

ARKADIUSZ BIŃCZYK

Uniwersytet Łódzki

WYBRANE ASPEKTY METODYCZNE REALIZACJI PROJEKTU BADAŃ PRZESTRZENNYCH

Abstract: Chosen Methodological Aspects on Realisation of Project of Spatial Research.

Spatially-oriented research projects are more and more popular. This is a natural consequence of constant development in GIS and information technology. Growing demand for knowledge in supporting every part of research process with digital data computing is impossible to deny. Computers are being used in all parts of research process – making research plan, data collecting and analysis, as well as visualization of results and finally desktop publishing for publication purposes. The article focuses on practical aspects of supporting spatially-oriented research projects in all of its parts. Author, after years of research and educational work in University of Lodz, tries to synthesize his knowledge and experiences in using digital information technology to serve the purposes of many research projects. Presented article is partially a case study. It shows some tasks that are necessary to realize in complex supporting of spatially-oriented research projects. Furthermore, the practical tasks, that were realized taking part in revitalization concept research project in Lodz has been briefly presented.

Wstęp

Współcześnie zauważa się w Polsce wzrastające zainteresowanie przestrzennym podejściem do badań. Wynika to m.in. z czynników w postaci decentralizacji władzy państwowej i rozwoju wolnorynkowej gospodarki w wyniku transformacji ustrojowej. Nowe warunki wymogły konieczność podjęcia problematyki zarządzania przestrzennego w nowych aspektach na różnych szczeblach administracji państwowej. Sektorowi prywatnemu dały natomiast możliwości rozwoju w warunkach konkurencji. Konkurencyjność wymaga obniżania kosztów. Można tego dokonać racjonalizując działania w przestrzennym kontekście. Sprzyja temu rosnąca dostępność wiedzy o GIS oraz narzędzi komputerowych wspomagających ogromną różnorodność aspektów przestrzennych zarządzania. Obecnie można brać pod uwagę aspekt przestrzenny, tam gdzie wcześniej nie było to możliwe z powodu ograniczeń w dostępności danych lub braku możliwości ich automatycznego przetwarzania.

W 2009 r. uruchomiono projekt badawczy, który podjął tematykę rewitalizacji wybranych terenów dużego miasta. Badania zorientowane na utworzenie spójnej koncepcji rewitalizacji terenów poprzemysłowych i powojkowych w Łodzi zainicjowano w 2009 r. (grant Prezydenta Miasta Łodzi). Projekt badawczy o interdyscyplinarnym charakterze podjęły równolegle trzy zespoły naukowe. Eksperci reprezentowali odmienne i uzupełniające się wzajemnie zakresy wiedzy specjalistycznej (m.in. gospodarka przestrzenna, urbanistyka, geografia miast, socjologia, GIS). Autor miał zaszczyt i przyjemność uczestniczyć w pracach zespołu ekspertów Instytutu Geografii Miast i Turyzmu na Wydziale Nauk Geograficznych Uniwersytetu Łódzkiego. Uczestnictwo obejmowało zadania związane z wykorzystaniem narzędzi, technik i metod cyfrowego zapisu oraz przetwarzania danych dla sprawnej realizacji badań.

Treść opracowania ma realizować dwa główne cele, znajdujące odbicie w układzie tekstu. Pierwszym z nich jest podjęcie próby zaprezentowania zadań, których realizacji wymagała kompleksowa obsługa projektu w zakresie implementacji narzędzi i technologii komputerowego przetwarzania danych. Zadania zaprezentowano wraz ze skrótowym opisem czynności wykonywanych w ramach każdego z nich. Kolejność omówienia jest zgodna z kolejnością realizacji zadań. Zakreślenie tej problematyki wydaje się zasadne z punktu widzenia wciąż rozwijającego się popytu na badania przestrzenne zorientowane aplikacyjnie. Im pełniejsza wiedza o zadaniach i czynnościach potrzebnych do realizacji projektów, tym dokładniej można przewidywać wymagania i potencjalne problemy, które mogą się pojawić w kolejnych etapach ich realizacji. Wymagania te powinny być możliwie dokładnie określone w odniesieniu do różnego rodzaju zasobów. Warto wspomnieć choćby o zasobach ludzkich (liczba osób, niezbędne kwalifikacje), czasie potrzebnym na realizację każdego etapu (jako części założonego ogólnie terminu realizacji projektu lub podstawy do ustalenia tego terminu), środkach finansowych czy zasobach materialnych w postaci np. specjalistycznego sprzętu. W efekcie pełniejszego przewidywania niezbędnych zasobów i potencjalnych problemów redukuje się ryzyko, a w drodze racjonalizacji gospodarowania zasobami ogranicza związane z nim koszty realizacji projektu.

Drugi cel pracy to zwrócenie szczególnej uwagi na zadanie zmierzające do utworzenia modelu danych, którego określenie powinno być realizowane w możliwie najpełniejszym zakresie już na etapie planowania badań. W toku pracy naukowo-dydaktycznej autor zajmował się problematyką wymagającą wiedzy z zakresu wykorzystania systemów informacji przestrzennej oraz geoinformatyki określanej jako dziedzina zajmująca się „zastosowaniem informatyki w dziedzinie informacji przestrzennej oraz integracji nauk o Ziemi” (Gaździcki 2006). Działalność w tym zakresie wyrażała się nie tylko udziałem w projektach badawczych, ale także udzielaniu bieżącej pomocy naukowcom, realizującym własne badania. Model danych to jeden z aspektów, które były najczęstszym przedmiotem pytań i wątpliwości zgłaszanych ze strony badaczy. Problemy z modelowaniem danych wynikają głównie z konieczności wykorzystania pewnej specyficznej wiedzy. Wiedza ta obejmuje pojęcia i metody funkcjonujące na

styku nauk o odmiennym charakterze. Przyjmując za podstawowe rozróżnienie nauk na formalne i empiryczne (Such 1987) można powiedzieć, że wspomniana specyfika polega na równoległym uwzględnianiu dorobku dziedzin, które wywodzą się z obydwu wspomnianych kategorii (informatyka, statystyka, geografia, socjologia, demografia *etc.*). Mając na względzie nieślabnące zapotrzebowanie na tego rodzaju „integrującą” nauki wiedzę, autor podjął próbę wskazania istotniejszych aspektów budowy modelu danych. W treści opracowania znalazły się zarówno odwołania do teorii (istotne pojęcia, podstawy teoretyczne), jak i praktyki – (odniesienie do przykładów z prac, w których autor uczestniczył). Całość opracowania można dodatkowo traktować jako uzupełnienie wyników pracy ekspertów o swego rodzaju raport z realizacji określonych prac.

