



Komisja Europejska

## RAPORTY REFERENCYJNE JRC



# Model koncepcyjny rozwijania specyfikacji interoperacyjności w infrastrukturach danych przestrzennych

Katalin Tóth, Clemens Portele, Andreas Illert,  
Michael Lutz, Maria Nunes de Lima

**2012**

Raport EUR 25280 EN

Wspólnotowe Centrum Badawcze

**Komisja Europejska**

Wspólnotowe Centrum Badawcze  
Instytut ds. Środowiska i Zrównoważonego Rozwoju

**Dane kontaktowe**

Katalin Tóth

Adres: Joint Research Centre [Wspólnotowe Centrum Badawcze], Via Enrico Fermi 2749, TP 262, 21027 Ispra (VA), Włochy

E-mail: [katalin.toth@irc.ec.europa.eu](mailto:katalin.toth@irc.ec.europa.eu)

Tel.: +39 0332 78 6491

Faks: +39 0332 78 6325

<http://ies.jrc.ec.europa.eu>

<http://www.jrc.ec.europa.eu>

Niniejsza publikacja jest Raportem Referencyjnym Wspólnotowego Centrum Badawczego Komisji Europejskiej.

**Informacja prawna**

Komisja Europejska ani żadna osoba działająca w imieniu Komisji nie ponosi odpowiedzialności za ewentualne wykorzystanie niniejszej publikacji.

Europe Direct to serwis, poprzez który mogą Państwo uzyskać odpowiedź na pytania dotyczące Unii Europejskiej.

Bezpłatny numer telefonu (\*): 00 800 6 7 8 9 10 11

(\*) Niektórzy operatorzy telefonii komórkowej nie udostępniają połączeń z numerami 00 800 lub pobierają opłaty za takie połączenia.

Wiele dodatkowych informacji na temat Unii Europejskiej jest dostępnych w Internecie. Dostęp do nich można uzyskać z serwera Europa <http://europa.eu/>.

JRC69484

EUR 25280 EN

ISBN 978-92-79-22552-9 (pdf)

ISBN 978-92-79-22551-2 (druk)

ISSN 1018-5593 (druk)

ISSN 1831-9424 (on-line)

doi:10.2788/21003

Luksemburg: Urząd Publikacji Unii Europejskiej, 2012 r.

© Unia Europejska, 2012 r.

Kopiowanie niniejszego dokumentu jest dozwolone pod warunkiem powołania się na źródło.

*Drukowano we Włoszech*

# Model koncepcyjny rozwijania specyfikacji interoperacyjności w infrastrukturach danych przestrzennych

Tóth, Katalin<sup>1</sup>  
Portele, Clemens<sup>2</sup>  
Illert, Andreas<sup>3</sup>  
Lutz, Michael<sup>1</sup>  
Nunes de Lima, Vanda<sup>1</sup>

<sup>1</sup>Komisja Europejska, Wspólnotowe Centrum Badawcze, <sup>2</sup>interactive instruments Gesellschaft für Software-Entwicklung mbH, <sup>3</sup>Bundesamt für Kartographie und Geodäsie

**Słowa kluczowe:** infrastruktura danych przestrzennych, interoperacyjność, podstawowy model koncepcyjny, rozwój specyfikacji danych

## Najważniejsze informacje

Współcześnie informacje geograficzne są gromadzone, przetwarzane i wykorzystywane w tak zróżnicowanych obszarach jak hydrologia, łagodzenie skutków klęsk żywiołowych, statystyka, zdrowie publiczne, geologia, obrona cywilna, rolnictwo, ochrona przyrody i wiele innych. Wyzwania związane z brakiem dostępności, jakością, organizacją, dostępnością i współdzieleniem informacji przestrzennej są powszechne w wielu politykach i działaniach, a natknąć się na nie można na różnych szczeblach władzy publicznej w Europie.

**Dyrektywa 2007/2/WE** Parlamentu Europejskiego i Rady, przyjęta dnia 14 marca 2007 r., podejmuje środki dotyczące tych wyzwań poprzez ustanowienie Infrastruktury informacji przestrzennej we Wspólnocie Europejskiej (**INSPIRE**) dla polityk środowiskowych lub polityk i działań mających wpływ na środowisko. Ponadto Infrastruktury Danych Przestrzennych (Spatial Data Infrastructures, SDI) są w coraz większym stopniu powiązane i zintegrowane z systemami rozwijanymi w kontekście e-Rządu. Istotnym czynnikiem stymulującym tę ewolucję jest Agenda cyfrowa dla Europy, która rekomenduje "ustanowienie wspólnego wykazu usług transgranicznych, które umożliwią przedsiębiorstwom i obywatelom niezależnie funkcjonowanie lub działalność w całej UE" oraz "utworzenie systemów wzajemnego uznawania tożsamości elektronicznej"<sup>1</sup>.

Niniejszy raport odnosi się do kwestii umożliwienia i usprawnienia ponownego wykorzystywania informacji geograficznych i środowiskowych stworzonych i przechowywanych przez różne organizacje w Europie. Podstawowa trudność związana z tym zadaniem dotyczy zaradzenia niejednorodności danych oraz ustanowienia przepływu informacji pomiędzy wspólnotami wykorzystującymi informacje geograficzne w różnych obszarach środowiskowych.

Niniejszy raport prezentuje wszechstronny ogląd komponentu danych w SDI, ze szczególnym uwzględnieniem głównych cech ram koncepcyjnych. Przewidujemy, że niniejszy dokument będzie przydatny dla:

- Organów decyzyjnych odpowiedzialnych za rozwój strategiczny SDI, które chciałyby zrozumieć korzyści płynące z wykorzystania ram koncepcyjnych i pragną ocenić poziom złożoności oraz zasobów związanych z takim działaniem;
- Wiodących urzędników służby cywilnej z organizacji działających w państwach członkowskich, które są prawnie umocowane do wdrażania INSPIRE;
- Naukowców zainteresowanych szybkim, wszechstronnym przeglądem kluczowych elementów komponentu danych w SDI.

**Rozdział 1** zawiera wprowadzenie do Infrastruktur Danych Przestrzennych (Spatial Data Infrastructures, SDI) i sposobu ich rozwijania jako logicznej konsekwencji postępu technologicznego oraz związanych z nim wyzwań społecznych i technologicznych. Wraz z rozwojem technologii informacyjnej i komunikacyjnej tradycyjne mapy papierowe zostały wyparte przez cyfrowe informacje geograficzne oraz usługi oparte na położeniu geograficznym. Ta nowa technologia cyfrowa może przyczynić się do poszerzenia zakresu ponownego wykorzystywania informacji geograficznej, jednak proces ten utrudniają takie czynniki jak: niekompletność dokumentacji, niekompatybilność zbiorów danych przestrzennych, niespójność zbiorów danych, a także bariery kulturowe, językowe, finansowe i organizacyjne. SDI stanowią propozycję rozwiązań organizacyjnych i technicznych umożliwiających wyszukiwanie, znajdowanie i ponowne wykorzystywanie danych przestrzennych zgromadzonych przez inne organizacje.

Jedną z podstawowych koncepcji SDI jest **interoperacyjność**, która "oznacza możliwość łączenia zbiorów danych przestrzennych oraz interakcji usług danych przestrzennych bez powtarzalnej interwencji manualnej, w taki sposób, aby wynik był spójny, a wartość dodana zbiorów i usług danych przestrzennych została zwiększona"<sup>2</sup>. Inicjatywa INSPIRE, służąca jako główna inicjatywa w zakresie SDI, z której niniejszy raport czerpie przykłady i najlepsze praktyki, jest skonstruowana w oparciu o istniejące standardy, systemy informacyjne i infrastruktury, praktyki profesjonalne i kulturowe 27 krajów członkowskich Unii Europejskiej we wszystkich 23 językach urzędowych, a także ewentualnie w językach mniejszościowych UE.

**Rozdział 2** koncentruje się na informacjach geograficznych i wyszczególnia **wyzwania i rozbieżności**, z jakimi borykają się użytkownicy SDI podczas prób łączenia lub ponownego wykorzystywania danych pozyskanych z różnych źródeł. Wyzwania te mają swe pierwotne źródło w różnorodności definicji danych geograficznych jako częściowego abstrahowania rzeczywistości. Dane geograficzne, podobnie jak wszelkie inne dane, są zawsze abstrakcyjne, zawsze cząstkowe i zawsze stanowią tylko jedną z wielu możliwych perspektyw. W konsekwencji np. rzeki będą reprezentowane w formie wielokątów w jednym zbiorze danych lub w formie linii w innym; linie reprezentujące drogi po obu stronach granicy państwa niekiedy nie łączą się; woda zaś może sprawiać wrażenie płynącej do góry, gdy łączy się zbiór danych hydrologicznych i wysokościowych. Takie i inne wyzwania związane z ponownym wykorzystywaniem danych w SDI są zilustrowane i wyjaśnione w omawianym rozdziale.

Zasadniczą część raportu zawierają **rozdziały 3, 4 i 5**, w których znajduje się opis ram rozwijania specyfikacji danych z uwzględnieniem wielu ze wspomnianych wyżej trudności. Specyfikacje te określają cele interoperacyjności i pożądaną sposób przekształcania istniejących danych dla realizacji takich celów. Rozdział 3 dzieli się na dwie zasadnicze części. Obie z nich są w znacznym stopniu oparte na doświadczeniach i najlepszych praktykach INSPIRE:

- Podstawowy Model Koncepcyjny (Generic Conceptual Model, GCM) definiuje 25 aspektów lub elementów istotnych dla osiągnięcia interoperacyjności danych w SDI oraz proponuje metody i narzędzia postępowania z nimi. Obejmują one między innymi rejestry, systemy odniesienia za pomocą współrzędnych, zarządzanie identyfikatorami, metadane, przechowywanie.
- Opis metodologii rozwijania specyfikacji danych na potrzeby interoperacyjności obejmuje szczegółowe omówienie zaangażowanych podmiotów, kroków oraz ogólnego obiegu zadań - poczynając od rejestracji wymogów użytkowników aż po dokumentowanie i testowanie specyfikacji powstałych w tym procesie.

Oba podrozdziały wyjaśniają aspekty organizacyjne i techniczne ustalania komponentu danych SDI oraz przyczynki uzgodnień z zakresu interoperacyjności, standaryzacji i harmonizacji danych do tego procesu. Od 2005 roku program INSPIRE jest pionierem we wprowadzaniu, rozwijaniu i stosowaniu ram koncepcyjnych określania komponentu danych w SDI. To doświadczenie pokazuje, iż ramy koncepcyjne określone w niniejszym raporcie są wystarczająco solidne dla wsparcia interoperacyjności w obrębie 34 specyfikacji danych opracowanych dla SDI. Ponadto ze względu na fakt, że ramy te są niezależne od platformy i tematu, radzą sobie z różnorodnością kulturową i są oparte o przykłady najlepszych praktyk z Europy i spoza Europy, mogą one zapewniać rozwiązania problemów dotyczących SDI również w innych środowiskach.

<sup>1</sup> <http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/10/200&format=HTML&aged=0&language=EN&guiLanguage=en>

<sup>2</sup> Artykuł 3 ust. 7 dyrektywy 2007/2/WE (INSPIRE)

## Spis treści

Najważniejsze informacje .....	1
Słowniczek pojęć .....	5
Przedmowa .....	6
1 Infrastruktury danych przestrzennych – wprowadzenie .....	7
1.1 Od map do infrastruktury danych przestrzennych .....	7
1.2 Przykłady inicjatyw SDI .....	8
1.3 Interoperacyjność i harmonizacja danych .....	11
2 Dane przestrzenne .....	13
2.1 Od świata rzeczywistego do danych przestrzennych .....	13
2.2 Kwestie niekompatybilności i niespójności danych przestrzennych .....	16
2.3 Przedmiot SDI .....	19
3 Ramy koncepcyjne modelowania danych w SDI .....	20
4 Podstawowy model koncepcyjny .....	25
4.1 Zasadnicze elementy .....	25
4.1.1 Wymagania .....	25
4.1.2 Model referencyjny .....	26
4.1.3 Wsparcie interoperacyjności danych przez architekturę .....	26
4.1.4 Terminologia .....	27
4.1.5 Tekst wielojęzyczny i możliwości adaptacji kulturowej .....	27
4.1.6 Użycie ontologii .....	28
4.1.7 Odniesienia za pomocą współrzędnych i jednostki miary .....	28
4.1.8 Akta i rejestry .....	29
4.2 Modelowanie danych .....	30
4.2.1 Odwoływanie się do obiektów .....	30
4.2.2 Aspekty przestrzenne i czasowe .....	30
4.2.3 Zasady dotyczące schematów zastosowań i katalogów obiektów .....	31
4.2.4 Wspólne schematy zastosowań .....	32
4.2.5 Skonsolidowane repozytorium modelu .....	32
4.2.6 Wiele reprezentacji .....	33
4.2.7 Punkty rozszerzeń .....	34
4.3 Zarządzanie danymi .....	34
4.3.1 Zarządzanie identyfikatorami .....	34
4.3.2 Spójność danych .....	35
4.3.3 Jakość danych i informacji .....	36
4.3.4 Metadane .....	36
4.3.5 Zgodność .....	37
4.3.6 Zasady rejestracji danych .....	37
4.3.7 Model przekształceń danych/wytyczne .....	38
4.3.8 Zasady przechowywania danych .....	38
4.3.9 Prezentacja .....	39
4.3.10 Dostarczanie danych .....	39
5 Metodologia rozwijania specyfikacji danych .....	40
5.1 Definicja zakresu tematów danych .....	40
5.2 Zasady rozwijania specyfikacji danych .....	41
5.3 Cykl rozwoju specyfikacji danych .....	43
5.4 Utrzymanie specyfikacji .....	46
5.5 Względy dotyczące stosunku kosztów do korzyści .....	47
5.6 Podmioty uczestniczące w procesie specyfikowania danych .....	48
5.7 Narzędzia wsparcia .....	49
6 Wnioski .....	50
Podziękowania .....	52
Bibliografia .....	53

Słownik pojęć	
AFIS	Amtliches Festpunktinformationssystem (Urzędowy system informacji z punktów stałych)
ALKIS	Amtliches Liegenschaftskataster Informationssystem (Urzędowy system informacji katastralnych dot. nieruchomości)
ATKIS	Amtliches Topographisch-Kartographisches Informationssystem (Urzędowy system informacji topograficzno-kartograficznych)
ATS	Abstract Test Suite (Pakiet testów abstraktów)
BAG	Bathymetry Attributed Grid (Siatka z atrybutami batymetrycznymi)
DNF	Digital National Framework (Krajowe ramy cyfrowe)
KE	Komisja Europejska
UE	Unia Europejska
GBIF	The Global Biodiversity Information Facility (Światowa Sieć Informacji o Bioróżnorodności)
GEOSS	Global Earth Observation System of Systems (Globalny System Systemów Obserwacji Ziemi)
GCM	Generic Conceptual Model (Podstawowy model koncepcyjny)
GIS	Geographic Information Systems (Systemy informacji geograficznej)
GML	Geography Markup Language (Język danych geograficznych)
HY	Hydrografia, hydrologia
ICAO	International Civil Aviation Organisation (Organizacja Międzynarodowego Lotnictwa Cywilnego)
INSPIRE	Infrastructure for Spatial Information in Europe (Infrastruktura informacji przestrzennej we Wspólnocie Europejskiej)
ISO	International Standards Organisation (Międzynarodowa Organizacja Normalizacyjna)
KML	Keyhole Markup Language (Język znaczników danych przestrzennych)
NUTS	Nomenclature des unités territoriales statistiques (Nomenklatura jednostek terytorialnych dla celów statystycznych)
OGC	Open Geospatial Consortium (Otwarte Konsorcjum Geoprzestrzenne)
OWL	Ontology Web Language (Język ontologii sieci)
SDI	Spatial Data Infrastructure (Infrastruktura danych przestrzennych)
SI	Système international d'unités (Międzynarodowy system jednostek)
SLD	Styled Layer Descriptor (Deskryptor warstwy stylu)
SKOS	Simple Knowledge Organization System (Prosty system organizacji wiedzy)
SDI	Global Spatial Data Infrastructure (Globalna infrastruktura danych przestrzennych)
TAPIR	Taxonomic Databases Working Group Access Protocol for Information Retrieval (Protokół dostępu do uzyskania informacji Grupy roboczej ds. taksonomicznych baz danych)
TC	Technical Committee (Komitet Techniczny)
THREDDS	Thematic Real-time Environmental Distributed Data Services (Tematyczne rozproszone usługi danych środowiskowych w czasie rzeczywistym)
TIFF	Tagged Image File Format (Format plików graficznych)
UK	Wielka Brytania
UML	Unified Modelling Language (Ujednolicony język modelowania)
UTC	Universal Time Coordinates (Uniwersalny czas koordynowany)
XML	Extensible Markup Language (Język znaczników)
WMS	Web Mapping Service (Usługa mapowania Sieci)

## Przedmowa

Informacja geograficzna, infrastruktury danych przestrzennych (SDI), interoperacyjność i wspólne systemy informacyjne to pojęcia, z którymi stykać się mogą na co dzień programiści zajmujący się technologiami informacyjnymi i komunikacyjnymi, decydenci odpowiedzialni za informacje w sektorze publicznym, naukowcy, inżynierowie i urzędnicy służby cywilnej, pracujący w takich dziedzinach jak: hydrologia, łagodzenie skutków klęsk żywiołowych, statystyka, zdrowie publiczne, geologia, obrona cywilna, rolnictwo, ochrona przyrody, czy też jednej z wielu innych dyscyplin.

Czy powinni się niepokoić? Czy potrafią w prosty sposób stawić czoła wyzwaniu polegającemu na zapoznawaniu się z coraz obszerniejszą, rozproszoną, a niekiedy bardzo zaawansowaną technicznie dokumentacją? Czy można zrozumieć podstawowe koncepcje bez dogłębnej analizy polityki, aspektów organizacyjnych, obiegów zadań, czy wreszcie bez uprzedniej znajomości przedmiotu i odnośnej technologii?

O ile odpowiedź na pierwsze z tych pytań jest zdecydowanie twierdząca, o tyle jeśli chodzi o dwa pozostałe pytania, w przypadku większości osób będzie ona przecząca. Niniejszy raport stawia sobie za cel udzielenie odpowiedzi na te pytania poprzez wyjaśnienie podstawowych koncepcji i zasad, podsumowanie znaczenia interoperacyjności dla dziedziny informacji geograficznej oraz wykazanie, w jaki sposób SDI mogą odegrać kluczową rolę w radzeniu sobie z odnośnymi wyzwaniami. Wszystkie te kwestie zostaną wyjaśnione z punktu widzenia danych przestrzennych, z pobieżnymi wzmiankami o innych komponentach SDI,<sup>3</sup> wyłącznie w celu zilustrowania powiązań.

Czytelnicy, którym koncepcja SDI jest znana, mogą zadawać sobie pytanie, dlaczego tak wiele uwagi poświęca się komponentowi danych, jeśli jest to najprawdopodobniej komponent, w odniesieniu do którego osiągnięcie interoperacyjności jest najtrudniejsze. Oto kilka powodów:

- Komponent danych najlepiej ilustruje potrzebę interoperacyjności;
- Dane przestrzenne stanowią majątek gromadzony przez długi okres czasu przez wiele różnych organizacji. Organizacje te żywią uzasadnione obawy co do wpływu SDI na swoją pracę. Zrozumienie ducha interoperacyjności może przyczynić się do wyjaśnienia ewentualnych nieporozumień;
- Aktualnie użytkownicy informacji geograficznej poświęcają 80% czasu na składanie i zarządzanie informacjami, a tylko 20% na ich analizę w celu rozwiązywania problemów i uzyskiwania korzyści (Panel Informacji Geograficznej, 2008 r.);
- Psychologia ludzka: nazwa "Infrastruktura danych przestrzennych" implikuje dane jako przedmiot.

Na całym świecie działa wiele inicjatyw typu SDI. Autorzy, wszyscy aktywnie zaangażowani w INSPIRE, bezspornie zaczerpnęli najwięcej odniesień z tej właśnie inicjatywy. Pragną oni jednak podkreślić te cechy INSPIRE, które najprawdopodobniej będą istotne także w innych środowiskach, uzupełniając je o odniesienia do innych inicjatyw.

Podstawowym celem niniejszego raportu jest wyjaśnienie ram koniecznych do rozwijania modeli informacji i specyfikacji interoperacyjności w SDI bez nadmiernego skupiania się na szczegółach technicznych; raport pozwala "świadomemu twórcy polityki", dysponującemu podstawowymi umiejętnościami z zakresu informatyki, zrozumieć istotę tych ram. Jako pomoc dla czytelników, podstawowe definicje wyróżniono na zielono, natomiast przykłady umieszczono w polach jasnobrązowych.

## 1 Infrastruktury danych przestrzennych – wprowadzenie

### 1.1 Od map do infrastruktur danych przestrzennych

Fakty, niezależne dane i informacje, niezależnie od poziomu ich dokładności, nigdy nie osiągną tego samego efektu jak w przypadku umieszczenia ich w kontekście czasu i miejsca, które to stanowią najczęściej wykorzystywane odniesienia danych.

Od tysięcy lat prowadzenia obserwacji przestrzennych<sup>4</sup> produktem końcowym tych wysiłków były mapy, prezentujące w formie graficznej kontekst przestrzenny (Klinghammer, I. (1995). Starożytne mapy służyły do realizacji najważniejszych misji państwa: marynarki, odkryć i kolonizacji nowych terytoriów, podatków, prowadzenia wojen itp. Posiadanie map wiązało się z możliwością posiadania monopolu i zyskania luksusów. Po rozpowszechnieniu się nowoczesnej typografii niektóre popularne produkty, takie jak plany miast, mapy drogowe, mapy turystyczne i atlasy geograficzne weszły do powszechniejszego użytku.

Większość map pozostawała jednak dostępna tylko dla specjalistów. Każdy rodzaj mapy odpowiadał danej linii produkcji i zakresowi tematycznemu. Możliwości ponownego wykorzystywania tych map były ograniczone. Tylko mapy topograficzne zyskały szerszą popularność, ponieważ stanowiły formę ogólnej charakterystyki powierzchni Ziemi oraz geometryczną podstawę do tworzenia map tematycznych.

**Analiza przestrzenna** to proces uzyskiwania lub opracowywania nowych informacji poprzez modelowanie, oszacowanie, zrozumienie i ocenę zjawisk naturalnych oraz społecznych w kontekście położenia geograficznego.

Wraz z rozwojem technologii informacyjnych i telekomunikacyjnych tradycyjne mapy papierowe były stopniowo wypierane przez cyfrową informację geograficzną pochodzącą z digitalizacji map, satelitów do obserwacji Ziemi, lokalnych czujników cyfrowych i globalnych systemów pozycjonowania. Mapy papierowe są nadal wykorzystywane do celów wizualizacji, jednak komputery i inne urządzenia<sup>5</sup> stały się główną platformą analizy przestrzennej, projektowania technicznego i usług opartych o położenie geograficzne.

Systemy informacji geograficznej (GIS) to zintegrowane zbiory oprogramowania komputerowego i danych, służące do przeglądania i zarządzania informacją geograficzną w celu analizy relacji przestrzennych i modelowania procesów przestrzennych (Wade, T. i Sommer, S. (red.) 2006). Początkowe wdrożenia GIS poniekąd powtarzały kroki analogowego przetwarzania danych, wykorzystując dane zgromadzone konkretnie dla danego problemu wymagającego rozwiązania, a tym samym bez możliwości czerpania korzyści z potencjalnego ponownego wykorzystania danych cyfrowych.

