

WALDEMAR IZDEBSKI*, MICHAŁ KURSA**

WYKORZYSTANIE DANYCH ADRESOWYCH W USŁUGACH LOKALIZACYJNYCH

Streszczenie

Numeracja adresowa należy do podstawowych i najczęściej wykorzystywanych danych przestrzennych. Jednym z jej zastosowań jest wyszukiwanie po adresie poprzez usługi lokalizacji (geokodowania), które dla danego adresu zwracają jego położenie w postaci współrzędnych. W pracy porównano 2 usługi: Usługę Lokalizacji Adresów firmy Geo-System i OpenLS (część geoportalu krajowego). Analiza obejmuje aspekty teoretyczne (ilość danych, ich aktualność, formaty zapytania i odpowiedzi) oraz praktyczne - w postaci testowania funkcjonowania obu usług.

Słowa kluczowe: adresy, PRG, geokodowanie, usługi lokalizacyjne

WPROWADZENIE

W pracy przedstawiono istniejące w Polsce usługi lokalizacyjne oparte o dane pochodzące z ewidencji miejscowości, ulic i adresów. W części pierwszej opisano podstawy prawne związane z zagadnieniem numeracji adresowej oraz podstawowe pojęcia związane z geokodowaniem. W drugiej części zaprezentowano dwie usługi ogólnopolskie wraz z ich porównaniem od strony teoretycznej i praktycznej. W ostatnim rozdziale omówiono wyniki i sformułowano wnioski.

* Politechnika Warszawska, Wydział Geodezji i Kartografii

** GEO-SYSTEM Sp. z o.o.

PODSTAWY PRAWNE PROWADZENIA NUMERACJI ADRESOWEJ

Numeracja adresowa w Polsce jest jednym z rejestrów urzędowych. Rejestr prowadzony jest przez wójta, burmistrza lub prezydenta miasta na podstawie ustawy "Prawo geodezyjne i kartograficzne". Wcześniej zagadnienie to regulowały różne przepisy, o których szczegółowo można przeczytać w pracy [1].

Wszystkie wcześniejsze akty prawne regulowały jedynie zasady oznaczania numerami nieruchomości, ale nie odnosiły się do sposobu i formy prowadzenia tej ewidencji. Z punktu widzenia technicznego numeracja była prowadzona w sposób niejednorodny w formie papierowej w postaci zeszytów i segregatorów, a także różnego typu map. Wraz z rozwojem techniki komputerowej część samorządów (głównie większe gminy i miasta) stopniowo przenosiła rejestry na nośniki elektroniczne, będące jednak najczęściej odzwierciedleniem wcześniejszego sposobu prowadzenia numeracji tj. zamieniono tabele z zeszytów na tabele w arkuszu kalkulacyjnym. Nieliczne jednostki w późniejszych latach decydowały się na dedykowane oprogramowanie do prowadzenia baz adresowych.

Wspomniany sposób prowadzenia ewidencji skutkował de facto jego zamknięciem i brakiem dostępności z zewnątrz (dla obywateli oraz instytucji). Nie istniały ani podstawy prawne, ani rozwiązania techniczne, służące do sprawnej wymiany informacji zgromadzonych w rejestrze gminnym. Prowadziło to do sytuacji, gdy instytucje centralne jak GUS, służby publiczne (pogotowie ratunkowe, straż pożarna) czy np. firmy zajmujące się tworzeniem nawigacji GPS, musiały do swoich celów tworzyć własne bazy. Każda z tych baz różniła się zawartością i aktualnością; każda stanowiła oddzielną „samotną wyspę”.

Znaczące zmiany nastąpiły dopiero po 2010 roku, kiedy to w Polsce zaimplementowano dyrektywę INSPIRE w postaci ustawy z 4 marca 2010 r. o "Infrastrukturze informacji przestrzennej", jak również znowelizowano ustawę "Prawo geodezyjne i kartograficzne". W roku 2012 na bazie tych regulacji wydano rozporządzenie Rady Ministrów z 10 stycznia 2012 r. w sprawie państwowego rejestru granic i powierzchni jednostek podziałów terytorialnych kraju oraz rozporządzenie Ministra Administracji i Cyfryzacji z 9 stycznia 2012 r. w sprawie ewidencji miejscowości, ulic i adresów. Wymienione akty prawne określały jednolite wymagania techniczne prowadzenia rejestru ewidencji adresowej i umiejscawiały ją w ramach krajowej infrastruktury danych przestrzennych [5].

