

УДК 004.62

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ИНТЕГРАЦИИ РАЗНОРОДНЫХ ПО СТРУКТУРЕ И ТЕМАТИКЕ ПРОСТРАНСТВЕННЫХ БАЗ ДАННЫХ В ЕДИНУЮ РЕГИОНАЛЬНУЮ БАЗУ ДАННЫХ

С.В. Павлов¹, О.А. Ефремова²

Уфимский государственный авиационный технический университет, Уфа, Россия

¹psvgis@mail.ru, ²efremova-oa@yandex.ru

Аннотация

Статья посвящена решению задачи интеграции разнородных по структуре и тематике пространственных баз данных в единую региональную базу данных для организации информационной поддержки принятия решений по управлению крупным промышленным регионом. Для осуществления такого рода интеграции на семантическом уровне авторами предложена онтологическая модель, позволяющая описать с единых методологических позиций структуру разнородных баз пространственных данных, выделив в их составе онтологии пространственных, атрибутивных и метаданных. Разработан алгоритм построения и встраивания онтологии в программное обеспечение интеграции разнородных по структуре и тематике пространственных баз данных. Апробация разработанного алгоритма осуществлена на примере создания единой региональной базы пространственных данных в составе Геоинформационной системы органов исполнительной власти Республики Башкортостан.

Ключевые слова: онтология, пространственные данные, базы данных, интеграция, промышленный регион, принятие решений.

Цитирование: Павлов, С.В. Онтологическая модель интеграции разнородных по структуре и тематике пространственных баз данных в единую региональную базу данных / С.В. Павлов, О.А. Ефремова // Онтология проектирования. – 2017. – Т. 7, №3(25). – С.323-333. – DOI: 10.18287/2223-9537-2017-7-3-323-333.

Введение

В процессе принятия решений по управлению крупным промышленным регионом огромную роль играет вопрос наличия и своевременного предоставления пространственной информации о территории региона (карт территории различных масштабов, снимков из космоса, ортофотопланов), о промышленных и других объектах, расположенных на его территории. На сегодняшний день, как в органах исполнительной власти, так и на предприятиях, существует множество различных источников пространственной информации [1]. В качестве основных особенностей данных источников необходимо отметить их неоднородность и динамическую пополняемость. Для организации информационной поддержки принятия решений по управлению крупным промышленным регионом необходимо объединить данные из разных источников в единую региональную базу данных (БД).

Часть источников представляет собой структурированные данные (реляционные или объектно-ориентированные БД), часть - неструктурированные данные (тексты, документы и др., содержащие в своём составе пространственные данные), а часть являются квазиструктурированными (например, различного вида отчёты и атласы).

Широкие возможности для создания интегрированных БД предоставляют современные семантически ориентированные технологии, одной из наиболее перспективных является технология представления знаний в виде онтологий [2-5].

1 Онтологическая модель процесса интеграции пространственных баз данных

Онтологическая модель интегрированной пространственной БД представляет собой совокупность онтологических моделей отдельных информационных систем, пространственных БД, текстов документов, отчетов и атласов, т.е. онтологий отдельных источников пространственной информации (см. рисунок 1).

