

## Zastosowanie zbiorów danych RDF do modelowania czasoprzestrzennego w geoinformatyce\*

The application of RDF datasets  
for spatiotemporal modelling in geoinformatics

Jaromar Łukowicz, Adam Iwaniak

Uniwersytet Przyrodniczy we Wrocławiu, Wydział Inżynierii Środowiska i Geodezji,  
Instytut Geodezji i Geoinformatyki

**Słowa kluczowe:** ontologie, zbiory danych RDF, grafy nazwane, ontologie formalne, model  
czasoprzestrzenny

Keywords: ontologies, RDF datasets, named graphs, formal ontologies, spatiotemporal model

### Wstęp

Jednym z wyzwań gospodarki przestrzennej, w tym planowania przestrzennego, jest poprawny opis zjawisk zachodzących w przestrzeni w celu prognozowania jej przyszłego stanu. Istotne jest widzenie środowiska i zagospodarowania przestrzennego jako układu dynamicznego. Jednocześnie opis ładu przestrzennego ma charakter interdyscyplinarny, wymagający łączenia wielu modeli dziedzinowych. Artykuł prezentuje rozwiązania służące gromadzeniu i porządkowaniu danych, opisujących stan przestrzeni w jej dynamicznej postaci. Wykorzystują one technologie sieci semantycznych (ang. *Semantic Web*), w tym zbiory danych RDF, służące udostępnianiu zasobów przestrzennych, jako danych powiązanych (ang. *Linked Data*) oraz ontologie, zarówno w postaci taksonomicznej jak i aksjomatycznej, oparte na logice opisowej (ang. *Description Logic*).

Odniesienie przestrzenne i czasowe obiektów występujących w rzeczywistej przestrzeni może być prezentowane jako cechy własne tych obiektów lub w formie kontekstu czasoprzestrzennego, w jakim informacja o tych obiektach została pozyskana. W pierwszym przypadku jest to geometria lub „historia życia” tych obiektów, w drugim jest to informacja o momencie pobrania danych, okresie ich aktualności oraz przestrzeni, do której się te dane odnoszą. Pierwsze podejście odnosi się więc do modelowania danych, a drugie do modelowania metadanych.

Autorzy prezentują strukturę zbiorów danych RDF (ang. *RDF Datasets*), która pozwala połączyć oba modele w spójny system (Kuhn i in., 2014). Dotąd w systemach infrastruktury

---

\* Opisane w artykule badania są sfinansowane przez Narodowe Centrum Nauki w ramach projektów UMO-2012/05/N/HS4/00642.

ry informacji przestrzennej dane i metadane były wyodrębnione, a często rozprowadzane odrębnymi kanałami (*vide* odrębne pliki metadanych, odrębna usługa CSW). Wspólny system, który łączy dane i metadane daje możliwość wykorzystania tych samych narzędzi do eksploracji metadanych (w celu wyszukiwania zasobów na ich podstawie), a następnie do eksploracji i przetwarzania samych zasobów. Ponadto pozwala uzupełniać brakujące informacje o odniesieniach czasowych i przestrzennych w opisie obiektów rzeczywistych, dzięki informacji zawartej w metadanych. Metadane zawierające kontekst czasowy i przestrzenny, w postaci tzw. „grafów nazwanych” tworzących strukturę RDF Datasets, mogą być użyte do interpretacji niekompletnych zasobów danych przestrzennych, tak aby uzyskać całościowy opis ich dynamicznego, czasoprzestrzennego charakteru. Może to być konieczne, kiedy integracja dotyczy heterogenicznych zasobów sieci WWW, do których stosuje się „założenie świata otwartego”, o różnej wiarygodności, kompletności i aktualności, a często wzajemnie sprzecznych (Kuhn i in., 2014).

W poprzednich artykułach (Iwaniak i in., 2013; Łukowicz i in., 2014) prezentowane były metody modelowania danych czasoprzestrzennych z użyciem ontologii formalnych do budowy baz wiedzy, w celu wnioskowania na temat tempa zmian stanu obiektów, odkrywania związków przyczynowo-skutkowych oraz skutków tych zmian dla klasyfikacji obiektów. W tym artykule koncentrujemy się na formie, w jakiej powinny być porządkowane semantyczne dane przestrzenne, na wejściu do tworzonych baz wiedzy.

## Obecny stan wiedzy

Technologie sieci semantycznych (Berners-Lee i in., 2001), obejmują cały kompleks rozwiązań i narzędzi, takich jak: reprezentacja danych w postaci grafów RDF (<http://www.w3.org/TR/rdf11-primer/>) i dane powiązane (ang. *Linked Data*) (Bizer i in., 2009), ontologie (Gruber, 1995) między innymi oparte na logice opisowej (*Description Logic*), a wśród nich ontologie formalne (Grenon, 2003; Husserl, 1901), zwane też ontologiami fundamentalnymi lub ontologiami wyższego rzędu (Herre, 2010). Technologie te dostarczają atrakcyjnych rozwiązań dla modelowania zagadnień gospodarki przestrzennej (Harvey i in., 2014; Iwaniak i in., 2011; Iwaniak i in., 2014). Umożliwiają one tworzenie kompletnego i spójnego obrazu rzeczywistości, uwzględniającego jej dynamikę oraz otwierającego możliwości analiz przekształceń przestrzeni z uchwyceniem związków przyczynowo-skutkowych, niezbędnych do odkrywania potencjalnych scenariuszy oraz tworzenia prognoz przyszłego stanu przestrzeni (Iwaniak i in., 2013; Łukowicz i in., 2014).

W zestawie technologii *Semantic Web* znajdują się rozwiązania służące reprezentacji aspektów przestrzennych odwzorowywanej rzeczywistości. Zalicza się do nich projekt GeoSPARQL (<http://www.opengeospatial.org/standards/geosparql>) będący rozszerzeniem języka SPARQL (<http://www.w3.org/TR/sparql11-query/>), oferujący również ontologię, słownik zawierający pojęcia i predykaty służące do opisu geometrii oraz wzajemnych relacji przestrzennych. W ramach *Semantic Web* zaproponowano również wiele rozwiązań służących reprezentacji zagadnień temporalnych, takich jak: temporalne grafy RDF (Gutierrez i in., 2007), „adnotowany” RDF, czyli aRDF (Udrea i in., 2008), dedykowane ontologie (Bittner i in., 2009) i innych (Li i in., 2008).