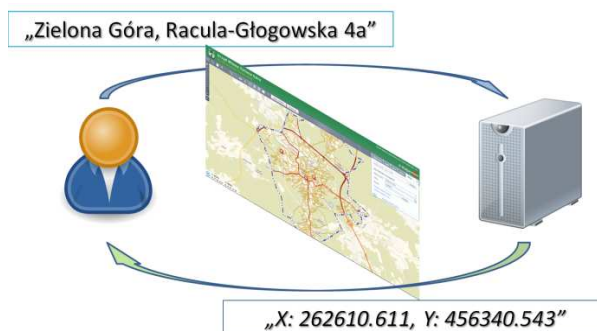
Nowe przepisy uregulowały formę prowadzenia rejestru (system komputerowy) oraz pozwoliły na udostępnianie danych w ustandaryzowanej formie (pliki GML i usługi sieciowe). Ponadto wprowadzono centralny rejestr PRG, który łączy w jednej bazie dane adresowe pochodzące z baz gminnych. Dzięki nowym rozwiązaniom prawnym każdy obywatel, firma i instytucja mogą pozyskać dane adresowe bezpośrednio z danej gminy lub z kompleksowej bazy PRG utworzonej z baz gminnych.

Przedstawione powyżej akty prawne opisują przekazywanie informacji adresowej na dwa sposoby: w postaci plików GML i za pomocą usług sieciowych. Pierwszy sposób co do zasady dotyczy całości bazy danych i ma zastosowanie w przypadku procedury eksportu i importu. Drugi sposób może mieć również zastosowanie w przypadku fragmentów takiej bazy, a do odbiorcy mogą być przekazane informacje odpowiednio przetworzone i zinterpretowane.

Przykładem usługi INSPIRE jest usługa przeglądania realizowana poprzez standard OGC WMS, która w najprostszym przypadku zwraca graficzny obraz mapy z zadanego obszaru.

USŁUGI GEOKODOWANIA

W przypadku adresów, które są podstawą lokalizacji wielu innych zbiorów danych, ważną rolę odgrywają także usługi geokodowania [6], które pozwalają na zamianę zapisu adresu w formie słownej (najczęściej w postaci „miejscowość, ulic, numer”) na współrzędne geograficzne [4]. Idea geokodowania przedstawiona została na rys. 1.



Rys. 1. Zasada działania geokodowania

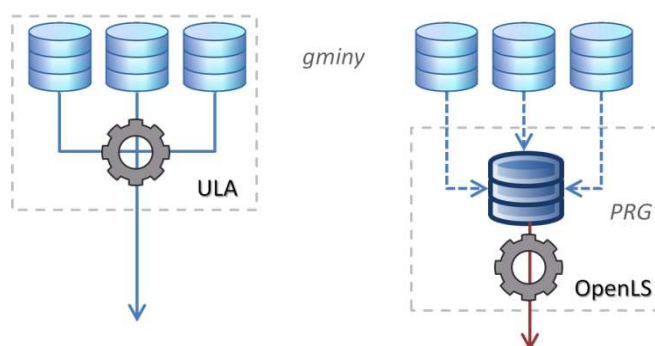
Fig. 1. Principle of geocoding

W Polsce można wyróżnić dwie główne usługi geokodowania: ULA (Usługa Lokalizacji Adresów) firmy Geo-System Sp. z o.o. [3] i OpenLS realizowana przez Główny Urząd Geodezji i Kartografii w ramach projektu Geoportal [7]. Oprócz nich istnieją też inne tego typu usługi udostępniane przez serwisy komercyjne (nie oparte jednak na danych urzędowych) oraz lokalne serwisy gminne i miejskie (ograniczone jedynie do wybranej jednostki). Ze względu na charakter i zakres danych nie są one przedmiotem dalszej analizy.

Wymienione usługi porównano pod względem teoretycznym i praktycznym. Analiza teoretyczna obejmowała ogólne zasady działania każdej z usług, jej dokumentację, sposób zadawania zapytań i format odpowiedzi. W analizie praktycznej przeprowadzono test działania obu usług na podstawie odpowiednio dobranych danych – zbadano wydajność, aktualność oraz jakość i trafność odpowiedzi.