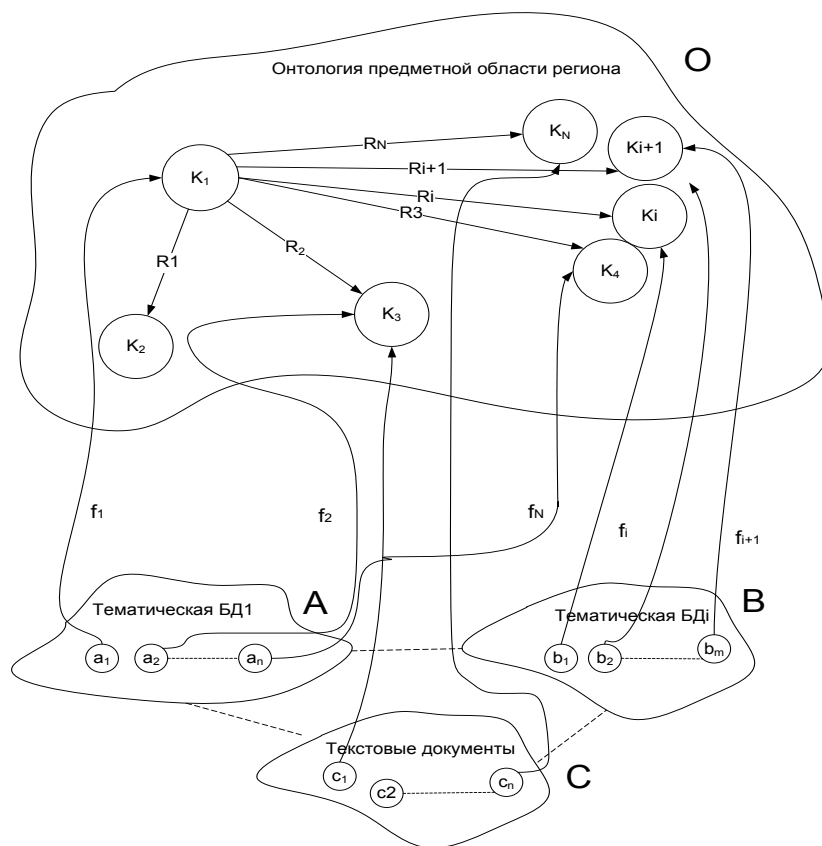


Рисунок 1 – Онтологическая модель интеграции пространственных баз данных

Интегрирующую онтологическую модель пространственных данных O можно представить в виде следующей совокупности объектов:

$$(1) \quad O = \langle K, R, F, T, I \rangle,$$

где $K = \{K_1, \dots, K_n\}$ – множество концептов или терминов (понятий) предметной области;

$R = \{R_1, \dots, R_n\}$ – множество отношений между концептами;

$F = \{f_1, \dots, f_i\}$ – множество функций отображения множества концептов отдельных источников пространственных данных во множество концептов единой онтологической модели рассматриваемой предметной области. Необходимо отметить, что функция F не является биективной, так как одному концепту результирующей онтологической модели K_i , может соответствовать некоторое множество концептов онтологических моделей источников пространственных данных;

$T = \{T_1, \dots, T_2\}$ – множество типов данных;

$I = \{I_1, \dots, I_l\}$ – множество источников пространственных данных.

$$F : A \rightarrow O, F : B \rightarrow O, F : C \rightarrow O,$$

где $A = \{A_1, \dots, A_j\}$ – множество концептов предметной области A ;

$B = \{B_1, \dots, B_r\}$ - множество концептов предметной области В;
 $C = \{C_1, \dots, C_f\}$ - множество концептов предметной области С.

Важной особенностью пространственных БД является то, что в их состав входят три типа данных – пространственные, атрибутивные и метаданные, соответственно онтологическая модель отдельного i-го источника пространственных данных – пространственных БД – может быть представлена в следующем виде:

$$(2) \quad M_{БД_i}^O = \langle O^M, O^{ПД}, O^{АД} \rangle,$$

где O^M – онтология метаданных;
 $O^{ПД}$ – онтология пространственных данных;
 $O^{АД}$ – онтология атрибутивных данных.

Онтология метаданных – это онтология верхнего уровня, которая содержит базовые понятия и отношения между ними, используемые в дальнейшем при построении онтологий пространственных и атрибутивных данных.

Онтология метаданных (см. рисунок 2) может быть представлена следующим соотношением:

$$(3) \quad O^M = \{K^M, R^M, S^M\},$$

где $K^M = \{K_1^M, \dots, K_s^M\}$ – множество концептов метаданных;
 $R^M = \{R_1^M, \dots, R_w^M\}$ – множество отношений между концептами метаданных;
 $S^M = \{S_1^M, \dots, S_m^M\}$ – множество свойств классов метаданных.