Odniesienia przestrzenne oraz czasowe mogą być realizowane zarówno na poziomie opisu cech odwzorowywanych obiektów rzeczywistych, jak również w odniesieniu do danych,

które je reprezentują. W pierwszym przypadku wykorzystujemy podstawowe konstrukcje składni RDF, stosując predykaty oparte na istniejących słownikach, takich jak: Dublin Core Dublin Core (<http://dublincore.org/>) (Ryan i in., 1998), schema.org (<http://schema.org/>) lub GeoSPARQ albo, jeżeli okażą się nie dość ekspresywne, możemy stworzyć własne predykaty, jak również własne typy danych (<http://www.w3.org/TR/swbp-xsch-datatypes/#sec-user-Defined>). W drugim przypadku, *de facto*, używamy metadanych. Odnosi się to do danych pozyskiwanych na przykład z monitoringu środowiska, choćby przez próbkowanie stanu określonych elementów przyrodniczych z rozmieszczonych w przestrzeni stacji pomiarowych (Cox, 2014). Mamy tu możliwość wykorzystania wyżej przytoczonego rozwiązania aRDF. Możemy również przekształcić asercje w grafie RDF, do postaci węzłów anonimowych (ang. *Blank Nodes*) typu `rdf:Statement` ([http://www.w3.org/TR/rdf-schema/#ch\\_reification](http://www.w3.org/TR/rdf-schema/#ch_reification)) i opisywać je jak każde inne zasoby RDF – jest to tzw. reifikacja wyrażenia (Berners-Lee, 2004).

Najbardziej obiecującą w zakresie adnotowania trójek RDF jest struktura zbioru danych RDF (ang. *RDF Dataset*; <http://www.w3.org/TR/rdf11-datasets/>; <http://www.w3.org/TR/rdf11-concepts/#section-dataset>) (Carroll i in., 2005). RDF Dataset to multigraf zbudowany z jednego grafu domyślnego (ang. *Default Graph* – może być grafem pustym, ale musi istnieć) oraz dowolnej liczby grafów nazwanych (ang. *Named Graphs*). Grafy nazwane posiadają identyfikatory w postaci URI, czyli wyglądają jak każdy zasób RDF.

Grafy nazwane są więc podzbiorami trójek wyodrębnionymi w ramach danego zbioru RDF (RDF Dataset) przez kontekst, w którym dane asercje są prawdziwe. Tego typu rozwiązanie zostało wprowadzone w wersji 1.1 specyfikacji RDF. Trójka RDF (ang. *triple*) w postaci  $\langle S, P, O \rangle$ , gdzie S to podmiot (ang. *Subject*), P – predykat, O – przedmiot (ang. *Object*) zostaje poszerzona o URI grafu nazwanego G, przyjmując postać  $\langle S, P, O, G \rangle$ . Ta struktura nosi miano czwórki (*quad* od ang. *quadruple*).

W założeniu autorów specyfikacji RDF 1.1, grafy nazwane służą partycjonowaniu (syndykacji) wielkich zbiorów danych, umożliwiając operowanie na nich, jako na pewnych całościach oraz umożliwiając zdefiniowanie meta-informacji o całych grafach i zawartych w nich asercjach. Graf nazwany może więc być użyty do określenia czasu, w którym badana trójka występuje i jest prawdziwa (ang. *temporal context*), do określenia pochodzenia informacji (ang. *pedigree, provenance, lineage*), z możliwością określenia łańcucha pochodzenia (następstwo czasowe, sekwencje). Kontekst może być użyty do określenia modalności informacji (możliwość, konieczność), prawdopodobieństwa (wyrażonej przez ułamek) lub wiarygodności. Grafy nazwane pozwalają na zarządzanie publikowanymi zbiorami danych przez określenie ograniczeń dostępu, ograniczeń użytkowania, własności intelektualnej itd. (Carroll i in., 2005). Dla nowych struktur RDF 1.1, W3C wprowadziło wiele nowych form serializacji uwzględniających koncepcję grafu nazwanego w formie grafu lub quadu (TriG – <http://www.w3.org/TR/trig/>; TriX – <http://www.w3.org/2004/03/trix/>; N-Quads – <http://www.w3.org/TR/n-quads/>; JSON-LD – <http://www.w3.org/TR/json-ld/>).

Powstają kolejne techniki porządkowania dużych zbiorów RDF, na przykład tagowanie trójek, jako tzw. zbiorów trójek (ang. *Triplesets*) (Mera Caraballo i in., 2014; Kiryakow, Momtchev, 2009). Tworzą one razem z „czwórką” (*quad, quadruple*), „piątkę” (ang. *quintuple*) w postaci  $\langle S, P, O, G, \{TS_1, \dots, TS_n\} \rangle$ .

Odrębnym zagadnieniem są koncepcje i metody modelowania zagadnień temporalnych z wykorzystaniem ontologii formalnych (Grenon, Smith, 2004), opisane w (Iwaniak i in., 2013; Łukowicz i in., 2014). Do tego celu używane były ontologie dziedziczne zbudowane na

bazie ontologii formalnych (BFO). Podejście to prezentuje zmiany przez opis obiektów trwałych (materialnych i niematerialnych), zwanych endurantami lub kontynuantami (ang. *continuants*), prezentując ich „historię życia” oraz zmiany stanu, jakości i typu oraz przez opis zdarzeń i procesów zwanych perdurantami lub okurentami (ang. *occurents*), które te zmiany wywołują (Grenon, Smith, 2004). Obecnie brak eleganckiej polskiej terminologii, która dobrze odpowiadałaby charakterowi tych kategorii. Przytoczone nazewnictwo odzwierciedla jednak charakter tych konstrukcji. „Enduranty”, czyli obiekty, które „są [trwale] w czasie ich trwania” lub, *vide* „kontynuant”, „są ciągle w swoim charakterze w czasie” – wystarczy pobrać próbkę ich cech w dowolnym momencie ich trwania, aby mieć pojęcie o ich charakterze. „Perduranty” natomiast „przebiegają przez czas” lub „mają miejsce”, albo „dzieją się” w czasie (ang. *to occur*) – aby je poznać trzeba je analizować w całym okresie, kiedy zachodziły, tak jak w przypadku jakiegoś procesu lub też procedury.