ANALIZA TEORETYCZNA

Usługa Lokalizacji Adresów firmy Geo-System uruchomiona została w 2010 roku. GUGiK przedstawił pierwszą wersję usługi OpenLS w roku 2013, a w roku 2015 oparł ją o dane z zasobu PRG. Podstawowa zasada działania obu usług jest taka sama – zamiana podanego adresu na współrzędne, ale różnią się one sposobem realizacji. W przypadku OpenLS można mówić o modelu centralnym – usługa operuje na jednej bazie PRG, a dla usługi ULA będzie to model rozproszony – operuje ona na wielu bazach gminnych [8]. Ze względu na różne modele obu usług spodziewać się należy różnic w aktualności danych. Usługa ULA z uwagi na bezpośredni dostęp do bazy w gminie zapewniać będzie natychmiastową aktualność, natomiast rozwiązanie GUGiK limitowane będzie przez opóźnienia wynikające z procedury przesyłania danych z gminy do PRG.



Rys. 2. Porównanie modeli działania usług ULA i OpenLS

Fig. 2. Comparison between ULA and OpenLS operating models

Usługa ULA obejmuje gminy, które są użytkownikami oprogramowania Internetowy Manager Punktów Adresowych firmy Geo-System – w chwili pisania pracy było to ok. 70% gmin w Polsce (tj. ok. 1700), a dla pozostałych gmin korzysta z bazy PRG. Usługa OpenLS oparta jest o bazę PRG i obejmuje teoretycznie wszystkie polskie gminy (te 1700 z IMPA również) – bo gminy bez względu na stosowaną aplikację zobowiązane są do zasilania PRG [2].

Tabela 1. Podstawowe różnice między usługami ULA i OpenLS

Table 1. Basic differences between ULA and OpenLS

	ULA	OpenLS
Typ usługi	usługa rozproszona	usługa centralna (PRG)
Rok uruchomienia	2010	2013
Aktualność	natychmiastowa	opóźnienie związane z importem do PRG

Usługa Lokalizacji Adresów jest autorskim rozwiązaniem firmy Geo-System i nie jest implementacją żadnego ze standardów OGC, opiera się na standardowych elementach i technologiach: HTTP GET, JSON i XML. OpenLS oparta jest na standardzie OGC Open Location Services Interface [9], rozszerzonym o tzw. profil GUGiK. OpenLS do komunikacji używa interfejsu HTTP POST i dokumentów XML.

Tab. 2. Podstawowe różnice między usługami ULA i OpenLS cd.

Tab. 2. Basic differences between ULA and OpenLS cont.

	ULA	OpenLS
Metoda zapytania	GET wywołanie adresu URL	POST przesłanie dokumentu XML
Odpowiedź	JSON, XML	XML
Układ współrzędnych w odpowiedzi	WGS 84, PUWG 92, 2000 (wybierany przez użytkownika)	WGS 84

Wybór komunikacji POST może stanowić utrudnienie w korzystaniu z usługi OpenLS, ponieważ zapytanie musi być wysłane w formie dokumentu XML z odpowiedniego formularza na stronie internetowej lub w aplikacji komputerowej. Metoda GET nie wymaga takich rozwiązań i może być stosowana bezpośrednio z przeglądarki internetowej.

Tab. 3. Porównanie możliwości usług ULA i OpenLS – zapytania

Tab. 3. Comparison of ULA and OpenLS capabilities – queries

	ULA	OpenLS
zapytania	słownikowe: - Lista miejscowości - Lista ulic - Lista punktów adresowych	tekstowe: “miejscowość, ulica, numer” + dodatkowe ograniczenia - Województwo - Powiat - Gmina - Miejscowość - Ulica - Kod pocztowy
	tekstowe: “miejscowość, ulica, numer” + dodatkowe ograniczenia - Gmina - Miejscowość - Ulica	

Pomimo różnic w metodach komunikacji obie usługi zapewniają podobne możliwości wyszukiwania z opcjonalnym ograniczeniem poszukiwań do zakresu wybranej jednostki, miejscowości lub ulicy. W przypadku usługi ULA zwraca uwagę dostępność zapytania słownikowego tj. zwracającego listę wszystkich adresów przy wybranej ulicy lub ulic w miejscowości.