Rozpowszechnienie się Internetu oraz umiejętności obsługi komputera spowodowały powstanie całkowicie nowego paradygmatu posługiwania się danymi przestrzennymi, opartego o promowanie współdzielenia danych w obrębie różnych społeczności i zastosowań. Ramy współdzielenia danych stanowią Infrastruktury danych przestrzennych (Spatial Data Infrastructures, SDI),<sup>6</sup> które można interpretować jako rozszerzenia komputerowego GIS (Craglia, M. (2010), w którym istnieje możliwość przeszukiwania, pobierania i wykorzystywania danych zgromadzonych przez inne organizacje, zgodnie ze ściśle określonymi zasadami polityki dostępu.

Zgodnie z podręcznikiem ("Cookbook") Stowarzyszenia Globalnej Infrastruktury Danych Przestrzennych - Global Spatial Data Infrastructure (GSDI) Association (Nebert, D. D. (red) 2004) "SDI zawiera dane geograficzne i atrybuty, wystarczający zakres dokumentacji (metadane), środki odkrywania, wizualizowania i oceny danych (katalogi i mapowanie sieciowe) oraz pewne metody zapewnienia dostępu do danych geograficznych. Ponadto istnieją dodatkowe usługi i oprogramowanie wspierające zastosowania tych danych. Aby infrastruktura SDI była funkcjonalna, musi obejmować także uzgodnienia organizacyjne niezbędne do jej koordynowania i administrowania w skali lokalnej, regionalnej, krajowej i/lub ponadnarodowej".

Opis GSDI klasyfikuje komponenty SDI jako dane, metadane, usługi (technologie) i uzgodnienia organizacyjne. Według Craglia i in. (2003), "Infrastruktury danych przestrzennych (SDI) zawierają w sobie polityki, uzgodnienia instytucjonalne i prawne, technologie oraz dane umożliwiające współdzielenie i efektywne wykorzystywanie informacji geograficznej". Ta definicja uzupełnia poprzednią o aspekt niezwyklej wagi, tj. efektywność wykorzystywania danych geograficznych, z którego wynika wymóg interoperacyjności.

Stopień rozwoju SDI jest ściśle skorelowany z rozwojem społeczeństwa informacyjnego w ogóle, wykorzystaniem technologii informatycznych przez społeczeństwo oraz z powszechnością Internetu. SDI może funkcjonować na szczeblu globalnym, ponadnarodowym, krajowym, regionalnym, transgranicznym lub lokalnym. W przypadku idealnym poziomy te są ze sobą powiązane i wzajemnie obejmują swoje istotne komponenty.

### 1.2 Przykłady inicjatyw SDI

Ustanowienie SDI wymaga współpracy wielu stron. Współpraca ta może być oparta o dobrowolne porozumienia pomiędzy zainteresowanymi stronami lub może być uregulowana w sposób bardziej formalny czy nawet narzucona prawnie, z zobowiązaniem docelowych organizacji do wypełniania przepisów

<sup>3</sup> Definicja i krótki opis SDI znajdują się w ust. 1.1.

<sup>4</sup> Historia kartografii jako nauki sięga czasów Eratostenesa i Ptolemeusza.

<sup>5</sup> Komputery osobiste i przenośne, telefony komórkowe oraz specjalistyczne urządzenia, np. wykorzystywane w marynarce, wyposażone są w aplikacje oparte o dane przestrzenne.

<sup>6</sup> Niekiedy SDI zwane są także "infrastrukturami informacji przestrzennej"; podkreśla się w ten sposób fakt, że zwykle zapewniają one dostęp do danych poprzez usługi wnoszące wartość dodaną. W niniejszym raporcie będziemy jednak stosować bardziej rozpowszechniony termin "infrastruktura danych przestrzennych".

aktów prawnych. Inicjatywy dobrowolne, np. GSDI oraz niektóre krajowe SDI, są często koordynowane przez stowarzyszenia międzynarodowe i krajowe lub organizacje typu parasolowego.

Według Longley i in. (2011), w literaturze opisano ponad 150 inicjatyw SDI. Poniższe przykłady dotyczą tylko tych inicjatyw, do których odnosimy się w kontekście niniejszego raportu. Dwie z tych inicjatyw funkcjonują na szczeblu globalnym, jedna na szczeblu krajowym i jedna na szczeblu ponadnarodowym w Unii Europejskiej.

#### GSDI

Global Spatial Data Infrastructure Association - Stowarzyszenie Globalnej Infrastruktury Danych Przestrzennych zostało założone w 1998 roku w celu "promowania współpracy i współdziałania międzynarodowego na rzecz rozwoju lokalnych, krajowych i międzynarodowych infrastruktur danych przestrzennych, umożliwiających poszczególnym narodom skuteczniejsze reagowanie na palące problemy społeczne, gospodarcze i środowiskowe". GSDI jako międzynarodowa organizacja, w której członkostwo jest dobrowolne, nie dąży do ustanowienia globalnej infrastruktury przestrzennej, a zamiast tego koncentruje się na podwyższeniu świadomości i wymianie przykładów najlepszych praktyk.

#### GEOSS

Global Earth Observation System of Systems - Globalny System Systemów Obserwacji Ziemi ma dostarczać narzędzia wspierające podejmowanie decyzji wielu różnym użytkownikom. GEOSS jako "system systemów" oparty jest o dotychczasowe obserwacje, systemy przetwarzania, wymiany i rozpowszechniania danych. Obejmuje on obserwacje lokalne, obserwacje z powietrza i z przestrzeni kosmicznej. W celu osiągnięcia interoperacyjności od podmiotów dostarczających informacje i dane oczekuje się przyjęcia niezbędnego poziomu koordynacji oraz uzgodnień technicznych, w tym między innymi specyfikacji gromadzenia, przetwarzania, przechowywania i rozpowszechniania współdzielonych danych, metadanych oraz produktów.

Interoperacyjność w GEOSS koncentruje się na obszarach rozdziału, dążąc do ograniczenia do minimum jakiegokolwiek oddziaływania na systemy składowe. W ramach dziesięcioletniego planu wdrażania (2005) GEOSS wykorzystuje istniejące komponenty infrastruktury danych przestrzennych w takich obszarach jak geodezyjne ramy odniesienia, wspólne dane geograficzne, standardowe protokoły. Zakres tematyczny GEOSS obejmuje "obszary korzyści społecznych" związane z klęskami żywiołowymi, ochroną zdrowia, energią, klimatem, rolnictwem, ekosystemami, bioróżnorodnością, wodami i pogodą.

#### UK Location Strategy - Brytyjska Strategia Lokalizacyjna

Brytyjska Strategia Lokalizacyjna została uruchomiona w 2008 r. Jej celem jest "maksymalizacja wykorzystania i korzyści dla społeczeństwa, rządu i przemysłu brytyjskiego z informacji geograficznej, a także stworzenie ram wsparcia inicjatyw europejskich, krajowych, regionalnych i lokalnych. Strategia stworzy infrastrukturę informacji o położeniach geograficznych, która będzie pomocna w procesach decyzyjnych dotyczących polityki, realizacji usług i decyzji operacyjnych" (Panel Informacji Geograficznej, 2008).

Dokument strategii zawiera galerię, w której informacje lokalne są stosowane w tworzeniu polityki publicznej, a także proponowane są działania strategiczne na rzecz lepszego wykorzystania informacji geograficznej. Zdefiniowano w nim również niewielką liczbę kluczowych zbiorów danych (Core Reference Geographies - Kluczowe geograficzne dane odniesienia), które stanowią będą wspólne ramy informacyjne definiowane, potwierdzane i wykorzystywane przez wszystkich posiadaczy danych w sektorze publicznym i prywatnym. Kluczowe geograficzne dane odniesienia obejmują ramy geodezyjne (w tym informacje o wzniesieniach terenu), nazwy geograficzne, adresy, ulice, własność gruntów i nieruchomości, hydrologię/hydrografię, granice statystyczne i granice administracyjne. W ramach Strategii Lokalizacyjnej zdefiniowano Krajowe Ramy Cyfrowe - Digital National Framework, DNF - jako mechanizm integrowania i współdzielenia brytyjskich informacji opartych o położenie geograficzne z wielu źródeł.

#### INSPIRE

INSPIRE to ważny przykład infrastruktury narzuconej prawnie. Dyrektywa INSPIRE Parlamentu Europejskiego i Rady (2007/2/WE z dnia 14 marca 2007 r.) ustanawia infrastrukturę informacji przestrzennej w Europie jako wsparcie dla polityki lub działań środowiskowych mogących mieć wpływ na środowisko.

Według Craglia (2011) INSPIRE charakteryzuje się pewnymi właściwościami, dzięki którym stanowi szczególne wyzwanie:

1. Infrastruktura jest budowana w 27 krajach członkowskich Unii Europejskiej, w ponad 23 językach<sup>7</sup>. Założenie to wymaga współistnienia i współpracy bardzo różnych systemów informacyjnych, praktyk profesjonalnych i kulturowych;
2. Zważywszy na ten poziom złożoności konieczne było przyjęcie procesu wypracowywania konsensusu, z udziałem setek ekspertów krajowych, w celu opracowania specyfikacji technicznych dla INSPIRE;
3. Istniejące normy należy przetestować w rzeczywistych, rozproszonych i wielojęzycznych warunkach;
4. Normy niedostatecznie dojrzałe lub takie, które pozostawiają zbyt dużo możliwości odmiennych interpretacji (ze względu na wdrożenie pod przymusem prawnym), należy udoskonalić;
5. Nieistniejące unormowania należy opracować;<sup>8</sup>
6. Należy odnieść się do niespójności i niekompatybilności danych oraz metadanych w obrębie 34 tematów mieszczących się w zakresie przedmiotowym Dyrektywy (por. Tabela 1).

Tematy danych INSPIRE podzielone są na bloki modułowe. \* Aneksy I i II koncentrują się na danych odniesienia, natomiast Aneks III skupia się na danych do analizy środowiskowej i oceny wpływu.

Aneks I	Aneks III
1. Systemy odniesienia za pomocą współrzędnych	14. Jednostki statystyczne
2. Systemy siatek geograficznych	15. Budynki
3. Nazwy geograficzne	16. Gleba
4. Jednostki administracyjne	17. Zagospodarowanie przestrzenne
5. Adresy	18. Zdrowie i bezpieczeństwo ludzi
6. Działki katastralne	19. Usługi użyteczności publicznej i służby państwowe
7. Sieci transportowe	20. Urządzenia do monitorowania środowiska
8. Hydrografia	21. Obiekty produkcyjne i przemysłowe
9. Obszary chronione	22. Obiekty rolnicze oraz akwakultury
	23. Rozmieszczenie ludności - demografia
	24. Gospodarowanie obszarem/strefy ograniczone/regulacyjne oraz jednostki sprawozdawcze
	25. Strefy zagrożenia naturalnego
	26. Warunki atmosferyczne
	27. Warunki meteorologiczno-geograficzne
	28. Warunki oceanograficzno-geograficzne
	29. Regiony morskie
	30. Regiony biogeograficzne
	31. Siedliska i obszary przyrodniczo jednorodne
	32. Rozmieszczenie gatunków
	33. Zasoby energetyczne
	34. Zasoby mineralne

Tabela 1: Tematy danych w INSPIRE

Dyrektywa nie wymaga gromadzenia nowych danych i nie nakłada na podmioty dostarczające dane obowiązku zmiany dotychczasowych obiegów zadań. Dzięki umożliwieniu interoperacyjności dane mogą być stosowane w sposób spójny niezależnie od tego, czy istniejący zbiór danych zostanie faktycznie przekształcony (zharmonizowany) trwale czy tylko tymczasowo przez usługę sieciową na potrzeby jego publikacji w INSPIRE.

Infrastruktura SDI przewidywana przez INSPIRE jest nadal w budowie. Proces legislacyjny stale ewoluuje, a Dyrektywa jest uzupełniana o "reguły wdrożeniowe", określające obowiązki państw członkowskich w konkretnym ujęciu technicznym i prawnym. Każdej regule wdrożeniowej towarzyszą wytyczne techniczne, które nie tylko zapewniają ogólne wsparcie implementacji, ale także mogą wyznaczać kierunki dalszego doskonalenia interoperacyjności.

Doświadczenie INSPIRE jest znaczące pod względem zasięgu i rezultatów. Obejmuje ono nie tylko nadzwyczaj dużą liczbę tematów danych i udział setek (a może i tysięcy) zainteresowanych organizacji na terenie Unii Europejskiej i poza nią, a nadto doprowadziło do zawarcia prawnie wiążących porozumień w

<sup>7</sup> <http://www.gsdi.org/>

<sup>8</sup> 23 języku urzędowe UE oraz języki mniejszościowe.

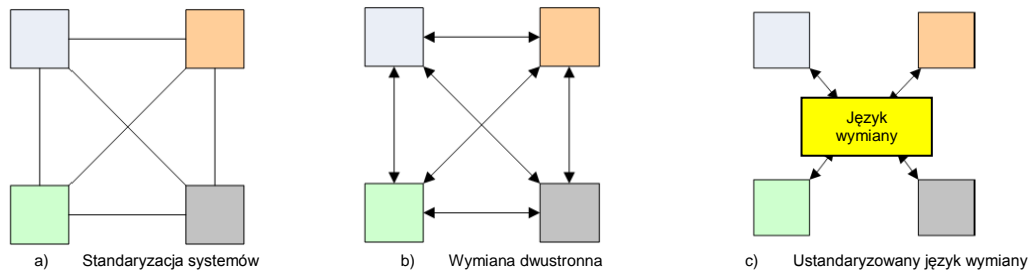
<sup>9</sup> Przykładowo, potrzebne są standardy usług "wywoływania" w przypadku łańcuchów usług lub specyfikacje docelowego poziomu interoperacyjności dla danych przestrzennych.

krajach członkowskich.

### 1.3 Interoperacyjność i harmonizacja danych

Cel efektywnego wykorzystywania wysuwa na czoło zagadnienie interoperacyjności. Zgodnie z 10-letnim Planem wdrożeniowym (GEOSS, 2005a) interoperacyjność dotyczy zdolności aplikacji do pracy w niekompatybilnych systemach.

Istnieją trzy podstawowe architektury systemów interoperacyjnych (Lasshuyt i van Hekken, 2001), przedstawione są one na Rys. 1.



Rysunek 1: Podstawowe architektury interoperacyjności (adaptacja na podstawie: Lasshuyt i van Hekken, 2001)

Jak uwidoczniono na Rys. 1a, gdy systemy są ustandaryzowane, komunikują się między sobą w sposób całkowicie interoperacyjny. W większości przypadków podejście takie jest nieefektywne, ponieważ każdy system jest tworzone zgodnie ze standardami, konwencjami lub najlepszymi praktykami konkretnej organizacji czy społeczności użytkowników.

W przypadku wymiany dwustronnej (Rys. 1b) wymagane są dedykowane interfejsy pomiędzy każdą parą połączonych systemów. Liczba interfejsów rośnie gwałtownie wraz z liczbą różnych systemów. Opcja trzecia (Rys. 1c) jest powszechnie uznawana za najbardziej praktyczne rozwiązanie na potrzeby interoperacyjności. Jest to elastyczny system systemów, do którego można dodawać nowe systemy bez konieczności adaptacji istniejących lub dodawania nowych interfejsów.

Choć brak jest jednoznacznej definicji systemu systemów, SDI z pewnością spełnia jego podstawowe kryteria (niezależność pod względem zarządczym i operacyjnym, rozwój ewolucyjny, zachowania emergentne, szeroki zasięg geograficzny). SDI łączy rozproszony geograficznie system różnych dostawców danych na poziomie lokalnym, regionalnym, krajowym, ponadnarodowym i globalnym. Każdy system działa niezależnie, zgodnie z zasadami lokalnymi; poszczególne systemy komunikują się ze sobą nawzajem z wykorzystaniem uzgodnionych standardów. Zgodnie z najlepszymi praktykami SDI powinny być ustanawiane i rozwijane etapowo, przy stale zwiększającej się liczbie uczestników i poszerzającym się zakresie. Zachowania emergentne (zdolność do wykonywania funkcji niezwiązanych trwale z komponentami) można stwierdzić poprzez bardziej efektywne podejmowanie decyzji, gdy informacja jest wbudowana w kontekst transgraniczny lub międzytematyczny.

Według INSPIRE, interoperacyjność "oznacza możliwość łączenia zbiorów danych przestrzennych oraz interakcji usług danych przestrzennych bez potrzebnej interwencji manualnej, w taki sposób, aby wynik był spójny, a wartość dodana zbiorów i usług danych przestrzennych została zwiększona". Definicja ta przesuwa punkt ciężkości ze sposobu interakcji pomiędzy systemami<sup>10</sup> na potencjalne korzyści dla użytkowników z eliminacji barier występujących powszechnie przy próbie łączenia danych z różnych źródeł.

W SDI interoperacyjność łączy ze sobą niejednorodność komunikujących się systemów na dwa sposoby:

1. Przekształcenie danych przestrzennych (z zastosowaniem technologii informacyjnych i komunikacyjnych); oraz
2. Harmonizacja danych zawartych w systemie.

Dane są przekształcane przez konkretne oprogramowanie w celu uzyskania ustandaryzowanej prezentacji danych. Przekształcenie może mieć miejsce on-line lub off-line. W procesie on-line dane są często przekształcane przez usługi sieci Web. W metodzie off-line tworzony jest interoperacyjny widok (kopia), który jest zapisywany z możliwością dostępu poprzez usługę pobierania. W obu przypadkach zachowywana jest pierwotna semantyka i struktura danych, aby spełnić pierwotne wymagania użytkownika, z myślą o których dane zostały stworzone.

Harmonizacja jest niezbędna, gdy rozwiązania techniczne nie zapełniają luki w interoperacyjności i konieczne są zmiany danych bazowych. Harmonizacja przybliża semantykę i strukturę danych oraz usuwa pozostałe rozbieżności, którym nie można zaradzić z wykorzystaniem dostępnej technologii. Zarówno rozwiązania z zakresu interoperacyjności, jak i harmonizacja prowadzą do standaryzacji uzyskanych informacji.

Harmonizacja danych to proces modyfikowania / precyzowania semantyki i struktury danych dla ułatwienia przestrzegania uzgodnień (specyfikacji, norm lub aktów prawnych) w skali transgranicznej i/lub pomiędzy społecznościami użytkowników.

Normy w dziedzinie geoprzestrzennej wprowadzane są głównie na poziomie krajowym i międzynarodowym. Komitet Techniczny (TC) 211 Międzynarodowej Organizacji Normalizacyjnej (ISO) oraz Otwarte Konsorcjum Geoprzestrzenne (Open Geospatial Consortium, OGC) definiują podstawy tworzenia informacji geoprzestrzennej, które muszą być spójne pomiędzy domenami. Normy ISO formułowane są we współpracy z krajowymi organami normalizacyjnymi, natomiast standardy OGC tworzone są przy wsparciu użytkowników i dostawców technologii. Obie organizacje gromadzą wiedzę na temat najwyższej klasy specjalizacji międzynarodowej, co usprawnia rozpowszechnianie tych norm na całym świecie.

Standaryzacja według tematów prowadzona jest w wielu organizacjach międzynarodowych, np. Międzynarodowej Organizacji Hydrograficznej (IHO), Pakcie Północnoatlantyckim (NATO), Światowej Organizacji Meteorologicznej (WMO) itp. Organizacje te współpracują ze sobą przy opracowywaniu interesujących wzajemnie zagadnień zarówno pod kątem formalnych procesów standaryzacji,<sup>11</sup> jak i inicjatyw SDI<sup>12</sup>, co prowadzi do dalszej konwergencji informacji geograficznej.

Poza wymienionymi wyżej standardami zgodnymi z treścią unormowań prawnych (de iure) przy osiągnięciu interoperacyjności można także brać pod uwagę najlepsze praktyki społeczności (normy faktyczne - de facto), np. GeoTIFF - fotografia z odniesieniami geograficznymi, GBIF i TAPIR - bioróżnorodność, THREDDS - dane środowiskowe czasu rzeczywistego, czy BAG - dane batymetryczne.

Ewidentne jest, że rozwiązania z zakresu interoperacyjności i harmonizacja danych współistnieją ze sobą w SDI; luka w interoperacyjności może być zapełniona tylko poprzez ich wzajemne wyważenie. Systemy interoperacyjne, niezależnie od zwiększonego potencjału efektywnego wykorzystania ponownego, muszą pozostać idealnie zdadne do celów, do których zostały stworzone.

## 2 Dane przestrzenne

### 2.1 Od świata rzeczywistego do danych przestrzennych

<sup>10</sup> Nie oznacza to, że INSPIRE pomija interoperacyjność systemów. Komponent usług sieciowych obejmuje także technologię informatyczną.

<sup>11</sup> [http://www.dgiwg.org/dgiwg/htm/activities/external\\_c\\_c.htm](http://www.dgiwg.org/dgiwg/htm/activities/external_c_c.htm)

<sup>12</sup> [http://www.iho.int/iho\\_pubs/CB/C-17\\_e1.1.0\\_2011\\_EN.pdf](http://www.iho.int/iho_pubs/CB/C-17_e1.1.0_2011_EN.pdf) i <http://www.ungiwg.org/contact.htm>



Dane przestrzenne to wszelkie dane mające bezpośrednie lub pośrednie odniesienie do określonego położenia lub obszaru geograficznego<sup>13</sup>. Informacja przestrzenna obejmuje dane przestrzenne o strukturze wynikającej z określonego celu. Poza opisywaniem położenia i rozkładu różnych zjawisk w naszym ziemskim środowisku informacja przestrzenna analizuje kontekst oraz relacje pomiędzy danymi przestrzennymi i innymi niż przestrzenne.

Należy zauważyć, że "każdy opis rzeczywistości jest zawsze abstrakcyjny, zawsze częściowy i zawsze stanowi tylko jedną z wielu możliwych perspektyw" (ISO TC 211 2005a). Zróżnicowane opisy (abstrakcje) prowadzą do wielości informacji dotyczących tego samego położenia geograficznego/przestrzennego. Proces abstrahowania może obejmować różne punkty widzenia, może być związany z różnymi punktami w czasie i może prowadzić do uzyskania różnych poziomów szczegółowości informacji na temat opisywanego obszaru<sup>14</sup>. Oto trzy metody prowadzące do

1. wiele widoków (widok wielotematyczny),
2. reprezentacje w wielu punktach w czasie,
3. reprezentacje w wielu skalach (rozdzielczości).

#### 1. Wiele widoków

Zależnie od kontekstu i punktu widzenia to samo zjawisko może być reprezentowane na różne sposoby. Każda społeczność podkreśla te właściwości zjawiska, które są istotne dla konkretnego obszaru lub zadania. Przykładowo, rzeka może być postrzegana jako część sieci hydrologicznej, środek transportu, część granicy państwa czy siedlisko chronionego gatunku. Każdy z tych opisów jest prawidłowy; odcinek rzeki jest ten sam, jednak zakres zgromadzonych danych i informacje uzyskane na podstawie tych danych są różne dla poszczególnych scenariuszy. Każdy punkt widzenia zarysowuje określone pole tematyczne. Określenie "temat danych przestrzennych" stosuje się często w odniesieniu do zbierania i klasyfikacji obiektów przestrzennych z tej samej perspektywy.