Reguły do budowy ontologii dziedzinowych z zakresu informacji przestrzennej są opisane w normie ISO: 19150-1:2012 Geographic information – Ontology – Part 1: Framework oraz 19150-2 Geographic information – Ontology – Part 2: Rules for developing ontologies in the Web Ontology Language (OWL) (ISO/TC 211, 2009).

## **Struktury danych służących reprezentacji uwarunkowań czasoprzestrzennych obiektów (zjawisk) geograficznych**

Uwzględnienie czynnika czasu w opisie rzeczywistej przestrzeni można dokonać na kilka sposobów. Odniesienie czasowe lub przestrzenne może być dokonywane zarówno na poziomie cech opisujących zjawiska rzeczywiste, jak i na poziomie metadanych. Pierwsze będą reprezentowane w postaci danych czasowych lub geometrycznych bezpośrednio odnoszących się do odwzorowywanych obiektów lub procesów. Są to cechy immanentne tych obiektów, które opisują ich kondycję lub mogą być podstawą ich klasyfikacji. Drugie będą odzwierciedlały sposób korzystania lub zarządzania zasobami informacyjnymi opisującymi rzeczywistość i są cechami transcendentnymi względem opisywanych obiektów. Nie określają one kondycji tych obiektów, ale opisują kontekst (czasowy, przestrzenny), w którym ta kondycja była faktem.

### **Opis bezpośredniego odniesienia czasowego i przestrzennego obiektów rzeczywistych**

Opisywanie cech obiektów rzeczywistych w grafach RDF jest dokonywane przez konstrukcję trójki (triple). Podmiotem trójki jest opisywany obiekt (obiekt przestrzenny, działka, użytek, budynek, teren planistyczny), predykatem – właściwość, a przedmiotem trójki – wartość tej właściwości. Podmiot i przedmiot są węzłami grafu, predykaty – krawędziami. Struktura RDF, zgodnie z podstawową specyfikacją, jest bardzo liberalna – podmiotami, predykatami, jak i przedmiotami mogą być dowolne „nazwane” zasoby, identyfikowane „globalnie” przez URI. Dodatkowo podmiotem może być zasób „nienazwany” (ang. *Blank Node*), którego identyfikacja jest lokalna, zapewniająca unikatowość w ramach danego grafu. Przedmiotem natomiast może być również „literal”, czyli wartość typu prostego, zgodna ze specyfikacją XML Schema, reprezentująca cechy obiektu, które wyrażone są numerycznie, tekstowo lub za pomocą dat, przedziału czasu itp. Aby uporządkować struktury RDF opraco-

wano meta-słownik, ograniczający swobodę konstruowania grafów. Jest nim RDF Schema (RDFS). W ramach danego modelu, danej dziedziny dyskursu, zasoby reprezentujące podmioty winny być instancjami „klas” typu `rdfs:Class` („klas” będących instancjami meta-klasy `rdfs:Class`). Predykaty są instancjami `rdfs:Property`. Natomiast przedmioty są instancjami „klas” typu `rdfs:Class` lub literałami. Instancję `rdfs:Property` można więc traktować jak funkcję (*de facto* „funkcję częściową” w rozumieniu matematycznym, będącą uogólnieniem pojęcia „funkcji”), która odwzorowuje obiekty ze zbioru (klasy), będącej jej dziedziną (zdefiniowaną przez właściwość RDFS `rdfs:domain`) na zbiór wartości (definiowaną przez `rdfs:range`). Zbiorem wartości może być klasa lub typ prosty danych wyrażony literałem. Język ontologii OWL zwiększa ekspresywność opisu grafu. Właściwość, której wartość (przedmiot) jest wyrażona literałem, czyli wartością typu prostego, jest typu `owl:DatatypeProperty`, będącej podklasą `rdfs:Property`. Natomiast właściwość tworząca relację do innego zasobu nazwanego (URI) lub węzła anonimowego jest typu `owl:ObjectProperty`, również będącej podklasą `rdfs:Property`. Język OWL, w tym jego najnowsza wersja OWL2, dodatkowo klasyfikuje `owl:ObjectProperty`, w zależności od tego czy te właściwości są symetryczne, spełniają kryteria „funkcji” w rozumieniu matematycznym (czyli każdemu podmiotowi z dziedziny przyporządkowują dokładnie jedną wartość ze zbioru wartości) itp... Określa również relacje pomiędzy właściwościami – czy są wzajemnie odwrotne, przechodnie itd... OWL wprowadza dużo innych konstrukcji, które pozwalają precyzyjnie opisywać relacje między obiektami i je klasyfikować. Dzięki temu, że struktura RDF używa terminologii zdefiniowanej w słownikach korzystających z RDF Schema oraz ewentualnie OWL, przez trójki można opisać wszelkie immanentne cechy obiektów, które są istotne z punktu widzenia modelowanej dziedziny.

**Historia życia obiektów (kontyuantów).** To podejście opiera się na opisie obiektów przestrzennych, zarówno materialnych (budynki, obiekty topograficzne, obiekty infrastruktury) lub niematerialnych (działki, użytki gruntowe), które mają cechę trwałości w czasie.

Historia obiektu zapisana jest przez właściwości bezpośrednio odnoszące się do rzeczywistego bytu, który ten obiekt reprezentuje. W najprostszej postaci musi być właściwość powstania obiektu oraz potencjalnie właściwość zakończenia istnienia obiektu. W przypadku określonego rejestru, takiego jak Ewidencja Gruntów i Budynków, jeżeli tej drugiej właściwości nie ma, a baza danych jest poprawnie prowadzona, to oznacza, że obiekt ciągle istnieje. Jest to jednak oparte na tzw. założeniu świata zamkniętego, który zmusza nas do przyjęcia, że odpowiednie zasoby zawierają kompletną informację na temat przechowywanych obiektów, zgodnie z założonym modelem danych. Gdy jednak przyjmiemy założenie świata otwartego, takiego wniosku nie można wysnuć – wiemy, że obiekt powstał, natomiast brak nam informacji co do jego zakończenia (może istnieje, ale nie udało się jej znaleźć). Przykład 1 (RDF w serializacji Turtle) ilustruje sposób opisu obiektu, który powstał i już zakończył swoje istnienie.

