże się rozróżnienie granic *bona fide* (opartych na fizycznych nieciągłościach) i *fiat* (ustanowionych umownie).

2. Wymiar kognitywny i językowy

Istotnym źródłem niepewności jest fakt, że pojęcia geograficzne są zakorzenione w języku naturalnym i kategoryzacji potocznej. Dyskretyzowanie świata do struktur bazodanowych wymusza arbitralne progi, co generuje przypadki granicz-

ne. Można się tu odwołać do paradoksu stosu (*Sorites*), aby pokazać, że nie da się precyzyjnie wyznaczyć momentu, w którym pojedyncze drzewa stają się lasem (Krukowski, 2021). Z punktu widzenia semantyki nie jest to błąd procesu, tylko strukturalna cecha opisu świata. W praktyce oznacza to, że twórca danych może spełnić kryteria dokładności geometrycznej, a mimo to pozostawić użytkownika z ryzykiem interpretacyjnym. Niepewność nie dotyczy wtedy tego, gdzie coś jest, lecz czym coś jest i według jakich kryteriów zostało wydzielone. Ten wymiar, obejmujący nieostrość poznawczą i nieostrość konceptualizacji, powinien być traktowany jako integralny element jakości danych.

ROZSZERZENIE KONCEPCJI JAKOŚCI DANYCH – INTEGRACJA NIEPEWNOŚCI I NIEOSTROŚCI

Dla rozwiązania zidentyfikowanego problemu zasadne wydaje się poszerzenie ram pojęciowych jakości danych geograficznych w taki sposób, aby włączały one niepewność w najszerszym sensie, obejmującą nieostrość i niejednoznaczność obok tradycyjnych miar dokładności. Taka propozycja ma charakter metodologiczno-epistemologiczny, co oznacza, że chodzi nie tylko o nowe techniki pomiarowe, ale przede wszystkim o zmianę sposobu myślenia o jakości danych

W pierwszej kolejności warto rozważyć modyfikację standardów metadanych tak, aby uwzględniały informacje o konceptualizacji danych. Metadane mogłyby opisywać znaczenie i definicje kategorii użytych w zbiorze danych oraz kontekst ich tworzenia zamiast jedynie podawać dokładność geometryczną. Korzystne wydaje się dodanie do dokumentacji danych elementów, które obejmują:

1. Kontekst organizacyjno-ekspercki – opis celów powstania danych oraz założeń przyjętych przy definiowaniu obiektów. Pozwala ujawnić epistemologiczne uwarunkowania informacji, np. fakt, że dane powstały na potrzeby inwentaryzacji gruntów rolnych według określonych wytycznych urzędowych z lat 90. XX w., co bezpośrednio wpływa na sposób klasyfikacji.
2. Kontekst polityki i zamówienia – informacje o uwarunkowaniach powstania danych, np. „klasa ‘bagno’ zdefiniowana na potrzeby ochrony mokradł według kryteriów hydrologicznych”. Tłó takie wyjaśnia, dlaczego dane przybrały taką formę, oraz pozwala zrozumieć różnice między pozornie podobnymi zbiorami.
3. Ontologia i definicje obiektów – formalny lub opisowy słownik pojęć użytych w danych, zawierający definicje każdej klasy tematycznej i ewentualnie różnice względem innych znanych klasyfikacji. Taki opis (np.

odwołanie do ontologii dziedzinowej czy standardu definicji) pozwoli użytkownikowi ocenić, na ile pojęcia w danych odpowiadają jego własnemu rozumieniu tych pojęć lub innym źródłom danych. W praktyce może to przyjąć formę rozszerzonego słownika metadanych lub załącznika z definicjami.

4. Informacja o doświadczeniach użytkowników – jeżeli dostępne są dane o wcześniejszym zastosowaniu danego zestawu danych (np. publikacje, raporty, w których go użyto), warto w metadanych zamieścić uwagi dotyczące zastosowań. Comber i in. (2004) sugerują, że metadane mogłyby zawierać relacje między jakością danych a niepewnością zaobserwowaną w konkretnych zastosowaniach, np. „analiza zasięgu bagien z użyciem tych danych wykazała rozbieżności na glebach torfowych – użytkownicy zgłaszali, że klasa ‘bagno’ pomija obszary bez wody stojącej”. Takie informacje pełniłyby rolę wskazówek dla kolejnych użytkowników co do ograniczeń danych.