1. ‘Cyfrowa’ realizacja projektu badań przestrzennych

Lata pracy naukowo-dydaktycznej w Uniwersytecie Łódzkim dały autorowi możliwość uczestniczenia w realizacji wielu projektów badawczych. Praca z ekspertami różnych dziedzin była niezwykle cenna z punktu widzenia poszerzania zarówno teoretycznej, jak i praktycznej wiedzy o wykorzystaniu cyfrowego przetwarzania danych w prowadzeniu badań, publikowaniu ich wyników (prace poligraficzne) oraz praktycznych aspektach organizacji prac zespołów badawczych. Tekst jest próbą dokonania pewnej syntezy zdobytych doświadczeń. Obsługa ‘komputerowej’ strony realizowania projektów badawczych zawsze obejmuje stały zakres zadań, niezależnie od szczegółowej tematyki. Przedmiot prezentowanego opracowania stanowią kategorie zadań, które są wspólne dla projektów badań przestrzennych: przygotowanie modelu danych (z uwzględnieniem danych przestrzennych), organizację badań terenowych, analizowanie danych, wizualizację wyników badań oraz przygotowanie materiałów do publikacji i sporządzenie raportu z wykonanych prac. Skrótowy opis wybranych zagadnień w ramach każdej z kategorii zadań zamieszczono poniżej.

1.1. Opracowanie modelu danych

Trudno obecnie wyobrazić sobie prowadzenie badań bez wykorzystania możliwości, oferowanych przez nowoczesne technologie. Zanim więc rozpocznie się gromadzenie danych powinno się zastanowić, jak opisać badaną rzeczywistość w postaci cyfrowej. Są to m.in. pytania: „Jak uporządkować wszystkie zgromadzone dane w pamięci komputera?”, „Jaki sposób cyfrowego zapisu będzie wskazany z punktu widzenia założonej koncepcji badań?”. Drugie pytanie rodzi inne, szczegółowe: „Jaki sposób zapisu będzie najbardziej odpowiedni z punktu widzenia gromadzenia danych za pomocą przyjętych metod i narzędzi badawczych?”, „Jaki sposób zapisu danych pozwoli wykorzystać programy komputerowe, które umożliwią wykonanie stosownych analiz i wizualizacji wyników?”. Pytania te są niezwykle ważne dla sprawniej realizacji

wszystkich kolejnych etapów badań. Często jednak nie bierze się ich pod uwagę przygotowując plan badawczy. Ilość danych, które trzeba zgromadzić i przetworzyć w badaniu otaczającej nas współcześnie rzeczywistości, daje właściwie gwarancję, że bez użycia komputera zrealizowanie projektu badań w założonym czasie będzie niemożliwe. Pytania o model danych pojawiają się zatem samoistnie i nie da się ich uniknąć. Pominiecie ich w trakcie planowania tylko odsuwa problem w czasie i to na mniej dogodny moment. Na podstawie własnych doświadczeń autor mógłby pokusić się o określenie tego momentu. Jest nim zazwyczaj chwila, w której badacz ma już zgromadzone dane. Świadomość konieczności wykorzystania komputera do badań powoduje, że dane zostały już w najlepszej wierze sprowadzone przez badacza do postaci cyfrowej (lub od razu w takiej postaci pozyskane). Okazuje się jednak, że zapis cyfrowy ma postać dość przypadkową i niejednorodną z punktu widzenia całości zbioru. Brak uporządkowania według kryteriów innych niż tylko „cyfrowość” zapisu powoduje, że w pewnym kontekście nie do końca można taki zbiór nazwać strukturą. Wykorzystanie zasobu w określonych komputerowych narzędziach analizy napotyka więc poważny problem. Aby wykorzystać możliwości konkretnych programów komputerowych należy dostarczyć im dane w odpowiedniej postaci. Jeżeli nie uwzględniono tego odpowiednio wcześniej zbiór danych okazuje się nagle całkowicie nieużyteczny. Pozbawia nas to możliwości wykonania planowanych analiz i (w konsekwencji) dalszych etapów badania. Perspektywy postępowania są dwie – spełnić wymagane warunki albo zrezygnować z przyjętych założeń realizacji badań. Wybieramy oczywiście opcję pierwszą – zapewnienie zgodności z wymaganiami odnośnie do danych. Pociąga to za sobą koszty w postaci nakładów naszego czasu i pracy lub ew. kosztów opłacenia specjalisty, który odpowiednio przygotuje dane. Zależnie od ilości danych, wymagań konkretnych narzędzi przetwarzania komputerowego co do postaci danych oraz – w konsekwencji – stopnia stwierdzonej niezgodności koszty te mogą się okazać czasem zadziwiająco wysokie. W pozytywnym scenariuszu udaje się ostatecznie przygotować dane – badania zostają zrealizowane z mniejszym lub większym naruszeniem założeń kosztorysu lub harmonogramu projektu. Jest też scenariusz mniej pozytywny – brak możliwości sprowadzenia danych do wymaganej postaci zapisu z jednoczesnym zachowaniem stopnia szczegółowości niezbędnego dla realizacji badań. Wiarygodność wyników może znaleźć się wówczas poniżej wymaganego poziomu. W tej sytuacji wracamy do etapu gromadzenia danych – trzeba pozyskać taki zbiór, który spełni wymagania badań.

1.2. Inwentaryzacja terenu

W badaniach o charakterze przestrzennym są wykorzystywane metody pozwalające na uzyskanie informacji o charakterze pierwotnym. W praktyce często odbywa się to w drodze inwentaryzacji terenu. Realizacja badań wymaga zazwyczaj udziału większej liczby osób, które będą pozyskiwać dane. Ważne jest zapewnienie odpowiedniej organizacji i przygotowania zarówno merytorycznego, jak i technicznego. Dotyczy to

osób organizujących oraz wykonujących inwentaryzację. W przypadku badań zespołu reprezentującego Wydział Nauk Geograficznych, wybrane tereny przemysłowe inwentaryzowało ok. 50 osób – studentów Wydziału z kierunku *Geografia i Gospodarka Przestrzenna*. Czynności przygotowawcze dotyczyły głównie opracowania materiałów roboczych dla celów pozyskiwania danych w terenie. Materiał wyjściowy stanowiły zarówno wzorce różnego rodzaju formularzy (wzór karty inwentaryzacyjnej budynku, wzór kwestionariusza), jak i opracowania kartograficzne (mapy każdego z inwentaryzowanych obszarów, podlegające aktualizacji w terenie). Ponadto, warto przygotować materiały informacyjno-szkoleniowe o charakterze ogólnym (wspólne dla wszystkich ekip). Obejmowały one: mapy całego zakresu przestrzennego badań z delimitacją poszczególnych obszarów badań i informacją o przydziale każdego z obszarów do konkretnej grupy (numeracja grup na mapie). Niezwykle ważne są instrukcje opisujące gromadzenie danych w zakresie wymogów zapisu. Chodzi o wymogi, które zostały uprzednio ustalone na etapie budowy modelu danych. Dane z terenu dostarczone w postaci zgodnej z wytycznymi modelu mogą być „od ręki” wprowadzone do gromadzonego zbioru i poddane wizualizacji (np. dla wstępnej oceny jakości) oraz poddane dowolnym z przewidzianych metod przetwarzania. Dzieje się tak dzięki temu, że odgórnie ustalono standard, którego trzymają się wszystkie. Nawiązano do tzw. obserwacji standaryzowanej, którą wykonuje badacz według ustalonej instrukcji i sposobu zapisu wyników (zapis jest kodowany).