**Przykład 1.** Historia życia obiektu określona za pomocą predykatów „dataUtworzenia” i „dataZakonczenia”.

```
@prefix egibvoc: <http://wogis2.igig.up.wroc.pl/egib_sloownik#> .
@prefix egib: <http://wogis2.igig.up.wroc.pl/egib_zasoby#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
egib:221106_2.0012.546 rdf:type egibvoc:dzialka_geod;
    egibvoc:dataUtworzenia "2013-05-13"^^xsd:date;
    egibvoc:dataZakonczenia "2014-09-19"^^xsd:date.
```

Predykaty `egibvoc:dataUtworzenia` i `egibvoc:dataZakonczenia` są zdefiniowane w określonej ontologii adresowanej, tu dla katastru, która uwzględnia proste cechy opisujące historię obiektów.

Alternatywnie można wykorzystać ontologię „Time Ontology” (`owl-time`, <http://www.w3.org/TR/owl-time/>), która dostarcza klas obiektów czasowych (`Instant`, `Interval`), relacji pomiędzy nimi wykorzystujących operatory Allena (Allen, 1984) oraz właściwości pozwalających dowiązać je do dowolnych obiektów świata rzeczywistego (np. `duration`).

Jeżeli obiekty zmieniają swój stan w sposób skokowy (dla działek może to być wydzielenie z jej powierzchni innej działki, bez likwidacji danej działki), to takie sytuacje standardowo są obsługiwane przez wersje obiektu (przykład 2).

**Przykład 2.** Historia życia obiektu opisana wersjami – wzajemnie odwrotne właściwości „`maWersje`” i „`jestWersjaObiektu`” wiążące obiekt działki z jego wersjami i na odwrót (RDF w serializacji Turtle)

```
egib:221106_2.0012.546 rdf:type egibvoc:dzialka_geod;
  egibvoc:maWersje egib:221106_2.0012.546v001;
  egibvoc:maWersje egib:221106_2.0012.546v002;
  egibvoc:maWersje egib:221106_2.0012.546v003;
  egibvoc:dataUtworzenia "2013-05-13"^^xsd:date.
egib:221106_2.0012.546v001 rdf:type egibvoc:dzialka_geod;
  egibvoc:jestWersjaObiektu egib:221106_2.0012.546;
  egibvoc:dataUtworzenia "2013-05-13"^^xsd:date;
  egibvoc:dataZakonczenia "2014-09-19"^^xsd:date.
egib:221106_2.0012.546v002 rdf:type egibvoc:dzialka_geod;
  egibvoc:jestWersjaObiektu egib:221106_2.0012.546;
  egibvoc:dataUtworzenia "2014-09-19"^^xsd:date;
  egibvoc:dataZakonczenia "2014-12-24"^^xsd:date.
egib:221106_2.0012.546v002 rdf:type egibvoc:dzialka_geod;
  egibvoc:jestWersjaObiektu egib:221106_2.0012.546;
  egibvoc:dataUtworzenia "2014-12-24"^^xsd:date.
```

Kolejność wersji jest wyznaczona historią ich życia. Aktualną wersją jest ta, która nie zakończyła swojego istnienia. Gdy obiekt kończy swoje życie do ostatniej wersji, jak i do obiektu głównego, zostanie dodana właściwość przechowująca tę samą datę końca istnienia.

**Kontynuanty jako obiekty zależne od okurentów.** Stan obiektów trwałych zależy od zdarzeń lub procesów. Podział działki wywołuje zmianę stanu obiektu trwałego. Może to być zmiana jego wersji lub jego zniszczenie i na jego miejsce powstanie nowego. W tym przypadku możemy powiązać obiekt trwały z obiektem zdarzenia lub procesu. Czyli do utworzenia asercji używamy `owl:ObjectProperty`, która tworzy relację z okurentem, mogącym zawierać wszelkie okoliczności oraz pełen opis zdarzenia lub procesu. W przypadku procesu, jego opis może zawierać tzw. profil procesu, definiujący na przykład tempo zmian jakiejś cechy opisującej kontynuant.

### Opis kontekstu czasowego i przestrzennego stanu przestrzeni

Przedstawione wcześniej podejścia odnoszą się do obiektów rzeczywistych – jeżeli są to obiekty trwałe, „kontynuanty” typu budynek, działka, użytek gruntowy, lub są to zdarzenia lub procesy, „okurenty” typu „podział działki”, „zmiana właściciela”, „zalesianie” itd.

Często mamy jednak do czynienia z informacją o kontekście czasowym podaną w innej formie. Pobierając dane statystyczne, na przykład na temat liczby ludności gmin, otrzymujemy parametr, nazwany jako „stan na”, zawierający datę. Uzyskujemy w ten sposób dwie dane odnoszące się do różnych poziomów informacyjnych. Jedną jest liczba ludności, która odnosi się do gminy, czyli opisuje obiekt rzeczywisty oraz datę, która określa jakiego momentu w czasie dana o liczbie ludności gminy dotyczy. Ta druga informacja nie dotyczy więc gminy, ale tej konkretnej wartości liczby ludności, czyli mówi nam, w którym momencie asercja „gmina” -> „liczba ludności” -> „wartość liczby ludności” jest prawdziwa. Jest więc ona metadana tej właściwości, a inaczej mówiąc, określa kontekst czasowy zarejestrowanego stanu obiektu rzeczywistego, jakim jest gmina. Jest to kontekst jednowymiarowy, ustawiający zestawy danych w szereg czasowy.

Analogicznie możemy postrzegać kontekst przestrzenny. Jeżeli rejestrowane są wartości dla obiektów rozciągających się na znacznym obszarze, na przykład rzeki, to pomiary poziomu wody dokonane w różnych punktach będą miały trzy dane: poziom wody odnoszący się bezpośrednio do rzeki, moment pomiaru oraz lokalizację wodowskazu. Moment i lokalizacja stanowią metadane dla konkretnej wartości poziomu wody w rzece. W tym przypadku kontekst jest wielowymiarowy.

**Zbiory danych RDF (RDF Datasets).** Zbiory danych RDF mają strukturę multigrafu, w którym grafy nazwane służą do porządkowania trójek, stosownie do kontekstu, w którym asercja kodowana przez trójkę jest prawdziwa. Takim kontekstem może być czas (okres czasu, moment) lub przestrzeń (lokalizacja, zasięg). W ten sposób graf nazwany obejmuje wszystkie trójki, które opisywały rzeczywistość w określonym czasie lub przestrzeni. Ponieważ grafy nazwane identyfikowane są przez URI i można na nich operować jak na „zwykłych” zasobach RDF, możliwe jest pokazywanie relacji pomiędzy grafami a innymi dowolnymi obiektami. Odbywa się to przez stworzenie trójek, w których węzłami będą URI grafów nazwanych, zawierające asercje na temat czasu, źródła pochodzenia lub prawdopodobieństwa odnoszącego się do tych kontekstów. W najprostszej postaci grafy identyfikowane przez URI opisane są predykatem określającym datę utworzenia grafu, czyli moment pozyskania informacji o obiektach przestrzennych (przykład 3).