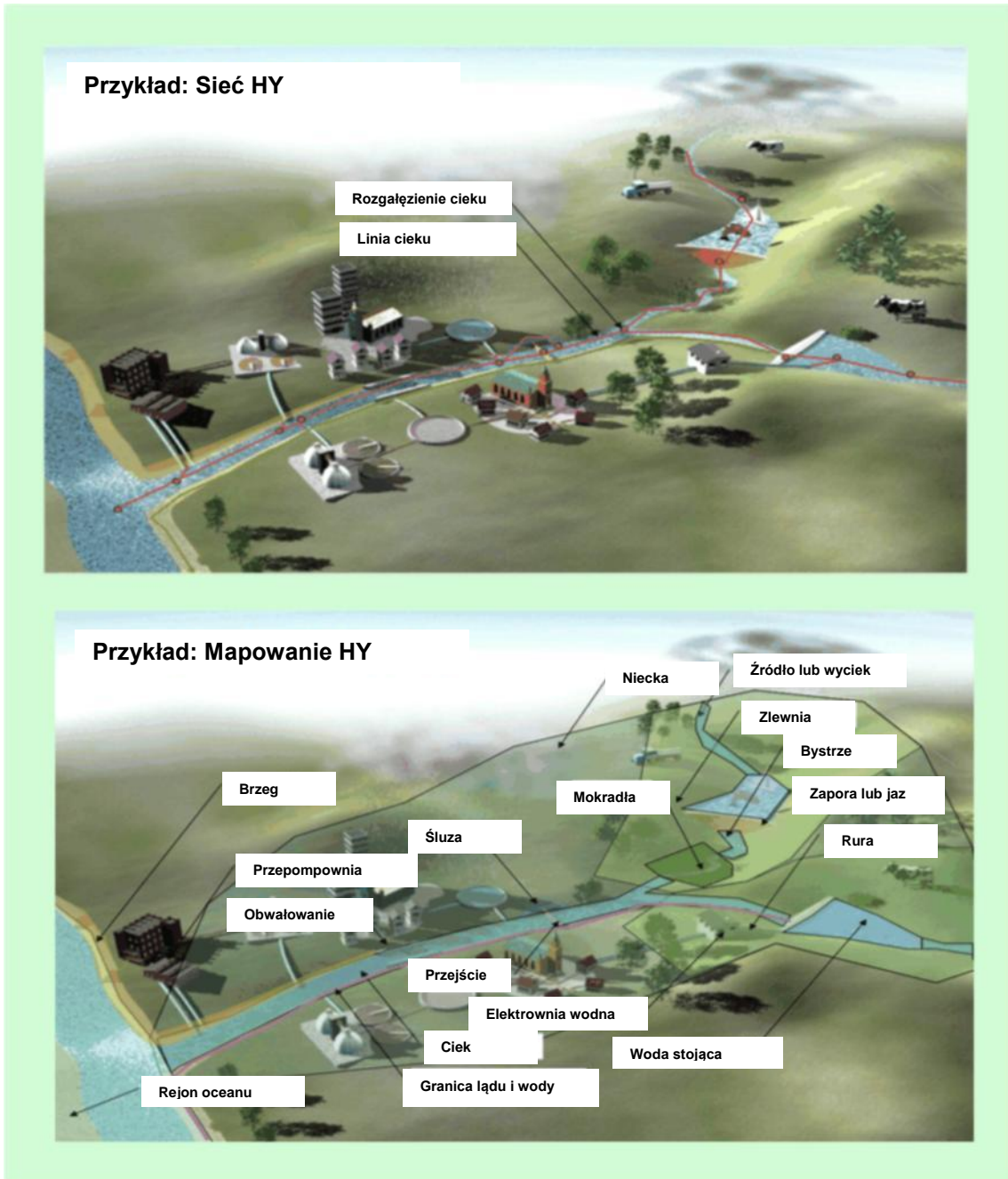
Dwa potencjalne widoki danych hydrologicznych uwidoczniono na Rysunku 2. "Sieciowy" widok hydrografii jest bardzo przydatny przy modelowaniu powodzi, natomiast widok "mapy" jest niezbędny przy planowaniu urządzeń technicznych.

Informacja geograficzna czy przestrzenna?  
Informacja geograficzna jest powiązana z konkretnym położeniem na powierzchni Ziemi. Informacja przestrzenna dotyczy położenia na powierzchni Ziemi (topografia), pod powierzchnią Ziemi (geologia) lub nad powierzchnią Ziemi (meteorologia). Ponadto dane przestrzenne mogą odnosić się do systemów lokalnych, będących niekiedy mikrosystemami (np. dane pochodzące z fotogrametrii bliskiego zasięgu).

Temat danych przestrzennych obejmuje wszystkie obiekty przestrzenne istotne przy opisie świata rzeczywistego z konkretnej perspektywy. Obiekty przestrzenne (warunki) to abstrakcyjne reprezentacje wybranych elementów świata rzeczywistego.

<sup>13</sup> Art. 3 ust. 2 Dyrektywy INSPIRE

<sup>14</sup> W normach ISO TC 211 stosuje się pojęcie "uniwersum dyskursu" w celu podkreślenia faktu, że w procesie modelowania uwagę skupia się tylko na pewnych wybranych podmiotach świata rzeczywistego.

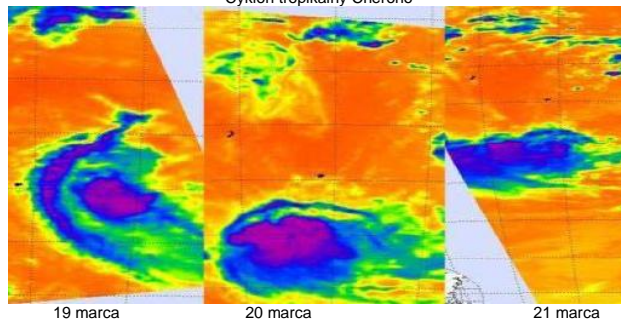


Rysunek 2: Wiele widoków w zakresie hydrologii

## 2. Reprezentacje w wielu punktach w czasie

Nasz świat zmienia się w czasie, co powinno być odzwierciedlone w opisach danych empirycznych. Reprezentacje w wielu punktach w czasie to zasada wielokrotności łącząca obiekt przestrzenny występujący w określonym momencie w czasie z jego poprzednikiem/poprzednikami i/lub następcą/następcami.

Cyklon tropikalny Cherocho



Rysunek 3: Reprezentacja w wielu punktach w czasie

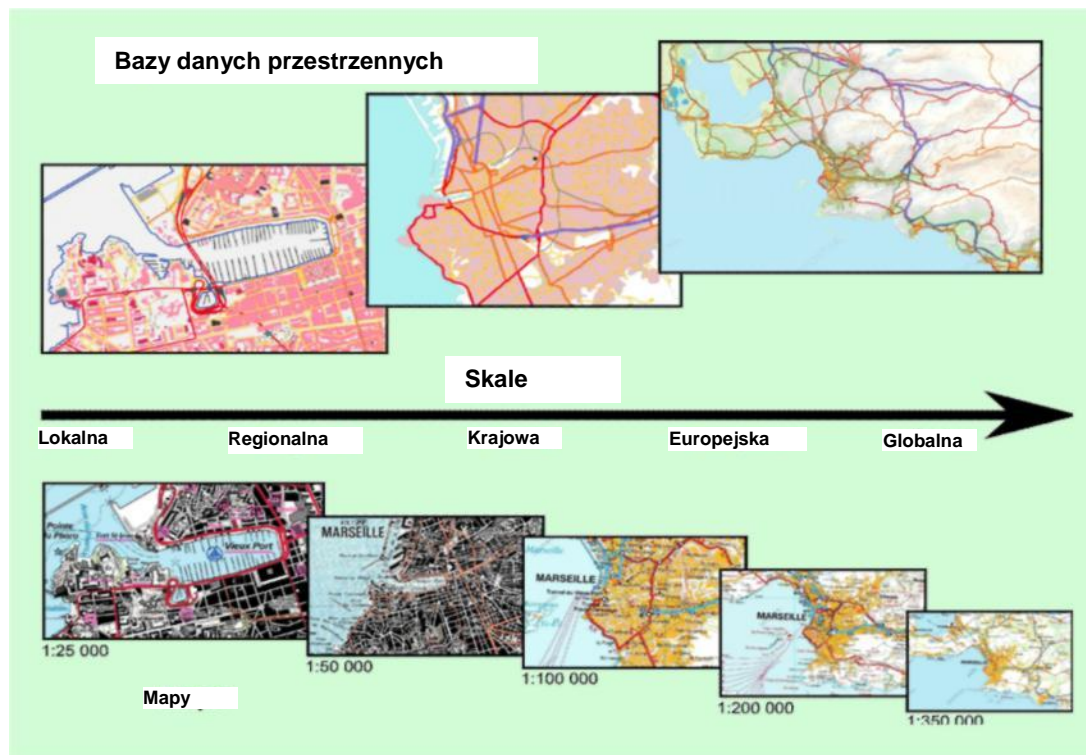
Gwałtownie zmieniające się zjawiska naturalne, np. cyklony pogodowe, śledzi się z wykorzystaniem szeregów czasowych zdjęć satelitarnych. W tym przypadku tożsamość cyklonu jest niezmienna, natomiast jego położenie, zasięg i właściwości fizyczne zmieniają się w czasie.

Częstotliwość rejestrowania danych może być bardzo duża, szczególnie przy wykorzystaniu automatycznych czujników. Informacje te mogą być agregowane z biegiem czasu tak, aby reprezentowały stan i/lub parametry danego zjawiska w wybranych momentach w czasie lub poprzez wartości średnie przez określony okres. Dane klimatyczne uzyskuje się poprzez agregowanie obserwacji meteorologicznych z różnych okresów czasu.

### 3. Reprezentacja w wielu skalach (rozdzielczości)

W ramach tematu danych elementy świata rzeczywistego mogą być opisywane na różnych poziomach szczegółowości. Proces generalizacji obejmuje ograniczenie poziomu szczegółowości w odwzorowaniu informacji. W przypadku opisu miejscowości - zob. Rys. 4 - bardzo szczegółowy opis może obejmować pojedyncze budynki i wszystkie ulice na danym terenie, mniej szczegółowy - tylko grupy budynków i główne drogi, a w małej skali wszystkie grupy budynków reprezentowane są jako jeden teren zabudowany. Mniej szczegółowe odwzorowania obejmują tylko niewielką liczbę najważniejszych właściwości tematycznych (np. punkt reprezentujący całą miejscowość i jej nazwę geograficzną).

Co do zasady, reprezentacje szczegółowe opisują obiekty z najlepszym przybliżeniem ich kształtu i rzeczywistego położenia, natomiast mniej szczegółowe odwzorowania umożliwiają upraszczanie, co jest istotne dla zachowania przejrzystości i czytelności informacji przestrzennej na mapach lub ekranach. Podejście łączące w sobie różne poziomy szczegółowości zwane jest reprezentacją wieloskalową lub wielu rozdzielczości, jednak często zwane są one też reprezentacjami wielorakimi.



Rysunek 4: Reprezentacja w wielu skalach i uogólnienie (źródło grafiki: [www.geoportail.fr](http://www.geoportail.fr))

Wielość informacji dotyczących tego samego miejsca czy tego samego zjawiska w różnych momentach w czasie zapewnia niezwykle potencjał uzyskania wyższego poziomu zrozumienia świata, w którym żyjemy, ponieważ przy pomocy analizy jednoczesnej lub porównawczej możemy badać nowe, bez nich ukryte przed nami aspekty.

Wieloraka informacja geoprzestrzenna bywa bardziej wymagająca pod względem przetwarzania i przechowywania danych ze względu na potencjalną niespójność reprezentacji. W kolejnym rozdziale opisano trudności integrowania informacji z różnych źródeł.

#### 2.2 Kwestie niekompatybilności i niespójności danych przestrzennych

Użytkownicy, którzy próbują integrować dane przestrzenne z odrębnych źródeł lub wykorzystywać ponownie informacje opracowane w innych systemach, borykają się często z problemem niekompatybilności i niespójności danych. Istotą problemu zasada się na różnych politycznych, gospodarczych, kulturowych i technicznych czynnikach tworzenia danych, wyrażanych w formie różnic składni, semantyki, odwzorowania przestrzennego i czasowego, a także braku uwzględnienia współzależności pomiędzy tematami.

**Składnia** to wewnętrzny wzór strukturalny języka naturalnego lub maszynowego. Najprostsze przykłady różnic składniowych to formaty zapisu plików stosowane przez różne programy oraz reguły gramatyczne języków ludzkich. Bez uzgodnionej składni lub pogłębionej wiedzy na temat języków kodowania, komunikacja pomiędzy systemami nie jest możliwa.

Różnice składniowe mogą być wyeliminowane przez rozwiązania technologiczne i organizacyjne. I tak na przykład, technologia zapewnia narzędzia programowe do przekształcania formatów plików zapisu. Zharmonizowaną prezentację danych można uzyskać poprzez porozumienie w sprawie stosowania konkretnego kodowania, najlepiej typu open source.

**Semantyka** to nauka o znaczeniu. Koncentruje się ona na relacji pomiędzy elementami znaczącymi, np. słowami, znakami i symbolami oraz przedmiotami, które elementy te reprezentują. Spójność semantyczna oznacza, że dowolne dwie osoby lub dowolne dwa systemy uzyskują te same wnioski z tych samych informacji. Różnorodność semantyczna informacji i danych geograficznych wynika z procesów abstrahowania, w drodze których różne społeczności w środowiskach wielonarodowych lub wielodyscyplinarnych opisują świat rzeczywisty w różny sposób.

Koncepcje stosowane do opisu elementów świata rzeczywistego mogą być rozbieżne pod względem treści (definicja), stopnia zagregowania (rozdzielczość semantyczna) i bogactwa opisu (liczba właściwości lub atrybutów), co prowadzi do różnic w klasyfikacji i/lub poziomie agregacji, uwidocznionych w Tabeli nr 2.

Przykłady różnic semantycznych		
Inny poziom agregacji		Ten sam podmiot ze świata rzeczywistego jest reprezentowany na różnych poziomach agregacji (domy i bloki)

Inne klasyfikacje		Ten sam podmiot jest klasyfikowany odmiennie po dwóch stronach granicy (strefa przemysłowa i teren zabudowany)
-------------------	---	--

Tabela 2: Przykłady różnic semantycznych danych przestrzennych

Różnice semantyczne można wyeliminować poprzez harmonizację koncepcji lub zastosowanie technologii opracowanych w kontekście sieci semantycznej<sup>15</sup>. Do narzędzi stosowanych przy publikowaniu uzgodnionych i zharmonizowanych koncepcji danych przestrzennych należą słowniki pojęć, taksonomie, schematy klasyfikacji, listy kodów itp.

**Reprezentacja przestrzenna** może stanowić kolejną trudność przy integrowaniu danych geograficznych. Rozbieżności występują często na poziomie odwzorowania graficznego, jak również mogą prowadzić do problemów z przetwarzaniem danych. Niektóre typowe przykłady przedstawiono w Tabeli 3.




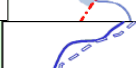
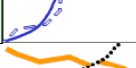
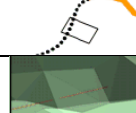
Różne odwzorowania przestrzenne		Ograniczone możliwości – nakładanie się odwzorowań rastrowych (ortobraz) i wektorowych (drogi)
Różne układy geometryczne odwzorowania (trójwymiarowe i dwuwymiarowe)		Ten sam budynek odwzorowany w geometrii trój- i dwuwymiarowej
Różne układy geometryczne odwzorowania na płaszczyźnie		Rzeka jest odwzorowana w formie wielokątu po jednej stronie granicy i osi środkowej po drugiej stronie
Różne granice		Możliwe przyczyny: brak uzgodnień pomiędzy organami władz, błędy pomiaru/przekształceń, różne uogólnienia
Nakładające się na siebie obiekty przestrzenne i przesunięcie geometryczne		Błędy wzdłuż granicy, prawdopodobnie spowodowane przez odmiennie pierwotne układy rzutowania
Niespójność tematów danych (cyfrowy model ukształtowania terenu i drogi)		Naruszenie naturalnych współzależności (droga przekracza powierzchnię terenu bez tunelu)

Tabela 3: Problemy z interoperacyjnością związane z reprezentacją przestrzenną

Rozwiązania z zakresu interoperacyjności i harmonizacja danych w SDI mają na celu wyeliminowanie niekompatybilności i niespójności danych, a tym samym zwolnienie użytkowników z obowiązku prowadzenia kłopotliwych manipulacji danymi przed rozpoczęciem korzystania z nich w określonych zastosowaniach. W kolejnych ustępach omówione są przykłady problemów z interoperacyjnością w związku z różnicami reprezentacji przestrzennej, zilustrowane w Tabeli nr 3.

Pierwszy przykład w Tabeli nr 3 przedstawia niekompatybilność przestrzenną wynikającą z odmiennych odwzorowań przestrzennych. Integracja danych pokrycia (rastrowych) i wektorowych<sup>16</sup> rzadko wykracza poza nakładanie i analizę wizualną ze względu na niekompatybilność algorytmów przetwarzania. Wprawdzie konwersja danych wektorowych na proste dane pokrycia jest stosunkowo prosta i może być przeprowadzona automatycznie, niemniej konwersja danych pokrycia na dane wektorowe może wymagać digitalizacji mapy.

W zależności od przewidywanego wykorzystania danych właściwości przestrzenne zjawisk ze świata rzeczywistego mogą być reprezentowane z wykorzystaniem różnych modeli geometrycznych. Zalicza się do nich objętość w modelach trójwymiarowych (3D) lub powierzchnie w modelach 2D. Dane dotyczące tych samych lub podobnych podmiotów modelowanych z wykorzystaniem różnych rodzajów geometrii wymagają modyfikacji w celu zintegrowania. Należy zauważyć, że bez dodatkowych informacji można zasadniczo przekształcać różne formy odwzorowań jedynie poprzez zmniejszenie wymiaru. I tak na przykład, rzeka może być odwzorowana jako powierzchnia lub oś środkowa, jak wskazano w Tabeli nr 3. Aby uzyskać wspólne, interoperacyjne odwzorowanie, należy zmniejszyć przestrzeń do osi środkowej, co można zrobić przy użyciu różnych algorytmów.

Położenie w świecie rzeczywistym elementów o charakterze społecznym i politycznym (np. granic administracyjnych, jednostek zarządzania itp.) musi być uzgodnione przez właściwe władze przed ich zaprezentowaniem w formie danych geograficznych. Brak takich uzgodnień może skutkować niespójnością odwzorowań sąsiadujących ze sobą i przecinających się obiektów przestrzennych wzdłuż granic takich elementów. Różnice położenia granic, w szczególności granic państw, mogą być skutkiem wykorzystania różnych systemów odniesienia i rzutowania<sup>17</sup>, co może uwidaczniać się formie nieuzasadnionego nakładania się lub nieciągłości, co widać w przykładzie piątym z Tabeli nr 3.

W opisie świata rzeczywistego z wykorzystaniem abstrakcyjnych reprezentacji z konkretnego punktu widzenia można pominąć naturalne zależności zjawisk ze świata rzeczywistego. Staje się to ewidentne przy integrowaniu danych z różnych źródeł. Jak widać na podstawie ostatniego przykładu z Tabeli nr 3, odwzorowanie drogi przecinającej powierzchnię cyfrowego ukształtowania terenu bez tunelu prowadzi do uzyskania niespójnego modelu rzeczywistości.

### 2.3 Przedmiot SDI

Jak wskazano w rozdziale 2.1, opis naszego środowiska z różnych punktów widzenia, w różnych momentach w czasie i na różnym poziomie szczegółowości prowadzi do zwielokrotnienia danych przestrzennych, gdzie każdy opis służy do określonego celu. Opisy mogą jednak zawierać elementy wspólne. Im bardziej pogłębiona będzie analiza konkretnych aspektów, tym mniej wspólnych elementów można ustalić. I odwrotnie: niektóre aspekty, np. metody opisu położenia przestrzennego, są wspólne we wszystkich zastosowaniach.

Gdzie jest miejsce SDI wśród niezliczonej liczby aplikacji wykorzystujących dane przestrzenne? SDI powinny obejmować wspólne aspekty przestrzenne ustanawiające standardowy kontekst położenia dla szerokiego zakresu zastosowań. Przykładowo, dane demograficzne mogą być powiązane z adresami lub mogą wykorzystywać ponownie położenie geometryczne jednostek administracyjnych. Wykorzystanie danych odniesienia jako punktu stałego, z którym łączy się inne informacje geograficzne czy gospodarcze, to jedna z kluczowych koncepcji brytyjskich Krajowych Ram Cyfrowych (Digital National Framework).

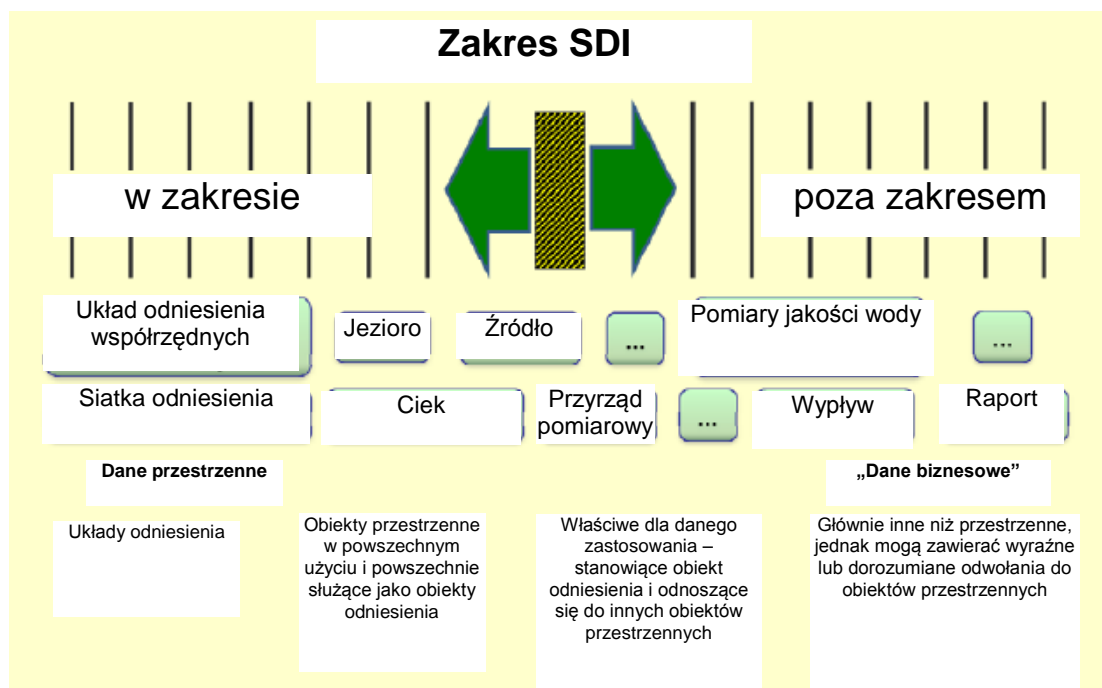
**Dane odniesienia** służą do ustalenia standardowego kontekstu położenia, który może być wykorzystany ponownie z innymi informacjami (tj. można się do niego odnosić).

Sposoby definiowania zakresu SDI przedstawione są na Rys. 5. Tematyczna infrastruktura SDI, jak np. INSPIRE, może obejmować standardowe koncepcje związane z docelowym polem tematycznym, np. dane przestrzenne dotyczące hydrologii. Zgodnie z zasadą opisaną powyżej infrastruktura powinna uwzględniać tylko takie obiekty przestrzenne, które wykazują się dużym potencjałem ponownego wykorzystania. Konkretnie zastosowania, np. dostarczające informacje gospodarcze, są poza zakresem przedmiotowym infrastruktury.

<sup>15</sup> Zob. rozdział 4.1.6

<sup>16</sup> Główne typy odwzorowań przestrzennych opisane są w rozdziale 4.1.7.

<sup>17</sup> Układy odniesienia określają ramy opisu położenia obiektów przestrzennych z wykorzystaniem współrzędnych. Konieczne jest wykonywanie rzutów w celu odwzorowania profilowanej powierzchni Ziemi na nośnikach płaskich (papier, ekran).



Rysunek 5: Zakres przedmiotowy SDI

Zamiast uwzględniania bardzo konkretnych szczegółów w SDI, potencjalni użytkownicy powinni wiedzieć, w jaki sposób można wykorzystać i/lub rozszerzać ramy przestrzenne zapewnione przez tę infrastrukturę do własnych celów.

