

Ireneusz Borowiecki, Marek Ślusarski

METODYKA OPISU JAKOŚCI BAZ DANYCH GEODEZYJNEJ EWIDENCJI SIECI UZBROJENIA TERENU

METHODOLOGY OF DESCRIPTION OF DATABASES THE SPATIAL REGISTRATION OF UTILITY INFRASTRUCTURE

Streszczenie

Obserwujemy obecnie dynamiczny rozwój procesów informatyzacji krajowego systemu informacji o terenie. Tworzone są geoinformacyjne bazy danych na poziomie krajowym, regionalnym i lokalnym. W celu budowy systemów spełniających stawiane przed nimi wymagania formalno-prawne konieczne jest opracowanie szczegółowych zasad metodologii oceny jakości danych przestrzennych.

W pracy przedstawiono propozycję opisu jakości danych geodezyjnej ewidencji sieci uzbrojenia terenu (GESUT). Procedurę oceny opracowano zgodnie z wytycznymi normy PN-EN-ISO 19113. Na podstawie badań dotyczących modelu oceny systemu informacji o terenie przeprowadzonych w latach poprzednich oraz bieżących analiz opracowano metodę oceny jakości danych GESUT. Zaproponowana metoda opisu jakości danych przestrzennych pozwala na sporządzenie precyzyjnego i wiarygodnego raportu oceny.

Słowa kluczowe: System informacji o terenie, jakość danych przestrzennych, geodezyjna ewidencja sieci uzbrojenia terenu

Summary

In the destination of construction geoinformatics systems meeting requirements put before them is necessary elaboration detailed principles of the assessment methodology of spatial data quality. The spatial registration of utility infrastructure is one of basic databases of the local system. This study focuses proposial the description of the quality cadastre data in accordance with guidelines of the norm PN-EN-ISO 19113

Key words: land information system, spatial data quality, the spatial registration of utility infrastructure

WSTĘP

Zachodzące w drugiej połowie XX wieku rewolucyjne zmiany na każdym polu ludzkiej działalności, a w szczególności w naukach technicznych, biologicznych, medycznych i ekonomicznych oraz szybkie wdrażanie osiągnięć naukowo technicznych sprawiły, że zwiększyła się rola wiarygodnej informacji i szybkiego do niej dostępu. Z tego powodu szczególnego znaczenia nabrała problematyka zbierania, przetwarzania, przechowywania i rozpowszechniania informacji szczególnie w aspekcie zastosowania technik komputerowych.

Procesy te dotyczą również informacji przestrzennej. Gwałtownie rośnie zapotrzebowanie na informacje przestrzenne wykorzystywane w procesach decyzyjnych dotyczących funkcjonowania państw i jakości życia obywateli. Zwiększa się również liczba instytucji gromadzących dane jak również ilość gromadzonych informacji. W Polsce krajowy system informacji przestrzennej budowany jest w na trzech zasadniczych poziomach: centralnym, regionalnymi i lokalnym. Poziom centralny to obszar całego kraju, region obejmuje województwo, a zasięg lokalny dotyczy obszaru powiatu.

Rozwój systemów lokalnych – w skali całego kraju – nie jest jednorodny, szczególnie w odniesieniu do rodzajów baz referencyjnych, jakości gromadzonych danych oraz stosowanych platform sprzętowych i programowych. W celu porównania systemów posiadających odmienną organizację oraz oceny poziomu realizacji zadań jakim służą konieczne jest tworzenie metodyki opisu jakości danych przestrzennych.

JAKOŚĆ ZBIORÓW DANYCH JAKO KOMPONENT METADANYCH

Kompendium infrastruktury danych przestrzennych: The SDI Cookbook [Nebert 2004] definiuje metadane jako dane o danych. W odniesieniu do zbioru danych przestrzennych, metadane zawierają informacje o tym zbiorze.