Prace terenowe grup rozpoczęto spotkaniem szkoleniowo-organizacyjnym. Omówiony został harmonogram prac (terminy na wykonanie prac oraz sposoby i terminy przekazania materiałów do odbioru). Ponadto, przeprowadzono szkolenie w zakresie wykorzystania otrzymanych materiałów roboczych (zapoznanie z materiałami roboczymi, szkolenie w zakresie poprawnego wypełniania dokumentów). Bardzo ważne jest zapoznanie wszystkich obecnych z określonymi wymogami technicznymi. Należy także dokonać omówienia niezależnie od tego, że jest ono zamieszczone w instrukcji technicznej. Najbardziej karygodnym błędem jest przeświadczenie o doskonałości własnej pracy i przekonanie, że przewidzieliśmy wszystko. Zawsze w trakcie omówienia wytycznych pojawiają się dodatkowe pytania i wątpliwości ze strony grup terenowych, a uwagi te są czasem nadzwyczaj cenne i pozwalają na wykluczenie nieprzewidzianych wcześniej problemów. Każda grupa zdawała materiał z terenu w formie analogowej i cyfrowej. Pozyskiwanie danych w terenie bezpośrednio do urządzeń elektronicznych nie jest jeszcze aż tak powszechne. Wynika to zresztą nie tylko z możliwości finansowych zamawiającego czy organizatora badań. Wchodzą tu w grę inne czynniki, choćby obawy ze strony badających teren przed utratą sprzętu w wyniku np. kradzieży. Wersja cyfrowa musi spełniać założenia i wymogi określone przyjętym modelem danych. Materiał analogowy służy przede wszystkim do weryfikacji w przypadku wątpliwości odnośnie do cyfrowej wersji. W miarę realizacji prac następuje odbiór danych z terenu połączony z weryfikacją ich zgodności z określonymi wcześniej wymogami. Przeprowadza się ponadto kontrolę pod kątem innych

kryteriów jakości danych. Ostatecznie (po ewentualnych korektach i uzupełnieniach) wprowadza się dane do zasobu.

1.3. Wykorzystanie modelu danych dla zaspokojenia potrzeb informacyjnych ekspertów dziedzinowych

Na początku tego etapu sporządzono opracowania kartograficzne z informacją o jakości danych z terenu. Chodziło głównie o czytelne zaprezentowanie ekspertom pokrycia badanych terenów danymi o własnościach klas encji. Przykładem może być mapa w postaci kartogramu, obrazująca dla każdego terenu przemysłowego odsektów budynków, dla których udało się pozyskać dane o dominującej funkcji.

Poza tym etap obejmuje głównie wykonanie różnego rodzaju analiz, w tym oczywiście przestrzennych (programy GIS). Trudno tutaj dokonać omówienia, gdyż temat analizy danych jest tak obszerny, że stanowi odrębny przedmiot niejednego opracowania w postaci książkowej. Wspomnieć jedynie należy, że na tym etapie wykonuje się wszelkie zestawienia i analizy, których wyniki są, lub mogą być potrzebne ekspertom w budowaniu wiedzy. Warto na tym etapie ponownie przedyskutować we wspólnym gronie problemy, które są przedmiotem badań i wspólnie zrewidować listę potrzeb informacyjnych oraz sporządzić listę wszystkich, nawet tylko potencjalnie przydatnych analiz. Każda uwaga może przyczynić się do pogłębienia wiedzy w zakresie problematyki badań. Wyniki analiz są ostatecznie przekazywane ekspertom do interpretacji i dalszych prac merytorycznych.

1.4. Opracowanie techniczne wyników prac badawczych

Są to m.in. czynności w zakresie przygotowania stosownej wizualizacji ostatecznych wyników badań (wnioski, ekspertyzy, opinie zespołu ekspertów). Ponadto, często obejmuje przygotowanie techniczne materiałów na potrzeby poligrafii (przygotowanie ilustracji i tekstów do druku w sposób zgodny z wymogami redakcji technicznej/drukarni). Etap ten jest dość różnorodny, podobnie jak poprzedni. Wszystko zależy od szczegółowych potrzeb podyktowanych wynikami badań i potrzebami przekazu wyników. Czynności w zakresie przygotowania do druku są dodatkowo uzależnione od wymogów redakcji technicznej. Jeżeli publikacja ma być przekazana w postaci gotowej do druku – wymogi określa wówczas drukarnia. Na ogół wymogi w zakresie poligrafii dotyczą zapewnienia wymaganego formatu plików (najczęściej drukarnie proszą o PDF). Format pliku musi również spełniać określone parametry, przynajmniej dotyczące rozdzielczości (zwłaszcza w przypadku ilustracji), odpowiedniego modelu kolorów (właściwym dla druku jest CMYK), zapewnienia poprawnej czerni tekstów (konwersja modeli kolorów i formatów plików nie zawsze daje w wyniku czerń 'czystą' czyli C:0; M:0; Y:0; K:100) oraz wielu innych, które nie będą tutaj omawiane. W efekcie otrzymuje się zestaw materiałów gotowych do publikacji drukiem.

1.5. Dokumentacja przebiegu prac

Ostatecznie wykonuje się opracowanie raportu z przebiegu projektu w zakresie wykonanych czynności. Raport ma formę stosowną do wymogów – w przypadku opisywanego projektu wymogi określała strona zamawiająca grant. Niezależnie od szczegółowych wytycznych raport powinien zawierać m.in.: opis metadanych¹. Powinien się tu znaleźć opis danych pierwotnych i wtórnych pod kątem: wykorzystanych źródeł, sposobów gromadzenia, oceny danych pod kątem jakości (np. aktualność, kompletność *etc.*). Raport może też opisywać strukturę zbioru danych pierwotnych. Opis może mieć charakter raczej ogólny (model konceptualny, czasami niektóre aspekty modelu logicznego)² lub też szczegółowy (model logiczny i fizyczny). Opis o charakterze ogólnym ma znaczenie dla oceny i ewentualnej weryfikacji wiarygodności wyników badań. Szczegółowy opis struktury zbioru danych jest natomiast wymagany zazwyczaj w specyficznych przypadkach. Przykładem może być projekt zamawiany, który jako przedmiot zamówienia wyraźnie określał nie tylko wyniki badań (np. ekspertyzy), ale i zgromadzone dane. Szczegółowe informacje o danych i sposobie ich organizacji i zapisu na poziomie fizycznym umożliwią wykorzystanie danych przez osoby nie biorące bezpośredniego udziału w badaniach. Raport uwzględnia także ogólny opis najistotniejszych z wykorzystanych metod analiz, wraz z ew. wzmianką na temat narzędzi komputerowych, za pomocą których wspomniane analizy wykonywano. Ponadto, mogą się w nim znaleźć informacje nt. opracowania wizualizacji wyników badań oraz ew. przygotowania materiałów do publikacji.