### 3 Ramy koncepcyjne modelowania danych w SDI

**Model danych przestrzennych** to konstrukt matematyczny formalizujący postrzeganie przestrzeni. **Model koncepcyjny** wykorzystuje semantykę (pojęcia) do klasyfikacji obiektów przestrzennych w zakresie opisu (uniwersum dyskursu). **Schemat zastosowania** uzupełnia semantykę zdefiniowaną w modelu koncepcyjnym o strukturę logiczną.

Dane przestrzenne odwzorowują zjawiska ze świata rzeczywistego w formie abstrakcyjnej, która może być strukturyzowana w modelach danych. W obrębie zainteresowanej społeczności koncepcje stosowanych modeli danych są dobrze znane, a niekiedy nawet formalnie uzgadniane. Osoby pracujące w dziedzinie rejestrów gruntów kierują się wspólnym rozumieniem działek katastralnych, specjaliści od ochrony przyrody wiedzą, czym jest wyznaczony obszar, a specjaliści od topografii nie potrzebują wyjaśnień na temat poziomicy. Podsumowując, każda społeczność kieruje się pewnymi podstawowymi uzgodnieniami na temat wykorzystywanych przez siebie modeli danych. Uzgodnienia te są często publikowane w formie regulaminów, norm, bądź też współdzielone w formie konwencji czy przykładów najlepszych praktyk.

Modelowanie danych i specyfikacje danych powiązane są przede wszystkim z gromadzeniem danych i dostarczaniem produktu danych. Jaką rolę odgrywają w SDI?

Interoperacyjność w SDI oznacza, że użytkownicy mają możliwość integrowania danych przestrzennych z różnych źródeł "bez powtarzalnej interwencji manualnej", tj. zbiory danych uzyskiwane przez nich z infrastruktury mają wspólną strukturę i semantykę. Jeden ze sposobów osiągania interoperacyjności w tej formie mógłby polegać na wyborze jednego zbioru danych i dostosowaniu do niego innych. Istnieje jednak nieskończona liczba sposobów możliwego łączenia ze sobą zestawów danych, w związku z czym każdorazowo przy wyborze zbioru danych jako modelu docelowego konieczne byłoby przekształcenie wszystkich pozostałych w celu zapewnienia zgodności z jego specyfikacjami. Wymagałoby to także publikowania modeli danych dla każdego źródłowego zbioru danych. Nie jest to rozwiązanie efektywne pod względem kosztowym, a dodatkowo nie zapewnia większej wartości niż rozwiązania dostępne już w komputerowych systemach GIS.

Zamiast doraźnego definiowania celów interoperacyjności co do zasady preferuje się uzgadnianie wspólnych celów interoperacyjności, formalizowanych i dokumentowanych dla każdego tematu danych w taki sposób, by mogły one być odczytane i wykorzystywane zarówno przez ludzi, jak i przez maszyny.

Specyfikacja danych w szerszym znaczeniu dotyczy zarówno specyfikacji produktu danych, wykorzystywanego do tworzenia określonego zbioru danych lub produktu, jak i specyfikacji docelowej interoperacyjności w SDI, służącej do przekształcania istniejących danych w taki sposób, by wykazywały one wspólne właściwości. W niniejszym raporcie pojęcie specyfikacji danych odnosi się do specyfikacji docelowej interoperacyjności.

Specyfikacja danych obejmuje model danych i inne odnośne uregulowania dotyczące danych, np. reguły rejestrowania danych, kodowania i dostarczania, wymagania dotyczące jakości danych, metadane do oceny i wykorzystania, spójność danych itp.

Krytycznym czynnikiem sukcesu dla każdej infrastruktury SDI jest jej akceptacja przez zainteresowanych. Podejście oddolne, w którym tworzy się warunki wspierające uczestnictwo w procesie rozwijania specyfikacji, przewiduje różnego rodzaju interakcje oraz przekazywanie informacji zwrotnych do społeczności zainteresowanych. A zatem potrzebny jest model oparty na współpracy, obejmujący zabezpieczenia niezbędne do procesów budowania konsensusu.

Ponieważ infrastruktura SDI składa się zwykle z wielu tematów danych, gdzie może być wymagana interoperacyjność międzytematyczna, należy ustanowić trwałe ramy danych. Pomysł ten został zaproponowany w Niemczech już w 1997 roku w formie zharmonizowanego podstawowego modelu koncepcyjnego ("AAA-Basischema") dla trzech krajowych baz danych: Urzędowego systemu informacji z punktów stałych (AFIS), Urzędowego systemu informacji katastralnych dot. nieruchomości (ALKIS), Urzędowego systemu informacji topograficzno-kartograficznych (ATKIS)<sup>18</sup>. Inicjatywa Niebieskiej Księgi Geoprzestrzennej - Geospatial Blue Book - w Stanach Zjednoczonych (2005), która miała na celu stworzenie "GIS dla Krajowego Modelu Danych"<sup>19</sup>, sugerowała utrzymywanie schematów zastosowań tematów danych we wspólnym systemie informatycznym, co przyczyniało się do spójnego traktowania wspólnych pojęć.

W Unii Europejskiej inicjatywa INSPIRE przyjęła ramy koncepcyjne składające się z dwóch zasadniczych części, jak przedstawiono na Rysunku 6:  
 - Podstawowy Model Koncepcyjny oraz  
 - Metodologia rozwijania specyfikacji danych

<sup>18</sup> <http://web.archive.org/web/19981206200623/http://www.adv-online.de/neues/oinhalt.htm>

<sup>19</sup> <http://support.esri.com/en/downloads/datamodel/detail/42>

Standardy	
Specyfikacja danych Temat 1	Specyfikacja danych Temat 2
Specyfikacja danych Temat (n-1)	Specyfikacja danych Temat n
Podstawowy model koncepcyjny	Metodologia opracowania specyfikacji
Ramy koncepcyjne	
Najlepsze praktyki	

Rysunek 6: Relacje ram koncepcyjnych

Podstawowa rola ram koncepcyjnych polega na zapewnieniu metodologii rozwijania powtarzalnych specyfikacji danych oraz ogólnych uregulowań dotyczących procesu specyfikowania danych, obowiązujących dla wszystkich tematów danych przestrzennych. Ramy koncepcyjne zarysowują etapowy, powtarzalny proces ustalania komponentu danych: prace powinny się rozpoczynać od zdefiniowania części wspólnych, po których powinno się przechodzić do zadań dotyczących konkretnego tematu. Innymi słowy, proces specyfikacji tematów danych może rozpocząć się dopiero po opracowaniu na odpowiednim poziomie ram koncepcyjnych.<sup>20</sup>

Wprowadzenie ram koncepcyjnych jest zgodne z zasadą ponownego wykorzystania. W kontekście SDI ponowne wykorzystanie dotyczy nie tylko współdzielienia danych w różnych zastosowaniach, ale także współdzielenia wiedzy, rozwiązań technicznych, narzędzi i komponentów. Standardy i przykłady dobrych praktyk dostawców danych przestrzennych i społeczności użytkowników stanowią podstawę do definiowania ram koncepcyjnych i procesu specyfikowania danych.

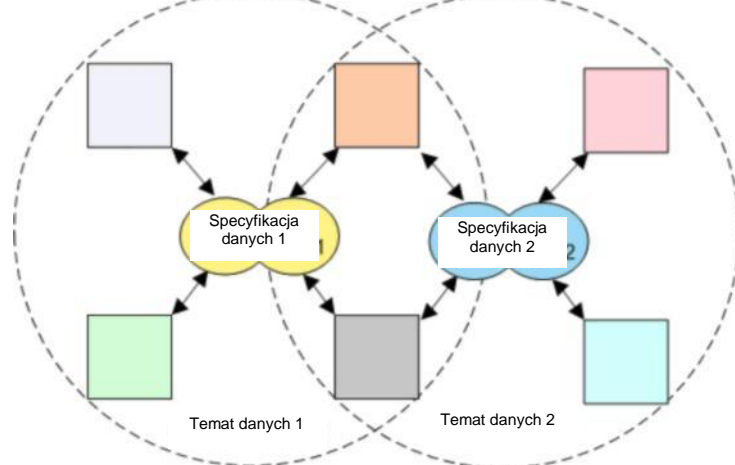
Złożoność związana z wypracowywaniem uzgodnień z zakresu interoperacyjności wzrasta wraz ze wzrostem liczby tematów danych oraz liczby uczestniczących podmiotów zainteresowanych. INSPIRE, z 34 tematami danych, setkami uczestniczących ekspertów oraz rygorystyczną dokumentacją, stanowi dobry przykład ilustrujący rolę ram koncepcyjnych. Dlatego też rozdziały 4 i 5 oparte są przede wszystkim na doświadczeniach INSPIRE i uzupełniane w stosownych miejscach przez doświadczenia z innych inicjatyw.

Jedno z głównych zadań inicjatywy INSPIRE polega na umożliwieniu interoperacyjności oraz - tam, gdzie jest to możliwe - harmonizacji zbiorów danych przestrzennych i usług danych w Europie. Należy zauważyć, że interoperacyjność musi wykraczać poza konkretne społeczności i uwzględniać różnorodne potrzeby informacyjne wielu społeczności (Portele C. (red.) 2010a).

Podstawowy model koncepcyjny (GCM) konkretyzuje pojęcia interoperacyjności i harmonizacji danych dzięki zastosowaniu zestawu elementów interoperacyjności. Elementy te pochodzą z wymagań i założeń infrastruktury, które dopasowywane są do odnośnych warunków technicznych technologii geoprzestrzennej i modelowania informacji.

Uzasadnione jest pytanie, czy komponent danych SDI może być ustalony bez podstawowego modelu koncepcyjnego. Podstawowy model koncepcyjny nie jest konieczny do osiągnięcia interoperacyjności w ramach jednego tematu danych, jeśli pojedyncza specyfikacja interoperacyjności zaradziłaby brakowi interoperacyjności. SDI składa się jednak z wielu tematów danych, które nie stanowią oddzielnych przepływów informacji. Interoperacyjność i harmonizacja są niezbędne, jeśli infrastruktura ma współdzielić semantykę, odwzorowanie przestrzenne i składnię pomiędzy tematami.

Na Rysunku 7 poszczególne pola odpowiadają należycie zdefiniowanemu elementowi schematu zastosowania, który może być semantycznym obiektem przestrzennym, odwzorowaniem geometrycznym, importowanym schematem, listą kodów itp. Ze względu na nakładanie się tematów danych oraz ograniczoną liczbę norm znajdujących zastosowanie, niektóre z tych elementów należy traktować w podobny sposób.



Rysunek 7: Interoperacyjność pomiędzy tematami w SDI (adaptacja na podstawie: Lasschuyt, E. & van Hekken, M., 2001)

GCM obejmuje wspólne koncepcje modelowania danych i rozwijania specyfikacji danych. Elementy zawarte w GCM nie powinny być specyfikowane w specyfikacjach danych dla poszczególnych tematów, i na odwrót - jeśli w specyfikacjach danych dwóch lub większej liczby tematów stwierdzone zostaną wspólne tematy, elementy te należy usunąć ze specyfikacji danych i uwzględnić w GCM.

Wprawdzie brytyjskie Krajowe Ramy Cyfrowe (DNF) nie są określane jako Podstawowy Model Koncepcyjny, jednak określają one zasady, pojęcia i metody zapewnienia wyższego poziomu integralności informacji przestrzennej. Odnoszą się one do zagadnień przekrojowych, np.:

- Powiązanie informacji z wielu źródeł z definitywnym odniesieniem położenia z wykorzystaniem niepowtarzalnych identyfikatorów,
- Usystematyzowana prezentacja i formalizacja wspierająca dzielenie i ponowne wykorzystywanie danych,
- Niezawodność i integralność danych,
- Elastyczność umożliwiającą wymianę informacji i zastosowania wielobranżowe.

Syntetyczna prezentacja elementów interoperacyjności zawartych w GCM jest przedstawiona w Tabeli nr 4. Większość przedstawionych elementów została zdefiniowana na samym początku prac technicznych nad INSPIRE. W dalszej kolejności zostały one uzupełnione o wyniki inicjatyw badawczych (np. wykorzystanie ontologii) oraz praktyczne doświadczenia z procesu rozwijania INSPIRE (dodanie skonsolidowanego repozytorium modelu i migracja przechowywania specyfikacji danych z GCM do metodologii rozwijania specyfikacji).

Podstawy	Modelowanie danych	Zarządzanie danymi
<ul style="list-style-type: none"> <li>- Wymagania</li> <li>- Model referencyjny</li> <li>- Wsparcie architektury dla interoperacyjności</li> <li>- Terminologia</li> <li>- Tekst wielojęzyczny i możliwości adaptacji kulturowej</li> <li>- Wykorzystanie ontologii</li> </ul>	<ul style="list-style-type: none"> <li>- Odniesienia do obiektów</li> <li>- Aspekty przestrzenne i czasowe</li> <li>- Zasady dotyczące schematów zastosowań i katalogów obiektów</li> <li>- Wspólne schematy zastosowań</li> <li>- Skonsolidowane repozytorium modelu</li> <li>- Odwzorowanie wielorakie</li> <li>- Punkty rozszerzenia</li> </ul>	<ul style="list-style-type: none"> <li>- Zarządzanie identyfikatorami</li> <li>- Spójność pomiędzy danymi</li> <li>- Jakość danych i informacji</li> <li>- Metadane</li> <li>- Zgodność</li> <li>- Reguły rejestracji danych</li> <li>- Wytyczne w zakresie przekształcania danych</li> <li>- Zasady przechowywania danych</li> </ul>

20

Ramy koncepcyjne mogą być rozwijane w drodze przeglądów przez społeczności zainteresowanych, testowanie i przechowywanie. Ten ostatni etap dotyczy modyfikacji wynikających z zastosowania ram koncepcyjnych w procesie rozwijania specyfikacji danych.

- Odniesienia za pomocą współrzędnych i jednostki miary - Rejestry i akta	- Prezentacja - Dostarczanie danych
--	--

Tabela 4: Elementy interoperacyjności dla komponentu danych SDI

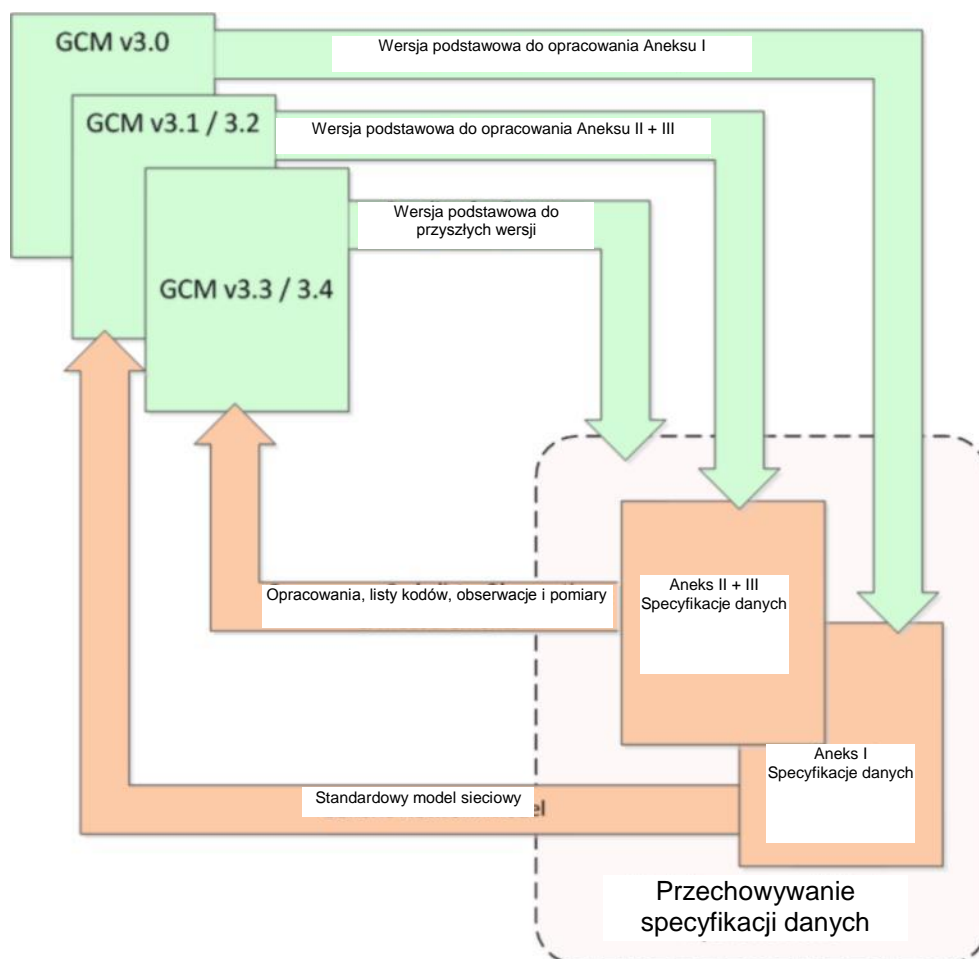
Pierwsza grupa elementów interoperacyjności określa punkt wyjścia dla procesu specyfikacji danych zarówno w teorii, jak i w zakresie narzędzi praktycznych. Grupa druga wspiera proces specyfikowania danych, a trzecia wspiera interoperacyjność z punktu widzenia zarządzania danymi.

Niektóre elementy, np. model referencyjny, wspólne schematy zastosowań, odniesienia za pomocą współrzędnych itp., muszą być modelowane, uzgadniane i publikowane. Inne muszą być obsługiwane i publikowane w rejestrach w ramach współdziałania przy współdzieleniu informacji na etapie rozwijania specyfikacji oraz w fazie eksploatacji infrastruktury (gdy użytkownicy mogą pozyskiwać dane zgodnie ze specyfikacjami interoperacyjności). Istnieją także elementy stanowiące wytyczne i przykłady najlepszych praktyk na rzecz spójnego wdrażania. Każdy element odnosi się do wszystkich tematów danych przestrzennych, jednak poziomy istotności są różne dla poszczególnych tematów.

GMC INSPIRE opracowywany jest metodą powtarzalną. Wersja pierwsza została opracowana przez Zespół ds. opracowywania specyfikacji danych zgodnie z wymaganiami Dyrektywy INSPIRE poprzez ich powiązanie z regulacjami technicznymi zawartymi w normach międzynarodowych i innych materiałach odniesienia, opisujących przykłady dobrych praktyk. Po udoskonaleniu projektu GCM na podstawie konsultacji z zainteresowanymi podmiotami wersja podstawowa została przekazana Tematycznym Grupom Roboczym odpowiedzialnym za rozwijanie specyfikacji danych dla tematów zawartych w Aneksie I.

GCM uaktualniano przez cały czas rozwijania specyfikacji danych z Aneksu I. Zasadnicza zmiana polegała na wprowadzeniu standardowego modelu sieciowego, ponieważ stwierdzono, że forma reprezentacji sieciowej była wykorzystywana w dwóch tematach. Tematyczne Grupy Robocze odpowiedzialne za rozwijanie specyfikacji danych dla tematów z Aneksu II i III rozpoczęły pracę na uaktualnionych GCM i wprowadziły w trakcie prowadzonych czynności inne wspólne elementy, np. schemat pokrycia i model obserwacji i pomiarów. Ponieważ rozwijanie specyfikacji danych dla Aneksów II i III trwa nadal, mogą być konieczne dalsze uaktualnienia GCM. Dalsze modyfikacje mogą wystąpić w trakcie procesu przechowywania specyfikacji, zaprezentowanego na Rysunku nr 8.

Podstawowy Model Konceptyjny INSPIRE zawiera standardowy model sieciowy, wprowadzony na etapie, gdy temat danych Hydrografia i Sieci Transportowe zaczynał modelowanie danych przestrzennych jako sieci. Standardowy model sieciowy zapewnia stosowanie tych samych zasad geometrycznych. Na dalszym etapie wykorzystano go ponownie w odniesieniu do tematu danych Media.



Rysunek 8: Iteracyjne rozwijanie Podstawowego modelu koncepcyjnego

W kolejnych rozdziałach przedstawiono dalsze szczegóły dotyczące poszczególnych elementów interoperacyjności, zawartych w Tabeli nr 4. Ze względu na charakter zagadnienia rozdziały te bezspornie muszą mieć bardziej techniczny charakter. Czytelnicy, których bardziej interesuje proces, mogą pominąć rozdział 4 i przejść bezpośrednio do rozdziału 5 - Metodologia rozwijania specyfikacji danych.

#### 4 Podstawowy model koncepcyjny

##### 4.1 Zasadnicze elementy

###### 4.1.1 Wymagania

Doświadczenie pokazuje, że wymagania oraz reguły wdrażania SDI mogą być rozproszone w różnych dokumentach polityki, aktach prawnych, studiach technicznych i innych dokumentach. W celu nakreślenia zakresu wymaganych czynności technicznych należy zgromadzić i usystematyzować te wymagania

oraz zasady. Katalog tych zasad nie jest wyczerpujący, a wśród nich wymienić można:

- Brak obowiązku gromadzenia danych na nowo: uzgodnienia dotyczą istniejących danych i przyszłych zbiorów danych, zakładanych przez właściwe organizacje zainteresowanych podmiotów.
- Zasada włączania: jakiegokolwiek dane są lepsze niż brak danych;
- Podejście uzależnione od użytkownika: ustalenie, co powinno być włączone w zakres infrastruktury (informacje geograficzne możliwe do ponownego wykorzystania) i jaki poziom opisu jest właściwy;
- Brak obowiązku zmiany istniejących obiegów zadań: tylko publikowanie danych według uzgodnionego docelowego poziomu interoperacyjności za pośrednictwem usług sieciowych;
- W miejsce reinżynierii wysoki priorytet nadawany jest przekształcaniu istniejących danych;
- Ponowne wykorzystywanie istniejących norm, konwencji i inicjatyw;
- Wykonalność pod względem technicznym i proporcjonalność (mimo że ograniczenia dotyczące komponentów oprogramowania nie są sprawą najważniejszą) dla zapewnienia możliwości dopasowania specyfikacji do infrastruktury ICT dostawców danych;
- Etapowe podejście do wdrażania;
- Proporcjonalność finansowa i względny koszt do korzyści dla zapewnienia optymalnego rozwiązania;
- Spójność danych/informacji dotyczących tego samego położenia przestrzennego, przedstawionych w różnych skalach i rozdzielczościach, a także wzdłuż granic (państwowych, regionalnych itp.).

Wyjaśnienie takich ogólnych wymagań stanowi pierwszy krok w definiowaniu GCM, ponieważ wymagania te przekładane są następnie na konstrukty modelowania i elementy specyfikacji.

#### 4.1.2 Model referencyjny

Model referencyjny określa, kiedy stosuje się standardy i w jaki sposób powinny one być wykorzystywane przy rozwijaniu komponentu danych SDI. Ponieważ standardy co do zasady charakteryzują się szerszym zakresem, konieczne jest uzgodnienie zasad ich dostosowania do konkretnego celu. Taki proces adaptacji zwany jest profilowaniem. Model referencyjny wyszczególnia także rodzaje usług informatycznych, jakie mogą być wykorzystywane przy dostępie, przetwarzaniu i udostępnianiu danych geograficznych oraz informacji powiązanych w infrastrukturze. Przykładem modelu referencyjnego jest ISO 19101 - Informacja geograficzna, Model referencyjny zawierający ogólny opis sposobu tworzenia informacji geograficznej i wzajemnego dopasowania standardów odnoszących się do tego zakresu.

Podstawowy model koncepcyjny INSPIRE można traktować jako konkretny model referencyjny, stanowiący podstawę do rozwijania specyfikacji danych. GMC może także służyć do opracowywania innych infrastruktur w innych kontekstach geograficznych i tematycznych.

#### 4.1.3 Wsparcie interoperacyjności danych przez architekturę

Osadzanie danych przestrzennych w infrastrukturze oznacza, że dostęp do danych jest wspierany przez inne elementy strukturalne SDI. Do takich elementów strukturalnych należą dane, metadane, usługi sieciowe oraz rozwiązania z zakresu udostępniania danych.