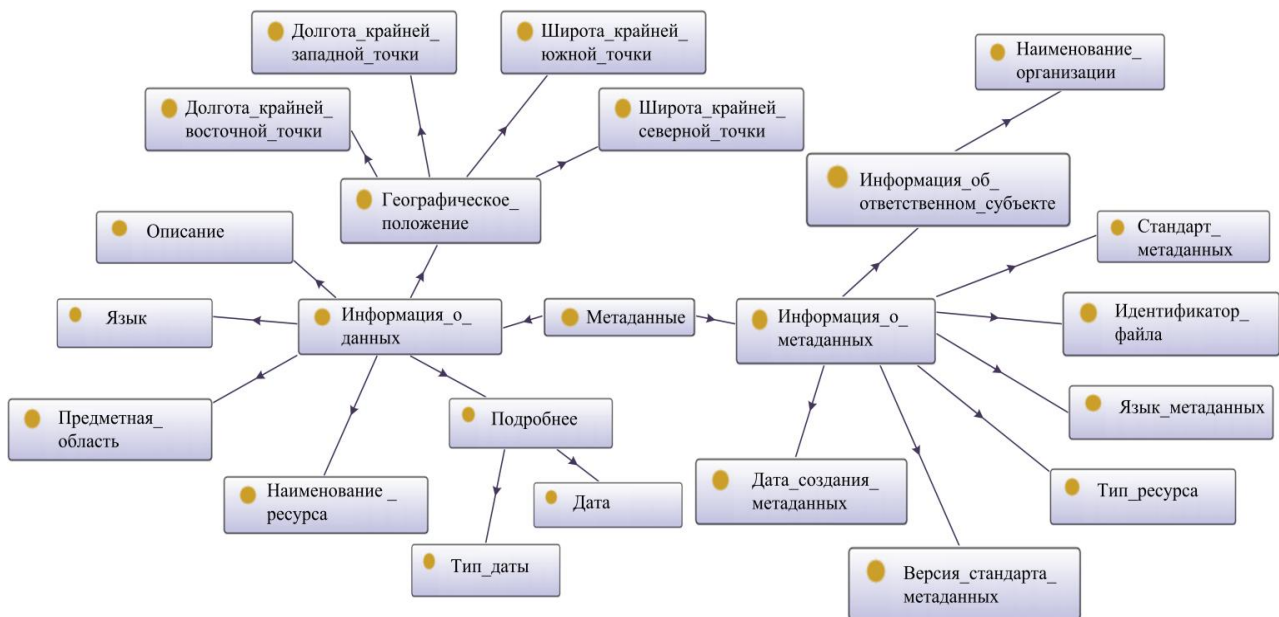


Рисунок 2 – Онтология метаданных базы пространственных данных

Онтология пространственных данных содержит описание основных типов пространственных объектов, их свойств, связей между ними, а также описание свойств используемой топологии. Пример онтологической модели пространственных данных представлен на рисунке 3. Онтология пространственных данных может быть представлена в следующем виде:

$$(4) \quad O^{ПД} = \{K^{ПД}, R^{ПД}, S^{ПД}\}$$

где $K^{ПД} = \{K_1^{ПД}, \dots, K_s^{ПД}\}$ – множество концептов пространственных данных;

$R^{ПД} = \{R_1^{ПД}, \dots, R_w^{ПД}\}$ – множество отношений между концептами пространственных данных;

$S^{ПД} = \{S_1^{ПД}, \dots, S_m^{ПД}\}$ – множество свойств классов пространственных данных.

Для всех трёх типов пространственных данных, исходя из того, что для пространственных объектов характерны свои виды отношений [4, 5], множество R можно представить в виде:

$$(5) \quad R^{ПД} = \{L, V, N, D, Y, Z\},$$

где L – отношение «*Касается*» (часть объекта из класса пространственных объектов 1 соприкасается с границей объекта из класса пространственных объектов 2, внутренние части объектов не пересекаются);

V – отношение «*Содержит*» (объект из класса пространственных объектов 1 полностью включает в себя объект из класса пространственных объектов 2);

N – отношение «*Пересекает*» (любая часть объекта из класса пространственных объектов 1 соприкасается с любой частью объекта из класса пространственных объектов 2);

D – отношение «*Внутри*» (объект из класса пространственных объектов 2 полностью включает в себя объект из класса пространственных объектов 1);

Y – отношение «*Пересекает*» (внутренняя часть объекта из класса пространственных объектов 1 в какой-либо точке соприкасается с внутренней частью или границей (в случае с полигоном) объекта из класса пространственных объектов 2);

Z – отношение «*Перекрывает*» (внутренняя часть объекта из класса пространственных объектов 1 частично перекрывает объект из класса пространственных объектов 2, сравниваться могут только объекты с одинаковой геометрией).