Definiowane standardy metadanych w swoich zbiorach podstawowych zawierają ocenę jakości danych przestrzennych. Cytowane powyżej kompendium infrastruktury danych przestrzennych (SDI) wyróżnia metadane rozpoznania, które pozwalają na ocenę jakości danych zbioru oraz określenie danych zbioru pod względem wymagań użytkownika. Główne elementy standardu CSDGM (US Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata) zawierają (wg kolejności ważności) [Longley i in. 2006]: podstawową informację o zbiorze danych, informację o jakości danych (ogólna ocena jakości danych w zbiorze), sposób uporządkowania danych przestrzennych w zbiorze i inne.

Rozporządzenie Ministra Rozwoju Regionalnego i Budownictwa z 2001 w sprawie szczegółowych zasad i trybu założenia i prowadzenia krajowego systemu informacji o terenie mówi m. in. o zakładaniu i prowadzeniu krajowych,

wojewódzkich i powiatowych baz metadach. Centralne, regionalne i lokalne zbiory metadanych mają zawierać informacje dotyczące m. in. nazwy systemu, zakresu tematycznego danych oraz stanu aktualności.

Parlament Europejski w 2007 przyjął dyrektywę ustanawiającą infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej (INSPIRE). Dyrektywa określa mechanizmy i procesy współdziałania, dostępu i użytkowania danych przestrzennych. Infrastruktura informacji przestrzennej rozumiana jest jako zbiór danych przestrzennych opisany metadanymi oraz usługi i procesy związane z tym przedsięwzięciem [Dyrektywa 2007/2/WE]. Metadane mają obejmować informacje dotyczące m. in. jakości i ważności zbiorów danych przestrzennych. Rozporządzenie wykonawcze do Dyrektywy [Rozporządzenie 1205/2008] określa zapis historii i rozdzielczość przestrzenną jako elementy metadanych opisujące jakość i wiarygodność danych przestrzennych.

Zespół do spraw krajowego profilu metadanych w zakresie geoinformacji opracował wzorcowe metadane dla standardowych danych przestrzennych tworzonych na poziomie centralnym, regionalnym i lokalnym [Augustynowicz, Soczewski 2008]. Zgodnie z tym wzorcem metadane mają zawierać informacje o jakości danych przestrzennych dla każdego rodzaju baz danych. Opracowane przez Zespół wytyczne techniczne dotyczące zasad tworzenia metadanych w zakresie geoinformacji [Baranowski i in. 2008] określają zasady wprowadzania informacji o jakości danych. Informacja o jakości danych może być prezentowana jako atrybut metadanych na dwa sposoby. Pierwszy to podanie informacji o historii powstania zbioru z uwzględnieniem danych źródłowych, ich jakości i sposobów przetworzenia. Drugi dotyczy utworzenia bazy danych zgodnie z opublikowanym standardem. W tym przypadku należy podać informacje o tym standardzie i opisać poziom zgodności zasobu informacyjnego ze standardem.

OPIS JAKOŚCI DANYCH PRZESTRZENNYCH WG NORM ISO

Charakteryzując jakość danych przestrzennych, można użyć kilku różnych własności. Według Gaździckiego [2008] jakość danych opisują następujące cechy: kompletność, zgodność logiczna, dokładność pozycyjna, dokładność czasowa, dokładność tematyczna oraz dokładność semantyczna i pochodzenie. Kompletność rozumiana jest jako występowanie wszystkich zamierzonych danych bez niedomiaru i nadmiaru. Zgodność logiczna to brak wewnętrznej sprzeczności w zbiorze danych. Dokładność pozycyjna dotyczy geodezyjnych dokładności – wyrażonych współrzędnymi- położenia obiektów. Dokładność czasowa związana jest ze zmianami danych w czasie, a tematyczna to prawidłowość określania np własności jakościowych. Dokładność semantyczna przedstawiana jako zbiór danych odtwarza przestrzeń rozważań (dziedzina problemu). Pochodzenie opisuje sposób i czas pozyskania danych oraz materiały źródłowe, metody i techniki.