**Przykład 3.** Grafy nazwane, które zawierają trójki opisujące stan przestrzeni w określonym momencie – migawki stanu rzeczywistości (RDF Dataset w serializacji TriG)

```
@prefix egibvoc: <http://wogis2.igig.up.wroc.pl/egib_slovník#> .
@prefix egib: <http://wogis2.igig.up.wroc.pl/egib_zasoby#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix gspq1: <http://www.opengis.net/ont/geosparql#>.
# ...
@prefix gr: <http://wogis2.igig.up.wroc.pl/ngraphs#> .
egib:res-reg {
  egib:221106_2.0012.143 rdf:type egibvoc:dzialka_geod.
  egib:221106_2.0012.257 rdf:type egibvoc:dzialka_geod.
  egib:221106_2.0012.546 rdf:type egibvoc:dzialka_geod.
  egib:221106_2.0012.578 rdf:type egibvoc:dzialka_geod.}
gr:graph-20130101 {
  egib:221106_2.0012.143 rdfs:label "143"^^xsd:string ;
  egibvoc:powierzHektar "0.050287890625"^^xsd:double;
  egibvoc:dataUtworzenia "2012-05-28"^^xsd:date;
  gspq1:rcc8ec egib:221106_2.0012.257.
  egib:221106_2.0012.257 rdfs:label "257"^^xsd:string ;
```

```

    egibvoc:powierzHektar "0.0678658877"^^xsd:double;
    egibvoc:dataUtworzenia "1984-05-28"^^xsd:date.
  egib:221106_2.0012.546 rdfs:label "546"^^xsd:string ;
    egibvoc:powierzHektar "0.0708124454"^^xsd:double.}
gr:graph-20130701 {
  egib:221106_2.0012.143 rdfs:label "32/1"^^xsd:string ;
  egibvoc:powierzHektar "0.043887890625"^^xsd:double;
  egibvoc:dataUtworzenia "2012-05-28"^^xsd:date;
  egibvoc:dataZakonczenia "2013-05-03"^^xsd:date;
  gspql:rcc8ec egib:221106_2.0012.578.}
gr:graph-20140101 {
  # ...}
gr:temp-graphs-reg {
  gr:graph-20130101 dct:created "2013-01-01"^^xsd:date.
  gr:graph-20130701 dct:created "2013-07-01"^^xsd:date.
  gr:graph-20140101 dct:created "2014-01-01"^^xsd:date.}

```

Przykład 3 zawiera 3 grafy nazwane, które gromadzą trójki prezentujące relacje między działkami w różnych momentach, w których została pobrana informacja o stanie przestrzeni. Opisuje to predykat ze słownika Dublin Core *dct:created* stwierdzający, w którym momencie ten graf powstał – graf odzwierciedla migawkę relacji pomiędzy obiektami oraz ich stanu na moment jego utworzenia. Podobny przykład jest zobrazowany na rysunku.

Trójki opisujące stan obiektów w różnym momencie, migawkach, dzięki grafom nazwanym posiadają trwały kontekst, w którym są prawdziwe. Kontekst, opisujący w tym przypadku moment pozyskania danych, przeciwdziała powstaniu sprzeczności między różnymi wykluczającymi się stwierdzeniami dotyczącymi tych samych obiektów, gdyż asercje te nie mogą być nigdy rozpatrywane w ramach jednego łącznego grafu. Ich prawdziwość może być rozpatrywana wyłącznie w kontekście opisanym za pomocą „czwórki”. Ale dzięki enkapsulacji w grafie nazwanym możemy porównywać jak zmieniły się trójki dla poszczególnych obiektów w przedziale czasu między migawkami.

Grafy nazwane mogą być opisywane w sposób bardziej ekspresywny, dzięki czemu można zawrzeć informację o relacjach między nimi, czyli między poszczególnymi momentami w badanym przedziale czasu. Predykaty do obrazowania relacji między grafami mogą być zdefiniowane w odrębnej ontologii. Przykład 4 prezentuje taką ontologię (*spattemporal\_onto.owl*), zapisaną w języku OWL (serializacja Turtle)

