

Dariusz Gotlib\*

## Wybrane aspekty modelowania wielorozdzielczych i wieloreprezentacyjnych baz danych topograficznych

### 1. Wprowadzenie

Pojęcie wielorozdzielczych, wieloreprezentacyjnych lub wieloskalowych baz danych przestrzennych pojawia się w literaturze światowej i polskiej od kilku lat. Pojęcia te nie są jednak dobrze zdefiniowane i nie zawsze są jednoznacznie rozumiane. W języku angielskim funkcjonuje przede wszystkim określenie (skrót) MRDB (*Multiresolution/Multirepresentation Database*). Pod tym określeniem najczęściej rozumie się taką konstrukcję bazy danych, w której wyróżnia się kilka poziomów uogólnienia danych LoD (*Level of Details*) i w przypadku każdego z poziomów przechowuje się nieco inną, odpowiednią dla niego reprezentację obiektu. Przy czym niezmiernie ważne jest istnienie odpowiednich powiązań między poziomami pozwalających na odniesienie wszystkich reprezentacji do jednego rzeczywistego, reprezentowanego obiektu.

W krajowych opracowaniach zwraca się szczególną uwagę na możliwość pozyskiwania i zarządzania w jednej bazie danymi o różnym poziomie szczegółowości, zależnie od potrzeb informacyjnych na danym terenie. Wskazuje się, że zastosowanie baz tego typu umożliwiłoby przyspieszenie tworzenia infrastruktury danych przestrzennych.

W literaturze polskiej pojawia się coraz więcej opracowań poruszających kwestię tworzenia tego typu opracowań. Szczególnie ciekawym polem do wdrożeń tej metody są bazy danych topograficznych, ponieważ powinny stanowić one referencje do wielu danych tematycznych o różnym stopniu szczegółowości i dokładności.

W opracowaniach [2, 3], zwrócono uwagę, iż warte przeprowadzenia dokładnych analiz i eksperymentów praktycznych wydaje się rozwiązanie polegające na opracowaniu baz danych o modelu pojęciowym pozwalającym na gromadzenie

---

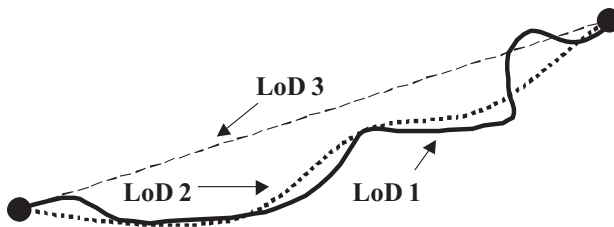
\* Instytut Fotogrametrii i Kartografii, Wydział Geodezji i Kartografii, Politechnika Warszawska, Warszawa.

danych na różnych poziomach szczegółowości w różnych częściach przestrzeni, ale z zachowaniem tej samej dokładności opisu geometrycznego obiektów. Umożliwiłoby to stosunkowo szybkie opracowanie bazy danych w skali całego kraju na poziomie dokładności właściwym np. dla bazy danych topograficznych i map topograficznych 1:10 000, ale o ograniczonym zakresie treści odpowiadającym np. mapie 1:50 000. W razie potrzeby teren przedstawiony w sposób uogólniony może ulec uszczegółowieniu wraz z pojawieniem się takich potrzeb i rozwojem bazy danych. Podejście to pozwala na szybki rozwój bazy danych już w pierwszym etapie, dzięki wydzieleniu obszarów kraju o różnych potrzebach informacyjnych. Tak skonstruowana baza danych umożliwia dodatkowo łatwe wydawanie użytkownikom danych uogólnionych, a także ułatwia prezentacje kartograficzne danych.