Rozszerzenie metadanych o powyższe elementy zbudowałoby pomost pomiędzy wiedzą producenta danych a potrzebami ich użytkowników. W literaturze podkreśla się konieczność wzmocnienia powiązań między raportowaniem jakości a oceną niepewności (Comber i in., 2004; Fisher, 2003). Proponowane rozwiązania wpisują się w ten nurt poprzez wyposażenie metadanych w wiedzę konceptualną, dzięki czemu użytkownik uzyskuje wiedzę nie tylko o dokładności technicznej, lecz również o rzeczywistym znaczeniu informacji, co ułatwia ocenę przydatności danych do konkretnego celu.

W drugiej kolejności za celowe uznaje się opracowanie spójnej typologii niepewności informacji geograficznej, która umożliwi badaczom i praktykom posługiwanie się wspólnym językiem przy opisie jakości danych. Ujednolicona ontologia pojęć mogłaby wyeliminować chaos terminologiczny wokół takich terminów jak: „niedoskonałość”, „niejednoznaczność”, „niespecyficzność” czy „nieokreśloność” i wiele innych. Dążenie do wyraźnego rozgraniczenia rodzajów niepewności pozwala na powiązanie ich z odpowiednimi metodami postępowania. Zaproponowano już pewne porządki, np. klasyfikacje Klira i Yuana (1995) czy Fishera (1999) dzielące niepewność według źródeł (błąd, niejasność, brak danych, konflikt informacji itp.). Krukowski (2021) zaproponował nową typologię niedoskonałości informacji przestrzennej, opartą o analizę filozoficzną i istniejące koncepcje. Dążenie do wyraźnego rozgraniczenia rodzajów niepewności pozwala na powiązanie ich z odpowiednimi metodami postępowania. W praktyce mogłoby to przybrać formę tabelarycznego schematu zawierającego rodzaj niepewności, jego opis oraz przykładową miarę lub technikę modelowania, co prezentuje tab. 1.

Tab. 1. Przykładowe typy niepewności informacji geograficznej i możliwe podejścia do ich kwantyfikacji (opracowanie własne na podstawie: Fisher, 1999; Krukowski, 2021)

Tab. 1. Example types of uncertainty in geographic information and possible approaches to their quantification (author's own elaboration based on Fisher, 1999; Krukowski, 2021)

Rodzaj niepewności	Opis	Przykładowe ujęcie/miara
Błąd	Różnica między danymi a rzeczywistością wynikająca z ograniczeń pomiaru lub modeli. Dotyczy obiektów „dobrze zdefiniowanych”	Statystyczne miary dokładności (RMSE, odchylenie standardowe, przedziały ufności)
Niekompletność	Braki danych lub niepełne pokrycie obiektów danej kategorii	Wskaźnik kompletności (odsetek brakujących obiektów) zgodnie ze standardem (ISO 19157)
Nieostrość (rozmytość)	Rozmyte granice obiektu lub stopniowalna przynależność do klasy. Typowe dla obiektów słabo zdefiniowanych	Zbiory rozmyte i funkcja przynależności w domenie od 0 do 1 dla każdego punktu przestrzeni (np. model rozmytych granic zasięgu)
Niejednoznaczność	Wieloznaczność przypisania obiektu do klas oraz różne możliwe interpretacje tej samej sytuacji	Analiza w oparciu o scenariusze lub teoria dowodów – przydzielenie wag prawdopodobieństwa do alternatywnych klasyfikacji obiektu (metoda Dempstera–Shafera)
Niezgodność	Sprzeczność między dwiema klasyfikacjami, gdzie obiekt spełnia kryteria więcej niż jednej klasy w różnych systemach	Macierz zgodności między dwiema klasyfikacjami; miara podobieństwa semantycznego klas (np. współczynnik <i>kappa</i> lub <i>F-score</i>)
Nieokreśloność/niespecyficznosc	Sytuacja, w której zbyt ogólna definicja klasy powoduje niepewność przypisania obiektu do konkretnej kategorii	Dokumentacja zakresu interpretacji lub zastosowanie zbiorów przybliżonych, podających granice pewnego przypisania
Nieprzejrzystość (<i>black-box</i>)	Brak wiedzy o procesie powstania danych, wynikający np. z użycia niejawnego modelu sztucznej inteligencji generującego dane	Informacja o rodowodzie i metodzie uzyskania danych oraz wskaźniki zaufania generowane przez model