**Przykład 4.** Fragment ontologii opisu właściwości grafów – tutaj przykład relacji porządkowych pomiędzy grafami

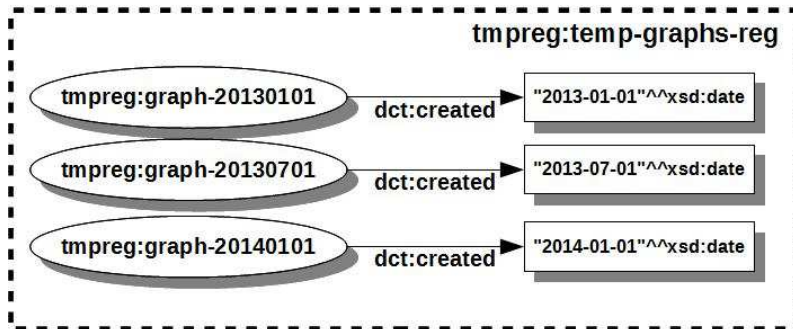
```

@prefix tmpvoc: <http://wogis2.igig.up.wroc.pl/spattemporal_onto#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
# ...
tmpvoc:poprzednik rdf:type owl:ObjectProperty.
tmpvoc:nastepca rdf:type owl:ObjectProperty.

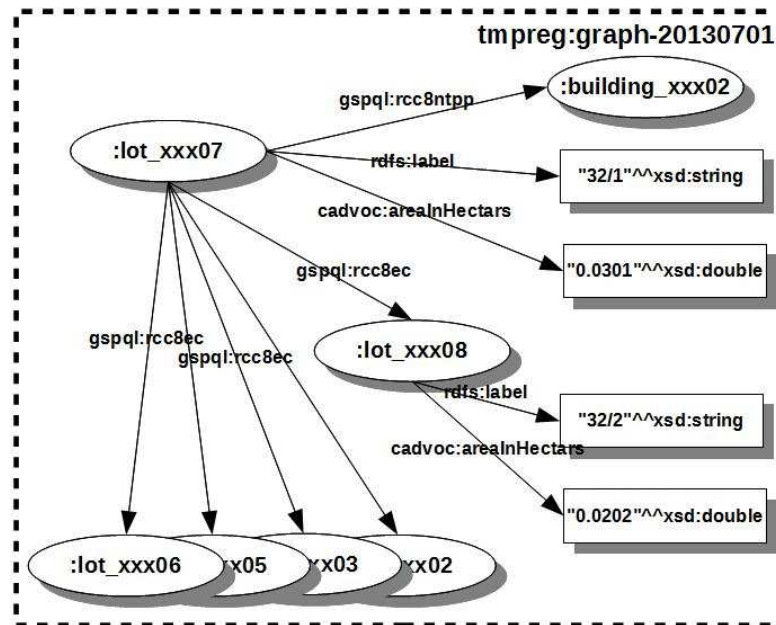
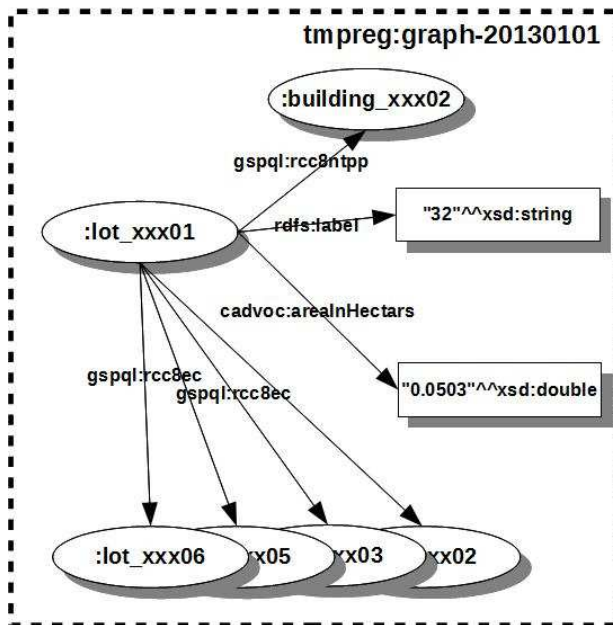
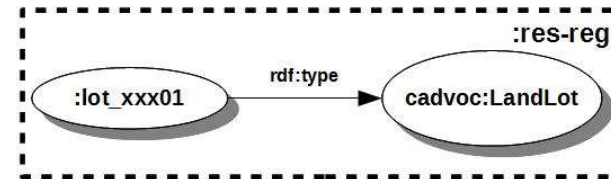
```

Korzystając ze słownika z przykładu 4, można zaprezentować relacje następstwa między grafami (przykład 5 – RDF Dataset w serializacji TriG).





@prefix : <http://wogis2.igig.up.wroc.pl/abox/sdss/toporel\_abox#> .  
 @prefix rdfs: <http://www.w3.org/2004/03/rdf/rdfs-1/> .  
 @prefix tmpreg: <http://wogis2.igig.up.wroc.pl/abox/sdss/tmpreg#> .  
 @prefix cadvoc: <http://wogis2.igig.up.wroc.pl/abox/spatlod/cadastre#> .  
 @prefix gspq: <http://www.opengis.net/ont/geosparql#> .



Rysunek. Grafy nazwane obrazujące migawki, zawierające opis rzeczywistości w kontekście czasowym

**Przykład 5.** Ontologia służąca do opisu właściwości grafów – tutaj przykład relacji porządkowych pomiędzy grafami

```
@prefix tmpvoc: <http://wogis2.igig.up.wroc.pl/spattemporal_onto#> .
# ...
gr:temp-graphs-reg {
  gr:graph-20130101 dct:created "2013-01-01"^^xsd:date;
  tmpvoc:nastepca gr:graph-20130701.
  gr:graph-20130701 dct:created "2013-07-01"^^xsd:date;
  tmpvoc:poprzednik gr:graph-20130101;
  tmpvoc:nastepca gr:graph-20140101.
  gr:graph-20140101 dct:created "2014-01-01"^^xsd:date;
  tmpvoc:poprzednik gr:graph-20130701.}
```

Grafy mogą być traktowane również jako instancje określonych klas. Ontologia `spattemporal_onto.owl` może definiować typ kontekstu czasowego opisanego przez graf (przykład 6).

**Przykład 6.** Inny fragment ontologii z przykładu 4 – klasyfikacja grafów ze względu na czas trwania (OWL/Turtle).

```
@prefix tmpvoc: <http://wogis2.igig.up.wroc.pl/spattemporal_onto#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
# ...
tmpvoc:Okres rdf:type owl:Class.
tmpvoc:Migawka rdf:type owl:Class;
  owl:subClassOf tmpvoc:Okres;
  owl:subClassOf [ a owl:Restriction;
    owl:onProperty tmpvoc:zakresCzasowy;
    owl:hasValue "P0D"^^xsd:duration.].
```

Przykład 7 użycia typów dla obiektów grafów w asercjach RDF.

**Przykład 7.** Zastosowanie klasyfikacji grafów z przykładu 6 do opisu typu grafów – typ grafu determinuje jego charakter zdefiniowany w ontologii (RDF Dataset w serializacji TriG)

```
gr:temp-graphs-reg {
  gr:graph-20130101 rdf:type tmpvoc:Migawka;
  dct:created "2013-01-01"^^xsd:date;
  tmpvoc:nastepca gr:graph-20130701.
  gr:graph-20130701 rdf:type tmpvoc:Migawka;
  dct:created "2013-07-01"^^xsd:date;
  tmpvoc:poprzednik gr:graph-20130101;
  tmpvoc:nastepca gr:graph-20140101.
  gr:graph-20140101 rdf:type tmpvoc:Migawka;
  dct:created "2014-01-01"^^xsd:date;
  tmpvoc:poprzednik gr:graph-20130701.}
```

Technika zaprezentowana w przykładzie 7 może być użyta do klasyfikacji grafów, jako instancji klas zdefiniowanych w ontologiach formalnych lub ontologiach dziedzinowych, tworzonych na bazie ontologii formalnych. Grafy nazwane, zawierające migawki czasowe, mogą być analizowane pod kątem zachodzących zmian z użyciem narzędzi do badania różnic w grafach, analogicznych do programu Diff dla kodu źródłowego. Opierają się one na badaniu izomorfizmu grafów (<http://www.w3.org/TR/rdf11-concepts/#graph-isomorphism>), a rezultaty są zapisywane w specjalnej ontologii Delta (Berners-Lee, Connolly, 2001). Klasyfikacja grafów może służyć do definiowania strategii odkrywania różnic.