## 2. Podstawowe cechy baz typu MRDB

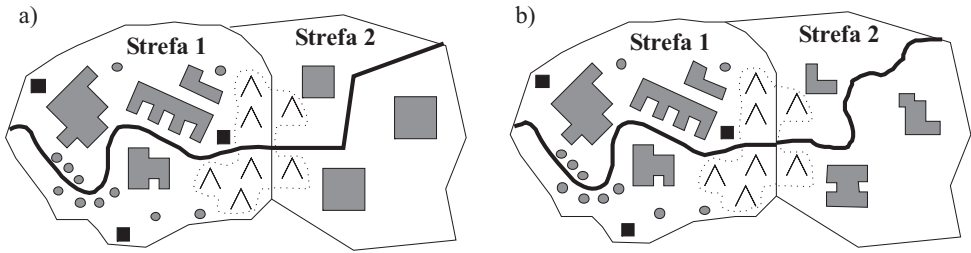
Jak wspomniano powyżej, pojęcie bazy typu MRDB może być rozumiane na różne sposoby. Dla uniknięcia niejednoznaczności proponuje się przyjąć następujące pojęcia i zdefiniować pomiędzy nimi następujące zależności:

- Wieloreprezentacyjna baza danych przestrzennych – baza danych, w której ten sam obiekt geograficzny posiada wiele reprezentacji geometrycznych, zależnych od przyjętych dla konkretnej bazy danych poziomów uogólnienia. Poszczególne reprezentacje różnią się stopniem generalizacji, ale są ściśle powiązane ze sobą i traktowane jako element opisu tego samego obiektu (rys. 1).



Rys. 1. Trzy reprezentacje geometryczne tego samego obiektu w bazie danych przestrzennych

- Wielorozdzielcza baza danych przestrzennych – baza danych, w której szczegółowość i dokładność opisu terenu jest różna w różnych miejscach przestrzeni, ale zachowana jest ciągłość przestrzenna i jednolitość modelu pojęciowego (rys. 2a). Mianem tym można określać również bazy danych, w przypadku których dokładność opisu terenu jest taka sama w całym obszarze opracowania, zmienna jest natomiast szczegółowość opisu – treść (rys. 2b).



Rys. 2. Ilustracja idei wielorozdzielczej bazy danych.

W wybranych obszarach dane szczegółowe (Strefa 1), w innych (Strefa 2) dane bardziej uogólnione, ale z zachowaniem spójnego modelu pojęciowego

- Wieloskalowa baza danych przestrzennych – baza danych, w której ten sam obiekt geograficzny posiada wiele reprezentacji geometrycznych, zależnych od przyjętych dla konkretnej bazy danych poziomów uogólnienia lub baza danych, w której szczegółowość i dokładność opisu terenu jest różna w różnych miejscach przestrzeni, ale zachowana jest ciągłość przestrzenna i jednolitość modelu pojęciowego (rys. 2a i b). Pojęcie wieloskalowości w kontekście baz danych nie wydaje się jednak najbardziej trafnym określeniem.

Relacje między pojęciami „wieloskalowa”, „wieloreprezentacyjna” i „wielorozdzielcza” baza danych przedstawiono na rysunku 3.



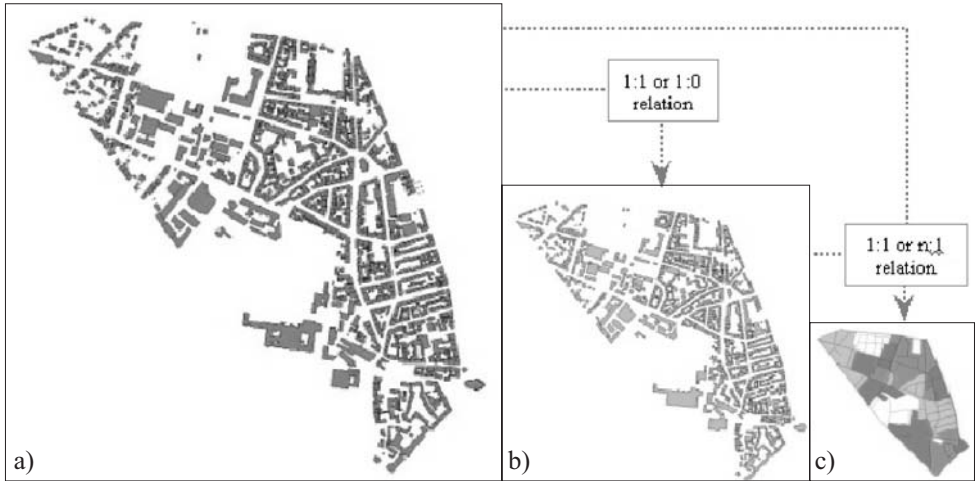
Rys. 3. Związki między pojęciami wieloskalowej, wielorozdzielczej i wieloreprezentacyjnej bazy danych