Zestawienie zamieszczone w tab. 1 stanowi jedynie ilustrację postulowanych zmian. Istotne wydaje się, aby w ramach oceny jakości danych dążyć do jawnego raportowania tych kategorii niepewności, które dotąd były pomijane. Użytkownik mógłby dowiedzieć się nie tylko o wartościach błędów lokalizacji, lecz także o specyfice definicji klas, co pozwoliłoby na uzyskanie pełniejszego obrazu ograniczeń informacji. Przykładem może być sytuacja, w której definicja klasy ‘las’ opiera się wyłącznie o kryterium powierzchni z koroną drzew, co oznacza, że młode zalesienia mogły zostać nieuwjęte w zbiorze, co stanowi specyficzny rodzaj nieokreśloności definicji. Taki opis jakości, łączący aspekt techniczny z konceptualnym, daje pełniejszy obraz ograniczeń informacji.

W trzeciej kolejności proponowane podejście przewiduje wdrożenie praktycznych metod uwzględniania niepewności w analizach i prezentacji danych. Same metadane to nie wszystko – niepewność trzeba komunikować i wykorzystywać. W geoinformacji istnieje nurt badań nad wizualizacją niepewności (Kinkeldey i in., 2014; MacEachren i in., 2005), postulujący używanie zmiennych wizualnych i sonifikacji (np. przezroczystości, rozmycia lub dźwięku) do przedstawiania nie-

pewnych informacji na mapach w różnych systemach informacyjnych (np. mapy pokazujące zarówno wartość atrybutu, jak i niepewność co do tej wartości). Inne możliwe podejścia obejmują analizy scenariuszowe, polegające na generowaniu wielu realizacji danych i sprawdzaniu stabilności uzyskiwanych wyników (Agumya i Hunter, 2002; Heuvelink, 1998). Z kolei Comber i in. (2006) sugerują integrację narzędzi GIS z mechanizmami wnioskowania z niepewną wiedzą, by wspomagać użytkownika w podejmowaniu decyzji z uwzględnieniem niepewności. Z perspektywy metodologicznej oznacza to przesunięcie paradygmatu od założenia o jednoznaczności wyników do podejścia, w którym wynik posiada określony rozkład wiarygodności, a dane niosą za sobą przedział możliwych interpretacji. Wychodzimy zatem od założenia, że „mamy dokładne dane, więc wynik jest jednoznaczny”, do podejścia, że „wynik ma pewien rozkład wiarygodności, a dane niosą przedział możliwych interpretacji”.

PRZYKŁAD PROBLEMU NIEPEWNOŚCI INFORMACJI GEOGRAFICZNEJ

Problem niepewności semantycznej, rozumianej jako brak jednoznaczności w definiowaniu kategorii obiektów geograficznych, znajduje odzwierciedlenie w badaniach nad dynamiką zmian pokrycia terenu. Za jedną z głównych przyczyn błędnych interpretacji danych przestrzennych uznaje się zjawisko ukrywania się istotnych zmian definicyjnych pod tymi samymi etykietami klas. W kartografii i geoinformacji obok niepewności pomiarowej istnieje fundamentalna niepewność pojęciowa, która wynika z prób binarnego opisu zjawisk o charakterze ciągłym.