Podsumowując – poziom szczegółowości raportu zależy m.in. od tego, czy badania mają charakter jednorazowy, czy też mają być powtarzane – wówczas trzeba zapewnić możliwość wykonania porównań. Perspektywa przeprowadzenia tych samych badań w przyszłości powinna sugerować większą szczegółowość raportu dotyczącą tych aspektów badań, które są istotne z punktu widzenia możliwości późniejszego porównywania uzyskanych wyników.

2. Projektowanie i budowa modelu danych dla projektu badań przestrzennych

Każde badanie, w szczególności o charakterze naukowym, wykonuje się według założonej *a priori* koncepcji. Błędem jest przekonanie, że na etapie planowania nie trzeba uwzględniać rozważań nad cyfrową reprezentacją badanej rzeczywistości. Poniekąd pomijanie tego aspektu wydaje się być wynikiem naturalnej kolei realizowania koncepcji badań, w której zastosowanie metod komputerowej analizy danych

¹ *Metadane* to najprościej mówiąc te, które dostarczają informacji o danych.

² Pojęcia omówiono w części tekstu poświęconej budowie modelu danych.

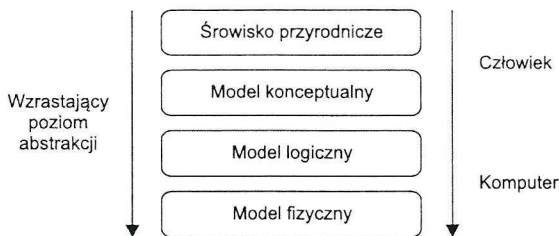
następuje dopiero po ich zgromadzeniu. Trudno się dziwić, że konieczność zorganizowania i opisu danych w tak wczesnym etapie projektu (przecież nie zaczęto nawet gromadzić danych) może nie być w pełni uświadomiona. Logika i zdrowy rozsądek nie musi przecież prowokować do rozważań o tym, jak ułożyć na półce książki, których nie mamy. Kiedy będziemy je mieć, wtedy będzie sens je układać. Podobnie nie musimy uchwycić potrzeby określenia struktury zbioru danych, których jeszcze nie posiadamy. Tutaj sytuacja jest jednak nieco bardziej złożona, a skutki przeoczenia poważniejsze. Wykorzystanie zaawansowanych narzędzi komputerowego przetwarzania danych wymusza odpowiednio wczesne przeprowadzenie tych rozważań, aby nie stanąć w obliczu problemu, o którym wspomniano w pierwszej części opracowania. Danych jest po prostu zbyt dużo, aby pozwolić sobie na początkową dowolność zapisu. Zmiana przyjętej raz postaci cyfrowej jest czasem bardzo kosztowna (jeśli w ogóle możliwa z zachowaniem wymaganego stopnia szczegółowości). Opracowanie modelu danych na etapie przygotowania projektu badawczego ma na celu utworzenie zasobu, który pozwoli na sprawne gromadzenie i przetwarzanie wszystkich danych, które uznano za istotne z punktu widzenia koncepcji badawczej. Zasób ten ma gwarantować możliwość zaspokojenia potrzeb ekspertów odnośnie do informacji, które są im niezbędne do budowania stosownej wiedzy na temat przedmiotu badań. Można powiedzieć, że taki zasób ma stanowić zarówno odwzorowanie potrzeb informacyjnych badacza na dane (etap przygotowania modelu), jak i podstawę wykonania procesu odwrotnego – uzyskiwania niezbędnych badaczom informacji z wykorzystaniem już zgromadzonych danych. Opisanie czym jest model danych, jak powstaje i czemu służy wymaga na wstępie zdefiniowania pewnych pojęć. użytymi już we wcześniejszej części tekstu określeniami są: *wiedza*, *informacja* i *dane*. Problematyka określania znaczeń i hierarchizacji pojęć mądrości, wiedzy, informacji i danych (ang. *DIKW* – *Data-Information-Knowledge-Wisdom*) często omawiana jest w odniesieniu do nauk o zarządzaniu wiedzą czy technologii informacyjnych (m.in. tematyka inżynierii systemów). Koncepcje i modele dotyczące definicji, określania relacji między wspomnianymi pojęciami oraz sposobów zapisu i przetwarzania opisano dość obszernie (patrz m.in.: Ackoff, Russel 1989; Davenport, Prusak 1998; Martin 2009; Syed, Shah 2006). Unikając nadmiernego rozbudowywania wykazu źródeł autor odsyła w tekście do kilku przykładów prac w omawianym zakresie. Zainteresowany czytelnik może w nich odnaleźć dalsze odwołania.

Jednym z określeń terminu *informacja* jest „znaczenie, jakie przy zastosowaniu odpowiedniej interpretacji przypisuje się wiadomościom” (Nowicki 1998). Wymienione źródło przyjmuje, że *wiadomość* to „każdy przekaz o określonym znaczeniu wysyłany i odbierany w świecie materii ożywionej i nieożywionej”. Wiadomości zaczynają stanowić informację wtedy, gdy w drodze interpretacji zostanie im przypisane konkretna treść. *Dane* natomiast, to „każda wiadomość lub wiadomości będące przedmiotem przetwarzania metodami tradycyjnymi lub automatycznymi”. Przedstawione określenia nawiązują do teorii budowy systemów. W odniesieniu do tematyki

badani z wykorzystaniem narzędzi GIS warto przytoczyć definicję Leksykonu Geomatycznego (Gaździcki 2002). Według tego źródła:

Dane to „reprezentacja informacji właściwa do komunikowania się, interpretacji lub przetwarzania. Dane występują w postaci znaków, (...) symboli, obrazów i innych – forma rejestracji zrozumiałych dla człowieka lub nadających się do przetwarzania komputerowego oraz transmisji. (...) Dane same w sobie nie mają znaczenia, dopiero w drodze interpretacji przez człowieka stają się informacją, która może być użyta do wzbogacenia wiedzy (...)”. *Informacja* to „wiedza uzyskiwana w zakresie interpretacji danych, która w ustalonym kontekście ma określone znaczenie i dotyczy obiektów, takich jak fakty, zdarzenia przedmioty, zjawiska, procesy i idee” a *informacja przestrzenna* to „wiedza uzyskiwana w drodze interpretacji danych przestrzennych”. Taka definicja informacji rozszerza znaczenie terminu w stosunku do poprzedniego podejścia – określa informację jako wiedzę. Definicji pojęcia samej wiedzy leksykon nie zawiera. Wydaje się jednak, że termin *wiedza* to nieco zbyt szeroka definicja, jak na potrzeby prezentowanego opracowania. Najwęższe z napotkanych przez autora określeń wiedzy to „ogół posiadanych informacji” (Syed, Shah 2006). W opracowaniach dotyczących znaczeń omawianych pojęć podkreśla się często odmiennosc tych znaczeń. Przykładem hierarchizacji terminów może być „piramida wiedzy” (Martin 2009). Najniżej w zakresie znaczeniowym stawia się sygnał (wiadomość). Zapis sygnału tworzy dane, które w drodze nadawania znaczenia (treści) przez interpretację przeradzają się w informacje, stanowiące z kolei punkt wyjścia do budowania wiedzy. Najwyższe miejsce w hierarchii przypisuje się pojęciu mądrości. Hierarchia pojęć zgodna jest z ujęciem Ackoffa: „dane stanowią pojedyncze symbole a przetworzone w sposób sprzyjający użytkownikowi i dające odpowiedź na pytania: „co?”, „kto?”, „gdzie?”, „kiedy?”, stają się informacjami. Wiedza stanowi aplikację informacji i koncentruje się na pytaniu „jak?”. Następny poziom stanowi odpowiedź na pytanie, „dlaczego?”, czyli rozumienie, które jest także funkcją komunikatywności, czyli jakości informacji. Najwyższy poziom procesu myślowego, według Ackoffa, stanowi mądrość, czyli ocena zrozumianych zasobów.” (Łukowska Mazur 2006).