### 3. Powiązania między obiektami

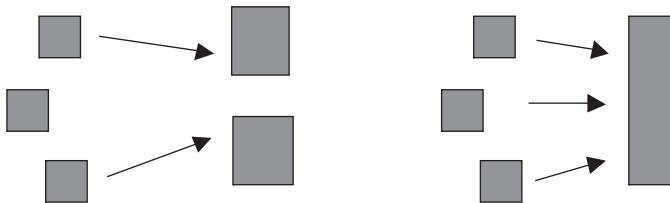
Przy konstruowaniu baz typu MRDB kluczowym zadaniem jest utworzenie odpowiednich powiązań między wszystkimi poziomami bazy, pozwalających na odniesienie wszystkich reprezentacji do jednego rzeczywistego, reprezentowanego obiektu.



modelu DCM<sup>1</sup>. W generalizacji modelu DLM<sup>2</sup> relacja ta nie występuje. Wprowadzony przez autorów formalizm matematyczny pozwala na zdefiniowanie więzów integralności w bazie danych przestrzennych.



Rys. 5. Przykład bazy MRDB z trzema LoD: a) oryginalne budynki; b) budynki uproszczone w sensie geometrycznym; c) obszary zabudowy [10]



Rys. 6. Dwie najczęściej występujące relacje pomiędzy obiektami dwóch poziomów w bazie MRDB – 1:1 i  $n:1$

#### 4. Reprezentacja geometryczna obiektów w wieloreprezentacyjnych bazach danych

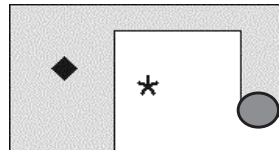
Przechowywanie kilku reprezentacji geometrycznych tych samych obiektów w wieloreprezentacyjnych bazach danych ma na celu uniknięcie budowy kilku niezależnych baz danych o różnych poziomach uogólnienia, a zamiast tego opracowanie jednej lub kilku ściśle powiązanych baz danych pozwalających na spójne przedstawienie terenu w różny sposób zależnie o potrzeb. Każdy obiekt będzie

<sup>1</sup> DCM – *Digital Cartographic Model*.

<sup>2</sup> DLM – *Digital Landscape Model*.

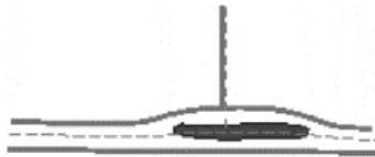
miał więc przypisanych kilka reprezentacji geometrycznych, mniej lub bardziej dokładnych. Poszczególne wersje geometrii obiektu mogą być wytworzone w automatycznych lub manualnych procesach generalizacji. Wynik tych procesów jest zapisywany w bazie danych. W obecnie dostępnych systemach zarządzania bazami danych geograficznych nie ma jednak mechanizmów specjalnie przygotowanych do tego typu rozwiązań. Aby móc przechowywać kilka geometrii obiektów w typowej relacyjnej bazie danych geograficznych, w przypadku każdego LoD, dla każdej klasy obiektów, należy np. zaprojektować oddzielną tabelę, w której poza geometrią obiektu zostanie zapisane powiązanie do obiektu źródłowego. Atrybuty opisowe obiektu są przechowywane raz, w tej klasie, która traktowana jest jako źródłowa. Może to być nawet zwykła relacyjna tabela niezawierająca współrzędnych obiektu.

Przykładem typowego obiektu posiadającego więcej niż jedną reprezentację geometryczną może być obiekt klasy „Budynek”. Na najbardziej dokładnym LoD przedstawiony może być w postaci powierzchniowej, natomiast na najmniej dokładnym jako punkt. Przy czym w najprostszym przypadku może to być środek geometryczny źródłowego obiektu powierzchniowego. Z punktu widzenia użyteczności danych o wiele lepszym rozwiązaniem może być świadomy wybór punktu reprezentacyjnego obiektu, np. wejście do budynku (rys. 7).