Klasycznym przykładem ilustrującym to zjawisko jest analiza przeprowadzona przez Fullera i współpracowników dotycząca brytyjskich map pokrycia terenu z lat 1990 i 2000 (Fuller i in., 1994, 2002). Wykazano w niej, że zmiana conceptualizacji klasy ‘bagno’ doprowadziła do pozornego i ogromnego przyrostu tego typu obszarów. W wersji mapy z 1990 r. stosowano kryteria florystyczno-hydrologiczne, oparte na obecności wody i roślinności bagiennej, co skutkowało sklasyfikowaniem zaledwie 12 pikseli zdjęcia satelitarnego jako bagniste. Dziesięć lat później przyjęto definicję glebową opartą wyłącznie na głębokości torfu przekraczającej 0,5 metra. W efekcie, bez fizycznej zmiany w terenie, powierzchnia bagien na mapie wzrosła z obszaru mniejszego niż jeden hektar do wartości 75 kilometrów kwadratowych (Fuller i in., 1994, 2002). Jak zauważają Comber i in. (2006), bez dostępu do metadanych semantycznych użytkownik jest skazany na błędny wniosek o radykalnej regeneracji ekosystemu, podczas gdy w rzeczywistości wystąpił jedynie akt zmiany definicji.

Innym, ale podobnym przykładem jest klasyfikacja terenu oraz określenie zasięgu przestrzennego lasu, który zmienia się radykalnie w zależności od przyjętego progu gęstości drzew. W szerokim rozumieniu klasa stanowi zbiór obiektów o jednakowych właściwościach, przy czym w procesie klasyfikacji często bezpodstawnie przyjmuje się, że obiekt posiada wartości typowe dla elementów danej kategorii i stanowi element wzorcowy. W kognitywizmie taki idealny wzorzec określa się mianem prototypu (Tabakowska, 2001), a jego istnienie powoduje efekt prototypu dodatkowo komplikujący proces kategoryzacji (Rosch, 1978). Nie wszystkie obiekty w danej klasie mają jednakowe wartości cech powszechnie akceptowane, co sprawia, że niektóre z nich znajdują się na skraju uznania za elementy zbioru. Sytuacja ta opiera się na domyślnym rozumowaniu poprzez porównanie obiektu klasyfikowanego z przyjętym wcześniej wzorcem, co niesie za sobą niepewność wyniku.

W polskich uwarunkowaniach las występuje w różnych znaczeniach, w zależności od celu definicji. Według ustawy o lasach z 1991 r. „las to grunt o zwartej powierzchni co najmniej 10 arów, pokryty roślinnością leśną lub teren przejściowo tej roślinności pozbawiony” (Ustawa, 1991). Z kolei rozporządzenie dotyczące bazy danych obiektów topograficznych BDOT10k określa las jako „teren leśny lub zadrzewiony, który reprezentuje tereny o zwartym zadrzewieniu, w tym lasy, zadrzewienia, zagajniki, parki oraz inne tereny porośnięte drzewami”, wprowadzając jednocześnie wymóg minimalnej szerokości obiektu wynoszący 20 metrów, długości 50 metrów, a minimalna powierzchnia to 2000 metrów kwadratowych (Rozporządzenie, 2021). Równolegle funkcjonują międzynarodowe standardy FAO (Food and Agriculture Organization of the United Nations), według których „lasem jest teren o powierzchni powyżej pół hektara, z drzewami o wysokości co najmniej 5 metrów i zwarciu koron przekraczającym 10%” (FAO, 2020). W tych definicjach nieostre pozostaje zwłaszcza kryterium zwartości lasu, co generuje niejednoznaczność przy próbach łączenia lub porównywania zbiorów danych pochodzących z różnych państw (Lund, 2018). Interesujące porównanie stanowi przedwojenna instrukcja Wojskowego Instytutu Geograficznego z 1925 r. wskazująca granicę lasu: „(...) tam, gdzie las stopniowo przechodzi w bardzo rzadki, należy przyjąć granicę jego na tym miejscu, gdzie przestaje on chronić przed obserwacją z góry” (WIG, 1925), co podkreśla orientację definicji na cele militarne.

Problemem wynikającym z istnienia nieostrości i niejednoznaczności jest wyciąganie bezkrytycznych wniosków na podstawie analiz przestrzennych, ignorując przy tym fakt, że funkcjonalność systemów informacyjnych zależy od różnic w klasyfikacjach stosowanych podczas gromadzenia i wykorzystywania danych. Zgodnie z opracowaniem Franka (2001) niepewność w klasyfikacji wynika nie tylko z błędu pomiaru, lecz przede wszystkim z wyboru cech istotnych, wyboru wartości progowych oraz błędów położenia granicy w przestrzeni i czasie. Wnioski

te wskazują, że dokumentacja jakości nie może ograniczać się do precyzji geometrycznej, ponieważ to niepewność co do znaczenia pojęcia decyduje o końcowej wiarygodności analiz geograficznych.