Podobnie jak dla kontekstu czasowego, można grupować stwierdzenia odpowiadające kontekstowi przestrzennemu. Grafy nazwane będą zawierały asercje opisujące stan rozległych zjawisk, wyodrębniając określone obszary. Te grafy również można opisać wzajemnymi relacjami i odpowiednio sklasyfikować.

### **Zastosowanie zasobów gromadzonych w RDF Datasets do budowy baz wiedzy**

Uporządkowane dane czasoprzestrzenne w formie zbiorów RDF pozwolą na kontrolowane zasilanie baz, wykorzystujących modele logiczne zdefiniowane w ontologiach formalnych i zbudowanych na ich podstawie ontologiach dziedzinowych. O ile składnia i struktura logiczna tych ontologii, oparta na języku OWL, nie uwzględnia struktur multigrafów z RDF Datasets, o tyle syndykacja i uporządkowanie danych pozwalają na metodyczne zasilanie baz wiedzy odpowiednimi zestawami danych. Klasyfikacja grafów może stanowić dodatkową meta-informację o charakterze pozyskiwanych obiektów, ułatwiając tworzenie formalnego systemu ich importu.

## **Podsumowanie i wnioski**

Wykorzystanie zbiorów danych RDF (RDF Dataset) daje zupełnie nowe, elastyczne możliwości strukturyzowania oraz opisywania złożonych i heterogenicznych danych, opisujących dynamiczne dane przestrzenne. Traktowanie grafów jako „normalnych” (z pewnymi zastrzeżeniami) zasobów RDF daje możliwość ich klasyfikowania, zgodnie ze wszystkimi regułami logiki opisowej. Dlatego wyłania się możliwość typizacji grafów, jako instancji klas zdefiniowanych w ontologiach formalnych lub ontologiach dziedzinowych, tworzonych na bazie ontologii formalnych.

Syndykacja danych za pomocą grafów może wykazać w pewnych sytuacjach przewagę nad agregacją danych za pomocą cech, gdyż izoluje wzajemnie wyrażenia, które mogłyby reprezentować sprzeczne asercje – pozostawienie historii obiektów w ramach jednej struktury (wspólnego grafu) wymaga wprowadzenia dodatkowych narzędzi, gwarantujących spójność pozyskiwanych, z różnych źródeł w różnym czasie, danych. Grafy nazwane gromadzą bowiem całe stwierdzenia (w postaci trójek), podczas gdy cechy odnoszące się do poszczególnych obiektów wyróżniają tylko te wskazane encje. Tak więc reprezentacji danych w RDF Datasets może dotyczyć wszelkich danych pozyskiwanych w postaci migawek (na przykład próbek pomiarów z monitoringu), pozyskiwanych cyklicznie lub reprezentujących szeregi czasowe zgodnie z harmonogramem sprawozdawczości (dane statystyczne).

Zamknięcie trójek opisujących stan obiektów w poszczególnych migawkach czasowych w odrębnych grafach nazwanych, pozwala na jednoznaczne zdefiniowanie kontekstu czasowego, w którym stwierdzenia reprezentowane przez te trójki były prawdziwe. Przeciwdziała to powstaniu sprzeczności między różnymi wykluczającymi się stwierdzeniami dotyczącymi tych samych obiektów, ale w różnych momentach, gdy dane o stanie rzeczywistości były pozyskiwane. Dodatkowo, dzięki enkapsulacji w grafie nazwanym możemy porównywać, jak zmieniły się trójki dla poszczególnych obiektów w przedziale czasu między migawkami. Różnice między grafami, dla poszczególnych momentów w czasie, będą rejestrowane w ontologii różnicowej Delta, tworzącej wykaz zmian stanu rzeczywistości.

Traktowanie grafów nazwanych jako zasobów RDF oraz ich klasyfikacja z użyciem ontologii odkrywa nowe możliwości ich opisu i analiz. Wbudowanie ich w „meta-grafy” RDF otwiera możliwość skonstruowania na ich bazie ontologii o charakterze aksjomatycznym, a co za tym idzie prowadzenie procedur wnioskowania z użyciem, tzw. silników wnioskowania (reasonerów). Służy to odkrywaniu nowych faktów dla całych struktur opisu rzeczywistości, uwzględniając ich kontekst czasoprzestrzenny.

### Literatura

- Allen J.F., 1984: Towards a general theory of action and time. *Artificial Intelligence* 23: 123-154.
- Mera Caraballo A.A., Moura Arruda Junior N., Pereira Nunes B., Rabello Lopes G., Casanova M.A., 2014: TRTML – A Triplet Recommendation Tool based on Supervised Learning Algorithms. The Semantic Web: ESWC 2014 Satellite Events May 25-29, 2014 (Anissaras, Crete, Greece)
- Berners-Lee T., Hendler J., Lassila O., 2001: The Semantic Web. *Scientific American* 501.
- Berners-Lee T., Connolly D., 2001: Delta: an ontology for the distribution of differences between RDF graphs. *MIT CSAIL* 2009/08/27. <http://www.w3.org/DesignIssues/Diff>
- Berners-Lee T., 2004: Reifying RDF (properly), and N3. *MIT CSAIL*. <http://www.w3.org/DesignIssues/Reify.html>
- Bizer Ch., Heath T., Berners-Lee T., 2009: Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*.
- Bittner T., Donnelly M., Smith B., 2009: A spatio-temporal ontology for geographic information integration. *International Journal of Geogr. Inf. Science*, Vol. 23: 765-798.
- Carroll J.J., Bizer C., Hayes P., Stickler P., 2005: Named graphs, provenance and trust. Proceedings of the 14<sup>th</sup> international conference on World Wide Web. ACM Press, (New York, NY, USA): 613-622.
- Cox S., 2014: Basic Observations and Sampling Feature Ontology. *Semantic Web Journal*. <http://www.semantic-web-journal.net/system/files/swj890.pdf>
- Grenon P., 2003: BFO in a Nutshell: A Bi-categorial Axiomatization of BFO and Comparison with DOLCE, *IFOMIS REPORTS*.
- Grenon P., Smith B., 2004: SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spat. Cog. Comp.* Vol. 4: 69-104.
- Gruber T., 1995: What is an Ontology? *International Journal of Human-Computer Studies*, Vol. 43, Issues 4-5: 907-928.
- Gutierrez C., Hurtado C.A., Vaisman A., 2007: Introducing Time into RDF. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 2: 207 p.
- Harvey F., Jones J., Scheider S., Iwaniak A., Kaczmarek I., Łukowicz J., Strzelecki M., 2014: Little Steps Towards Big Goals. Using Linked Data to Develop Next Generation Spatial Data Infrastructures (aka SDI 3.0). Proceedings of the AGILE' 2014 ICGIS (Castellón), ISBN: 978-90-816960-4-3.
- Herre H., 2010: General Formal Ontology (GFO) A Foundational Ontology for Conceptual Modelling. *Media*, Vol. 2, Issue: 15: 1-50.
- Husserl E., 1901: *Badania logiczne (Logische Untersuchungen)*. Trans. A.Póltawski. PWN 2006.
- Kuhn W., Kauppinen T., Janowicz K., 2014: Linked Data – A Paradigm Shift for Geographic Information Science. 8th IC, GIScience 2014, Vienna, Austria. Springer. *Lecture Notes in Computer Science* Vol. 8728: 173-186.
- Li Y.D., Tong X.H., Liu M.L., 2008: A Unified Version-Based Spatio-Temporal Data Model. *Archives*: 103-108.
- Iwaniak A., Kaczmarek I., Kubik T., Łukowicz J., Paluszynski W., Kourie D., Cooper A.K., Coetzee S., 2011: An intelligent geoportal for spatial planning. 25th International Cartographic Conference ICC 2011. (Paris).