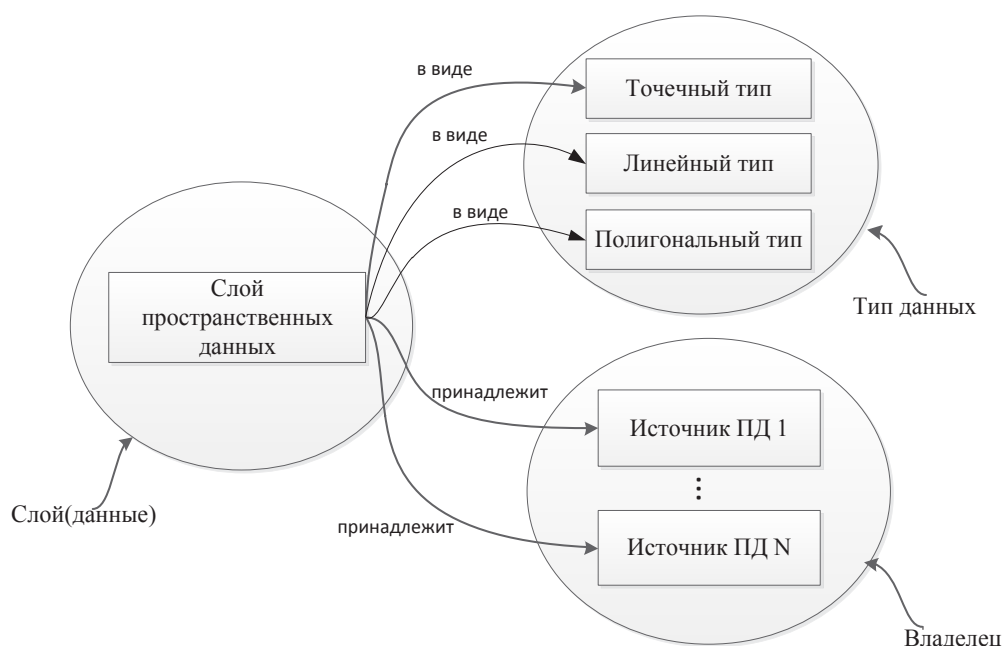


Рисунок 3 – Пример онтологической модели пространственных данных

Онтология атрибутивных данных содержит описание основных объектов, их свойств, связей между ними. Онтологию атрибутивных данных предлагается представить следующим соотношением:

$$(6) \quad O^{АД} = \{K^{АД}, R^{АД}, S^{АД}\},$$

где $K^{AD} = \{K_1^{AD}, \dots, K_s^{AD}\}$ – множество концептов атрибутивных данных;

$R^{AD} = \{R_1^{AD}, \dots, R_w^{AD}\}$ – множество отношений между концептами атрибутивных данных.

$S^{AD} = \{S_1^{AD}, \dots, S_m^{AD}\}$ – множество свойств классов атрибутивных данных.

Для атрибутивных данных могут быть выделены следующие виды отношений из множества R :

$$(7) \quad R^{AD} = \{P, H\},$$

где P – множество парадигматических отношений: причинно-следственные, отношения обобщения и агрегации;

H – синтагматические отношения, заданные дескриптивной логикой.

2 Алгоритм построения и встраивания онтологии в программное обеспечение для интеграции пространственных баз данных

Круг применения онтологий в программном обеспечении (ПО) различного назначения довольно широк [6-9]. Однако вопросам построения и внедрения онтологии в ПО для интеграции пространственных БД уделено недостаточно внимания.

В большинстве случаев современные онтологии строятся одинаково [6,10,11]. Так, в первую очередь определяются область и масштабы создаваемой онтологии и перечисляются важные термины и базовые понятия в онтологии. Затем осуществляется определение классов и отношений между ними, свойств классов и наложение ограничений на их значения. Заключительным этапом является создание отдельных экземпляров или индивидов классов и приращение значений атрибутам и свойствам. (см. рисунок 4).