**Rys. 7.** Przechowywanie w bazie danych dwóch reprezentacji geometrycznych obiektu: obrys budynku i punkt reprezentacyjny (inny od środka geometrycznego). Gwiazdką oznaczono środek geometryczny obiektu wyznaczony na podstawie MBR (*Maximum Bounding Rectangle*)

Innym przykładem może być przedstawienie geometrii sieci drogowej na dwóch poziomach LoD: jako osi jezdni oraz jako osi drogi (rys. 8). Oś jezdni i oś drogi są to inne obiekty w sensie pojęciowym, posiadające inne atrybuty i zasady reprezentacji w bazie danych, pozostające jednak w ścisłej relacji ze sobą.



**Rys. 8.** Przechowywanie w bazie danych w spójny sposób zarówno obiektów reprezentujących jezdnie, jak i reprezentujących osie dróg (dla dróg dwujezdniowych)

Kolejne przykłady obiektów, w przypadku których może być istotne przechowywanie kilku reprezentacji geometrycznych obiektów, podano poniżej:

- lotnisko, lądowisko: powierzchnia pasa startowego, oś pasa startowego, środek pasa startowego lub punkt przecięcia pasów startowych;
- stacja kolejowa: powierzchnia peronu, punkt reprezentujący peron (np. położony najbliżej budynku stacji);
- dworzec autobusowy: kompleks dworcowy, linie reprezentujące poszczególne stanowiska postojowe autobusów, punkt reprezentujący środkowe stanowisko postojowe autobusów;
- stacja benzynowa: kompleks stacji benzynowej, punkt reprezentujący zespół dystrybutorów;
- miejscowość: obszar wyznaczony granicą administracyjną, punkt w zależności od charakteru miejscowości umieszczony np. na głównym skrzyżowaniu lub w miejscu gdzie zlokalizowany jest rynek, budynek urzędu miasta lub gminy itd.

Podczas tworzenia kilku reprezentacji geometrycznych tego samego obiektu, niezwykle ważne jest ustalenie wzajemnych relacji między różnymi wersjami ich opisu geometrycznego. W przypadku obiektów punktowych na każdym poziomie uogólnienia może być zachowana dokładnie ta sama lokalizacja punktu. W niektórych sytuacjach może zostać jednak zmieniona i punkty reprezentujące obiekt na różnych LoD mogą być od siebie oddalone.

W przypadku obiektów powierzchniowych geometrie obiektu na różnych LoD rzadko będą się ze sobą pokrywały. Na kolejnych poziomach uogólnienia obiekt będzie często przewiększany, zwykle jednak będzie zawierał w sobie poligon reprezentujący obiekt na niższym poziomie. Oczywiście obiekt reprezentowany powierzchniowo na najdokładniejszym LoD może być reprezentowany na bardziej uogólnionym LoD przez punkt.

W przypadku obiektów liniowych na wszystkich poziomach uogólnienia powinny zostać zachowane współrzędne końców odcinków reprezentujących obiekt oraz współrzędne tych miejsc, które mogą być traktowane jako miejsca charakterystyczne, mające znaczenie semantyczne, które można traktować jako tzw. niezmienniki przestrzenne (np. ujście jednego cieką do cieką nadrzędnego). W takich miejscach przy tworzeniu bazy danych należy prowadzić segmentację linii. Punkty pośrednie (werteksy) linii mogą się różnić w przypadku różnych wersji geometrii danego obiektu. W innym podejściu można się starać, aby poszczególne punkty pośrednie otrzymywały atrybuty informujące o ich ważności w danej skali i przynależności do określonego LoD. Wymaga to jednak zastosowania skomplikowanych, niestosowanych modeli danych i technologii, a tym samym jest znacznie trudniejsze w realizacji.



## 5. Hierarchia pojęciowa obiektów w wielorozdzielczych bazach danych

W celu zrealizowania idei wielorozdzielczej bazy danych konieczne jest sklasyfikowanie obiektów w sposób hierarchiczny, tak aby poszczególne poziomy hierarchii odpowiadały różnym poziomom szczegółowości dla danego modelu danych.