- Iwaniak A., Lukowicz J., Strzelecki M., Kaczmarek I., 2013: Ontology Driven Analysis of Spatio-temporal Phenomena, Aimed At Spatial Planning And Environmental Forecasting, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Archives* Vol. XL-7/W2, ISPRS2013-SSG, (Antalya, Turkey): 119-124.
- Iwaniak A., Kaczmarek I., Strzelecki M., Lukowicz J., 2014: Publikowanie danych przestrzennych jako Linked Open Data. *Roczniki Geomatyki* t.12, z. 1(63): 67-79, PTIP, Warszawa.
- ISO/TC 211 Geographic information/Geomatics. 2009: Report from stage 0 Project 19150 Geographic information – Ontology. ISO/TC 211/WG 7/PT 19150.
- Kiryakow A., Momtchev V., 2009: Triplesets: Tagging and Grouping in RDF Datasets. <http://www.w3.org/2009/12/rdf-ws/papers/ws24.pdf>
- Li Y.D., Tong X.H., Liu M.L., 2008: A Unified Version-Based Spatio-Temporal Data Model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Archives* Vol. XXXVII. Part B2. (Beijing): 103-108.
- Lukowicz J., Iwaniak A., Strzelecki M., Kaczmarek I., 2014: Modelowanie dynamiki zjawisk w planowaniu przestrzennym z wykorzystaniem ontologii formalnych. *Roczniki Geomatyki* t.12, z.1(63): 81-91, PTIP, Warszawa.
- Lukowicz J., Kaczmarek I., Iwaniak A., 2012: Semantic metadata in SDI for decision support systems in spatial planning. Global Geospatial Conference 2012 (Québec City, Canada). GSDI Association.
- Ryan N., Pascoe J., Morse D., 1998: Enhanced Reality Fieldwork: the Context Aware Archaeological Assistant. Kent, CT2 7NF, UK. <http://www.cs.kent.ac.uk/pubs/1998/616/content.html>
- Udrea O., Reforgiato Recupero D., Subrahmanian V.S., 2008: Annotated RDF. University of Maryland IACS.

### **Streszczenie**

*Istotnym wyzwaniem dla gospodarki przestrzennej jest poprawny opis zjawisk zachodzących w przestrzeni i prognozowanie przyszłego jej stanu. Potrzebne są systemy służące gromadzeniu i porządkowaniu danych opisujących stan przestrzeni w jej dynamicznej postaci. Takich rozwiązań mogą dostarczyć technologie sieci semantycznych (Semantic Web), w tym grafy RDF, służące udostępnianiu zasobów przestrzennych, jako danych powiązanych (Linked Data) oraz ontologie, zarówno w postaci taksonomicznej, jak i aksjomatycznej. Obrazowanie odniesienia przestrzennego i czasowego może być realizowane w bezpośrednim odniesieniu do obiektów odzwierciedlających przedmioty i zjawiska występujące w rzeczywistej przestrzeni albo w formie kontekstu czasoprzestrzennego, w jakim informacja o tych obiektach została pozyskana. Pierwsze podejście odnosi się do modelowania danych, a drugie do modelowania metadanych. Autorzy prezentują strukturę, która, dzięki wykorzystaniu grafów RDF w formie tzw. zbiorów danych RDF (RDF Datasets), łączy oba modele w jeden spójny system. Wspólny system, który daje możliwość wykorzystania tych samych narzędzi do wyszukiwania zasobów na podstawie metadanych, a następnie do eksploracji i przetwarzania samych zasobów, które są nimi opisane. Dodatkowo pozwala uzupełniać brakujące informacje o lokalizacji lub trwaniu obiektów rzeczywistych, dzięki informacji zawartej w metadanych.*

### **Abstract**

*The crucial challenge for spatial development is proper description of phenomena occurring in real space and forecasting their future state. We need systems for gathering and organizing data describing the state of space in its dynamic form. Semantic Web technologies could be considered as promising tools to acquire such an aim. Among them, RDF graphs in the form of Linked Data and ontologies (in the form of taxonomic as well as axiomatic ontologies) seem to be very useful for such purpose. Presentation of spatial and temporal reference of reflected real objects can be performed by direct object description and by exposing spatiotemporal context of information about object from which this information was retrieved. First approach is related to data modeling, second to metadata modeling. The authors present the structure of RDF Datasets which makes it possible to combine both models in*

*a coherent system. The joint system makes it possible to reuse common tools for searching of resources using metadata, and then for exploration and processing of the spatial data found. In addition, such approach gives us a chance to fill the gaps in retrieved data, referring to duration or location of objects, through the information contained in metadata.*

mgr inż. Jaromar Łukowicz  
jaromar.lukowicz@struktura.eu

dr inż. Adam Iwaniak  
adam.iwaniak@up.wroc.pl