Tab. 4. Porównanie możliwości usług ULA i OpenLS – odpowiedzi

Tab. 4. Comparison of ULA and OpenLS capabilities – responses

	ULA	OpenLS
Informacje zawarte w odpowiedzi	<ul style="list-style-type: none"> - Liczba znalezionych obiektów - Określenie stopnia trafności - Nazwa miejscowości + SIMC - Nazwa województwa - Nazwa powiatu - Nazwa gminy - Zakres przestrz. miejscowości - Nazwa ulicy + ULIC - Zakres przestrzenny ulicy - Numer punktu - Współrzędne punktu 	<ul style="list-style-type: none"> - Współrzędne punktu - Identyfikator punktu - Adres (ulica + numer) - Nazwa miejscowości - Typ miejscowości - Województwo + TERYT - Powiat + TERYT - Gmina + TERYT - Określenie stopnia trafności

Zakres informacji zwracanych w odpowiedzi przez usługi jest podobny, jedyną różnicę stanowią identyfikatory SIMC i ULIC w usłudze ULA. Poniżej przedstawiono przykład wywołania obu usług dla adresu „Marki, Polna 1”:

- ULA <http://www.punktyadresowe.pl/lokalizacja.php?adres=Marki, Polna 1>
- OpenLS <http://mapy.geoportal.gov.pl/openLSgp/geocode>

```
<xls:XLS version="1.2" xmlns:xls="http://www.opengis.net/xls"
xmlns:gugik_ols="http://www.geoportal.gov.pl/schema/ols"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
<xls:RequestHeader/>
<xls:Request methodName="GeocodeRequest" version="1.0.0" requestID="ID_1">
<gugik_ols:GeocodeRequest>
<gugik_ols:AddressPoint countryCode="PL">
<gugik_ols:freeFormAddress>Marki, Polna 1</gugik_ols:freeFormAddress>
</gugik_ols:AddressPoint>
</gugik_ols:GeocodeRequest>
</xls:Request>
</xls:XLS>
```

Rys. 3. Zapytanie POST do usługi OpenLS w formacie XML

Fig. 3. OpenLS POST request in XML format

W przykładzie wyraźnie widać duży narzut spowodowany wykorzystaniem dokumentu XML w przesłanym do usługi OpenLS zapytaniu. Szukany adres ma jedynie 15 znaków, całe zaś żądanie liczy 508 znaków.

Odpowiedzi zwrotne dla szukanego adresu wyglądają następująco:

- ULA:

- Format JSON

```
{ "wyniki": [{"trafnosc": 1, "wdrozenie": "marki", "miejscowosc": "Marki", "simc": "0920901", "województwo": "mazowieckie",
"powiat": "wołomiński", "gmina": "Marki", "zakres_miejscowosci": [21.0769113, 52.3033986, 21.1645596, 52.3701611],
"ulica": "Polna", "zakres_ulicy": [21.10029, 52.320641, 21.101549, 52.32216], "ulica_id": "70472", "ulic": "17011",
"punkt_na_ulicy": [21.1009, 52.32142], "punkt": "1", "wspolrzedne_punktu": [21.100425, 52.321211], "punkt_id": "277845"}]}
```

Rys. 4. Odpowiedź usługi ULA w formacie JSON

Fig. 4. ULA response in JSON format

- Format XML

```
<WynikiWyszukiwania liczbaPozycji="1">
<Pozycja>
<Trafnosc>1</Trafnosc>
<Miejscowosc>Marki</Miejscowosc>
<SIMC>0920901</SIMC>
<Wojewodztwo>mazowieckie</Wojewodztwo>
<Powiat>wołomiński</Powiat>
<Gmina>Marki</Gmina>
<ZakresMiejscowosci>21.0769113,52.3033986,21.1645596,52.3701611</ZakresMiejscowosci>
<Ulica>Polna</Ulica>
<ZakresUlicy>21.10029,52.320641,21.101549,52.32216</ZakresUlicy>
<ULIC>17011</ULIC>
<PunktNaUlicy>21.1009 52.32142</PunktNaUlicy>
<Punkt>1</Punkt>
<WspolrzednePunktu>21.100425,52.321211</WspolrzednePunktu>
</Pozycja>
</WynikiWyszukiwania>
```