Oznacza to dość nietypowe w odniesieniu do baz danych podejście, w którym w jednej bazie danych mogłyby współistnieć obiekty należące do klas na różnych poziomach hierarchii. Pomimo, iż klasa „Teren zadrzewiony” jest nadklasą w stosunku do klas „Las liściasty” i „Las iglasty”, to w jednej bazie danych mogą wystąpić obiekty zaklasyfikowane zarówno jako „Las liściasty”, „Las iglasty”, jak i „Teren zadrzewiony”.

W razie potrzeby i możliwości uszczegółowienia bazy danych na danym obszarze w trakcie jej rozwoju możliwe jest zaklasyfikowanie obiektów „Teren zadrzewiony” do niższego poziomu klasyfikacyjnego.

W przypadku tworzenia bazy na niższym poziomie szczegółowości możliwe będzie zrezygnowanie całkowicie z wprowadzania niektórych klas obiektów, np. niemożliwych do wprowadzenia bez przeprowadzenia aktualizacji terenowej lub mniej istotnych z punktu widzenia funkcji bazy danych, typu fontanna, głąz itp.

Podstawą modelu jest wyróżnienie hierarchii pojęciowej oraz hierarchii encji. Kompletna hierarchia encji reprezentuje przestrzeń w wielu skalach. Na konkretne potrzeby możemy wybrać elementy drzewa hierarchii do odpowiedniej „głębokości”.

## 6. Metadane

Opracowując model wielorozdzielczych i wieloreprezentacyjnych baz danych, należy zwrócić szczególną uwagę na zapewnienie odpowiedniej struktury dla metadanych. Prawidłowe wykorzystanie tych baz danych nie jest możliwe bez dostępu do informacji na temat dokładności i szczegółowości danych w poszczególnych fragmentach przestrzeni. Odpowiednia informacja o znaczeniu pojęciowym danej reprezentacji geometrycznej może być istotna dla użytkowników zarówno w procesach ekstrakcji wybranych informacji z bazy danych, w procesach analiz przestrzennych, jak i w procesie produkcji map z bazy danych.

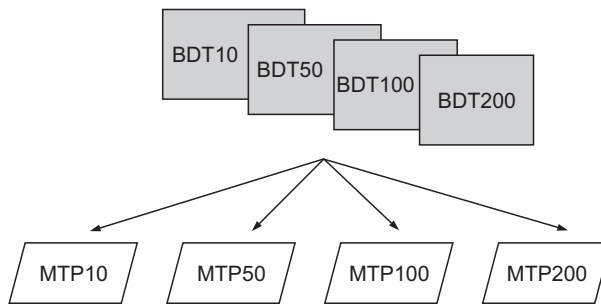
Wśród typowych atrybutów o charakterze metadanych można wymienić: sposób reprezentacji geometrycznej obiektu, dokładność danych, źródło danych, ale również w taki sposób można traktować numer LoD przypisany do obiektu.



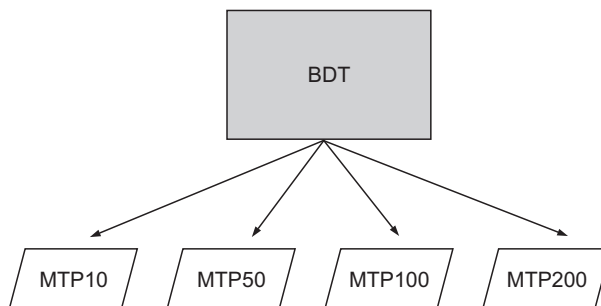
## 7. Aspekty wdrożeniowe w służbie geodezyjnej i kartograficznej w Polsce

Zarys pierwszych koncepcji dotyczących tworzenia bazy danych topograficznych w Polsce jako bazy wielorozdzielczej pojawił się w kilku opracowaniach, m.in. w projekcie badawczym prowadzonym na Politechnice Warszawskiej [3, 4], ekspertyzach wykonywanych na zlecenie Głównego Urzędu Geodezji i Kartografii, artykułach przedstawianych na szeregu konferencji, w tym ICA [7].