Definicja pojęcia *Model danych* stanowi, że jest to: „1) abstrakcja świata realnego, która uwzględnia wybrane jego elementy i jest opisana danymi, 2) opis danych i ich powiązań odzwierciedlający strukturę informacyjną w danej organizacji lub dziedzinie, 3) wzorzec struktury danych w bazie danych” (Gaździcki 2002). Model danych można ogólnie opisać jako „uporządkowany cyfrowy opis służący do reprezentacji wybranych cech świata rzeczywistego”, inaczej „zestaw struktur, które służą do opisu i reprezentacji wybranych części rzeczywistości w postaci cyfrowej” (Longley *et al.* 2006). Badania otaczającej nas rzeczywistości nie są w stanie objąć wszystkich jej aspektów. Eksperci badając daną dziedzinę wybierają zatem te elementy rzeczywistości i takie ich własności, które są istotne w odniesieniu do określonego problemu. Inne, mniej istotne, są pomijane. Proces ten nosi miano abstrahowania. Określenie modelu danych odbywa się w drodze przechodzenia przez kolejne pozio-



Ryc.1. Poziomy abstrakcji związane z modelami danych GIS

Źródło: Longley *et al.* 2006.

my abstrakcji. Proces określania modelu danych w opisywanym projekcie zostanie odniesiony do trzech kluczowych poziomów abstrakcji, odpowiadających zarazem kluczowym etapom projektowania baz danych przestrzennych. Zobrazowany schemat (ryc. 1) pokazuje cztery poziomy, dołączając ‘środowisko przyrodnicze’ czyli rzeczywistość, podlegającą reprezentacji cyfrowej w systemach GIS.

Model konceptualny (pojęciowy) – „reprezentacja wybranych właściwości przestrzeni rozważań przy użyciu encji oraz zależności między encjami”, gdzie *encja* oznacza „każdy konkretny lub abstrakcyjny obiekt (byt) rozpatrywany w przestrzeni rozważań. stosuje się również nazwy *jednostka*, *całość*” (Gaździcki 2002). Encję może więc stanowić dowolny element, którego własności opisujemy lub mamy zamiar opisać za pomocą danych. Własność encji nazywana jest *atrybutem*. Atrybut to każda własność encji, która posiada swoją nazwę i dziedzinę. Przykładem encji może być zarówno obiekt fizyczny (np. budynek), jak i niematerialny w postaci pojęcia lub zdarzenia (np. konferencja). Atrybutem encji „budynek” może być liczba kondygnacji lub stan techniczny, atrybutem encji „konferencja” – temat lub organizator konferencji. Model konceptualny jest więc etapem, w którym określa się te obiekty, pojęcia lub zdarzenia, których cyfrowe reprezentacje będą potrzebne do realizacji celu badań. Ponadto, dokonuje się określenia atrybutów encji stosownie do potrzeb informacyjnych. Ostatecznie powstaje lista tzw. *klas encji*³.

Dodatkowo określa się jeszcze zależności między encjami (klasami encji). Polega to na opisie zależności w odniesieniu do tzw. *liczebności* i *uczestnictwa*. Liczebność określa ile encji (instancji danej klasy) uczestniczy w zależności między klasami. Przykłady liczebności to *jeden do jednego*, *jeden do wielu*, *wiele do wielu*. Określa się je dość łatwo na podstawie własności świata rzeczywistego. Odnosząc do przykładu: w jednym lesie (instancji klasy lasów) jest wiele drzew (instancji klasy drzew). Liczebność zależności lasy → drzewa będzie więc określona jako *jeden do wielu*.

³ Mianem *klasy* (także warstwą tematyczną – w odniesieniu do mapy cyfrowej) określa się najniższy poziom uporządkowania danych przestrzennych, którym jest zorganizowany zbiór danych tematycznych (encji jednego rodzaju, opisanych wierszami pojedynczej tabeli w bazie). Pojedynczy wiersz to encja (instancja klasy encji). W przypadku warstwy tematycznej GIS reprezentacją encji jest pojedynczy obiekt warstwy, np. jeden budynek, jeden teren poprzemysłowy.

Podobnie określa się uczestnictwo, definiujące status encji w danym związku – może być wymagane lub opcjonalne. Wracając do przykładu: uczestnictwo jest takie, że las musi zawierać drzewa więc uczestnictwo encji drzewo w encji las jest wymagane. Nie każda grupa drzew musi być lasem, zatem uczestnictwo rozpatrywane w odwrotnym kierunku jest opcjonalne. Do określania związków wykorzystywane są dodatkowe pola w tabelach, tzw. pola – klucze, zawierające identyfikatory instancji klas encji.

W odniesieniu do reprezentacji informacji przestrzennej (GIS) model pojęciowy określa ponadto ogólną koncepcję cyfrowej reprezentacji danych przestrzennych. Dwa podstawowe modele konceptualne cyfrowej reprezentacji danych przestrzennych odnoszą się do podziału na obiekty dyskretne i pola ciągłe⁴ (Longley *et al.* 2006).

Studium przypadku – wybrane czynności na poziomie konceptualnym

W opisywanym przykładzie przedmiotem badań zespołu były: podatność terenów przemysłowych Łodzi na procesy rewitalizacji oraz funkcje, których wprowadzenie w drodze rewitalizacji danego terenu byłoby pożądane. Przyjętą miarą podatności był potencjał tych terenów z punktu widzenia możliwości przeprowadzenia rewitalizacji oraz przewidywanej atrakcyjności inwestowania na danym terenie z punktu widzenia miasta, mieszkańców oraz potencjalnego inwestora. Celem badań było zatem określenie tego potencjału oraz opracowanie zalecanego na danym terenie katalogu funkcji. W etapie wstępnym eksperci dokonali określenia koncepcji badań własnych, połączonych wspólną tematyką projektu. Ze względu na odmienną specjalizację dziedzinową, a co za tym idzie nieco odmienną orientacji badań poszczególnych członków zespołu, zostały również sprecyzowane cel i przedmiot badań cząstkowych, potrzeby informacyjne ekspertów oraz metody i narzędzia badawcze, które miały być wykorzystane⁵ (tab. 1).