Rys. 4. Odpowiedź usługi ULA w formacie XML

Fig. 4. ULA response in XML format

- OpenLS

```
<xls:XLS xmlns:gugik_ols="http://www.geoportal.gov.pl/schema/ols" xmlns:xls="http://www.opengis.net/xls/" xmlns:
<xls:ResponseHeader/>
<xls:Response requestID="ID_1" version="1.0.0">
<gugik_ols:GeocodeResponse>
<gugik_ols:GeocodeResponseList xmlns:gml="http://www.opengis.net/gml" numberOfGeocodedAddresses="1">
<gugik_ols:GeocodedAddress>
<gml:Point>
<gml:pos>21.10874683127732 52.32941784428723</gml:pos>
</gml:Point>
<gugik_ols:AddressPoint countryCode="PL" id="http://geoportal.gov.pl/PZGIK/dane/PL.PZGIK.200/7397260">
<xls:StreetAddress>
<xls:Street>Polna 1</xls:Street>
</xls:StreetAddress>
<xls:Place type="Municipality">Marki</xls:Place>
<xls:PostalCode/>
<gugik_ols:AdministrativeUnit hierarchyLevel="wojewodztwo" id="14">mazowieckie</gugik_ols:AdministrativeUnit>
<gugik_ols:AdministrativeUnit hierarchyLevel="powiat" id="1434">wołomiński</gugik_ols:AdministrativeUnit>
<gugik_ols:AdministrativeUnit hierarchyLevel="gmina" id="1434021">Marki</gugik_ols:AdministrativeUnit>
</gugik_ols:AddressPoint>
<xls:GeocodeMatchCode accuracy="1"/>
</gugik_ols:GeocodedAddress>
</gugik_ols:GeocodeResponseList>
</gugik_ols:GeocodeResponse>
</xls:Response>
</xls:XLS>
```

Rys. 5. Odpowiedź usługi OpenLS w formacie XML

Fig. 5. OpenLS response in XML format

Powyższy przykład również pokazuje rozbudowaną formę odpowiedzi z usługi OpenLS (ok. 1400 znaków), mimo podobnej treści, co w ULA, która odpowiada w bardziej zwartej formie (<600znaków).

ANALIZA PRAKTYCZNA

Analizę praktyczną przeprowadzono w kilku aspektach i na różnych zestawach danych. W celu określenia wydajności przygotowano listę 100 adresów z różnych gmin, dla których kolejno odpytywano obie usługi. Nie rozpatrywano czy szukany adres został znaleziony, a jedynie mierzono czas oczekiwania na

odpowiedź. Procedurę przeprowadzono w różnych porach dnia oraz z różnych łączy internetowych (stałe, wifi, komórkowe). Wyniki przedstawiono w tabeli:

Tab. 5. Czas odpowiedzi na zapytanie

Tab. 5. Response time

	Średnia	σ		Średnia	σ
ULA	343,6 ms	78,6	OpenLS	64,4 ms	21,8

Widoczna jest przewaga wydajnościowa usługi OpenLS, którą można tłumaczyć jej centralną architekturą i przeszukiwaniem tylko jednej bazy, w porównaniu do ULA, łączącej i przeszukującej poszczególne bazy gminne (rys. 2).

Szybkość usługi ULA nie stanowi problemu i w codziennych sytuacjach jest zadowalająca. Daje się też zauważyć różny czas odpowiedzi z poszczególnych baz, na co może mieć wpływ wielkość danej bazy.

Trafność i skuteczność wyszukiwania przetestowano na 100 adresach z gmin, które posiadają oprogramowanie iMPA i przekazują dane do PRG, a więc powinny być w obu usługach. Wybrane adresy zostały zapisane w różnych formach (np. „Piłsudskiego” / „Józefa Piłsudskiego” / „Marszałka Józefa Piłsudskiego” itp.). Wyniki przedstawiono w poniższej tabeli.

Tab. 6. Trafność wyszukiwania

Tab. 6. Search accuracy

	ULA	OpenLS
Trafność	95%	87%

Usługa ULA wykazała się znacznie większą skutecznością wyszukiwania. Przypadki nieznaledzenia adresu dotyczyły głównie sytuacji, gdy w zapytaniu stosowano niestandardową kolejność zapisu wielocłonowej nazwy ulicy (np. „Piłsudskiego Józefa”).

Usługa OpenLS nie zwróciła wyniku aż w 13% przypadków, z czego większość dotyczyła punktów, które zostały wstawione w przeciągu kilku tygodni przed wykonaniem testu, mimo że raporty aktualizacji PRG potwierdzały wgranie danych z tego okresu. Problematiczne były również zapytania o adresy z pełnymi nazwami ulic (nie udało się np. znaleźć punktów przy ulicy „Prymasa Stefana Wyszyńskiego”, konieczne było wyszukanie „Wyszyńskiego”).

Podczas testowania usługi OpenLS zauważono także dodatkowe problemy (nie uwzględniono ich jednak przy ocenie trafności wyszukiwania) – wielokrotnie zwracane były podwójne wyniki, różniące się jedynie identyfikatorem IdIIP oraz położeniem; zauważono także całkowity brak wyników dla zapytań dotyczących miasta Lublin.