W opracowaniu [6] przedstawiono trzy potencjalne warianty podejścia do procesu tworzenia baz danych topograficznych w Polsce, spośród których dwa zilustrowano na rysunkach 9 i 10. Pierwszy z nich to baza danych wieloreprezentacyjnych, w której kilka zbiorów danych dla różnych poziomów skalowych (kilka reprezentacji geometrycznych obiektów), jest ściśle ze sobą powiązanych poprzez odpowiednie odwołania pomiędzy obiektami na poziomie bazy danych.



**Rys. 9.** Koncepcja bazy danych topograficznych o charakterze wieloreprezentacyjnym – kilka powiązanych baz danych o różnym poziomie generalizacji obiektów (oznaczenie BDT) umożliwiających generowanie map pochodnych w kilku skalach (oznaczenie MTP)



**Rys. 10.** Koncepcja „źródłowej” bazy danych topograficznych – przechowywanie w jednej bazie obiektów właściwych różnym skalom, o tym samym poziomie dokładności geometrycznej

W drugim przypadku zaproponowano opracowanie bazy danych topograficznych Polski na podstawie jednej bazy źródłowej, która zawierałaby wszystkie obiekty przedstawiane na dotychczasowych mapach topograficznych. Już na etapie tworzenia źródłowej bazy danych definiowane byłyby wszystkie elementy charakterystyczne opisu geometrycznego obiektów niezbędne do prezentacji obiektu w opracowaniach pochodnych (np. mapach w skalach średnich i małych). Warunkiem realizacji tego celu jest posiadanie wysoce dokładnych danych dla całego obszaru opracowania. W ten sposób można zbudować bazę danych, która zawiera wszelkie niezbędne informacje do wykonania opracowań pochodnych. Podejście to wydaje się najlepsze ze względów merytorycznych, ale z przyczyn pragmatycznych trudne jest obecnie do zrealizowania, ponieważ jest dość czasochłonne i wymaga dużych nakładów finansowych.

Rozwinięciem tej koncepcji była propozycja wprowadzenia dwóch poziomów szczegółowości bazy danych topograficznych [3, 4]. Została ona uszczegółowiona [5] w projekcie celowym No 6T12 2005C/06552 *Metodyka i procedury integracji, wizualizacji, generalizacji i standaryzacji baz danych referencyjnych dostępnych w zasobie geodezyjnym i kartograficznym oraz ich wykorzystania do budowy baz danych tematycznych*, w którym zdefiniowano drugi poziom uogólnienia bazy danych topograficznych (TBD2). Idea ta wpasowuje się w ideę tworzenia wielorozdzielczej bazy danych opisanej w niniejszym artykule (rys. 2b). We wspomnianym projekcie wykonano także szereg prac badawczych i wdrożeniowych przygotowujących do realizacji wieloskalowej bazy danych topograficznych (WBDT). Przyjęto, że pewne obiekty przedstawiane będą w różny sposób w różnych powiązanych z WBDT bazach danych (również na różnych poziomach skalowych), ale posiadać będą jedną definicję i będą identyfikowane w jednym wybranym zasobie danych, który można nazwać zasobem odniesienia. Na przykład podstawową bazą odniesienia w zakresie reprezentacji miejscowości jest państwowy rejestr nazw geograficznych (PRNG). W tym zasobie utrzymywany jest identyfikator miejscowości oraz współrzędne charakterystyczne miejscowości. W bazach ewidencji gruntów i budynków (EGiB) zapisana jest geometria granicy miejscowości (jednostki ewidencyjnej). W państwowym rejestrze granic (PRG) przedstawiony jest dokładny opis zasięgu granic administracyjnych. W przypadku miast będących gminami możemy mówić o ich przedstawieniu zarówno w EGiB, jak i w PRG. Miejscowość jest także reprezentowana w bazach tematycznych (np. SOZO i HYDRO) oraz w postaci punktowej w zgeneralizowanych bazach pochodnych. Wspólnym elementem łączącym te bazy powinien być identyfikator PRNG (zsynchronizowany z identyfikatorem GUS). Bazy zintegrowane z PRNG aktualne nazwy powinny pobierać z bazy PRNG w postaci identyfikatora PRNG. Jedna zdefiniowana pojęciowo w PRNG miejscowość może więc mieć różne reprezentacje geometryczne w różnych bazach danych przestrzennych: EGiB, TBD, bazy SOZO i HYDRO, GUS itd.