PODSUMOWANIE

Analiza problematyki jakości w systemach informacji przestrzennej prowadzi do wniosku, że dominujący dotychczas paradygmat, oparty na tzw. wielkiej piątce parametrów technicznych, obejmującej dokładność pozycyjną, dokładność tematyczną, spójność logiczną, kompletność oraz rodowód, jest niewystarczający w obliczu rosnącej złożoności współczesnych analiz. Tradycyjne metadane, choć precyzyjnie raportują błędy pomiarowe czy rodowód danych, milczą na temat najważniejszego źródła niepewności, jakim jest rozbieżność między konceptualizacją rzeczywistości przez producenta a potrzebami i wiedzą użytkownika. Skupienie się wyłącznie na mierzalnych parametrach pomija wymiar niepewności epistemicznej, związanej z tym, co i jak dane reprezentują. W geoinformacji, będącej pomostem między światem realnym a jego modelami, ignorowanie tej sfery prowadzi do nadmiernej pewności siebie w interpretacji wyników, co można podsumować stwierdzeniem, że wiemy, iż nic nie wiemy dokładnie.

Współczesne kierunki standaryzacji, reprezentowane przez normę ISO 19157-1:2023 (ISO, 2023), zdają się dostrzegać tę lukę. Nowy standard wprowadza większą elastyczność, definiując proces tworzenia dodatkowych, dziedzinowych komponentów jakości, co stanowi szansę na formalne włączenie do raportowania aspektów semantycznych i ontologicznych. Zamiast ograniczać się do statystycznej poprawności, systemy metadanych mogą być uzupełniane o sformalizowane definicje pojęć, np. w formie ontologii, co pozwala na realne wdrożenie paradygmatu adekwatności do zamierzonego celu (*fitness for purpose*). Dzięki temu ocena przydatności danych przestaje być subiektywnym domysłem użytkownika, a staje się procesem opartym na porównywaniu sformalizowanych uniwersów dyskursu (Vukalić i in., 2024). Prace nad uaktualnieniem standardów zmierzają ku lepszej interoperacyjności, wprowadzając procedury opisu spójności konceptualnej, które pozwalają systemom ontologicznym wykrywać, że np. klasa ‘bagno’ w zbiorze A nie jest tożsama z klasą ‘bagno’ w zbiorze B.

Kwestia ta staje się jeszcze bardziej paląca w dobie dynamicznego rozwoju geoprzestrzennej sztucznej inteligencji (GeoAI). Współczesne modele uczenia maszynowego i sztucznej inteligencji obiecują automatyzację procesów klasyfikacji i ekstrakcji obiektów, ale kwestia ta staje się szczególnie istotna w dobie ich dynamicznego rozwoju (Mai i in., 2025). Masowe wykorzystanie modeli głębo-

kiego uczenia do klasyfikacji pokrycia terenu wprowadza nowy rodzaj niepewności, określanej jako nieprzejrzystość czarnej skrzynki (*black-box uncertainty*; zob. Hüllermeier i Waegeman, 2021). Modele te potrafią osiągać wysoką dokładność geometryczną (Ali, 2025), niosą jednak za sobą ryzyko ukrycia niepewności semantycznej pod wysokimi wskaźnikami ufności statystycznej, podczas gdy mechanizm ich decyzji klasyfikacyjnych oraz granice semantyczne często pozostają niejawne.