Ze względu na zapewnienie efektywnego działania elementy strukturalne SDI muszą być wzajemnie powiązane, co wymaga ich skoordynowania i uszczegółowienia w odniesieniu do wzajemnych funkcjonalności i właściwości technicznych. Ten komponent interoperacyjności zestawia także reguły i technologie stosowane przy publikowaniu informacji niezbędnych dla zrozumienia i interpretowania informacji geograficznej.

Dostęp do danych przestrzennych w SDI realizowany jest przez Internet, za pośrednictwem usług zapewniających określone funkcje, takie jak: wykrywanie, dostęp, mapowanie, przekształcanie i inne sposoby przetwarzania. Komponent danych SDI musi uwzględniać charakterystykę techniczną **Usług sieciowych**. Przykładowo, usługa podglądu może wymagać przekazywania danych w określonym systemie odniesienia za pomocą współrzędnych lub zdefiniowanych stylach prezentacji/wizualizacji. Powinno to być odwzajemnione w komponencie danych.

**Metadane** dostarczają informacji o zbiorach danych i usługach wbudowanych w infrastrukturę. Podstawowa funkcja metadanych polega na współdziałaniu przy wykrywaniu istniejących danych i usług, a także pomocy przy ocenie ich zdolności do zamierzonego celu. Metadane do oceny i wykorzystania są ściśle sprzężone z modelami danych i innymi elementami specyfikacji. Struktury danych, semantyka, kodowanie, ewentualne wymagania jakościowe i inne właściwości techniczne są określone w specyfikacjach danych, przekazywanych użytkownikom w formie metadanych. W sytuacji idealnej tworzenie danych i metadanych powinno odbywać się równolegle.

Celem **udostępniania danych i usług** jest zapewnienie zharmonizowanych warunków dostępu dla różnych grup użytkowników. W idealnej infrastrukturze SDI wszystkie warunki korzystania są jasne, kompletne, dostępne publicznie i opublikowane w Internecie w różnych językach, w kontekście globalnym. Uprawnienia przyznawane różnym grupom użytkowników w SDI są zarządzane przy użyciu funkcji kontroli dostępu.

**Usługi rejestrów** zapewniają dostęp do rejestrów<sup>21</sup>. Ponieważ odgrywają one istotną rolę w procesie rozwijania specyfikacji danych, są uwzględnione w podstawowym modelu koncepcyjnym jako element interoperacyjności.

#### 4.1.4 Terminologia

Spójność pod względem językowym ma kluczowe znaczenie dla interoperacyjności semantycznej. Infrastruktura SDI wymaga narzędzia referencyjnego do współdzielenia terminów i ich definicji. Glosariusze wraz ze Słownikami pojęć obiektów przyczyniają się do spójnego opracowywania dokumentów technicznych (specyfikacji, stron internetowych), poprawiają poziom ich spójności i umożliwiają zainteresowanym lepsze zrozumienie danych oraz usług obecnych w infrastrukturze. Aby były łatwiej dostępne, konieczne jest ich wdrożenie w formie rejestrów.

#### 4.1.5 Tekst wielojęzyczny i możliwości adaptacji kulturowej

SDI mogą rozciągać się ponad granicami językowymi i kulturowymi oraz obszarami kompetencji społeczności. Dlatego też konieczne jest ustanowienie mechanizmów eliminacji wszelkich problemów w wypracowywaniu wspólnego rozumienia terminów.

"Rozwiązaniem problemów wielojęzycznych nie jest przetłumaczenie wszystkiego na wspólny język (np. angielski). Często wystarczy uzyskać zasoby w języku, w którym zostały stworzone, zamiast ich tłumaczenia" (Europejski Komitet Normalizacyjny, 2011). Ze stwierdzenia tego wynikają dwa problemy:

- Co powinno się tłumaczyć,
- Kiedy i w jaki sposób powinny odbywać się tłumaczenia.

Ze względu na zapewnienie czytelności dla maszyn, wykorzystanie tekstu językowego w SDI powinno być ograniczone do minimum, w szczególności w obrębie specyfikacji technicznych. Idealnym rozwiązaniem byłoby przechowywanie terminów w centralnych (wielojęzycznych) słownikach, gdzie byłyby one tłumaczone na wszystkie języki odnośnych użytkowników<sup>22</sup>. Takie zarządzane centralnie słowniki mogłyby być wykorzystywane przez ludzi lub przez narzędzia do tłumaczenia maszynowego, a tym samym byłyby one pomocne przy likwidacji konieczności wykonywania doraźnych tłumaczeń przez użytkowników nie zaznajomionych z terminologią techniczną. Na potrzeby dostępu do danych oraz pomocy w ich zrozumieniu przydatne jest opracowanie strategii pozyskiwania informacji z różnych języków. Dlatego też wykazy kodów, słowniki pojęć obiektów i katalogi obiektów zgodne ze standardami ISO powinny być wielojęzyczne.

Reguły dotyczące nazw geograficznych różnią się od reguł dotyczących tekstu językowego. Ponieważ nazwy geograficzne stanowią pośrednie odniesienia przestrzenne, stosowane powszechnie przy wyszukiwaniu innych informacji przestrzennych, sprawą o zasadniczym znaczeniu jest zapewnienie przekazywania nazw i związanych z nimi egzonimów<sup>23</sup> w językach większościowych i mniejszościowych; żadna z tych nazw geograficznych nie może być zastąpiona przez tłumaczenia.

#### 4.1.6 Użycie ontologii

Ontologie to formalne odwzorowania semantyki, które mogą przyczyniać się do podwyższenia poziomu adaptacji kulturowej i dialogu pomiędzy różnymi grupami zainteresowanych. Odniesienie do Prostej Organizacji Wiedzy - Simple Knowledge Organization System, SKOS - stanowi standardową,

<sup>21</sup> Więcej szczegółów na ten temat znajduje się w rozdziale 4.1.8.

<sup>22</sup> W INSPIRE wykorzystywane są wszystkie języki urzędowe Unii Europejskiej.

<sup>23</sup> Nazwa geograficzna stosowana w konkretnym języku do opisu obiektu przestrzennego położonego poza terytorium, na którym mówi się danym językiem, np. angielskie słowo "Brussels" jest egzonimem Bruxelles i Brussel (Brukseli).



tanią możliwość migracji służącej przenoszeniu istniejącej wiedzy z różnych systemów - np. tezaursów, taksonomii, schematów klasyfikacyjnych itp. - do Sieci Semantycznej (Semantic Web) w oparciu o podobieństwa ich struktury. Może być wykorzystywane samodzielnie lub w połączeniu z formalnymi językami reprezentacji wiedzy, np. Ontology Web Language (OWL).

Ontologie są pomocne przy rejestracji aspektów wielokulturowości tylko wówczas, gdy są na tyle obszerne, by uwzględniały informacje kontekstowe niezbędne do osiągnięcia wspólnego zrozumienia przez różne społeczności. Ten komponent interoperacyjności zapewnia wytyczne dotyczące rozwijania ontologii w SDI.

Należy zauważyć, że chociaż możliwości operacyjnego wykorzystania ontologii w SDI, w tym w INSPIRE, są ograniczone, projekty badawcze i pojawiające się technologie Sieci Semantycznej otwierają nowe perspektywy ich stosowania.

#### 4.1.7 Odniesienia za pomocą współrzędnych i jednostki miary

Położenie przestrzenne może być zdefiniowane przez wartości współrzędnych punktów geometrycznych reprezentujących obiekt przestrzenny. Układ odniesienia jest potrzebny do zdefiniowania współrzędnych. Ponadto do odzwierciedlenia nierównej powierzchni Ziemi na nośnikach płaskich (mapy papierowe, ekrany itp.) konieczny jest system rzutowania. Wybór systemów odniesienia za pomocą współrzędnych oraz odwzorowań zależy od kraju (ze względu na ograniczenie do minimum związanych z tym błędów) i od społeczności (w celu zoptymalizowania analizy przestrzennej i odwzorowań odpowiednio do ich wykorzystania).. W celu zintegrowania danych zdefiniowanych pierwotnie w innych układach odniesienia i/lub odwzorowaniach konieczne jest przekształcenie danych na wspólny system.

Wspólne układy odniesienia i odwzorowania wybierane dla zapewnienia interoperacyjności, powinny być precyzyjnie opisane. Współistnienie różnych układów odniesienia wymaga ich rejestracji wraz z konkretnymi parametrami przekształceń niezbędnymi do przejścia z jednego systemu na drugi.

W INSPIRE do współrzędnych poziomych wykorzystywany jest międzynarodowy przestrzenny układ odniesienia ITRS (International Terrestrial Reference System) wraz z jego europejską wersją (ETRS), natomiast do współrzędnych pionowych - europejski pionowy układ odniesienia EVRS (European Vertical Reference System). Rekomendowanymi odwzorowaniami są: odwzorowanie azymutalne równopowierzchniowe (ETRS89-LAEA), odwzorowanie płaskoziemne wielopółkowe Lamberta (ETRS89-LCC) oraz odwzorowanie UTM (ETRS89-TMzn).

Model GCM powinien także regulować jednostki miar. Na podstawie międzynarodowych inicjatyw normalizacyjnych preferowany jest międzynarodowy układem jednostek SI, za wyjątkiem kątów, które opisywane są zwykle w stopniach. Poza systemami liniowymi mogą być również wykorzystywane systemy parametryczne lub inne niż oparte o długość<sup>24</sup> -

#### 4.1.8 Akta i rejestry

SDI zawiera wiele pozycji wymagających jednoznacznego opisu i możliwości odniesienia. Akta rejestracyjne przypisują identyfikatory do pozycji i ich definicji i/lub opisów. Często są one realizowane jako rejestry, tj. systemy informacyjne do prowadzenia akt rejestracyjnych. Rejestry to narzędzia udostępniania informacji i wiedzy. W ramach wsparcia ponownego wykorzystywania koncepcji i komponentów na etapie rozwijania infrastruktury są one włączone w zakres GCM. W przypadku operacyjnych SDI pomagają one użytkownikom w zrozumieniu semantyki i struktury danych.

Oto niewyczerpująca lista przykładów rejestrów odnoszących się do SDI:

- **Glosariusz:** dokumentacja terminologii stosowanej w infrastrukturze;
- **Słownik pojęć obiektów:** Jest to ustanowienie zestawu pojęć dotyczących obiektów (nazwa, definicja, opis), które można wykorzystać do opisu informacji geograficznej;
- **Rejestr katalogu obiektów:** Ten rejestr, oparty o katalogi obiektów ISO 19110, zawiera definicje i opisy typów obiektów przestrzennych, ich właściwości i związanych z nimi komponentów występujących w jednym lub większej liczbie zbiorów danych, wraz ze wszystkimi operacjami, które mogą być zastosowane;
- **Skonsolidowane repozytorium modelu:** Zbiór wszystkich modeli danych w wybranym języku schematu koncepcyjnego, umożliwiającą zarządzanie wzajemnymi zależnościami pomiędzy modelami;
- **Rejestr list kodów:** Kontrolowane słownictwo z możliwością rozbudowy, opisujące domeny wartości wybranych właściwości w schemacie zastosowania, zarządzane odrębnie w ramach własnego słownika;
- **Rejestr systemów odniesienia za pomocą współrzędnych:** Rejestr systemów odniesienia za pomocą współrzędnych, danych, układów odwzorowania i działań na współrzędnych, wykorzystywanych w infrastrukturze;
- **Rejestr jednostek miary:** Rejestr jednostek miary, który może być wykorzystywany w zbiorach danych przestrzennych;
- **Rejestr przestrzeni nazw:** Odpowiada on za niepowtarzalność przestrzeni nazw, które mogą być wykorzystywane ponownie, np. do identyfikatorów obiektów zewnętrznych w obrębie infrastruktury;
- **Rejestr prezentacji:** Rejestr obsługujący konfigurację usług podglądu i udostępniania stylów definiowanych przez użytkownika;
- **Rejestr schematów kodowania:** Służy do gromadzenia specyfikacji kodowania danych wykorzystywanych w infrastrukturze.

## 4.2 Modelowanie danych

### 4.2.1 Odwoływanie się do obiektów

W miejsce bezpośredniego przyporządkowywania współrzędnych można definiować położenie zjawiska w stosunku do istniejącego obiektu przestrzennego. Takie odniesienia pośrednie są możliwe poprzez:

- wskazanie odniesień do innych obiektów przestrzennych,
- wykorzystanie identyfikatora geograficznego z gazetera.

Odwołania do obiektów wykorzystują ponownie współrzędne geograficzne obiektu przestrzennego, którego dotyczą, określając możliwy sposób powiązania nowych informacji z istniejącymi współrzędnymi. Przykładowo, w przypadku odniesień liniowych można wykorzystać istniejący obiekt liniowy (np. odcinek drogi) do zlokalizowania innego obiektu przestrzennego (np. przystanku autobusowego) poprzez określenie odległości od punktu początkowego odcinka.

Gazeter umożliwia klientowi wyszukiwanie i pobieranie elementów słownictwa odniesień geograficznych. Ta alternatywna metoda odniesień jest szczególnie przydatna w przypadku nazw geograficznych i adresów.

### 4.2.2 Aspekty przestrzenne i czasowe

Dostępne są dwa sposoby opisu zasięgu przestrzennego lub rozkładu obiektów przestrzennych: reprezentujące dane w postaci zbiorów danych wektorowych lub "pokrycia".

Tradycyjne podejście geograficzne zakłada, że świat składa się z identyfikowalnych struktur posiadających obiektywne właściwości. Na podstawie tego założenia uzyskuje się dane wektorowe, gdzie każde zjawisko wyobrażone jest jako odrębny obiekt przestrzenny z odrębną tożsamością. Obiekty te reprezentowane są jako punkty, powierzchnie lub elementy kubaturowe (w odwzorowaniach trójwymiarowych). Właściwości takich obiektów określa się jako atrybuty. Dane wektorowe odpowiadają na pytanie: *Gdzie znajdują się obiekty przestrzenne należące do określonego typu i jakie są ich właściwości?*

Inną metodą opisu świata jest metodologia pola ciągłego, w której zjawisko jest reprezentowane przez wiele zmiennych, z których każda jest możliwa do zmierzenia w dowolnym punkcie na powierzchni Ziemi. Wartości te zmieniają się w przestrzeni i/lub w czasie (Longley, P. A. i in., 2011). Ta metoda odwzorowania, zwana często "pokryciem", jest bardzo powszechnie stosowana w obserwacjach i pomiarach, w tym obserwacjach Ziemi. Z matematycznego punktu widzenia pokrycie to funkcja odpowiadająca na pytanie: *Jaka jest wartość (określonego parametru) w określonej lokalizacji?* Przypisane wartości odpowiadają często rozkładom, np. temperatury, ukształtowania terenu, rozmieszczeniu ludności. Najczęściej wykorzystywane pokrycia to siatki zawierające zestaw wartości, z których każda jest skojarzona z jednym z elementów w regularnym układzie punktów lub komórek.

Obie formy reprezentacji przestrzennej są niezbędne, ponieważ "wyrażają [...] świat w postaci przestrzeni, w której znajdują się rzeczy, lub w postaci przestrzeni, w obrębie której właściwości podlegają zmianom" (Woolf i in., 2010). Należy zauważyć, że forma reprezentacji przestrzennej nie jest zdefiniowana przez zawartość danych. W zakresie tego samego zastosowania poszczególne formy mogą być przekształcane na siebie nawzajem. Przykładowo, stereoskopowa para cyfrowych zdjęć powierzchniowych lub satelitarnych (pokrycie) może być wykorzystana do uzyskania danych dotyczących ukształtowania terenu, a te z kolei mogą być następnie ukształtowania terenu, rozmieszczeniu ludności. Najczęściej wykorzystywane pokrycia to siatki zawierające zestaw wartości, z których każda jest skojarzona z jednym z elementów w regularnym układzie punktów lub komórek.

W przypadku odniesień czasowych wymagane jest określenie strefy czasowej i wykorzystywanego kalendarza. Ogólne stosowanie kalendarza gregoriańskiego wraz z wybraną i uzgodnioną strefą czasową usprawnia obsługę danych. W przypadku międzynarodowych i globalnych SDI uzasadnione jest stosowanie standardu uniwersalnego czasu koordynowanego (UTC). Do interoperacyjności przyczyniają się ponadto jednoznaczne i prawidłowo zdefiniowane metody reprezentacji dat i godzin zgodnie z ISO 8601 - Elementy danych i formaty wymiany - Wymiana informacji - Reprezentacja dat i godzin.

#### 4.2.3 Zasady dotyczące schematów zastosowań i katalogów obiektów

Jak już wskazano powyżej, schemat zastosowania to model danych koncepcyjnych opracowany dla danego zastosowania (w produkcji danych) lub do ustalenia docelowego poziomu interoperacyjności dla tematu danych w SDI. Zawiera on typy obiektów przestrzennych, ich relacje i atrybuty, a także ewentualne ograniczenia dotyczące elementów modelu. W SDI każdy temat danych obejmuje przynajmniej jeden schemat zastosowania. Większa liczba schematów zastosowania może być wprowadzona w następujących sytuacjach:

1. Gdy temat danych jest zbyt obszerny i możliwy jest jego podział logiczny według różnych punktów widzenia. Do sytuacji takiej doszło w obrębie tematu danych INSPIRE "Sieci transportowe", dla którego opracowano oddzielne schematy zastosowań dla transportu drogowego, kolejowego, wodnego, powietrznego oraz tras kablowych;
2. Gdy temat danych zawiera bazowy model danych, prawnie obowiązujący przy wdrożeniu, a także jeden lub większą liczbę rozszerzonych modeli danych, które są rekomendowane, ale nie obowiązkowe;
3. Gdy konieczne jest jednoznaczne modelowanie różnych poziomów agregacji (różne skale lub rozdzielczości).

Reguły modelowania koncepcyjnego regulują sposób reprezentacji świata rzeczywistego w formie schematów zastosowań. Wspólny Słownik Pojęć Obiektów, prowadzony dla wszystkich tematów danych, przyczynia się do spójności danych i eliminuje nadmiar.

Zasady dotyczące schematów zastosowań obejmują konstrukty modelowania wykorzystywane w budowie schematów zastosowań. Prostsze modele jednorodnie usprawniają zarówno proces specyfikowania, jak i wdrażanie specyfikacji przez dostawców danych.

Zastosowanie wspólnego języka schematów koncepcyjnych<sup>25</sup> do dokumentacji formalnej modeli danych umożliwia automatyczną obsługę schematów zastosowań. Współcześnie najczęściej stosowanym językiem schematów koncepcyjnych jest Unified Modeling Language (UML) - ujednolicony język modelowania. Podmioty zainteresowane SDI mogą uzgodnić profil UML, tj. ewentualne ograniczenia co do wykorzystywanych elementów UML.

Katalog obiektów stanowi analogiczną reprezentację informacji w schemacie zastosowania. Katalogi obiektów odgrywają ważną rolę, ponieważ:

- Wspierają konwersję informacji ze schematu zastosowania na tekst czytelny dla ludzi;
- Wspierają wielojęzyczność, ponieważ są tłumaczone na języki zainteresowanych (schemat zastosowania powinien być zarządzany tylko w jednym wspólnym języku);
- Usprawniają wyszukiwanie oraz dostęp do indywidualnych elementów schematu zastosowania, przez użytkowników-ludzi oraz przez oprogramowanie, ponieważ są publikowane za pośrednictwem usługi rejestru.

#### 4.2.4 Wspólne schematy zastosowań

Ten element interoperacyjności danych gromadzi modele komponentów, możliwe do ponownego wykorzystania, znajdujące zastosowanie w wielu schematach. Na Rys. 7, str. 24, komponenty do ponownego wykorzystania znajdują się na przecięciu dwóch tematów danych. Takie schematy mogą być definiowane dla infrastruktury lub importowane z innych inicjatyw.

Niewielkim, ale szeroko stosowanym modelem jest schemat niepowtarzalnych identyfikatorów. Struktura niepowtarzalnych identyfikatorów opisana jest w rozdziale 4.3.1. Kolejnym przykładem jest wspomniany już standardowy model sieciowy. Schemat zastosowania "Observacje i pomiary" jest wspólny dla wielu tematów danych INSPIRE, w tym Urządzenia do monitorowania środowiska, Warunki oceanograficzno-geograficzne, Warunki atmosferyczne, Warunki meteorologiczno-geograficzne, Gleby, Geologia.

Wspólne schematy zastosowań są istotnymi narzędziami wzmacniania spójności między tematami oraz interoperacyjności. Uzasadnione jest zatem sprawdzenie istniejących schematów zastosowań przed opracowaniem nowego tematu danych w SDI. Skonsolidowane repozytorium modelu, opisane w następnym rozdziale, zapewnia bezpośredni dostęp do wszystkich schematów zastosowań opracowanych w kontekście danego SDI. Jeśli repozytorium takie nie istnieje, opracowujący zobowiązani są do sprawdzenia norm i innych materiałów odniesienia, opisanych w punkcie rozdziału 5.3 dotyczącym analizy stanu obecnego.

#### 4.2.5 Skonsolidowane repozytorium modelu

W kontekście SDI, gdzie różne grupy tematyczne rozwijają i utrzymują modele danych, kluczowe znaczenie ma dostępność wszechstronnego, ale krótkiego przeglądu wszystkich uzgodnień i wyników procesu modelowania danych. Konieczne jest posiadanie konkretnego narzędzia do zapewnienia takiego przeglądu, tym samym umożliwiającemu konsekwentne wielokrotne wykorzystywanie modeli opracowanych przez inne grupy.

W procesie specyfikowania danych w INSPIRE przyjęto **skonsolidowane repozytorium modelu**, zawierające uzgodnione modele fundamentalne (np. ISO i inne normy), standardowy model koncepcyjny oraz schematy zastosowań dla tematów danych. Wprowadzenie skonsolidowanego repozytorium modelu było jedynym możliwym sposobem opracowywania spójnych modeli danych i schematów zastosowań dla 34 tematów danych przestrzennych, ponieważ umożliwiało ono grupom ekspertów pracujących nad modelami danych tematycznych wzajemne śledzenie prowadzonych przez siebie prac oraz stwierdzanie podobieństw, pokrywających się obszarów i luk w metodologiach modelowania. Doświadczenie INSPIRE wykazało istotną wartość tego podejścia, którą można podsumować następująco:

Po pierwsze, modele fundamentalne są rozproszone w obrębie różnych norm i prezentowane są zwykle w formie statycznych grafik lub wykresów. Skonsolidowane repozytorium modelu udostępnia je wszystkie w jednym miejscu w formie umożliwiającej ponowne wykorzystywanie. Przy pomocy konkretnego oprogramowania do modelowania informacji można pracować bezpośrednio na modelach danych zawartych w tych normach, importując ich odpowiednie komponenty (profile) do tematycznych modeli danych. W rezultacie standardy wdrażane są w poszczególnych tematach w podobny sposób.

Po drugie, każdy obiekt przestrzenny - niezależnie od schematu zastosowania lub tematu, w obrębie którego został stworzony - może być przedmiotem odwołania przez inne schematy zastosowań (w innych tematach). Jest to krok o kluczowym znaczeniu dla poprawy poziomu spójności pomiędzy modelami danych w obrębie różnych tematów, a tym samym dla interoperacyjności.

Po trzecie, prezentacja modeli danych w języku schematów koncepcyjnych (np. UML z zastosowaniem ISO/TS 19103:2005) i w formie graficznej (np. jako wykresy UML) zapewnią szybkość, zrozumiałą prezentację danych, które są ponadto czytelne dla maszyn. Prezentacje opisowe schematów (katalogi obiektów), a także elementy Słownika Pojęć Obiektów można uzyskać automatycznie na podstawie dokumentacji modeli danych w skonsolidowanym repozytorium. Opcja ta pomaga uniknąć niespójności w opisowej dokumentacji specyfikacji.