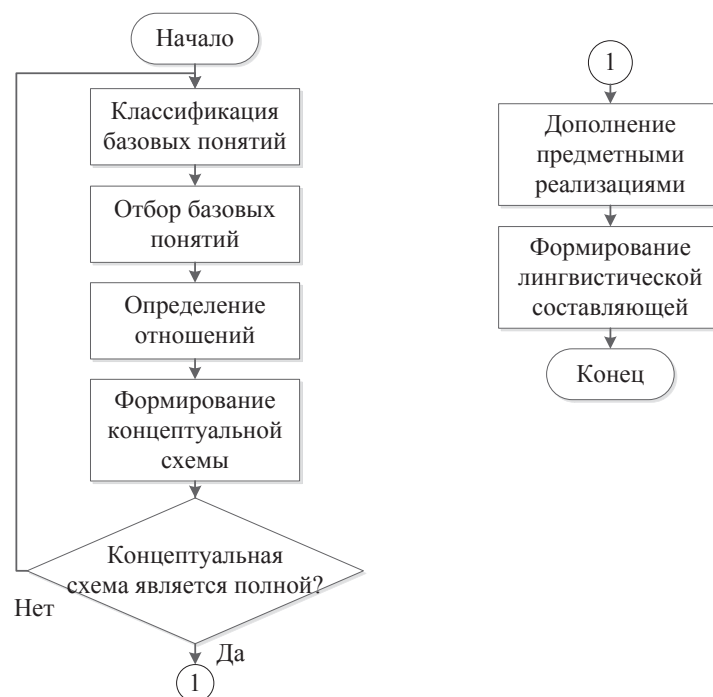


Рисунок 4 – Алгоритм построения онтологии

Для решения задачи интеграции разрозненных пространственных БД и организации поиска пространственных данных из различных источников необходимо включить процесс создания онтологии в процедуру разработки ПО. Для этого необходимо выполнить следующие этапы.

- 1) организовать обработку пространственных данных в составе БД с учётом требований для обеспечения организации интеллектуальной поддержки принятия решений по управлению регионом.
- 2) создать и дополнить онтологию пространственных данных, проверить на достаточность, корректность и непротиворечивость данных.
- 3) разработать типовые сценарии обработки пространственных данных в текущей онтологии пространственных данных.

Алгоритм построения и встраивания онтологии в ПО для интеграции пространственных БД представлен на рисунке 5.