W tym kontekście problem nieprzejrzystości modeli łączy się bezpośrednio z wyzwaniem, które Goodchild i Li (2021) określają jako specyfikę replikacji w przestrzeni. W przeciwieństwie do nauk eksperymentalnych, gdzie replikacja w czasie jest standardem, w geoinformacji mamy do czynienia z wysoką heterogenicznością przestrzenną. Oznacza to, że modele trenowane w jednym regionie mogą generować wyniki obciążone głęboką niepewnością kontekstową przy próbie ich transferu na inne obszary. Bez kwantyfikacji tej niepewności oraz ujawnienia „logiki pojęciowej” algorytmu wysoki wskaźnik jakości technicznej może jedynie maskować fundamentalne błędy semantyczne i uprzedzenia geograficzne (Mai i in., 2025). Integracja tradycyjnych miar statystycznych z metodami uczenia maszynowego świadomego niepewności (Valle i in., 2023) oraz wykorzystanie metod wyjaśnialnej sztucznej inteligencji (XAI; zob. Dahal i Lombardo, 2023) stanowią jedyną drogę do budowania zaufania w erze zautomatyzowanej analizy.

Z metodologicznego punktu widzenia konieczna jest zmiana postrzegania GIScience – nie tylko jako zbioru technik, lecz także jako dyscypliny epistemologicznej. Taka perspektywa integruje techniczne aspekty jakości z refleksją nad naturą informacji geograficznej. Świadomość istnienia niepewności informacji, a zwłaszcza jej nieostrości, powinna być powszechna, aby uniknąć częstych nadinterpretacji danych przestrzennych. Konieczne jest budowanie kultury geoinformacyjnej, w której pytanie o granice wiarygodności danych będzie rutynowe. Powinniśmy pytać nie tylko o to, jaka jest dokładność, lecz także o to, na ile ufamy definicjom i założeniom stojącym za danymi.

Podsumowując, jakość danych przestrzennych to problem wielowymiarowy, w którym wymiar niepewności konceptualnej jest równie istotny jak dokładność geometryczna. Rozwiązaniem jest holistyczne podejście – rozszerzenie metadanych o kontekst semantyczny, transparentność algorytmiczna GeoAI oraz ewolucja programów kształcenia geoinformacyjnego i geograficznego w stronę krytycznego myślenia o danych. Tylko wtedy, gdy producenci danych poddadzą rewizji swoje podejście do jakości, uwzględniając nowoczesne modele niepewności, standardy staną się kompletne i użyteczne. W geoinformacji precyzja pomiaru musi iść w parze z refleksją poznawczą, aby mapa i model wiernie służyły zrozumieniu świata, a nie ułudzie pewności.