Repozytorium umożliwi wreszcie automatyczne generowanie modeli z zastosowaniem schematu kodowania GML/XML<sup>26</sup>. Zaleca się udostępnianie zarówno modeli UML, jak i schematów kodowania GML/XML w formie rejestrów w obrębie infrastruktury w celu wsparcia pobierania i wdrażania modeli. Przykładowo, zainteresowani mogą wykorzystywać modele UML jako podstawę do tworzenia rozszerzeń spełniających wymagania właściwe dla domeny lub kraju/regionu. Mogą one być także wykorzystywane przez zainteresowanych do automatycznego generowania innych rodzajów kodowania.

Ze względów wdrożeniowych kluczowe znaczenie ma dostępność schematów kodowania związanych z konkretną specyfikacją danych, np. w celu umożliwienia automatycznej walidacji. Gdy modele i schematy są aktualizowane w ramach procedury przechowywania, ogromne znaczenie ma dostępność różnych wersji modelu danych i kodowania ze względu na możliwość stwierdzenia ich statusu (prawidłowe, pogorszone itp.).

#### 4.2.6 Wiele reprezentacji

Jak wspomniano w rozdziale 2.1, zjawiska ze świata rzeczywistego mogą być opisywane na różnych poziomach szczegółowości. Są one wyrażane w formie poziomów agregacji pojęć stosowanych przy abstrahowaniu (pojedyncze domy a teren zabudowany) i/lub przy odwzorowaniach przestrzennych (rzeka reprezentowana jako powierzchnia lub oś środkowa). Skala/rozdzielczość jest wybierana każdorazowo jako funkcja konkretnych wymagań użytkownika.

W razie wystąpienia potrzeby zastosowania innych skal/rozdzielczości w odniesieniu do określonego tematu w SDI, możliwe jest modelowanie różnych

<sup>25</sup>

Język schematów koncepcyjnych to formalny język oparty na formalizmie koncepcyjnym do celów reprezentacji schematów koncepcyjnych (ISO 19101:2005). Jest on zwykle czytelny dla maszyn, co jest pomocne w procesie przejścia do schematów kodowania.

<sup>26</sup>

Kwestia kodowania omówiona jest bardziej szczegółowo w rozdziale 4.3.10.

poziomów szczegółowości bezpośrednio z zastosowaniem oddzielnych schematów zastosowań, zapewniających wieloraką reprezentację świata rzeczywistego. Aby zachować spójność reprezentacji, schematy zastosowań muszą być wzajemnie powiązane. Proces agregacji przestrzennej powinien być wspierany przez hierarchie generalizacji-specjalizacji modelu. Przykładowo, obiekt przestrzenny zdefiniowany jako grupa domów w odwzorowaniu w małej skali powinien być powiązany z domami w odwzorowaniu w dużej skali poprzez zastosowanie stosunku agregacji. Praktyka ta ma korzystny wpływ na utrzymanie danych, ponieważ wspiera automatycznie przenoszenie aktualizacji z dużej skali do małej. I tak, w powyższym przykładzie, powierzchnia bloku zmieni się automatycznie na podstawie liczby domów powiązanych z tym blokiem.

Reprezentacja wieloraka zwiększa poziom złożoności schematów zastosowań. Dlatego też użycie tej metodologii powinno być uzasadnione konkretnymi wymaganiami użytkowników. Co do zasady zalecane jest modelowanie możliwie najmniejszej liczby poziomów szczegółowości. Doświadczenie z INSPIRE pokazuje, że w przypadku znacznej większości tematów danych udało się pozostać przy jednym standardowym schemacie zastosowania.

#### 4.2.7 Punkty rozszerzeń

Specyfikacje interoperacyjności są rozwijane przy uwzględnieniu wymogów wspólnych dla wielu użytkowników. W celu wsparcia konkretnych zastosowań lub powiązania informacji biznesowych, użytkownicy mogą być zainteresowani rozszerzeniem specyfikacji danych w infrastrukturze. Rozszerzenia takie mogą stanowić cenny wkład w dalszy rozwój infrastruktury, pod warunkiem, że dane rozszerzenie:

- nie zmieni niczego w specyfikacji docelowego poziomu interoperacyjności, a tylko będzie się do niej odnosiło normatywnie ze wszystkimi swymi wymaganiami, oraz
- nie doda żadnych wymagań naruszających wymagania specyfikacji docelowego poziomu interoperacyjności lub podstawowego modelu koncepcyjnego.

Rozszerzenia mogą dodawać nowe schematy zastosowań, nowe typy obiektów przestrzennych i danych, nowe ograniczenia schematów zastosowań, a także definiować dodatkowe zasady prezentacji itp. Może być także poszerzana lista kodów, o ile infrastruktura nie będzie jej identyfikowała jako listy kodów zarządzanych centralnie.

### 4.3 Zarządzanie danymi

#### 4.3.1 Zarządzanie identyfikatorami

Niepowtarzalne identyfikatory (unique identifiers, UID) są niezbędne do odnoszenia nowych obiektów przestrzennych do istniejących oraz pozyskiwania danych geograficznych. Można wyróżnić dwa typy identyfikatorów: zewnętrzne identyfikatory obiektów, które w unikalny sposób identyfikują wyabstrahowany obiekt przestrzenny, oraz identyfikatory tematyczne, służące do unikalnej identyfikacji zjawisk ze świata rzeczywistego.

Identyfikatory zewnętrzne powinny spełniać następujące warunki:

- Unikalność: żadne dwa obiekty przestrzenne nie mogą mieć tego samego identyfikatora,
- Trwałość: identyfikator nie zmienia się przez cały okres życia obiektu przestrzennego i nigdy nie jest przypisywany na nowo,
- Identyfikowalność: istnieje mechanizm wyśledzenia obiektu przestrzennego w infrastrukturze w oparciu o jego identyfikator,
- Wykonalność: UID może być stworzony w infrastrukturze na podstawie UID przechowywanych przez inne organizacje.

Identyfikatory przypisane w obrębie aplikacji GIS nie spełniają kryterium unikalności, ponieważ nie ma gwarancji, że ta sama sekwencja znaków alfanumerycznych nie jest wykorzystywana w innym miejscu lub zastosowaniu. Dlatego też niepowtarzalne identyfikatory muszą być niezależne i składać się z dwóch elementów:

- Przestrzeń nazwy do identyfikacji źródła danych. Przestrzeń nazwy jest własnością dostawcy danych i powinna być zarejestrowana w Rejestrze przestrzeni nazw,
- Identyfikator lokalny, przypisywany przez dostawcę danych. Identyfikator lokalny jest unikalny dla przestrzeni nazw, tj. żaden inny obiekt przestrzenny nie nosi tego samego unikalnego identyfikatora.

Tematyczne identyfikatory obiektów (np. identyfikatory położenia ICAO dla lotnisk czy kody NUTS dla jednostek statystycznych) noszą w sobie zakodowaną wiedzę istotną dla SDI. Jednakże w większości przypadków nie można ich uznać za identyfikatory zewnętrzne, głównie z powodu niespełnienia wszystkich czterech opisanych wyżej warunków. Dlatego też powinny one stanowić atrybuty tematyczne obiektów przestrzennych.

Identyfikatory tematyczne mogą służyć do ustalania relacji pomiędzy obiektami przestrzennymi w różnych zbiorach danych odnoszących się do tego samego obiektu rzeczywistego. Przykładowo, obiekty ze zbioru danych zawierającego informacje o geometrii sieci rzek mogą być zintegrowane z obiektami z innego zbioru danych, zawierającego informacje o jakości wód, jeśli oba zbiory wykorzystują ten sam identyfikator tematyczny, np. identyfikator rzeki (odcinka), zgodnie z przepisami środowiskowymi lub rejestrem. Z tego powodu prowadzony jest również system identyfikatorów tematycznych dla obiektów rzeczywistych, np. w ramach brytyjskich działań na danych otwartych (Rada Głównego Urzędu ds. Technologii - Chief Technology Officer Council 2011).

#### 4.3.2 Spójność danych

Po przekształceniu<sup>27</sup> danych zgodnie ze specyfikacjami interoperacyjności mogą pozostać pewne różnice szcążkowe,<sup>28</sup> jeśli dane są integrowane z różnych źródeł. Ze względów spójności dostawcy danych muszą uzgadniać swoje dane w oparciu o wzajemne porozumienia dotyczące klasyfikacji i/lub położenia odnośnych obiektów przestrzennych.

Ten element interoperacyjności zapewnia wskazówki na temat zakresu zastosowania dopasowywania danych i sposobu organizacji tego procesu. Niektórych tematów w infrastrukturze, takich jak np. warunki atmosferyczne, warunki meteorologiczno-geograficzne, warunki oceanograficzno-geograficzne czy regiony morskie itp., komponent ten dotyczy w mniejszym stopniu ze względu na ich transgraniczny, przejściowy lub nieostry charakter. Dopasowywanie danych pozycyjnych nie odnosi się do niejednoczesnych zbiorów danych. "Różbieżności" związane z różnicami czasowymi nie są klasyfikowane jako niezgodności w dosłownym tego słowa znaczeniu.

Gdy dopasowywanie danych jest uzasadnione, np. wzdłuż granic, dostawcy danych powinni uzgodnić "faktyczne" położenie obiektów przestrzennych, których dopasowywanie dotyczy, lub zasady prowadzenia procesu dopasowywania. Spójność pomiędzy różnymi tematami powinna być wymagana wyłącznie w obrębie tego samego lub bardzo podobnego poziomu szczegółowości.

Jeśli różne informacje geograficzne dotyczą tego samego położenia, należy ująć naturalne zależności. I tak na przykład droga i rzeka nie mogą się przecinać bez mostu, tunelu lub połączenia promowego. Wstępny wykaz wzajemnych zależności pomiędzy tematami wypracowany jest w procesie ustalania zakresów<sup>29</sup>.

#### 4.3.3 Jakość danych i informacji

Jakość danych stanowi istotny aspekt w przypadkach, gdy użytkownicy muszą zdecydować o zdolności danych do użytku. Dla wygody użytkowników prezentacja jakości danych powinna być w miarę możliwości zbliżona w obrębie wszystkich tematów.

Z punktu widzenia SDI słaba jakość danych może narazić na szwank interoperacyjność. Nie powinno się jednak wykluczać z infrastruktury żadnych danych na podstawie słabej jakości. "Słabe" dane są lepsze niż brak danych. W konsekwencji należy starannie oszacować, które wymagania są niezbędne do prawidłowego funkcjonowania infrastruktury. Przykładowo, z punktu widzenia interoperacyjności wymagania spójności logicznej (definiujące semantykę i strukturę danych) są "ważniejsze" niż wymagania dokładności pozycyjnej.

W kontekście SDI zamiast ustalania apriorycznych wymogów jakości danych bardziej zasadne byłoby rekomendowanie docelowych wyników. Docelowe wyniki zależne są także od rodzaju danych - surowsze parametry dotyczą danych referencyjnych, służących jako odniesienia dla obiektów.

Celem tego elementu interoperacyjności jest zbudowanie modelu koncepcyjnego dla odpowiednich elementów jakości danych, zdefiniowanych w odnośnych normach<sup>30</sup>, a także progów żądanych wyników docelowych w testach zgodności<sup>31</sup>. Celem ostatecznym jest zapewnienie użytkownika końcowego w pewien sposób o miarodajności informacji z wykorzystaniem identyfikowalnych wskaźników<sup>32</sup> lub parametrów jakości danych dla wybranych

<sup>27</sup> O przekształcaniu danych mowa jest w rozdziale 4.3.7.

<sup>28</sup> Zob. przykłady w rozdziale 2.2.

<sup>29</sup> Zob. rozdział 5.1.

<sup>30</sup> ISO 19113, ISO/TS 19138, która zostanie zastąpiona przez ISO 19158

<sup>31</sup> Zob. szczegóły zawarte w rozdziale 4.3.5.

<sup>32</sup> Zob. QA4EO GEOSS

elementów jakości danych (takich jak kompletność, spójność, aktualność, poprawność itp.) lub zgodności zbioru danych jako całości.

#### 4.3.4 Metadane

Metadane dostarczają "informacji na temat identyfikacji, zasięgu, jakości, schematu przestrzennego i czasowego, odniesienia przestrzennego oraz rozkładu cyfrowych danych geograficznych" (ISO TC 211, 2003a). Metadane opisujące zasoby geoprzestrzenne są ściśle związane z reprezentowanymi przez nie danymi. Dlatego też idealny cykl rozwoju usprawnia oba obszary.

Ze względów organizacyjnych metadane oraz rozwijane specyfikacje danych są niekiedy rozdzielone poprzez rozgraniczenie metadanych wyszukiwania i metadanych do celów oceny oraz wykorzystywania. Jest to uzasadnione przewidywaniem udostępniania danych w obrębie infrastruktury, również wówczas, gdy dane są niezgodne z docelowymi specyfikacjami interoperacyjności. Dlatego też metadane wspierające wyszukiwanie i ocenę pierwszego stopnia (tj. opis podstawowych właściwości technicznych, takich jak: skala/rozdzielczość, zasięg geograficzny, forma reprezentacji przestrzennej itp.) są publikowane w celu uzupełnienia lub uszczegółowienia przy pomocy metadanych pochodzących z procesu specyfikacji danych.

Metadane stanowią podstawowe zasoby dostarczające informacji na temat faktycznej jakości danych. W przeciwieństwie do danych apriorycznych, metadane dotyczące jakości danych zapewniają ocenę ex post, która w kontekście interoperacyjnego wykorzystania danych w SDI uzależniona jest od dwóch głównych czynników:

- jakości danych wejściowych oraz
- pomyślnego przebiegu procesu przekształcenia, niezbędnego do osiągnięcia interoperacyjności.

Po przekształceniu danych na potrzeby infrastruktury metadane dotyczące danych wyjściowych mogą utracić ważność. Ściśle mówiąc, powinny one zostać poddane ponownej ocenie lub przekształceniu, jeśli przekształcenia danych powodują systematyczne zmiany ich jakości. Jest to dodatkowo obciążeniem dla dostawców danych i może mieć niższy priorytet przy budowie infrastruktury. Tymczasowe rozwiązanie może polegać na publikacji oryginalnych metadanych z opisem kroków procesu przekształcenia w celu zapewnienia użytkownikom wystarczającego zakresu informacji na temat jakości danych.

Użytkownicy mogą także ocenić stosowność i użyteczność danych na podstawie metadanych. Niepewność związana z "obiektywnymi" parametrami jakości danych oraz potencjalnie subiektywnymi opisami użyteczności tworzy niekiedy więcej barier niż wsparcia dla użytkowników. Rozwiązaniem przyjaznym dla użytkowników może być w takim przypadku certyfikacja i etykietowanie produktów. Inicjatywa "GEO Label" służyć ma znakowaniu jakości produktów z zakresu obserwacji Ziemi na podstawie należycie zdefiniowanych parametrów oceny jakości danych lub informacji z systemu (GEO Komitet Zadaniowy ST-09-02 2010).

#### 4.3.5 Zgodność

Norma ISO 19105 definiuje zgodność jako spełnienie określonych wymogów. Z oczywistych powodów zgodność danych w SDI wymaga weryfikacji względem specyfikacji docelowego poziomu interoperacyjności. Zakres oceny zgodności może dotyczyć pojedynczego elementu specyfikacji (np.: schematu zastosowania, reguł rejestracji danych, wybranych elementów jakości danych itp.) lub może być zagregowany do poziomu specyfikacji jako całości.

Każdy produkt mający zachowywać zgodność ze specyfikacjami jako całość musi przejść pomyślnie wszystkie testy opisane w pakietach testów abstraktów (ATS), które odwołują się do badanych wymogów i zawierają wykaz odnośnych testów, parametrów jakości i odpowiednich wartości progowych.

Zbiór danych może być jednocześnie zgodny z jedną lub większą liczbą specyfikacji. Aby dostarczyć użytkownikom pełne informacje na temat zgodności danych, zaleca się deklarowanie zgodności ze wszystkimi specyfikacjami, względem których dane są testowane.

#### 4.3.6 Zasady rejestracji danych

Zasady rejestracji danych stanowią wytyczne co do wyboru zjawisk rzeczywistych, które powinny zostać uwzględnione w temacie danych. Są to również główne elementy wykorzystywane przy określaniu docelowego poziomu szczegółowości. Do typowych kryteriów wyboru należą: minimalna powierzchnia, długość lub parametry funkcjonalne.

Ponieważ SDI dotyczą zwykle danych istniejących, ustalenie metod rejestracji danych (np. metody topograficzne i pomiarowe, odpowiednie typy czujników itp.) nie dotyczą niniejszego komponentu.

#### 4.3.7 Model przekształceń danych/wytyczne

W dobrej infrastrukturze SDI wszyscy dostawcy danych publikują swoje dane zgodnie z uzgodnionymi specyfikacjami interoperacyjności. Cel ten można osiągnąć poprzez utrzymanie zgodności danych ze specyfikacjami interoperacyjności w zakresie bezpośredniego dostępu poprzez usługę pobierania. Możliwości zastosowania tego rozwiązania są ograniczone. Z jednej strony zainteresowane społeczności wykazują ściśle określone wymogi przestrzegania ich własnych specyfikacji. Z drugiej strony, w szczególności w przypadku infrastruktur ponadnarodowych, niemal zawsze konieczne jest przekształcenie współrzędnych odwzorowania.

Dlatego też najlepszym rozwiązaniem teoretycznym jest utrzymanie pierwotnej struktury danych i publikowanie danych SDI poprzez przekształcanie. Przekształcenie pomiędzy źródłowym a docelowym schematem zastosowania stanowi kluczowy typ przekształcenia, mogą być jednak wymagane również inne przekształcenia (np.: przekształcenie współrzędnych, dopasowanie krawędzi, tłumaczenie na inny język, przekształcenie formatu itp.).

Aby dane mogły być udostępnione za pośrednictwem usługi pobierania, są one zwykle przekształcane off-line do widoku statycznego, zgodnego ze specyfikacją docelowego poziomu interoperacyjności. Alternatywnie dane mogą być przekształcane na bieżąco w obrębie usługi pobierania, zgodnie ze zdefiniowanymi uprzednio regułami mapowania. Trzecia możliwość polega na wykorzystaniu odrębnej usługi przekształcania, wykonującej zdefiniowane wcześniej lub definiowane przez użytkownika mapowanie. Obowiązkiem każdego dostawcy danych powinno być wybranie metody i zapewnienie niezbędnego przekształcenia danych według własnego uznania.

#### 4.3.8 Zasady przechowywania danych

Ponieważ infrastruktura jest oparta o dane istniejące, przechowywanie zbiorów danych odbywa się u źródła, czyli u dostawców danych, według ich własnych procesów biznesowych. Istnieją dwa problemy wymagające rozwiązania z punktu widzenia infrastruktury:

- Zapewnienie, aby aktualizacje były terminowo przekazywane do podmiotów publikujących dane, zgodnie z obowiązującymi ich specyfikacjami interoperacyjności;
- Zapewnienie mechanizmu rozróżniania pomiędzy danymi bieżącymi i historycznymi.

Pierwszy problem zostanie rozwiązany automatycznie, gdy dane będą przechowywane przez dostawców danych w sposób zgodny ze specyfikacjami interoperacyjności lub gdy przekształcenia prowadzące do osiągnięcia interoperacyjności są zautomatyzowane. W takim przypadku dane w SDI są uaktualniane przy minimalnym wkładzie wysiłku ludzkiego.

Gdy przekształcanie danych odbywa się off-line, należy zwrócić szczególną uwagę na problem przenoszenia aktualizacji na prezentowane dane zgodnie ze specyfikacjami interoperacyjności. I tak, należy uzgodnić lub uregulować maksymalny czas opóźnienia we wprowadzaniu zmian.

Co do zasady, możliwość uaktualniania danych zależy od dostępności informacji na temat cyklu życia w schemacie zastosowania, dokumentujących czas, w jakim wstawiono nowe obiekty przestrzenne bądź też zmieniono lub usunięto istniejące obiekty przestrzenne. Informacje dotyczące cyklu życia mogą być wykorzystywane w kwerendach przy wyborze tylko tych obiektów przestrzennych, których dotknęły zmiany zachodzące od momentu wskazanego przez użytkownika.

#### 4.3.9 Prezentacja

Prezentacja graficzna informacji geograficznej zależy od wielu czynników, takich jak treść informacji, medium reprezentacji<sup>33</sup>, ewentualne konwencje prezentacyjne w społecznościach zainteresowanych itp. W SDI główny nacisk kładzie się na ponowne wykorzystywanie i łączenie danych z różnych źródeł, przez co tworzy się nieskończoną różnorodność danych, które muszą ze sobą współistnieć w trakcie analizy przestrzennej. Harmonizacja zasad prezentacji jest zatem zadaniem złożonym.

Zgodnie z zasadą wdrażania w etapach, na pierwszym etapie można dążyć do obsługi wyłącznie usługi podglądu, wykorzystywanej w fazie wykrywania. Takie podejście zastosowano w INSPIRE, gdzie do prezentacji podchodzi się z perspektywy pojedynczych tematów. Schematy reguł prezentacji oraz symbolika obiektów geograficznych określają podstawowe zasady (struktura warstwowa) oraz ustandaryzowany zestaw domyślnych stylów.

Najczęściej stosowane metody wizualizacji oparte są o deskryptor warstw OGC Styled Layer Descriptor (SLD), umożliwiający ustalanie symboli i barw

definiowanych przez użytkownika podczas wyświetlania danych w sieciowej usłudze mapowania - Web Mapping Service (WMS).

Keyhole Markup Language (KML)<sup>34</sup> to język XML, koncentrujący się na wyświetlaniu/wizualizacji danych geograficznych, w tym opisach map i zdjęć. Wizualizacja geograficzna odwzorowuje dane graficzne na kuli ziemskiej i prowadzi nawigację użytkownika pod względem kierunku ruchu i patrzenia.

Aby uniknąć zderzeń stylów stosowanych w obrębie różnych tematów, konieczne jest wprowadzenie pewnego poziomu podstawowej harmonizacji. Gdy nie ma harmonizacji, np. ta sama niebieska linia może odzwierciedlać batymetrię, ciekę i granice obszarów morskich. Współdzielenie deskryptorów SLD w rejestrze może korzystnie oddziaływać na ten proces harmonizacji, np. poprzez umożliwienie kwerend stylów zdefiniowanych dla różnych tematów danych. Można także wykorzystać rejestr do udostępniania stylów definiowanych przez użytkowników (np. do określonych celów, takich jak mapowanie strefy przybrzeżnej).

#### 4.3.10 Dostarczanie danych

Do wymiany danych przestrzennych niezbędne są skuteczne metody kodowania i dostarczania danych. Reguła kodowania określa typ danych do konwersji, a także składnię, strukturę i schematy kodowania. Prezentuje ona dane w formacie zdatnym do transportu i przechowywania. Jednoznaczne zdefiniowanie formatów danych jest pomocne przy zapewnieniu interoperacyjności pod względem składniowym.

Ze względu na różnorodność danych występujących w obrębie infrastruktury (wektorowe, rastrowe itp.) nie można zezwolić na unikalną regułę kodowania i strukturę danych wyjściowych. I tak, każda specyfikacja danych powinna określać przynajmniej jedną regułę kodowania obowiązującą dla danego tematu.

Elastyczność w zakresie obsługi dodatkowych reguł kodowania jest podejściem prawidłowym, niemniej równie istotna jest harmonizacja i ograniczenie rozrzutu reguł kodowania. Rozsądne jest prowadzenie wykazu dopuszczalnych reguł kodowania i schematów struktury danych wyjściowych w rejestrze. Reguły kodowania powinny być oparte o standardy międzynarodowe, najlepiej otwarte, a także powinny być zgodne z normą ISO 19118 - Informacja geograficzna - Kodowanie.