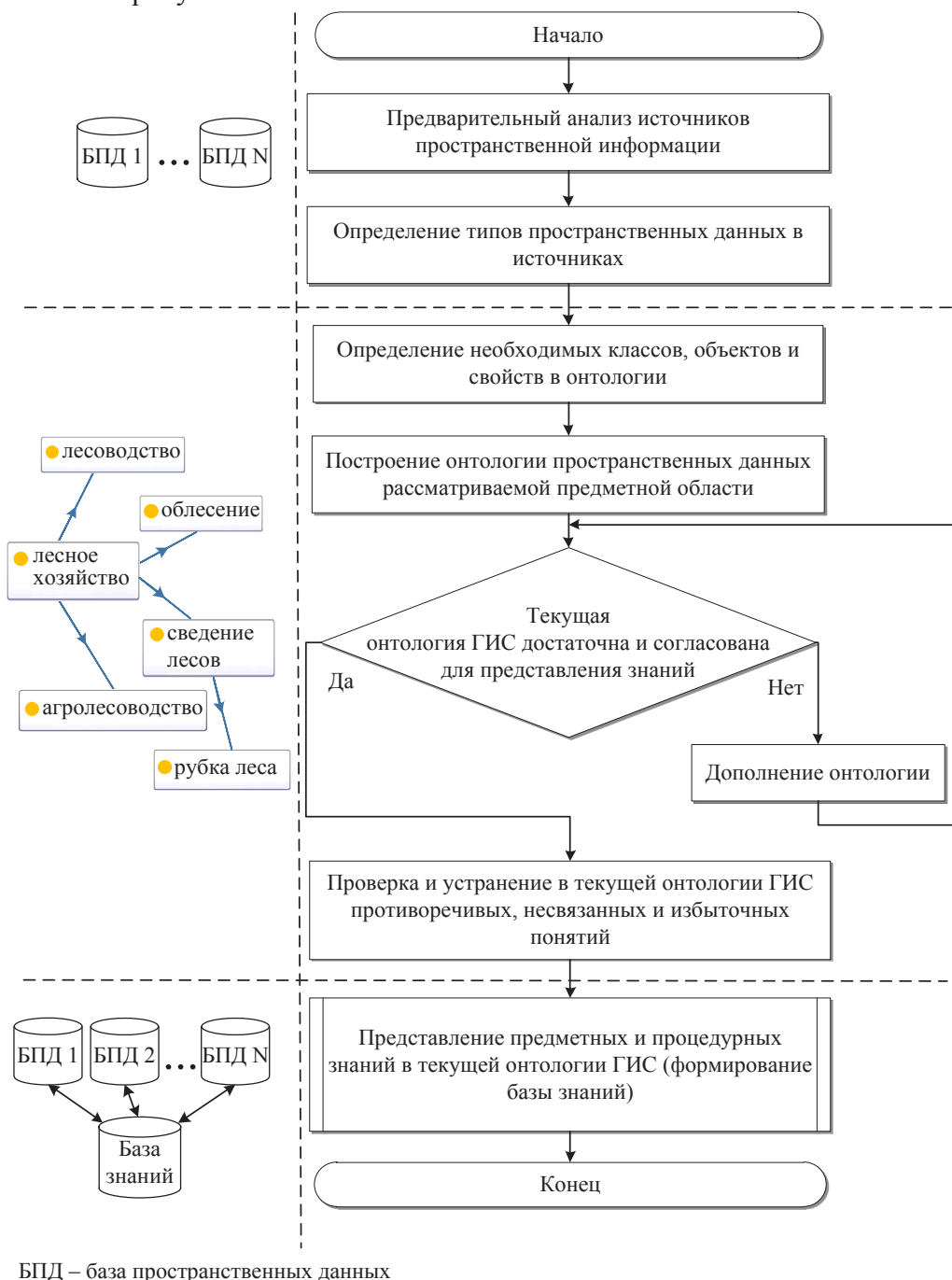


Рисунок 5 – Алгоритм интеллектуализации ГИС на основе онтологии

Применение данного алгоритма при разработке ПО позволяет обеспечить интеграцию пространственных данных из различных источников, а также повысить качество результатов запросов и упростить поиск интересующей пространственной информации.

3 Пример интеграции разнородных по структуре и тематике пространственных баз данных с использованием онтологической модели

Одним из примеров реализации онтологической модели интеграции пространственных БД и алгоритма её встраивания в ПО для интеграции пространственных БД является Геоинформационная система органов исполнительной власти республики Башкортостан (ГИС ОИВ РБ). Основной целью создания данной системы является обеспечение органов исполнительной власти республики актуальной, достоверной и комплексной геопространственной информацией для оперативного исследования, оценки и обоснования управленческих решений. Для интеграции пространственных БД различных органов исполнительной власти в единую региональную БД была разработана онтологическая модель, на основе которой был разработан каталог ресурсов системы, позволяющий осуществить интеграцию разрозненных пространственных БД [5]. Логическая структура ГИС ОИВ РБ представлена на рисунке 6.

Пример интеграции в ГИС ОИВ РБ пространственных БД двух органов исполнительной власти - Министерства связи и массовых коммуникаций и Министерства природопользования и экологии Республики Башкортостан представлен на рисунке 7. Результаты интеграции пространственных БД этих двух ведомств позволяют выявить места размещения базовых станций операторов сотовой связи, расположенные на территории особо охраняемых природных территорий, а также организовать поддержку принятия решений по размещению объектов связи на территории республики.

Заключение

Предложенная онтологическая модель интеграции разрозненных по структуре и тематике БД позволяет описать с единых позиций структуру разнородных пространственных БД и тем самым обеспечить возможность их дальнейшей интеграции в единую региональную базу. Использование данной модели на практике позволяет снизить затраты за счёт отсутствия необходимости в постоянных корректировках программ, моделей данных, разработки различных конвертеров, а также дублирования информации в различных информационных системах.