## 8. Wnioski

Jak dotąd nie opracowano od strony technicznej powszechnie dostępnych narzędzi ułatwiających zarządzanie bazami wielorozdzielczymi i wieloskalowymi. Nie ma jednak żadnych ograniczeń technicznych uniemożliwiających wykonanie tego typu oprogramowania.

Stosowanie omawianych rozwiązań może przynosić szereg korzyści:

- zapewnienie automatycznego zasilania (propagacji) bazy danych na wielu poziomach skalowych na podstawie aktualizacji modelu podstawowego,
- automatyzacja procesu zasilania danymi przestrzennymi systemów produkcji map topograficznych na różnym poziomie skalowym,
- zapewnienie wieloskalowej analizy danych przestrzennych.

Wdrożenie koncepcji wielorozdzielczych baz danych może mieć istotne znaczenie dla sukcesu programu budowy bazy danych topograficznych w Polsce ze względu na możliwość znacznie krótszego niż przy standardowym podejściu czasu dojścia do poziomu operacyjnego bazy danych pozwalającego na samofinansowanie jej dalszej realizacji.

## Literatura

- [1] Bobzien M., Morgenstern D.: *Abstracting Model Generalization: A Comprehensive Approach Towards Automated Generalization*. Materiały Międzynarodowej Konferencji Kartograficznej ICA, Durban (RPA) 2003.
- [2] Gotlib D.: *Możliwość wykorzystania baz danych o różnym stopniu szczegółowości do budowy krajowej infrastruktury danych przestrzennych*. Konferencja „Infrastruktura danych przestrzennych w Polsce i Europie – strategia, standardy, metadane i generalizacja”, Wrocław 2004.
- [3] Gotlib D.: *Modelowanie pojęciowe danych topograficznych*. [w:] Makowski A. (red.), *System informacji topograficznej kraju – teoretyczne i metodyczne opracowanie koncepcyjne*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2005.
- [4] Gotlib D.: *Możliwości zarządzania danymi topograficznymi na różnych poziomach uogólnienia*. [w:] Makowski A. (red.), *System informacji topograficznej kraju – teoretyczne i metodyczne opracowanie koncepcyjne*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2005.
- [5] Gotlib D.: *Opracowanie pełnej koncepcji bazy referencyjnej TBD na drugim poziomie uogólnienia (TBD2) jako elementu wielorozdzielczej referencyjnej bazy danych topograficznych (WTBD)*. Dokumentacja projektu celowego Nr 6 T 12 2005C/06552 *Metodyka i procedury integracji, wizualizacji, generalizacji i standaryzacji baz danych referencyjnych dostępnych w zasobie geodezyjnym i kartograficznym oraz ich wykorzystanie do budowy baz tematycznych*, 2008.

- 
- [6] Gotlib D., Olszewski R.: *Procesy generalizacji w ramach systemu informacji topograficznej – zarys koncepcji*. [w:] Makowski A. (red.), *System informacji topograficznej kraju – teoretyczne i metodyczne opracowanie koncepcyjne*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2005.
- [7] Gotlib D., Iwaniak A., Olszewski R.: *SDI in Poland – concept of topographic reference system for thematic, harmonized data bases*. Materiały Międzynarodowej Konferencji Kartograficznej ICA, La Coruna (Hiszpania) 2005.
- [8] Grünreich D.: *Development of Computer-Assisted Generalization on the Basis of Cartographic Model Theory*. GIS and Generalization – Methodology and Practice, Taylor & Francis, London 1995.
- [9] Grünreich D., Powitz B.M., Schmidt C.: *Research and Development in Computer-Assisted Generalization of Topographic Information at the Institute of Cartography*. Materiały Konferencji GIS, vol. 1, Hanover University, Monachium 1992.
- [10] Hampe M., Anders K., Sester M.: *MRDB Applications For Data Revision And Real-Time Generalisation*. Materiały Międzynarodowej Konferencji Kartograficznej ICA, Durban 2003.