Różnorodność cech wykorzystywanych do opisu jakości danych przestrzennych przyczyniła się do próby uregulowania zasad określania kwalifikacji geodanych w normach serii ISO. Rewolucyjny rozwój systemów informacji przestrzennej wiąże się z koniecznością budowy infrastruktury danych przestrzennych począwszy od projektów lokalnych do ponad państwowych. Wypracowane rozwiązania stają się obowiązującym prawem poprzez włączenie ich do systemu norm ISO.

Normy europejskie serii ISO 19100 dotyczą szerokiego zakresu pojęć informacji geograficznej. Charakteryzują się bogatym aparatem pojęciowym, są formalną dokumentacją rozwoju systemów informacyjnych. Ze względu na rewolucyjny rozwój systemów informacji przestrzennej normy te nie posiadają wieloletniego doświadczenia w ich stosowaniu, a zawarte w nich pojęcia i definicje nie zawierają powszechnie akceptowalnej terminologii.

Metodykę opisu jakości danych prezentuje norma PN-EN-ISO 19113 „Informacja geograficzna – Podstawy opisu jakości”. Zgodnie z zapisem normy jakość to „całość charakterystyk produktu, które zależą od jego zdolności do zaspokajania określonych i potencjalnych potrzeb”. Kompletny opis – identyfikacja informacji o jakości – powinien zawierać nieilościowe i ilościowe informacje o jakości oraz parametr metajakość. Propozycję zbioru cech identyfikujących informacje o jakości – zgodną z wytycznymi ISO 19113 – przedstawiono w tabeli 1.

Tabela 1. Identyfikacja informacji o jakości
Table 1. Identification information about the quality

NIEIŁOŚCIOWE INFORMACJE O JAKOŚCI DANYCH Data quality information – no quantitative description
– przeznaczenie – pochodzenie – wykorzystanie
IŁOŚCIOWE INFORMACJE O JAKOŚCI DANYCH Data quality information - quantitative description
<u>Elementy jakości danych</u> – kompletność – spójność logiczna – dokładność położenia – dokładność czasowa – dokładność tematyczna
<u>Elementy podrzędne jakości danych</u> – zakres jakości danych – miara jakości danych – procedura oceny jakości danych – wynik jakości danych – typ wartości jakości danych – jednostka wartości jakości danych – data jakości danych
JAKOŚĆ INFORMACJI O JAKOŚCI (metajakość) Information quality assurance (metaquality)

Prezentowane cechy pozwalają na opracowanie kompleksowej charakterystyki jakości danych każdego systemu geoinformacyjnego. Próbę scharakteryzowania jakości danych lokalnego systemu geoinformacyjnego przeprowadzono badając bazy danych geodezyjnej ewidencji sieci uzbrojenia terenu (GESUT).

PROPOZYCJA METODY OPISU JAKOŚCI DANYCH GESUT

Zespół do spraw krajowego profilu metadanych w zakresie geoinformacji opracował wzorcowe metadane dla standardowych danych przestrzennych tworzonych dla baz danych na poziomie powiatu [Augustynowicz, Soczewski 2008]. Opracowanie to w sposób ogólny opisuje wymagania dotyczące procedury oceny jakości lokalnego systemu geoinformacyjnego. Nie określono np. atrybutów zakresu informacji jakościowej.

W celu budowy systemów spełniających stawiane przed nimi wymagania formalno-prawne konieczne jest opracowania szczegółowych zasad metodologii oceny jakości danych przestrzennych. Jedną z podstawowych baz danych krajowego systemu informacji o terenie zakładanego i prowadzonego na poziomie lokalnym jest geodezyjna ewidencja sieci uzbrojenia terenu. Na podstawie własnych badań dotyczących modelu oceny systemu informacji o terenie przeprowadzonych w latach poprzednich [Ślusarski 2005] oraz bieżących analiz opracowano zarys metody oceny jakości danych GESUT. Wynik analiz przedstawiono w tabeli 2.