WNIOSKI I PODSUMOWANIE

Przeprowadzone testy i analizy nie wyczerpują całkowicie tematu, natomiast wskazują problematykę i obszary, którym można poświęcić kolejne badania. Na podstawie wykonanych porównań można stwierdzić, że obie omawiane usługi oferują podobne możliwości teoretyczne, ale w działaniu praktycznym daje się zauważyć pomiędzy nimi istotne różnice. Rozproszony model działania usługi ULA wpływa na jej niższą wydajność. Usługa OpenLS, choć szybsza, ma niższą skuteczność w znajdowaniu adresów, a wyniki są mniej aktualne. Oprócz opóźnienia wynikającego z procesu importu do PRG, problem stanowią też trudne do oszacowania opóźnienia wewnątrz infrastruktury Geoportalu (brak adresów nadanych tuż przed aktualizacją PRG) oraz niejasna struktura i jakość danych w PRG (podwójne adresy; całkowity brak adresów z niektórych miast/gmin).

Aby móc w pełni porównać obie usługi należałoby najpierw zbadać wskazane niejasności i zweryfikować bazę PRG. Następnie, gdy dane w bazie centralnej będą się pokrywać z bazami gminnymi, będzie można przystąpić do dalszych testów skuteczności wyszukiwania. Na podstawie poruszonych w pracy przykładów wydaje się wskazane, by szerzej zanalizować trafność algorytmów wyszukiwania pod kątem przybliżonego dopasowywania ciągów tekstowych (tj. dla różnych kolejności zapisu elementów adresu, słów z literówkami itp.).

LITERATURA

1. IZDEBSKI W., MALINOWSKI Z.: 2016. Dlaczego trzeba poprawić polskie przepisy dotyczące numeracji adresowej? *Magazyn Geoinformacyjny GEODETA* grudzień 2016 s. 32-37.
2. IZDEBSKI W., MALINOWSKI Z.: 2015. Analiza stanu numeracji adresowej w Polsce i możliwości jej wykorzystania przez obywateli i administrację. *Journal of Civil Engineering, Environment and Architecture*; t. XXXII, z. 62 (4/15), 2015, s. 141-154.
3. IZDEBSKI W.: 2014. Koncepcja standaryzacji usług lokalizacji przestrzennej adresów i działek katastralnych. *Magazyn Geoinformacyjny GEODETA* luty 2014 s. 14-18.
4. CICHOCIŃSKI P.: 2014. Problematyka geokodowania zdarzeń drogowych. *Roczniki Geomatyki*; t. XII, z. 2(64), s. 205-216
5. LEOŃCZYK M.: 2012. Ewidencja miejscowości, ulic i adresów. *Roczniki Geomatyki*; t. X, z. 6(56), s. 63-76
6. GOLDBERG D.: 2008. *A Geocoding Best Practices Guide*. University of Southern California, GIS Research Laboratory.

7. GUGiK: 2015. Dokumentacja profilu GUGiK usługi OpenLS. http://www.geoportal.gov.pl/documents/10179/26435/G2_Przewodnik_uzytkownika_1.04_Uzupelnienie_iMapLiteAPI+i+OpenLS+Zalacznik+1_Profil+OLS+GUGiK+1.01.pdf [dostęp: 2016-06-08].
8. GEO-SYSTEM: 2013. Specyfikacja techniczna usługi ULA. http://www.punktyadresowe.pl/UsługaLokalizacji-2_02.zip [dostęp:2016-06-08].
9. OPEN GEOSPATIAL CONSORTIUM: 2008. OpenGIS Location Service Implementation Specification. <http://www.opengeospatial.org/standards/ols> [dostęp: 2016-06-08].

USE OF THE ADDRESS REGISTER DATA IN LOCATION-BASED SERVICES

S u m m a r y

Address register is one of the basic and most commonly used spatial data records. One of its uses is finding addresses in location-based services (geocoding), which for the specified address (city, street, number) return the position in the form of coordinates. The paper presents a comparison of two services: ULA (Usługa Lokalizacji Adresów from Geo-System) and OpenLS (available as a part of the Polish national geoportal). The analysis includes both theoretical aspects (amount of data in the resource, its timeliness, documentation, request and response formats), as well as practical - in the form of testing the performance and accuracy of both services.

Key words: addresses, PRG, geocoding, location-based services