Bazując na ustaleniach w zakresie koncepcji badań i potrzeb informacyjnych ekspertów dokonano wyboru klas encji. Sprecyzowano ponadto własności, które będą brane pod uwagę oraz powiązania zachodzące między klasami. Wybrane ustalenia na poziomie konceptualnym zestawiono w tabeli 2. Należy zwrócić uwagę, że potrzebne są nie tylko dane o klasach encji bezpośrednio podlegających badaniu. Oprócz nich za istotne uznaje się również te, które będą wykorzystane w procesie analizy danych lub choćby wizualizacji wyników.

Do takich ‘pomocniczych’ klas zaliczono m.in. ulice, tereny kolei i tereny zielone. Reprezentacje siatki ulic były pomocne w procesie delimitacji badanych terenów (granice terenów badanych wyznaczył przebieg linii ulic), ocenie dostępności komu-

⁴ Model obiektów dyskretnych to reprezentacja za pośrednictwem obiektów o wyraźnie określonych granicach – obiekty istnieją w niektórych miejscach przestrzeni, w innych nie. Pola ciągłe to cyfrowa reprezentacja rzeczywistości wyrażonej zmienną, która może przyjmować wartość w dowolnym punkcie przestrzeni.

⁵ Treść tabeli nie ma za zadanie wyczerpywać tematyki przedstawianych w niej zagadnień. Autor przedstawia tylko przykłady, które mają zaprezentować praktyczny wymiar przejścia od koncepcji do modelu danych.

Tabela 1

Podstawowe informacje dotyczące koncepcji badań w omawianym projekcie

Eksperci	Sylvia Kaczmarek Jacek Kaczmarek	Ewa Szafrńska	Jolanta Latosińska
Dziedzina wiedzy eksperckiej	Urbanistyka, procesy rewitalizacji miast	Socjologia	Geografia miast, geografia turystyki
Zakres przestrzenny	Badanie o charakterze częściowym objęło wybrane tereny przemysłowe, głównie w wewnętrznej części miasta, ograniczonej linią kolei obwodowej. Eksperci wybrali zbiór obiektów do badań w drodze celowego doboru próby.		
Przedmiot badań eksperta	Ocena podatności terenów pofabrycznych w Łodzi na proces rewitalizacji. Wskazanie pożądanych dla miasta kierunków przemian funkcjonalnych, które proces rewitalizacji powinien realizować na badanych terenach.	Spółeczna percepcja rewitalizacji oraz wybranych terenów przemysłowych	Rozmieszczenie badanych terenów w kontekście: historycznego rozwoju, zabudowy mieszkaniowej oraz komunikacji miejskiej i miejskich terenów zielonych
Cel badań eksperta	Dokonanie oceny potencjału inwestycyjnego - atrakcyjności terenu z punktu widzenia możliwości jego rewitalizacji; Zaproponowanie pożądanych kierunków funkcjonalnych w postaci katalogu funkcji, które na danym terenie będą najbardziej odpowiednie (korzystne dla miasta/mieszkańców)	Rozpoznanie opinii mieszkańców Łodzi na temat rewitalizacji terenów przemysłowych zlokalizowanych w mieście.	Rozpoznanie czynników lokalizacyjnych oraz ocena walorów historycznych i/lub kulturowych terenów przemysłowych

Potrzeby informacyjne	Jakie jest położenie badanych terenów w przestrzeni miasta (strefa centralna – peryferia) Charakter morfologiczno-funkcjonalny terenów przemysłowych Jaki jest kontekst przestrzenny terenów przemysłowych (charakter funkcjonalny bezpośredniego otoczenia badanych terenów)	Czy mieszkańcy miasta dostrzegają potrzebę rewitalizacji terenów przemysłowych Kto w opinii mieszkańców powinien odpowiadać za rewitalizację badanych terenów Jakie funkcje badanych terenów byłyby najbardziej pożądane Czy i w jaki sposób mieszkańcy chcą partycypować w procesie rewitalizacji terenów przemysłowych Czy i w jaki sposób chcą być informowani o procesie rewitalizacji Jakie cechy społeczno-demograficzne determinują opinie na temat tego procesu	Jakie jest rozmieszczenie obiektów przemysłowych Czy/ i w obrębie których z badanych terenów znajdują się obiekty o szczególnej wartości historycznej i/lub kulturowej Czy /i w obrębie których z badanych terenów znajdują się obiekty figurujące w rejestrze zabytków Czy obszary przemysłowe mogą stanowić potencjał turystyczny
Metody, narzędzia, techniki badawcze	Inwentaryzacja urbanistyczna, inwentaryzacja kartograficzna (aktualizacja map roboczych), dokumentacja fotograficzna	Badania kwestionariuszowe wśród mieszkańców Łodzi, technika standaryzowanego wywiadu bezpośredniego	Inwentaryzacja standaryzowana, karta inwentaryzacyjna obiektów przemysłowych, karta inwentaryzacyjna obiektów zabytkowych

Źródło: Opracowanie własne na podstawie ustaleń zespołu ekspertów.

Wybrane ustalenia konceptualne

Eksperti	Sylwia Kaczmarek Jacek Kaczmarek	Ewa Szafrąńska	Jolanta Latosińska
Przedmiot badań	Ocena podatności terenów pofabrycznych w Łodzi na proces rewitalizacji. Wskazanie pożądanych dla miasta kierunków przemian funkcjonalnych, które proces rewitalizacji powinien realizować na badanych terenach.	Spółeczna percepcja rewitalizacji oraz wybranych terenów przemysłowych	Rozmieszczenie badanych terenów w kontekście: historycznego rozwoju, zabudowy mieszkaniowej oraz komunikacji miejskiej i miejskich terenów zielonych
Przykłady wybranych klas encji	Tereny przemysłowe Budynki Budowle Ulice Tereny kolei	Respondenci	Tereny przemysłowe Budynki Budowle Ulice Tereny kolei Tereny zielone
Przykłady instancji klas encji	Teren przemysłowy Budynek Budowla Ulica	Pojedyncza osoba poddana badaniu	Teren przemysłowy Budynek Budowla Ulica
Niektóre własności wybranych klas encji	Tereny przemysłowe: odległość od centrum miasta, odległość od innych terenów przemysłowych, odsetek powierzchni zabudowanej w powierzchni terenów przemysłowych) Budynki: dominująca funkcja, liczba kondygnacji, stan techniczny, powierzchnia budynku (obrys podstawy), identyfikator terenu przemysłowego, do którego budynek przynależy	Wiek, płeć, i inne dane z metryczki, Wariant odpowiedzi udzielonej na pytanie z kwestionariusza	Tereny przemysłowe: wartość historyczna, wartość kulturowa Budynki: wpis do rejestru zabytków, wpis do ewidencji zabytków

Źródło: Opracowanie własne.

nikacyjnej, oraz wizualizacji wyników (reprezentacja ulic i ich nazw na mapach). Tereny kolei stanowiły natomiast element pomocniczy na etapie wyboru próby do badań oraz wizualizacji wyników. Ocena położenia badanych terenów przemysłowych w stosunku do lokalizacji parków lub innych terenów zielonych była jednym z czynników branych pod uwagę w waloryzacji obiektów badań.