Tabela 2. Propozycja opisu jakości danych GESUT
Table 2. Proposed description of data quality GESUT

NIEIŁOŚCIOWE INFORMACJE O JAKOŚCI DANYCH		
Nazwa	Opis	Przykład
Przeznaczenie	Informacje o powodach utworzenia bazy i przewidywanym wykorzystaniu	Numeryczne zbiory danych utworzono w celu wypełnienia zobowiązań nałożonych przez ustawę Prawo geodezyjne (art. 28). Cyfrowa baza będzie wykorzystana w procesie projektowania zamierzeń inwestycyjnych. GESUT uporządkuje stan wiedzy o sieciach i poprawi jakość zbioru danych
Pochodzenie	Informacje o źródłach danych wykorzystanych do utworzenia bazy oraz metodach i technikach realizacji przedsięwzięcia	Bazy danych poszczególnych sieci technicznego uzbrojenia terenu powstały na podstawie: mapy zasadniczej 1:1000 (35%), operatów inwentaryzacji powykonawczej (60%) i dokumentacji technicznej gestorów sieci (5%).
Wykorzystanie	Informacje o zastosowaniu bazy	Dane o rozmieszczeniu sieci uzbrojenia technicznego terenu wykorzystywano w procesach inwestycyjnych.

ILOŚCIOWE INFORMACJE O JAKOŚCI DANYCH		
Elementy jakości danych		
Nazwa	Opis	Przykład
Kompletność	Informacja o wielkość niedomiaru lub nadmiaru danych w zbiorze	1. Szacowany procent długości sieci nie wykazanej w bazie (2% na podst. informacji zarządców sieci). 2. Szacowany procent liczby przyłączy nie wykazanych w bazie (5% na podst. informacji zarządców sieci).
Spójność logiczna	Informacje o stopniu zgodności struktury bazy z logicznymi regułami	Przestrzenne zbiory danych przestrzennych poprawne topologicznie. Zależności hierarchiczne typów sieci określone poprawnie
Dokładność położenia	Informacja o dokładności określenia pozycji w przestrzeni obiektów bazy	Dokładność wyznaczenia współrzędnych kartezjańskich poszczególnych odcinków przewodów sieci i armatury określona jest atrybutem-źródło danych - zgodnie z Instrukcją techniczną G7
Dokładność czasowa	Aktualność danych w stosunku do wymaganej częstości aktualizacji	Baza aktualizowana na bieżąco danymi z geodezyjnych inwentaryzacji powykonawczych

ILOŚCIOWE INFORMACJE O JAKOŚCI DANYCH		
Elementy podrzędne jakości danych		
Nazwa	Opis	Przykład
Zakres jakości danych	Informacja o zasięgu przestrzennym obiektów bazy	Sieci technicznego uzbrojenia terenu występujące na obszarze obrębu geodezyjnego
Miara jakości danych	Informacja o metodzie	Procent poprawnych atrybutów opisujących obiekty sieci uzbrojenia terenu
Procedura oceny jakości danych	Opis procesu badania jakości danych. Określenie cech jakości. Określenie miar jakości. Wybór i zastosowanie metody badawczej. Obliczenie współczynników oceny globalnej	Przeprowadzono analizy jakości danych w oparciu o trzy własności: dokładność, kompletność i aktualność. Obliczono średnią absolutną i współczynnik zmienności średnich.
Wynik jakości danych	Prezentacja współczynników opisujących jakość danych otrzymanych na podstawie wykonanego procesu badania jakości bazy.	Średnią absolutną (S_a) = 76 Współczynnik zmienności średnich (A) = 0.15
Typ wartości jakości danych	Informacja o typach wartości służących do przedstawienia jakości danych.	Średnią absolutną – współczynnik z zakresu od 0 do 100 punktów (najgorszy – najlepszy) Współczynnik zmienności średnich z zakresu wartości od 0 do 1 (najlepszy – najgorszy)
Jednostka wartości jakości danych	Informacja o jednostkach wartości służących do przedstawienia wyniku jakości danych	Jednostka średniej absolutnej - niemianowana Jednostka współczynnika zmienności średnich - niemianowana
Data jakości danych	Informacja o zakresie czasowym do którego odnosi się miara jakości danych	Data planowanej, następnej analizy jakości