BIBLIOGRAFIA

- Agumya, A., Hunter, G.J. (2002). Responding to the Consequences of Uncertainty in Geographical Information. *International Journal of Geographical Information Science*, 16(5), 405–417. DOI: <https://doi.org/10.1080/13658810210137031>
- Ali, R. (2025). Artificial Intelligence for Land Cover and Land Use Classification in Remote Sensing: Review Study. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-G-2025, 115–122. DOI: <https://doi.org/10.5194/isprs-archives-XL-VIII-G-2025-115-2025>
- Bielecka, E. (2010). *Metodyka oceny jakości bazy danych obiektów topograficznych*. Warszawa: Wojskowa Akademia Techniczna.
- Bielecka, E., Burek, E. (2019). Spatial Data Quality and Uncertainty Publication Patterns and Trends by Bibliometric Analysis. *Open Geosciences*, 11(1), 219–235. DOI: <https://doi.org/10.1515/geo-2019-0018>
- Burrough, P.A. (1996). Natural Objects with Indeterminate Boundaries. W: P.A. Burrough, A.U. Frank (Eds.), *Geographic Objects with Indeterminate Boundaries* (s. 3–28). London: Taylor and Francis. DOI: <https://doi.org/10.1201/9781003062660>
- Chrisman, N.R. (1991). The Error Component in Spatial Data. W: D.J. Maguire, M.F. Goodchild, D.W. Rhind (Eds.), *Geographical Information Systems: Principles and Applications* (s. 165–174). London: Longman.
- Comber, A., Fisher, P., Wadsworth, R. (2004). Integrating Land-Cover Data with Different Ontologies: Identifying Change from Inconsistency. *International Journal of Geographical Information Science*, 18(7), 691–708. DOI: <https://doi.org/10.1080/13658810410001705316>
- Comber, A.J., Fisher, P.F., Wadsworth, R.A. (2005). What Is Land Cover? *Environment and Planning B: Urban Analytics and City Science*, 32(2), 5229–5236. DOI: <https://doi.org/10.1068/b31135>
- Comber, A.J., Fisher, P.F., Harvey, F., Gahegan, M., Wadsworth, R.A. (2006). Using Metadata to Link Uncertainty and Data Quality Assessments. W: *Progress in Spatial Data Handling: Proceedings of the 12th International Symposium on Spatial Data Handling* (s. 279–292). Berlin: Springer. DOI: https://doi.org/10.1007/3-540-35589-8_18
- Dahal, A., Lombardo, L. (2023). Explainable Artificial Intelligence in Geoscience: A Glimpse into the Future of Landslide Susceptibility Modeling. *Computers & Geosciences*, 176. DOI: <https://doi.org/10.1016/j.cageo.2023.105364>
- Devillers, R., Jeansoulin, R. (Eds.). (2006). *Fundamentals of Spatial Data Quality*. ISTE. DOI: <https://doi.org/10.1002/9780470612156>
- Duckham, M., Mason, K., Stell, J., Worboys, M.F. (2001). A Formal Approach to Imperfection in Geographic Information. *Computers, Environment and Urban Systems*, 25(1), 89–103. DOI: [https://doi.org/10.1016/S0198-9715\(00\)00040-5](https://doi.org/10.1016/S0198-9715(00)00040-5)
- FAO (2020). *Global Forest Resources Assessment 2020: Terms and Definitions*. FAO. Online: <https://www.fao.org/3/i8661en/i8661en.pdf> (dostęp: 13.11.2025).
- Federal Geographic Data Committee (FGDC). (1998). *Content Standard for Digital Geospatial Metadata*. FGDC.
- Fisher, P.F. (1999). Models of Uncertainty in Spatial Data. W: D.J. Maguire, M.F. Goodchild (Eds.), *Geographical Information Systems: Principles, Techniques, Applications, and Management* (s. 191–205). New York: John Wiley & Sons.
- Fisher, P.F. (2000). Sorites Paradox and Vague Geography. *Fuzzy Sets and Systems*, 113(1), 7–18. DOI: [https://doi.org/10.1016/S0165-0114\(99\)00009-3](https://doi.org/10.1016/S0165-0114(99)00009-3)
- Fisher, P.F. (2003). Data Quality and Uncertainty: Ships Passing in the Night. W: M.F. Goodchild, P.F. Fisher (Eds.). *Proceedings of the 2nd International Symposium on Spatial Data Quality*. Hong Kong: The Hong Kong Polytechnic University.