W INSPIRE, o ile dla danego tematu danych nie określono inaczej, rekomendowanym kodowaniem jest OGC Geography Markup Language (GML), również zdefiniowany w ISO 19136. W przypadku dużej ilości danych pokrycia, np. sporządzania ortobrazów czy symulacji komputerowych (np. prognoz pogody), można zdefiniować inne, bardziej efektywne kodowanie na poziomie plików (np. geoTIFF) jako domyślny język kodowania. Takie schematy kodowania są powszechnie obsługiwane i mogą być wstawiane w większości GIS.

W SDI dane przestrzenne są dostępne za pośrednictwem usługi pobierania i podglądu. Ten komponent interoperacyjności obejmuje także usługi wykorzystywane przy dostarczaniu danych oraz odniesienie do formatów kodowania stosowanych w wymianie danych pomiędzy systemami.

## 5 Metodologia rozwijania specyfikacji danych

### 5.1 Definicja zakresu tematów danych

Definiowanie tematów w INSPIRE rozpoczęło od analizy wymogów europejskiego prawodawstwa w obszarze środowiska. Wstępna lista sporządzona w dokumencie Tematycznej Grupy Koordynacji - Środowisko była szeroko omawiana przed jej zdefiniowaniem w załącznikach do Dyrektywy. Ze względu na zmiany wprowadzone w procesie konsultacji konieczne było ponowne zajęcie się opisami tematów przed rozpoczęciem procesu specyfikowania danych. Czynności te zostały przeprowadzone przez Zespół ds. Sporządzania Specyfikacji Danych w dokumencie pn. "Definicja tematów i zakresu przedmiotowego Załącznika". Dokument ten, definiujący wzajemne zależności pomiędzy tematami, stanowił istotny wkład w przebieg procesu specyfikowania danych.

Definiowanie zakresu tematów danych i infrastruktury wymaga starannego rozważenia oraz wypracowania konsensusu pomiędzy zainteresowanymi, w tym użytkownikami danych, ich twórcami, dostawcami technologii oraz politykami odpowiedzialnymi za rozwój strategiczny w odnośnym obszarze. Wśród przykładów instrumentów, jakie mogą być wykorzystane w tym procesie, wymienić można: operaty, badania stanu obecnego, pisemne opinie formalne, konsultacje przez Internet, wysłuchania publiczne.

Ze względów historycznych i organizacyjnych gromadzeniem oraz przechowywaniem danych przestrzennych zajmuje się wiele organizacji. Ich działania nie zawsze są skoordynowane, mogą więc występować obszary nakładające się lub luki w treści danych. Ponieważ redundancja stanowi ważne źródło niespójności danych, konieczne jest zarysowanie granic pomiędzy tematami danych. Wyraźne określenie zakresu przedmiotowego tematów danych pomoże zainteresowanym w ocenie tego, w jaki sposób powstająca infrastruktura może wpłynąć na ich interesy i w których punktach mogą być zmuszeni do współdziałania.

Przy stwierdzeniu nakładających się obszarów pomiędzy dwoma lub większą liczbą tematów należy podjąć następujące decyzje:

- Czy wyraźnie nakładające się części są uzasadnione z koncepcyjnego punktu widzenia? Czy obiekty przestrzenne opisują **różne** abstrakcje tego samego zjawiska rzeczywistego (np. odcinek rzeki jako element hydrografii / odcinek rzeki jako część elementu nawigacji śródlądowej)? Jeżeli tak, takie obiekty przestrzenne powinny być modelowane w obrębie obu tematów. W przeciwnym wypadku należy zdecydować, który z tematów jest najważniejszy dla danego obiektu przestrzennego.
- Jeżeli rozgraniczenie jest uzasadnione pod względem koncepcyjnym, w jaki sposób można uwidocznić różnicę (wybór terminologii), jakie są punkty krytyczne wyznaczające tę różnicę i czy należy ustalić relację pomiędzy dwoma pojęciami (np. identyfikacja hydrologicznych odcinków rzek, którym odpowiadają odcinki rzek w ujęciu transportu wodnego)?

Należy zauważyć, że ramy koncepcyjne nie uwzględniają ograniczeń organizacyjnych i nie radzą sobie z nimi (np. która organizacja duplikuje informacje w przypadku nieuzasadnionego nakładania się obszarów?). Wskazują one tylko miejsca wymagające podjęcia działań w zakresie koordynacji. Koordynacja jest równie niezbędna w przypadku wykrycia wzajemnych zależności pomiędzy dwoma lub większą liczbą tematów danych.

Na podstawie istotnych przypadków praktycznych i materiałów referencyjnych proces ustalania zakresu wyszczególnia również możliwą zawartość tematów danych pod względem kluczowych typów obiektów przestrzennych i ich atrybutów. Taka niewyczerpująca lista nie powinna być uznawana za próbę zdefiniowania całej zawartości, a raczej za ilustrację poglądową. Właściwa analiza odniesień oraz definiowanie wymogów dotyczących danych powinny odbywać się w trakcie rozwijania specyfikacji danych. Głównym efektem procesu ustalania zakresu jest prawidłowo zdefiniowany punkt wyjścia dla procesu specyfikowania danych.

### 5.2 Zasady rozwijania specyfikacji danych

W ramach koncepcyjnych metodologii rozwijania specyfikacji kieruje procesem w taki sposób, by przestrzegane były ogólne zasady SDI, takie jak ponowne wykorzystanie, wykonaność i proporcjonalność. Metodologia dostarcza wytyczne co do działań, które należy podejmować na różnych etapach procesu.

Proces rozwijania specyfikacji może być stymulowany przez dostawców i/lub użytkowników danych. W wersji uzależnionej od dostawców podstawowa zasada polega na znalezieniu wspólnego mianownika dla istniejących zbiorów danych, należących do określonego tematu. Jednak bez zewnętrznych wzorców wymagania interoperacyjności mogą w przypadku takiego podejścia pozostawać niejasne, co może prowadzić do wystąpienia następujących problemów:

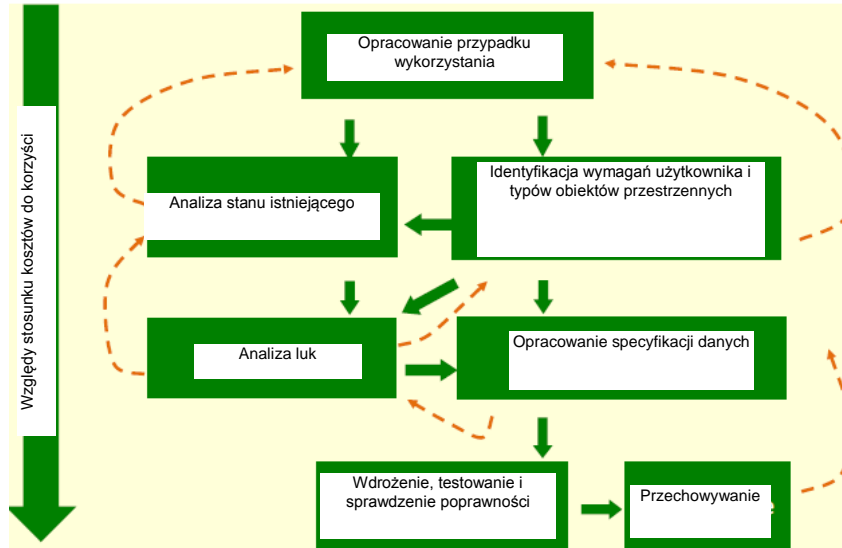
- Dane dostarczone zgodnie z rozwiązaniami w zakresie interoperacyjności nie są zgodne z wymaganiami użytkowników;
- Zamiast dążyć do optymalizacji poziomu interoperacyjności, najważniejsi zainteresowani mogą promować własne rozwiązania w celu ograniczenia do minimum potencjalnych przekształceń/zmian w obrębie wytwarzanych przez siebie zbiorów danych.

W wersji zależnej od użytkowników wzorce zewnętrzne wynikają z wymogów użytkowników, które są starannie analizowane i formalizowane na początku procesu rozwijania specyfikacji. Takie podejście może wiązać się z następującymi obszarami ryzyka:

- Trudność z określeniem z góry szczegółowych wymogów użytkowników;
- Wyrażone wymagania mogą być zbyt ambitne, powodując nadmierne koszty lub niemożność wdrożenia w oparciu o istniejące dane;
- Zamiast koncentrować się na wielokrotnym wykorzystywaniu, proces specyfikowania może doprowadzić do uzyskania specyfikacji produktu spełniającego potrzeby "silnego" użytkownika.

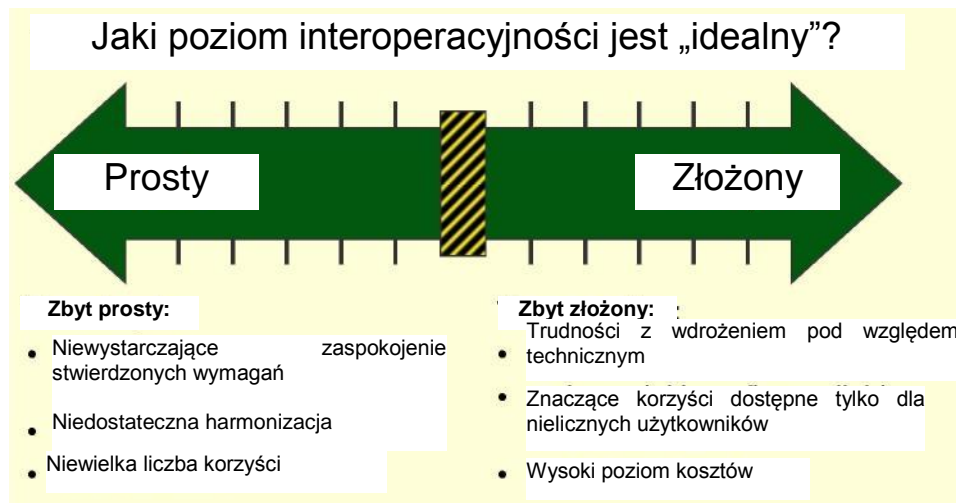
Doświadczenie pokazuje, że w praktyce najczęściej stosuje się połączenie tych dwóch metod, tj. zbilansowanie aspiracji i możliwości technicznych oraz finansowych.

Metodologia opisana w niniejszym rozdziale dotyczy szczegółów procesu rozwijania specyfikacji danych stosowanego w INSPIRE. W metodologii tej uwzględniono wyniki i doświadczenia z projektów badań naukowych<sup>35</sup> oraz najlepsze praktyki rozwijania SDI. Ponadto metodologia ta została formalnie opisana i przetestowana oraz uzyskano materialne wyniki dla każdego z 34 tematów danych uwzględnionych w INSPIRE. INSPIRE stosuje metodę iteracyjną w stosunku do rozrastającej się stopniowo infrastruktury SDI, opartą na zaangażowaniu zainteresowanych. Taki przewidywalny i powtarzalny model procesu rozwoju pozwala na wypracowanie wykonalnych i wzajemnie satysfakcjonujących rozwiązań systemowych. Główne kroki procesu wskazano na Rys. 9.



Rysunek 9: Kroki cyklu specyfikowania danych

Ta metodologia jest pomocna w zbilansowaniu ambicji i wykonalności. Jeśli ambicje są zbyt duże, może to prowadzić do opracowania złożonych specyfikacji, których wdrożenie będzie trudne i kosztowne. Ponadto jeśli specyfikacje są zbyt złożone, istnieje ryzyko, że nie będą obsługiwane przez społeczności dostawców danych i nie zostaną przyjęte przez użytkowników. Jednak zbyt uproszczone specyfikacje danych mogą prowadzić do uzyskania niewystarczającego poziomu interoperacyjności i nieosiągnięcia masy krytycznej zapewniającej zasadność podejmowanych wysiłków, wskutek czego nie będzie można stwierdzić wymiernych korzyści z infrastruktury. Główne trudności do rozwiązania przedstawione są na Rys. 10.



Rysunek 10: Trudności w znalezieniu równowagi w procesie specyfikacji danych

Prawidłowe podejście do zbilansowania polega na zastosowaniu dwóch zasad:

1. Działania powinny koncentrować się na tworzeniu spójnych informacji przestrzennych (i czasowych) do szerszego wykorzystania, z pominięciem informacji dotyczących realizacji procesów biznesowych, symulacji naukowych lub konkretnych wymogów sprawozdawczych.
2. Należy zapewnić mechanizmy rozbudowy modeli oraz wykazać możliwość powiązania innych aspektów przestrzennych i nieprzestrzennych z modelami.

W kolejnych rozdziałach zawarto bardziej szczegółowe opisy działań, jakie powinny być podejmowane w procesie specyfikowania danych.

### 5.3 Cykl rozwoju specyfikacji danych

**Analiza przypadku użytkowego** definiuje zestaw zorientowanych na cel interakcji pomiędzy uczestnikami a badanym systemem. Analiza przypadku użytkowego pomaga zrozumieć wymagania użytkowników i definiuje dane niezbędne do spełnienia tych wymagań.

#### Gromadzenie i rozwijanie "analizy przypadku użytkowego"

Faza ustalania zakresu infrastruktury określa, jakie potrzeby użytkowników muszą być spełnione. Poddawane są one następnie dalszemu uszczegółowieniu i dokumentowane we wstępnej fazie procesu specyfikowania danych. Analizy przypadków użytkowych są szeroko stosowane w technologii informatycznej przy formalizowaniu opisów interakcji użytkowników z budowanym systemem. Przy rozwijaniu SDI ilustrują one możliwe sposoby wykorzystywania danych.

Analiza przypadku użytkowego może dotyczyć kilku tematów danych. Przykładowo,

analiza przypadku użytkownika opisująca analizę zagrożenia powodziowego na określonym obszarze może wymagać danych z obszaru hydrografii, wzniesień, meteorologii itp., a w jej wyniku mogą powstać dane wejściowe dla tematu "Naturalne strefy ryzyka". Analizy popularnych przypadków użytkowych pomagają także w wyjaśnianiu ewentualnych współzależności pomiędzy tematami. W związku z powyższym przypadki użytkowe uwzględniane w infrastrukturze powinny odzwierciedlać wielorakość wykorzystywania danych.

Przez wzgląd na prawidłową hierarchizację wymagań konieczne jest uszeregowanie przypadków użytkowych według priorytetów. Wysoki priorytet powinien być nadawany tym analizom przypadków użytkowych, które stanowią część wielu scenariuszy udziału użytkowników lub mają krytyczne znaczenie pod względem czasowym (postępowanie w razie klęski żywiołowej, powodzi itp.). Są to obszary "szybkich korzyści", w których korzyści z SDI przynoszą natychmiastowe i namacalne rezultaty.

W praktyce jednak zgromadzenie analiz przypadków użytkowych od zainteresowanych może być trudne. Użytkownicy danych są mniej świadomi korzyści płynących z SDI lub inicjatyw rozwijania SDI. Nie powinno to jednak narażać na szwank procesu rozwijania specyfikacji. Rozwijanie specyfikacji może się rozpocząć od wstępnych analiz przypadków użytkowych dostarczonych przez dostawców danych, ponieważ zdają sobie oni zwykle sprawę z zadań, do jakich dane są wykorzystywane przez ich klientów. Użytkownicy danych mogą być uaktywniani równolegle. Konsultacje uwzględnione na dalszych etapach cyklu rozwijania specyfikacji danych mogą dostarczać niezbędnych informacji zwrotnych na potrzeby ulepszeń i zbieżności z potrzebami użytkowników.

#### Identyfikacja wymagań użytkowników i typów obiektów przestrzennych

Analizy przypadków użytkowych służą do identyfikacji wymagań w zakresie danych przestrzennych w pierwszej wersji modelu danych. Model ten zawiera listę kandydatów - typów obiektów przestrzennych, projekty definicji i opisów oraz wstępny zbiór innych elementów specyfikacji danych. Każdy z tych elementów jest definiowany według poziomu szczegółowości ustalanego w oparciu o wymagania użytkowników. Koncepcje typów obiektów przestrzennych powinny być wspólne i zharmonizowane pomiędzy poszczególnymi tematami. Pomocnym narzędziem w tym kontekście jest Słownik Pojęć Obiektów<sup>36</sup>.

#### Analiza stanu obecnego

Zgodnie z zasadą, że infrastruktura SDI powinna łączyć dane istniejące, wymagania dotyczące danych pochodzące z analizy przypadków użytkowych powinny być porównane z sytuacją istniejącą. Taka analiza określa, czy żądane dane mogą być dostarczone przez dostawców danych. Jeżeli tak, analiza ta wykazuje również złożoność odnośnych czynności z zakresu przekształceń. Jeśli brak jest prostej relacji jeden-na-jeden pomiędzy proponowanym schematem zharmonizowanym a tematycznymi zbiorami danych, nadal może być wymagana integracja danych na poziomie źródeł danych lub przez użytkowników. Analiza stanu obecnego prowadzona jest często równoległe z analizą luk.

#### Analiza luk

Analiza luk określa wymagania użytkowników, które nie mogą być spełnione przez aktualnie dostępne dane. Luki występują w dwóch rodzajach. Luki techniczne można wypełnić poprzez integrację danych z dowolnego właściwego zbioru danych lub poprzez przekształcenie danych, natomiast luki treściowe mogą być wypełnione wyłącznie poprzez gromadzenie danych. Istniejące badania stanu obecnego mogą stanowić podstawę do porównań.

Wypełnienie luk technicznych tworzy niekwestionowaną wartość dla użytkowników, może się jednak wiązać ze znaczącymi kosztami dla podmiotów tworzących dane. Pomocne mogą być metodologie stabilne pod względem technicznym i oszczędne pod względem kosztowym, takie jak automatyczne narzędzia integracji i przekształcania danych. Jednak na obecnym poziomie zaawansowania technologii takie narzędzia przekształceń nie zawsze są dostępne. Dlatego też należy przyjąć podejście ostrożne, oparte na zestawieniu korzyści z potencjalnymi kosztami.

#### Rozwijanie specyfikacji danych

Pierwsze wersje modeli danych oraz inne wstępne elementy specyfikacji danych, wypracowane w drodze analizy wymagań, należy dostosować według wyników analizy stanu obecnego i analizy luk. Ze względu na wykonalność techniczną i finansową treść specyfikacji danych może być oznaczona do wdrożenia obowiązkowego lub fakultatywnego.

W INSPIRE elementy obowiązkowe definiowane są jako "wymagania", natomiast elementy fakultatywne definiowane są jako "rekomendacje". Zastosowano profile, np. w obrębie tematów: Obszary chronione i Budynki.

Zgodnie z praktyką INSPIRE modele danych powinny być implementowane w całości; nie należy pominąć żadnego typu obiektów przestrzennych. Jeśli istnieje potrzeba rozróżnienia pomiędzy "ważniejszymi" i "mniej ważnymi" rodzajami obiektów przestrzennych, obie grupy należy ująć w oddzielnych modelach danych, zwanych także "profilami". Obiekty przestrzenne niezbędne do zaspokojenia kluczowych wymagań umieszczone są w modelu zasadniczym. Modele rozszerzone mogą opisywać wdrożenia dobrowolne oraz etapowe, spójne rozwijanie infrastruktury poprzez określanie celów dla dalszego gromadzenia i przechowywania danych.

Poza elementami technicznymi specyfikacje danych mogą również zawierać wyjaśnienia i przykłady pomocne w lepszym zrozumieniu i bardziej efektywnej implementacji.

#### Wdrożenie, testy, sprawdzenie poprawności

Specyfikacje muszą być poddawane przeglądowi i testom szerszej grupy zainteresowanych w celu stwierdzenia, czy specyfikacje danych są zgodne z celami infrastruktury i czy zawierają wystarczającą ilość informacji dla wsparcia wdrożenia.

Testy specyfikacji można przeprowadzić w celu uzyskania informacji zwrotnych na temat wykonalności lub podatności do użytku. Testy wykonalności oceniają działania podejmowane przez dostawców danych w celu przekształcenia ich danych na format zgodny ze specyfikacją docelowego poziomu interoperacyjności. W ten sposób uzyskuje się informacje zwrotne dotyczące wykonalności technicznej oraz kosztów związanych z wdrożeniem.

Proces testowania zastosowań służy do oceny tego, w jakim stopniu interoperacyjność usprawnia pracę użytkowników. Taki test wykonywany jest przez użytkowników danych w celu ustalenia, czy dane przekazane w sposób zgodny ze specyfikacjami docelowego poziomu interoperacyjności usprawniają efektywność ich pracy. Wyniki testów i konsultacji wśród zainteresowanych można wykorzystać przy powtarzaniu procesu specyfikowania danych od dowolnego etapu, najprawdopodobniej od etapu analizy stanu obecnego i analizy luk. Powtórzenia można prowadzić do momentu osiągnięcia konsensusu. Po zakończeniu tego procesu walidacji specyfikacje są publikowane, dzięki czemu mogą być wykorzystywane przez ogół społeczeństwa.

W INSPIRE przeprowadzono trzy powtórzenia. Po pierwszym powtórzeniu specyfikacje danych poddane zostały przeglądowi przez Tematyczne Grupy Robocze. Podstawowym celem tej fazy jest wyeliminowanie niespójności pomiędzy specyfikacjami w obrębie poszczególnych tematów danych. Drugie powtórzenie obejmowało przegląd oraz fazę testów, w której mogły uczestniczyć wszystkie społeczności zainteresowanych. Dla przyspieszenia procesu zwoływano spotkania konsultacyjne - "warsztaty rozpatrywania uwag" w celu rozstrzygnięcia pomiędzy rozbieżnymi opiniami zainteresowanych. Na podstawie wyników tych działań specyfikacje poddawano kolejnej rewizji i publikowano w formie wytycznych wdrożeniowych w trzecim powtórzeniu. Wybrane części wytycznych zostały uwzględnione w aktach prawnych nakazujących wdrożenie INSPIRE przez państwa członkowskie Unii Europejskiej.

Rozporządzenie Komisji nr 1089/2010 zakresie interoperacyjności zbiorów i usług danych przestrzennych zawiera podzbiór Podstawowego Modelu Koncepcyjnego INSPIRE oraz specyfikacji danych. Dla każdego tematu obowiązuje jedna specyfikacja danych, jednak wymagane prawnie części zebrane są w jedną regulację "wdrażającą".

W przypadku SDI wymaganych prawnie konieczne jest uwzględnienie dodatkowego kroku. Projekty techniczne powinny być przekształcone na akty prawne spełniające wymagania legislacyjne przy zachowaniu treści technicznej. Jednym ze sposobów zapewnienia wykonalności prawnej jest ujęcie jako obowiązujących tylko parametrów usług, za pośrednictwem których dane są udostępniane w infrastrukturze, pozostawiając modele semantyczne w sferze wytycznych. Inną opcją jest wybór podzbioru specyfikacji danych, obejmującego model semantyczny, w oparciu o kwestie wykonalności technicznej oraz stosunku korzyści do kosztów. W tym przypadku specyfikacje danych z pełną zawartością techniczną służą jako wytyczne dla

zainteresowanych i umożliwiając dalsze spójne rozwijanie infrastruktury.

#### 5.4 Utrzymanie specyfikacji

Zmiany wymagań lub stanu obecnego<sup>37</sup> mogą prowadzić do weryfikacji specyfikacji danych oraz powiązanych z nimi rejestrów, dokumentów i narzędzi niezbędnych do wsparcia działań technicznych oraz dokumentacji. Wniosek o zmianę specyfikacji danych może wystąpić pod wpływem następujących czynników:

- Problemy stwierdzone na dalszym etapie procesu etapowego specyfikowania danych oraz w fazie wdrażania,

<sup>36</sup> Zob. rozdział 4.1.8

<sup>37</sup> Zob. rozdział 5.3.