Model logiczny – Stanowi szczegółowy projekt struktur danych. Opisuje model pojęciowy w postaci schematu, który przedstawia encje wraz z przypisanymi im

atributami, opisuje typy danych, którymi będą opisane atrybuty oraz związki, które zachodzą między encjami należącymi do różnych klas. Typy danych determinują czy dane w bazie opisujące atrybut będą przechowywane jako liczby, tekst, data, wartość logiczna *etc.* Ponadto, w zakresie organizacji danych przestrzennych w GIS model logiczny określa przyjęte dla danych układy odniesienia przestrzennego oraz logiczny wymiar reprezentacji elementów w GIS. Warto rozróżnić dwa modele logiczne reprezentacji: rastrowy i wektorowy. Wynikają w dużym stopniu z koncepcji wybranych na poziomie modelu pojęciowego. Wektorowy model odpowiada obiektom dyskretnym a rastrowy – obiektom ciągłym. Unikając formalnej definicji można powiedzieć, że model wektorowy to taki, który w praktyce pracy badawczej w programie GIS będziemy widzieć w postaci dwóch rodzajów informacji: graficznej – prezentującej encje w postaci kształtów geometrycznych (punktów, linii lub wielokątów) oraz atrybutowej – w postaci tabeli. Tabela jest reprezentacją danych atrybutowych klasy encji. Pojedyncza encja to pojedynczy wiersz (zwany rekordem tabeli). Pojedynczy atrybut (wspólny dla klasy encji) to pojedyncza kolumna (zwana polem tabeli). W modelu rastrowym natomiast reprezentacja jest siatką kwadratów (pikseli), które prezentują pojedynczą wartość w postaci koloru. Przykładem rastrowego modelu są zdjęcia lotnicze i satelitarne powierzchni Ziemi (chodzi konkretnie o ortofotomapy – zdjęcia po odpowiednim przetworzeniu, obejmującym m.in. zorientowanie w określonym układzie odniesień przestrzennych). Decyzje i wybory w zakresie modelu logicznego są cenną informacją odnośnie do postaci danych, które będziemy wykorzystywać w badaniach – właśnie one decydują o tym, jak zorganizować zbiór danych, który zapewni możliwość sprawnej realizacji kolejnych po zgromadzeniu danych etapów badań.

Studium przypadku – wybrane czynności na poziomie logicznym

Kolejną czynnością w budowie modelu danych było wykonanie kategoryzacji danych z punktu widzenia kryterium planowanego wykorzystania danych w zasobie. W opisywanym przypadku określeniem było udzielenie odpowiedzi na pytanie, czy dane o klasie encji będą wykorzystane w postaci warstw tematycznych GIS, czy też z punktu widzenia planowanych badań nie ma takiej konieczności. Za takie dane, dla których nie stwierdzono takiej konieczności uznano dwie grupy danych. Pierwszą były zdjęcia cyfrowe, które miały być wykonane w ramach dokumentacji fotograficznej. Drugą grupą danych były dane z badań kwestionariuszowych – tutaj istotny był jednak aspekt atrybutowy. Pozostałe dane wymagały reprezentacji w postaci warstw tematycznych GIS.

Dane z dokumentacji fotograficznej terenu wymagały określenia struktury katalogów i konwencji nazewnicznej plików zdjęć. Celem postępowania było umożliwienie łatwego odnajdywania pojedynczych zasobów (plików zdjęć) dowolnego obiektu przez dowolnego użytkownika zasobu. Rozwiązanie problemu wymagało ustalenia pewnych reguł zapisu zdjęć. Każdy obiekt (badany teren pofabryczny, budynek, budowla) musi mieć swój własny kod (identyfikator). Kod nadaje się w taki sposób, aby

żaden inny obiekt nie miał takiego samego. Mówiąc inaczej – nie może być dwóch różnych budynków, budowli czy też terenów pofabrycznych o takim samym kodzie (mówimy w takim przypadku, że kod ma być unikatowy w obrębie jednej klasy obiektów). Struktura zapisu dokumentacji fotograficznej została zatem zaprojektowana według następująco przyjętych reguł:

- Cała dokumentacja dla danego terenu pofabrycznego znajdzie się w folderze nazwanym przyjętym identyfikatorem terenu. Przykładowo dla obiektu oznaczonego identyfikatorem 14, nazwa folderu również będzie brzmieć „14”.
- Zdjęcia budynków znajdują się w podfolderze „Budynki_foto”, zdjęcia budowli w podfolderze „Budowle_foto”.
- Nazwy plików fotografii muszą jednoznacznie wskazywać obiekt, który się na nich znajduje. Jest to konieczne, aby móc odnaleźć konkretny budynek wśród innych zdjęć – ktoś, kto nie był w terenie musi mieć możliwość odnalezienia zdjęcia według identyfikatora budynku/budowli zamieszczonego na mapie. Nazwę każdego zdjęcia tworzymy zatem według poniższego wzoru:

id budynku_trzycyfrowy kolejny numer fotografii

Przykład – pierwsze zdjęcie budynku oznaczonego na mapie kodem 102112 będzie nosić nazwę 102112_001. Kolejne zdjęcia tego budynku analogicznie: 102112_002, 102112_003... 102112_009, 102112_010, 102112_011 itd.

Dla danych ankietowych (kwestionariusz) należało określić strukturę kolumn tabeli i typy danych. Ustalono je stosownie do kryteriów określania struktury zapisu cyfrowego odpowiedzi kwestionariusza. Jednym ze szczegółowych kryteriów branych zazwyczaj pod uwagę przy wyznaczaniu struktury tabeli danych ankietowych jest podział pytań na otwarte, półotwarte i zamknięte. Ponadto, bierze się pod uwagę tzw. skale wartości atrybutów. Określają one sposób zapisu wartości cechy lub zjawiska – są to skale: nominalna, porządkowa, interwałowa, ilorazowa.

Znaczna część danych w postaci warstw tematycznych GIS z wektorową reprezentacją obiektów może (i powinna) być pozyskana jako dane wtórne. Trudno wyobrazić sobie, aby w ramach każdego badania przestrzennego dokonywać kartowania wszystkich obiektów badań i tworzyć dla nich od podstaw mapy cyfrowe elementów terenu, które uznano za istotne. W obrębie terenów pofabrycznych podlegających badaniom w opisywanym projekcie liczba samych tylko budynków wyrażała się w tysiącach – to daje pewien obraz nakładów pracy potrzebnych do utworzenia stosownego zasobu od podstaw. Zadanie można oczywiście uprościć, wektoryzując ortofotomapy lub skany map tradycyjnych. Metody tego typu są jednak również bardzo czasochłonne i często obciążone błędami niedokładności opracowanej reprezentacji obiektów. Ze wszech miar pożądane jest więc korzystanie ze źródeł istniejących. Są nimi przede wszystkim różnego szczebla ośrodki dokumentacji geodezyjnej i kartograficznej. Urzędy miast, gmin, powiatów czy województw również mogą być w posiadaniu określonych cyfrowych zasobów danych przestrzennych. Istnieją też inne źródła, któ-

re umożliwiają pozyskanie danych przestrzennych w formie cyfrowej, często nawet (zwłaszcza dla celów badań naukowych) nieodpłatnie. Można w nich znaleźć zasoby danych o różnorodnej treści, (np. granice podziałów administracyjnych, typy użytkowania ziemi), różnym poziomie szczegółowości (dokładność właściwa określonym skalom map) i różnych układach odniesień przestrzennych. Dokładniejsze omawianie źródeł cyfrowych danych przestrzennych nie jest jednak celem opracowania. Projekt badawczy w skali niedużych części jednego miasta wymaga dość szczegółowych danych. W tym przypadku najlepiej odwołać się jednak do zasobów geodezyjnych.