WNIOSKI

Współczesne systemy goinformacyjne powinny zawierać komponenty metadanych, służące do opisu baz danych systemu. Kluczowym elementem zbioru metadanych są cechy opisujące jakość danych przestrzennych. Zaproponowany w pracy system oceny jakościowej geodezyjnej ewidencji sieci uzbrojenia terenu – oparty na normie PN-EN-ISO 19113 – umożliwia sporządzenie precyzyjnego i wiarygodnego raportu oceny.

Wiarygodny raport oceny powinien zawierać atrybuty opisujące: przeznaczenie, pochodzenie, wykorzystanie, kompletność, spójność i dokładność danych. Dodatkowo należy opisać: zakres i miarę jakości danych, procedurę oceny jakości oraz typ i jednostkę wartości jakości. Wynikiem jakości danych powinny być współczynniki, takie jak średnia absolutna oraz współczynnik zmienności średnich.

Opracowana metoda pozwala na porównywanie informacji zgromadzonych w bazach geoinformacyjnych zróżnicowanych pod względem organizacyjnym i utrzymywanych przez odmiennie zorganizowane systemy informatyczne.

BIBLIOGRAFIA

- Augustynowicz A., Soczewski P. *Wzorcowe metadane dla standardowych danych przestrzennych tworzonych przez służbę geodezyjną i kartograficzną w Polsce – poziom powiatowy*. GUGiK, Warszawa 2008.
- Baranowski M. *Wytyczne techniczne. Zasady tworzenia metadanych w zakresie geoinformacji*. GUGiK, Warszawa 2008.
- Dyrektywa 2000/2/WE Parlamentu Europejskiego i Rady z 2007 r. ustanawiająca infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej (INSPIRE).
- Gaździcki J. 2008. *Leksykon geopatyczny*. 2008, www.ptip.org.pl.
- Longley P. i in. *GIS teoria i praktyka*. PWN, Warszawa 2006.
- Nebert D. (red.) *The SDI Cookbook*. 2004, www.gsdi.org.
- Norma PN-EN-ISO 19113. *Informacja geograficzna – Postawy opisu jakości*. PKN, Warszawa 2005.
- Rozporządzenie Ministra Rozwoju Regionalnego i Budownictwa z 2001 r. w sprawie szczegółowych zasad i trybu założenia i prowadzenia krajowego systemu informacji o terenie (Dz.U. Nr 80, poz. 866).
- Rozporządzenie Komisji (WE) nr 1205/2008 z 2008 r. w sprawie wykonania dyrektywy 2007/2/WE Parlamentu Europejskiego i Rady w zakresie metadanych.
- Ślusarski M. *Propozycja ogólnej oceny systemu informacji o terenie budowanego na poziomie powiatu*. Mat. Konf. Kataster, Fotogrametria i Geoinformatyka: Nowoczesne technologie i perspektywy rozwoju. Kraków 2005.

Dr inż. Ireneusz Borowiecki
Dr inż. Marek Ślusarski
Katedra Geodezji
Uniwersytet Rolniczy w Krakowie
ul. Balicka 253A
30-198 Kraków
iborowiecki@op.pl
rmslusar@cyf-kr.edu.pl

Recenzent: *Prof. dr hab. Ryszard Żróbek*