- Zmiany ram prawnych mające wpływ na wymagania dotyczące danych przestrzennych,
- Nowe inicjatywy i programy wpływające na rozwijanie SDI (pojawiające się inicjatywy SDI na wyższym szczeblu, eAdministracja itp.),
- Potrzeba harmonizacji z normami międzynarodowymi i innymi inicjatywami,
- Nowe odnośne wymagania użytkowników i analizy przypadków użytkowych,
- Zmiany obecnej sytuacji zainteresowanych i postęp technologiczny,
- Błędy lub niejednoznaczności w dokumentach,
- Nieśpójność z innymi elementami składowymi infrastruktury,
- Względny stosunek kosztów do korzyści.

Z organizacyjnego punktu widzenia procedura utrzymywania powinna być równie otwarta i oparta na zasadach uczestnictwa, jak proces rozwijania specyfikacji, co zagwarantuje spójność w obrębie wdrażania, rozwoju i utrzymania. I tak, należy zdefiniować osoby i organizacje, które będą zaangażowane w proces, a także metody i obiekty zadań.

Proces utrzymywania realizuje zasadniczo trzy sposoby dokonywania zmian. Metoda "naprawiania i dostosowywania" służy do poprawiania błędów i wprowadzania lub przywracania zgodności z pozostałymi komponentami lub składowymi infrastrukturą. Metoda "amortyzowania" służy do odrzucania elementów<sup>38</sup>, które nie są już wykorzystywane lub które zostały zastąpione nowymi elementami, natomiast metoda "dodawania" pozwala na wprowadzanie nowych elementów.

Niewielkie korekty umożliwiają wsteczną weryfikację zgodności, tj. wszystkie zestawy danych zgodne z poprzednią wersją zachowują zgodność z wersją zmienioną. Natomiast duże rewizje wprowadzają istotne zmiany. Jeżeli jest to wykonalne i stosowne, większe rewizje powinny zachowywać kompatybilność wsteczną. Ten typ rewizji jest dopuszczalny wówczas, gdy jest bezwzględnie konieczny dla danej domeny, np. w celu wprowadzenia do tematu znaczącej liczby dodatkowych typów obiektów przestrzennych lub fundamentalnego zaktualizowania Podstawowego Modelu Koncepcyjnego czy specyfikacji danych.

W ramach wsparcia procesu utrzymywania zaleca się korzystanie z systemów kontroli wersji repozytoriów, zarówno w odniesieniu do skonsolidowanego modelu danych, jak i dokumentacji technicznej.

### 5.5 Względy dotyczące stosunku kosztów do korzyści

Poza tym, że rozwiązania z zakresu interoperacyjności są oparte o względy wykonalności technicznej, powinny one ponadto opierać się o dokładną analizę odnośnych kosztów i korzyści, jak wykazano na Rys. 10 na str. 43. Analiza stosunku kosztów do korzyści w procesie rozwijania specyfikacji danych powinna być prowadzona przez cały czas trwania procesu specyfikacji.

W analizie kosztów i korzyści przewidywane koszty i korzyści przekształca się na porównywalne jednostki (zwykle są to wartości pieniężne). Przeprowadzenie ścisłej analizy kosztów i korzyści jest dość trudne w przypadku SDI, szczególnie pod względem korzyści. Korzyści odnoszą zazwyczaj użytkownicy oraz społeczeństwo jako takie. Ponadto korzyści z SDI przed uwidocznieniem potrzebują często czasu na osiągnięcie dojrzałości, tj. okresu przejściowego, w ciągu którego następuje przekształcenie masy krytycznej zbiorów danych aż do osiągnięcia interoperacyjności.

Względny stosunek kosztów do korzyści ujmują w sposób ogólny kryteria oceny ilościowej i jakościowej dla SDI. Zamiast podejmowania prób przeliczenia każdego aspektu kosztów i korzyści na porównywalne jednostki (pieniężne), zawierają one deklaracje dotyczące następujących obszarów:

- Gdzie i w jakiej formie najprawdopodobniej wystąpią koszty i korzyści,
- W jaki sposób unikać lub ograniczać koszty poprzez podejmowanie właściwych decyzji i działań technicznych,
- W jaki sposób uwidocznić potencjalne korzyści i zapewnić ich widoczność dla zainteresowanych.

Podstawowy sposób wykrywania potencjalnych kosztów związanych z wdrażaniem specyfikacji interoperacyjności to proces testowania, w ramach którego dostarczyciele danych mogą rejestrować nakłady niezbędne do osiągnięcia interoperacyjności pod kątem specjalistycznej wiedzy, czasu, nowego oprogramowania i sprzętu oraz potrzeb szkoleniowych. W INSPIRE tego typu testy zwane są "testami transformacji".

Inny typ testów, czyli testy praktyczne, pomagają w dokonaniu ujęcia ilościowego korzyści dla użytkowników poprzez zestawienie czasu potrzebnego na wykonanie określonego zadania z wykorzystaniem danych zgodnych ze specyfikacjami interoperacyjności oraz danych dostarczonych w formie pierwotnej. Jeżeli dane zgodne ze specyfikacjami interoperacyjności usprawnią wykonywanie zadań przez użytkowników, korzyści z infrastruktury staną się widoczne. Korzyści można ujmować ilościowo według skrócenia czasu realizacji zadań, konieczności zaangażowania mniejszej liczby wykwalifikowanego personelu do realizacji zadań itp.

W celu uzyskania szerszej perspektywy oglądu kosztów i korzyści z infrastruktury, dla INSPIRE przeprowadzono poszerzoną ocenę wpływu oraz bezpośrednie badanie wśród zainteresowanych. W Tabeli 5 zestawiono główne aspekty związane z uwarunkowaniami kosztów i korzyści SDI.

KOSZTY	Bezpośrednia wartość/korzyść użytkownika	KORZYŚCI dla Korzyści operacyjne dla instytucji
<ul style="list-style-type: none"> <li>- Koszty związane z rozwijaniem specyfikacji</li> <li>- Koszty reinyżynierii baz danych</li> <li>- Alternatywne koszty rozwoju mapowania schematów ze starych do nowych specyfikacji</li> <li>- Koszty sprzętu i oprogramowania, gdyby wymagane były nowe systemy</li> <li>- Koszty prowadzenia/sprawdzenia/weryfikowania poprawności przekształcenia</li> </ul>	<ul style="list-style-type: none"> <li>- Większa dostępność danych</li> <li>- Większa łatwość korzystania</li> <li>- Szersze możliwości udostępniania danych</li> <li>- Mniejszy koszt integracji danych</li> </ul> <p><b>Wartość społeczna</b></p> <ul style="list-style-type: none"> <li>- Umożliwia poprawę poziomu podejmowania decyzji</li> <li>- Ogranicza bariery pomiędzy organizacjami</li> <li>- Zwiększa efektywność instytucji</li> <li>- Propaguje bardziej efektywne wykorzystywanie środków (z podatków)</li> </ul>	<ul style="list-style-type: none"> <li>- Promowanie współpracy wewnątrz instytucji</li> <li>- Promowanie współpracy pomiędzy instytucjami</li> <li>- Ograniczenie kosztów integracji danych pomiędzy instytucjami</li> <li>- Promowanie wielokrotnego wykorzystywania istniejących zbiorów danych</li> <li>- Zmniejszenie kosztów informatyki/zarządzania informacjami</li> <li>- Ogólne oszczędności kosztów w zakresie zarządzania informacjami</li> <li>- Unikanie kosztów (w przeciwieństwie do oszczędności)</li> <li>- Działanie na rzecz ściślejszych relacji roboczych</li> <li>- Wsparcie ulepszonych procesów decyzyjnych</li> <li>- Wsparcie innej infrastruktury informacyjnej</li> </ul>

Tabela 5: Aspekty analizy kosztów i korzyści SDI

### 5.6 Podmioty uczestniczące w procesie specyfikowania danych

Struktura organizacyjna ustanawiania komponentu danych SDI zdefiniowana jest przez następujące warunki:

1. Proces powinien być oparty o budowanie konsensusu,
2. Ustanowienie i prowadzenie SDI w dążeniu do interoperacyjności pomiędzy zakresami tematycznymi wymaga zaangażowania wielu organizacji,
3. Interoperacyjność pomiędzy tematami wymaga narzędzi i środków organizacyjnych zapewniających ciągły przepływ informacji pomiędzy zainteresowanymi podmiotami.

Z warunków tych wynika potrzeba koordynacji dla zapewnienia komunikacji, planowania, dostarczania i utrzymywania narzędzi podczas procesu specyfikacji.

Im więcej tematów danych jest ujętych w infrastrukturze, tym większe jest zapotrzebowanie na należyte usystematyzowany proces. Podejście modułowe zapewnia większą swobodę z organizacyjnego punktu widzenia. Mogą wystąpić trudności w zaangażowaniu niezbędnych zasobów do rozwijania specyfikacji interoperacyjności dla wielu tematów danych równoległe. Gdy moduły są rozplanowane w odpowiednim porządku, wiedzę zgromadzoną na początku można wykorzystywać na dalszych etapach. Warto rozpocząć proces od danych referencyjnych, gdzie uczestnicy są "uświadomieni przestrzennie".

Znaczące rozmowy ze społecznościami zainteresowanych mogą być prowadzone wyłącznie w oparciu o dobre propozycje. Techniczne projekty specyfikacji docelowego poziomu interoperacyjności muszą być proponowane przez właściwy organ. Zgodnie z zasadą uczestnictwa rządząca SDI, za najlepsze formy organizacji uznaje się grupy ekspertów technicznych, w skład których wchodzi przedstawiciele zainteresowanych. Zakres specjalistycznej wiedzy tych grup obejmuje:

- Wiedzę specjalistyczną w zakresie modelowania informacji geograficznej oraz odnośnych unormowań,
- Wiedzę tematyczną (w określonych domenach) (znajomość danych przewidzianych przewidzianych do wykorzystania w analizach reprezentatywnych przypadków użytkowych),

<sup>38</sup> Przewzgląd na potrzebę identyfikowalności nie powinno się usuwać żadnego elementu.



- Wiedzę specjalistyczną na temat SDI: wiedzę na temat polityk leżących u podstaw infrastruktury oraz standardowej architektury SDI,
- Wiedzę specjalistyczną w obszarze usług sieciowych (wiedza na temat dostępu danych),
- Wiedzę specjalistyczną z zakresu oprogramowania: wiedzę na temat implementacji i wdrażania odnośnych specyfikacji.

Organ koordynujący w INSPIRE zwany jest "Zespołem ds. Konsolidacji", w skład którego wchodzi pracownicy Komisji Europejskiej. W odniesieniu do komponentu danych rozróżnia się dwa rodzaje grup eksperckich: Zespół ds. Sporządzania Specyfikacji Danych, który odpowiada za rozwijanie i utrzymywanie ram koncepcyjnych, oraz Tematyczne Grupy Robocze, które odpowiadają za rozwijanie specyfikacji docelowego poziomu interoperacyjności dla poszczególnych tematów danych. Członkowie tych grup eksperckich są delegowani przez społeczność zainteresowanych. Zainteresowani uczestniczą także w przeglądach i testach. Prawnie obowiązująca część specyfikacji jest przyjmowana przez Komitet INSPIRE, w skład którego wchodzi urzędowny przedstawiciel krajów członkowskich Unii Europejskiej.

Ze względu na zapewnienie efektywnej organizacji pracy, w grupach eksperckich przewidziane są konkretne role. **Lider grupy** planuje pracę, rozdziela zadania pomiędzy członków oraz pośredniczy w rozmowach z ekspertami w obrębie grupy, a także z partnerami zewnętrznymi. Na etapie opracowywania ram koncepcyjnych grupa powinna dobrze orientować się z zmianach rozwojowych w SDI oraz wykazywać dobre przygotowanie w dziedzinie modelowania i standaryzowania informacji. W fazie specyfikowania danych szczególny nacisk kładzie się na specjalistyczną wiedzę z określonych dziedzin oraz na znajomość ram koncepcyjnych.

Wyniki prac specyfikacyjnych dokumentowane są przez **redaktora** zgodnie ze zdefiniowanymi wcześniej szablonami. Redaktor musi sprawnie tworzyć dokumenty techniczne, odpowiada za sporządzenie dokumentacji opisowej oraz opanowuje wybrany język schematu koncepcyjnego do prezentacji modeli danych w formacie czytelnym dla maszyn.

#### 5.7 Narzędzia wsparcia

W fazie rozwijania specyfikacji danych udział bierze wielu różnych zainteresowanych. Wynik ich pracy musi być porównywalny. Każda specyfikacja danych powinna mieć tę samą strukturę w dokumentacji, co usprawnia komunikację pomiędzy grupami ekspertów oraz akceptację przez społeczność użytkowników. Grupy ekspertów odpowiedzialne za

tworzenie dokumentacji technicznej powinna odnajdywać pomoc w narzędziach i szablonach kierujących ich pracą, utrzymujących spójność rezultatów i pomocnych przy współdzieleniu wiedzy od samego początku procesu.

Narzędzia można zaklasyfikować do następujących grup: wspólne szablony dokumentów, repozytoria dokumentów, internetowe fora dyskusyjne i rejestry. Wspólne szablony dokumentów przyczyniają się do harmonizacji dokumentacji i zapewniają analogiczne uwzględnienie niezbędnych aspektów. Najbardziej wyrazistym przykładem szablonów w INSPIRE jest szablon specyfikacji danych oparty o ISO 19131. W celu usprawnienia pracy można stosować także inne szablony i listy kontrolne (np. do opisu analizy przypadków użytkowych oraz do analizy materiałów referencyjnych).

Repozytoria dokumentów są pomocne w udostępnianiu materiałów referencyjnych i projektów roboczych, głównie wśród członków grup ekspertów. Zapewnienie widoczności projektów we wszystkich grupach pomaga zapewnić spójność pomiędzy tematami danych. Systemy kontroli wersji repozytoriów dokumentów umożliwiają powrót do poprzedniej propozycji w dowolnym momencie. Dodatkowo prowadzenie rejestrów zmian umożliwia śledzenie i zapewnienie przejrzystości procesu.

## 6 Wnioski

Bogactwo cyfrowych danych przestrzennych zgromadzonych przez ostatnie 30 – 40 lat oraz postępy technologii informacyjnych i komunikacyjnych otwierają nowe perspektywy dla analizy naszego otoczenia fizycznego i społecznego. W analizie przestrzennej, wsparciu procesu decyzyjnego oraz usługach opartych o położenie geograficzne często wykorzystuje się ponownie dane stworzone pierwotnie do innych celów, co pozwala uzyskać znaczące oszczędności w zakresie rozwijania systemu.

Integrację danych przestrzennych z różnych źródeł utrudniają często ograniczenia w udostępnianiu danych oraz brak interoperacyjności. Infrastruktury danych przestrzennych stanowią sposób na zaradzenie tym trudnościom, ponieważ zapewniają usługi on-line do wykrywania, oceny, pozyskiwania i przekształcania danych. Jedną z przyczyn ograniczonej interoperacyjności jest niespójność i niekompatybilność. W większości przypadków konieczne jest przekształcenie danych dla uzyskania wspólnych parametrów, a tym samym osiągnięcia interoperacyjności.

Bez SDI przekształcenia takie dokonywane są przez użytkowników doraźnie. W SDI interoperacyjność jest zapewniona u źródła; dostawcy danych powinni dostarczać dane zgodnie ze zdefiniowanymi i uzgodnionymi normami. Techniczną formą takich norm są specyfikacje docelowego poziomu interoperacyjności, zwane często specyfikacjami danych.

Luki w interoperacyjności w kontekście danych przestrzennych można wypełnić na dwa sposoby: korzystając z rozwiązań z zakresu interoperacyjności, do których należą rozwiązania technologiczne i organizacyjne, oraz poprzez harmonizację danych. W SDI preferowane jest pierwsze rozwiązanie, ponieważ w tym przypadku dostawcy danych nie muszą zmieniać pierwotnej struktury swoich danych. Mogą wdrażać technologię (np. przekształcać masowych lub bieżących przekształceń danych) w celu spełnienia wymogów dotyczących interoperacyjności. Jednak bieżąca technologia nie zawsze wypełnia całkowicie luki w interoperacyjności. Harmonizacja danych zapewnia ściślejsze wzajemne dostosowanie struktur danych różnych dostawców. Doświadczenie pokazuje, że połączenie tych dwóch metodologii stanowi rozwiązanie najkorzystniejsze.

SDI jest zbiorem różnych tematów danych. Dla każdego z nich należy zdefiniować docelowy poziom interoperacyjności w formie specyfikacji interoperacyjności (lub danych). Do uzyskania interoperacyjności pomiędzy poszczególnymi tematami wymagane są solidne ramy wspierające powszechne środki techniczne, efektywną wymianę informacji oraz ustandaryzowaną metodologię rozwijania specyfikacji danych w obrębie całej infrastruktury. Są to tzw. ramy koncepcyjne. W oparciu o doświadczenia INSPIRE, ramy te dzielą się na dwa komponenty: podstawowy model koncepcyjny oraz metodologię rozwijania specyfikacji.

Podstawowy model koncepcyjny (GCM) przekształca rozwiązanie z zakresu interoperacyjności i harmonizacji danych na zestaw elementów interoperacyjności poprzez ich dopasowanie do odnośnych elementów modelowania informacji i technologii geoprzestrzennych. GCM zawiera wspólne koncepcje oraz stanowi główne narzędzie wprowadzania interoperacyjności w obrębie wszystkich tematów danych zawartych w infrastrukturze.

Metodologia GCM była rygorystycznie wdrażana w INSPIRE, a szczególny nacisk kładziono na bieżące udostępnianie rezultatów prac technicznych. Publicznie dostępne rejestry oraz wykorzystanie skonsolidowanego repozytorium modelu stanowi o innowacyjności podejścia do ustalania komponentu danych w SDI. "Przewiduje się, że ten model koncepcyjny będzie w przyszłości wpływał w wielu przypadkach na działania w zakresie modelowania danych przestrzennych na szczeblu krajowym, ponieważ zapewnia wartość dodaną do krajowej infrastruktury danych przestrzennych i upraszcza konwersję na specyfikacje danych INSPIRE" (Portele C. (red.), 2010a). Konwergencja technologiczna dostawców danych jest kluczowym elementem inicjatyw SDI.

W ramach koncepcyjnych metodologii rozwijania specyfikacji zapewnia przestrzeganie ogólnych zasad infrastruktury, takich jak ponowne wykorzystanie, wykonalność i proporcjonalność. Wbudowane w proces zabezpieczenia zapewniają podejmowanie wszystkich niezbędnych kroków i działań w obrębie każdego z tematów ujętych w infrastrukturze. Metodologia ma zapewnić przewidywalny i powtarzalny proces rozwoju, w wyniku którego uzyskuje się wykonalne i wzajemnie satysfakcjonujące rozwiązania. Metodologia powinna ponadto określać role odgrywane przez zainteresowane podmioty na różnych etapach procesu.

Ramy legislacyjne INSPIRE stanowią ważki precedens dla rozszerzanej przyrostowo infrastruktury SDI w oparciu o zaangażowanie zainteresowanych. To doświadczenie wskazuje, że metodologia taka może przynieść materialne rezultaty również w przypadku, gdy zakres przedmiotowy SDI jest szeroki, zaangażowane są setki uczestników z ponad 30 krajów,<sup>39</sup> a czynności techniczne muszą być przeprowadzone w stosunkowo krótkim czasie<sup>40</sup>. Dlatego też metodologia specyfikacji danych zaproponowana w INSPIRE została przyjęta w Infrastrukturze Danych Przestrzennych ONZ (Atkinson, R. i Box, P., 2008).

Szczególna wartość ram koncepcyjnych opisanych w niniejszym raporcie polega na tym, że stanowią one zbiór najlepszych praktyk z realizowanych na bieżąco inicjatyw. Zarówno metodologię rozwijania specyfikacji, jak i podstawowy model koncepcyjny przetestowano w warunkach rzeczywistych w ramach rozwijania specyfikacji danych. Wprawdzie w procesie rozwijania opracowano 9 gotowych i 25 projektowanych specyfikacji interoperacyjności, lecz należy zauważyć, że proces wdrażania nadal trwa, a korzyści dla użytkowników mogą być prawidłowo oszacowane dopiero w przyszłości.

Specyfikacje danych, które zostały poddane dokładnym przeglądom, testom i akceptacji przez społeczności zainteresowanych, dowodzą prawidłowości przyjętego podejścia, krystalizując zbiorową wiedzę z Europy i spoza Europy. Coraz szersze kręgi uczestnictwa w procesie, postępy w egzekwowaniu w przepisach prawnych, a także obszerne informacje zwrotne otrzymane w wyniku procesu testowania i wdrażania oznaczają, że podobne ramy koncepcyjne

<sup>39</sup>

Poza państwami członkowskimi Unii Europejskiej do procesu przystąpiły także zainteresowane strony z Europejskiego Obszaru Gospodarczego, Szwajcarii, Stanów Zjednoczonych oraz krajów ubiegających się o przyjęcie do UE.

<sup>40</sup>

Prace techniczne nad komponentem danych INSPIRE rozpoczęły się w 2005 roku, a ich zakończenie przewiduje się w kwietniu 2012 r.

mogą służyć jako czynnik sukcesu również w innych inicjatywach.

#### **Podziękowania**

Autorzy pragną podziękować wielu ekspertom i zainteresowanym podmiotom, którzy brali udział w opracowywaniu wytycznych technicznych i regul wdrażania Dyrektywy INSPIRE. W szczególności doceniamy pracę Zespołu ds. Opracowywania Specyfikacji Danych INSPIRE. Systematycznie i pracowicie układał on i dokumentował ramy koncepcyjne INSPIRE. Dziękujemy także podmiotom zainteresowanym INSPIRE, które poprzez swoje uwagi, prowadzone przez siebie testy i kwerendy wносиły nieustanny wkład w doskonalenie prac ekspertów.

Równie serdecznie dziękujemy naszym recenzentom z JRC, którymi byli: Max Craglia, Michel Millot i Katalin Bódis; pomogli oni autorom nie zgubić się w szczegółach technicznych. Ich uwagi były bardzo pomocne przy rozstrzygnięciu założeń i niejasności, a także w wypełnianiu luk i informacjami, przy pomocy których ten trudny techniczny temat staraliśmy się uczynić łatwiej przyswajalnym. Osoby te znakomicie sprawdziły się w roli "świadomych twórców polityki"!

Nasi recenzenci zewnętrzni, czyli Siri Jodha Khalsa, Zdzisław Kurczyński, Stefano Nativi, i Daniele Rizzi, pomogli nam wyabstrahować od posiadanych danych, przyjąć perspektywę INSPIRE i umieścić raport w takiej szerszej perspektywie. Osoby te odegrały decydującą rolę w nadaniu ostatecznego kształtu niniejszemu raportowi, który jest dzięki temu lepiej usystematyzowany i ujednolicony.



Misja wewnętrznego organu badawczego Komisji Europejskiej, czyli Wspólnotowego Centrum Badawczego, polega na zapewnieniu niezależnego, opartego na dowodach wsparcia naukowego i technicznego dla polityk UE przez cały cykl realizacji polityki.

JRC w ścisłej współpracy z Dyrektoratami Generalnymi odpowiedzialnymi za politykę odpowiada na najważniejsze wyzwania społeczne, a jednocześnie stymuluje innowacyjność poprzez opracowywanie nowych standardów, metod i narzędzi, a także dzielenie się swoją wiedzą i przekazywanie jej państwom członkowskim i społeczności międzynarodowej.

Do kluczowych obszarów polityki zalicza się: środowisko i zmiany klimatu; energetykę i transport; rolnictwo i bezpieczeństwo żywności; ochronę zdrowia i ochronę konsumentów; społeczeństwo informacyjne i cyfryzację; bezpieczeństwo i ochronę, w tym także nuklearną, przy czym wszystkie te obszary poddawane są analizie z zastosowaniem przekrojowej, wielodyscyplinarnej metodologii.