- Frank, A.U. (2001). Tiers of Ontology and Consistency Constraints in Geographic Information Systems. *International Journal of Geographical Information Science*, 15(7), 667–678. DOI: <https://doi.org/10.1080/13658810110061144>
- Fuller, R., Groom, G.B., Jones, A.R. (1994). The Land Cover Map of Great Britain: An Automated Classification of Landsat Thematic Mapper Data. *Photogrammetric Engineering and Remote Sensing*, 60(5), 553–562.
- Fuller, R.M., Smith, G.M., Sanderson, J.M., Hill, R.A., Thomson, A.G. (2002). The UK Land Cover Map 2000: Construction of a Parcel-Based Vector Map from Satellite Images. *The Cartographic Journal*, 39(1), 15–25. DOI: <https://doi.org/10.1179/caj.2002.39.1.15>
- Goodchild, M.F., Li, W. (2021). Replication across Space and Time Must Be Weak in the Social and Environmental Sciences. *PNAS*, 118(35), 1–8. DOI: <https://doi.org/10.1073/pnas.2015759118>
- Guptill, S.C., Morrison, J.L. (Eds.). (1995). *Elements of Spatial Data Quality*. Oxford: Pergamon. DOI: <https://doi.org/10.1016/C2009-0-14900-0>
- Heuvelink, G.B.M. (1998). *Error Propagation in Environmental Modelling with GIS*. London: CRC Press. DOI: <https://doi.org/10.4324/9780203016114>
- Hüllermeier, E., Waegeman, W. (2021). Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. *Machine Learning*, 110, 457–506. DOI: <https://doi.org/10.1007/s10994-021-05946-3>
- ISO (2023). ISO 19157-1:2023 *Geographic Information – Data Quality – Part 1: General Requirements*. International Organization for Standardization.
- Kinkeldey, C., MacEachren, A.M., Schiewe, J. (2014). How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualization User Studies. *The Cartographic Journal*, 51(4), 372–386. DOI: <https://doi.org/10.1179/1743277414Y.0000000099>
- Klir, G.J., Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River: Prentice Hall.
- Krukowski, M. (2021). *Nieostrość w modelowaniu kartograficznym*. Lublin: Wydawnictwo UMCS.
- Krukowski, M. (2025). Boundaries That Do Not Exist: Vagueness in Spatial Representation. *Prace Geograficzne*, 181, 143–158. DOI: <https://doi.org/10.4467/20833113PG.25.030.22984>
- Lund, H.G. (2018). *Definitions of Forest, Deforestation, Afforestation, and Reforestation*. Gainesville: Forest Information Services. DOI: <https://doi.org/10.13140/RG.2.2.31426.48323>
- MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E. (2005). Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32(3), 139–160. DOI: <https://doi.org/10.1559/1523040054738936>
- Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Yao-Yi, Jiao, J. (2025). Towards the Next Generation of Geospatial Artificial Intelligence. *International Journal of Applied Earth Observation and Geoinformation*, 136. DOI: <https://doi.org/10.1016/j.jag.2025.104368>
- Moellering, H. (1987). *A Draft Proposed Standard for Digital Cartographic Data*. Columbus: National Committee for Digital Cartographic Data Standards.
- Pawlak, Z. (1982). Rough Sets. *International Journal of Computer & Information Sciences*, 11(5), DOI: <https://doi.org/10.1007/BF01001956>
- Rosch, E. (1978). Principles of Categorization. W: E. Rosch, B.B. Lloyd (Eds.), *Cognition and Categorization* (s. 27–48). Hillsdale: Lawrence Erlbaum Associates.
- Rozporządzenie (2021). Rozporządzenie Ministra Rozwoju, Pracy i Technologii z dnia 27 lipca 2021 r. w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych, a także standardowych opracowań kartograficznych (Dz.U. 2021, poz. 1412).
- Salgé, F. (1995). Semantic Accuracy. W: S.C. Guptill, J.L. Morrison (Eds.), *Elements of Spatial Data Quality* (s. 139–151). Oxford: Pergamon. DOI: <https://doi.org/10.1016/C2009-0-14900-0>

- Smith, B., Mark, D.M. (2003). Do Mountains Exist? Towards an Ontology of Landforms. *Environment and Planning B: Planning and Design*, 30(3), 411–427. DOI: <https://doi.org/10.1068/b1282>
- Tabakowska, E. (2001). *Kognitywne podstawy języka i językoznawstwa*. Kraków: Universitas.
- Ustawa (1991). Ustawa z dnia 28 września 1991 r. o lasach (Dz.U. 1991, nr 101, poz. 444).
- Valle, D., Izbicki, R., Vieira Leite, R. (2023). Quantifying Uncertainty in Land-Use Land-Cover Classification Using Conformal Statistics. *Remote Sensing of Environment*, 295. DOI: <https://doi.org/10.1016/j.rse.2023.113682>
- Veregin, H. (1999). Data Quality Parameters. W: P.A. Longley, M.F. Goodchild, D.J. Maguire, D.W. Rhind (Eds.). *Geographic Information Systems: Principles and Technical Issues* (s. 17–189). New York: John Wiley & Sons.
- Vukalić, A., Triglav Čekada, M., Petrović, D. (2024). OpenStreetMap Data Quality Assessment According to ISO 19157-1:2023. *Abstracts of the ICA*, 7, 182. DOI: <https://doi.org/10.5194/ica-abs-7-182-2024>
- WIG (1925). *Instrukcja topograficzna Wojskowego Instytutu Geograficznego, Cz. 2: Techniczna*. Warszawa: Wojskowy Instytut Geograficzny.
- Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control*, 8(3), 338–353. DOI: [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)

PUBLICATION INFO		
SUBMITTED: 2025.12.17	ACCEPTED: 2026.01.11	PUBLISHED ONLINE: 2026.01.13