Model fizyczny – ostatni etap prac w zakresie wykonania cyfrowego zasobu danych. Model nawiązuje do wykorzystania danych z uwzględnieniem konkretnych standardów, określonego oprogramowania i określonych formatów plików. Model fizyczny jest (upraszczając) przeniesieniem modelu logicznego do konkretnego programu komputerowego. Prosty przykład: klasa encji ‘respondenci’ opisująca osoby ankietowane w postaci tabeli w programie Microsoft Access lub choćby Excel. Uwzględnia się tutaj dodatkowe parametry techniczne wynikające ze specyfiki oprogramowania. Nie są one jednak przedmiotem opracowania. Model fizyczny również warto utworzyć na etapie przygotowania do badań. Ważne jest to zwłaszcza w przypadku, gdy planowane są badania terenowe realizowane przez większą liczbę osób. Konstruując model fizyczny stworzymy standard dla wszystkich zespołów terenowych. Zapewni to ogromną oszczędność czasu choćby we wprowadzaniu danych po zgromadzeniu ich w terenie. Ponadto, pozwoli uniknąć problemu ze standardami kodowania zmiennych i wymagań odnośnie do technicznych warunków i standardów wprowadzania danych przez poszczególne grupy terenowe. Przykład problemu z brakiem standardu w nadaniach inwentaryzacyjnych: jedna z grup badających teren pozostawia puste wiersze w tabeli atrybutów budynków, gdy budynku z mapy roboczej nie było w terenie (wyburzony). Inna grupa przyjmie, że pustą wartością oznaczy te budynki, które były na terenie ogrodzonym i nie było można uzyskać danych o ich własnościach. Obydwa zestawy danych trafiają do celowo do jednego zbioru – jeśli nie zostaną wychwycone i skorygowane odpowiednio wcześniej staną się przyczyną poważnego obniżenia wiarygodności wyników.

Studium przypadku – wybrane czynności na poziomie modelu fizycznego

Na potrzeby opisywanego projektu wykorzystano dane przestrzenne w postaci warstw tematycznych reprezentujących: siatkę ulic, tereny kolei oraz budynki w badanych terenach. Warstwy zapisane w formacie *shapefile*, przetwarzano wykorzystując oprogramowanie ArcMap firmy ESRI. Jediną utworzoną od podstaw na potrzeby projektu warstwą tematyczną był zbiór reprezentujący tereny pofabryczne wybrane do badań – zostały narysowane jako obiekty wektorowe (powierzchnie) na podstawie delimitacji dokonanej przez ekspertów. Gdy otrzymujemy dane o charakterze wtórnym (np. warstwy tematyczne GIS pozyskane na etapie przygotowania badań) część modelu fizycznego zostaje w ten sposób zrealizowana. Fizyczna postać, w pewien sposób narzucona formatem otrzymanych warstw GIS może czasem wymuszać zachowanie pew-

nych standardów w pozostałych zestawach danych w zasobie. Wspomniano wcześniej, że dane przestrzenne z reprezentacją graficzną ulic, budynków *etc.* miały postać plików w formacie *shapefile*. Taka postać danych wymusiła przygotowanie warstwy GIS z reprezentacją terenów badań w tym samym formacie (*shapefile*) i w zgodnym układzie odniesienia przestrzennego. Realizacją fizycznego modelu danych było w projekcie utworzenie wektorowej warstwy tematycznej z terenami przemysłowymi oraz tabel wraz z wewnętrzną strukturą (nagłówki kolumn, nadanie typów danych, pola-klucze, identyfikatory obiektów *etc.*). Zakończenie prac na fizycznym poziomie abstrakcji owo- cuje gotową strukturą danych. Opracowanie i dokumentacja zasobu kończy pierwszy etap prac. Dane wtórne zgromadzone przed rozpoczęciem badań oraz wybrane części dokumentacji zasobu danych zostały wykorzystane w realizacji badań terenowych.

Decyzje w zakresie budowy modelu danych są determinowane przyjętą przez ekspertów koncepcją badań. Szczególnie istotne są: nadrzędny cel, przedmiot, zakres przestrzenny oraz przyjęte metody badań i narzędzia badawcze, którymi badacze będą się posługiwać. Prace należy więc zacząć od utworzenia opisu koncepcji badań. Stosownie do określonych poziomów abstrakcji można określić stopień wymaganego udziału wiedzy badacza w zakresie badań do udziału wiedzy o charakterze technicznym. W modelu conceptualnym największy jest udział wiedzy badacza. W miarę wzrostu poziomu abstrakcji, zwiększa się również poziom udziału wiedzy technicznej w określaniu modelu przy spadającym jednocześnie udziale wiedzy w zakresie konkretnego przedmiotu badań.

Literatura

- Ackoff, Russell L., 1989, *From Data to Wisdom*. Journal of Applied Systems Analysis, t. 16, s. 3-9.
- Davenport T. H., Prusak L., 1998, *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press.
- Gaździcki J., 2002, *Leksykon geomatyczny*. Wyd. „Wieś Jutra” Sp. z o.o., Warszawa.
- Gaździcki J., 2006, *Zakres tematyczny dziedziny geoinformacji jako nauki i technologii*. Roczniki geomatyki, t. IV, z. 2, PTIP, Warszawa.
- Longley P. A., Goodchild M. F., Maguire D. J., Rhind D. W., 2006, *GIS. Teoria i praktyka*. Wyd. Naukowe PWN, Warszawa.
- Łukomska-Mazur K., 2006, *Strategiczne zasoby informacyjne przedsiębiorstwa*. Świętokrzyskie Centrum Edukacji na Odległość, Zeszyty Naukowe 2/2006, s. 107-114.
- Martin J. N., 2009, *Knowledge Generation in the Enterprise Using Information and Data Systems*. Second International Symposium on Engineering Systems, Cambridge.
- Nowicki A., 1998, *Informatyka dla ekonomistów*. Wyd. Naukowe PWN, Warszawa.
- Such J., 1987, *Klasyfikacja nauk*, [w:] *Filozofia a nauka. Zarys encyklopedyczny*. Ossolineum, Wrocław, s. 297-305.
- Syed A., Shah A., 2006, *Data, Information, Knowledge, Wisdom: a Doubly Linked Chain?* International Conference on Information and Knowledge Engineering, s. 270-278.