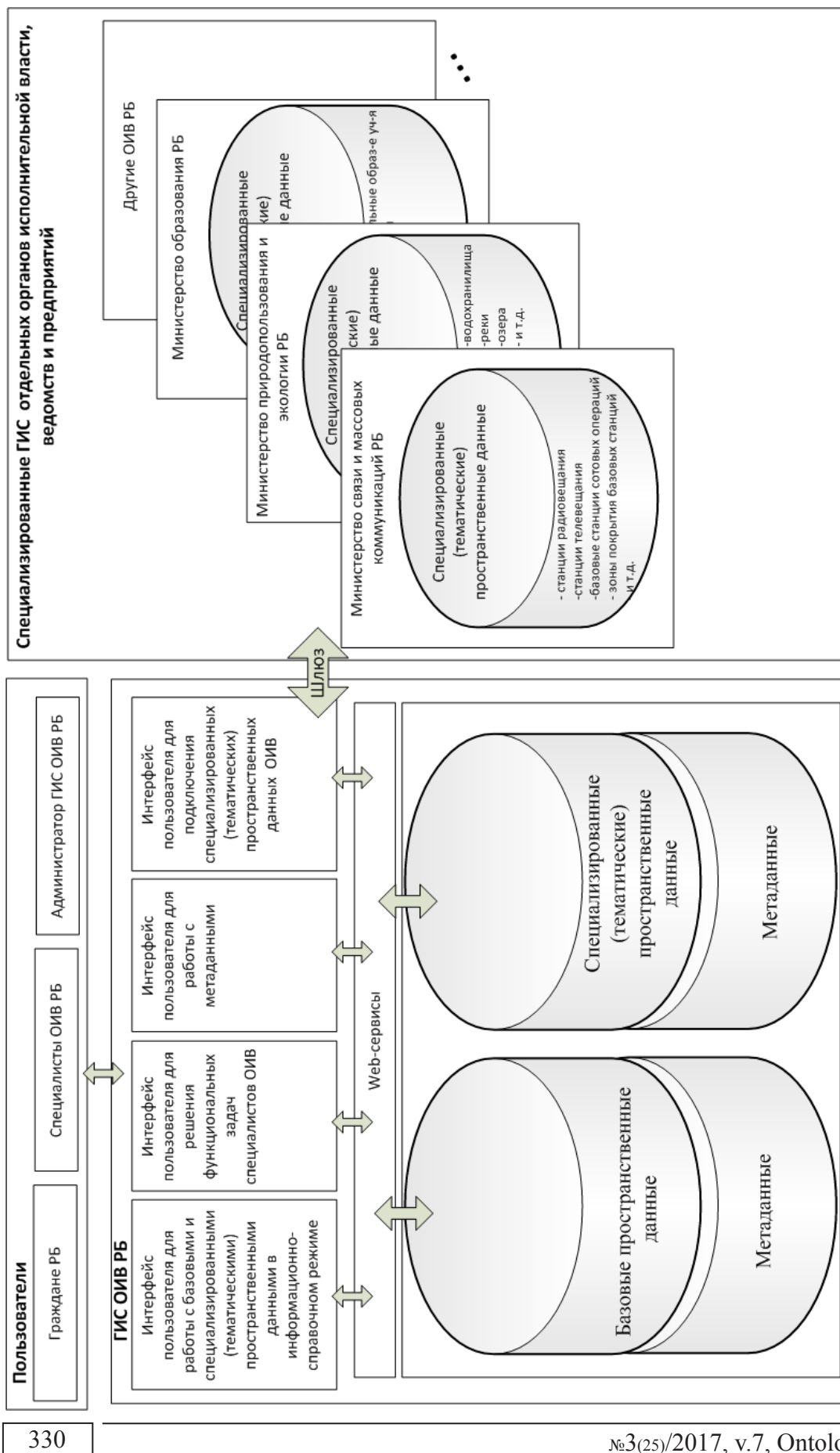


Рисунок 6 – Логическая структура ГИС ОИВ РБ

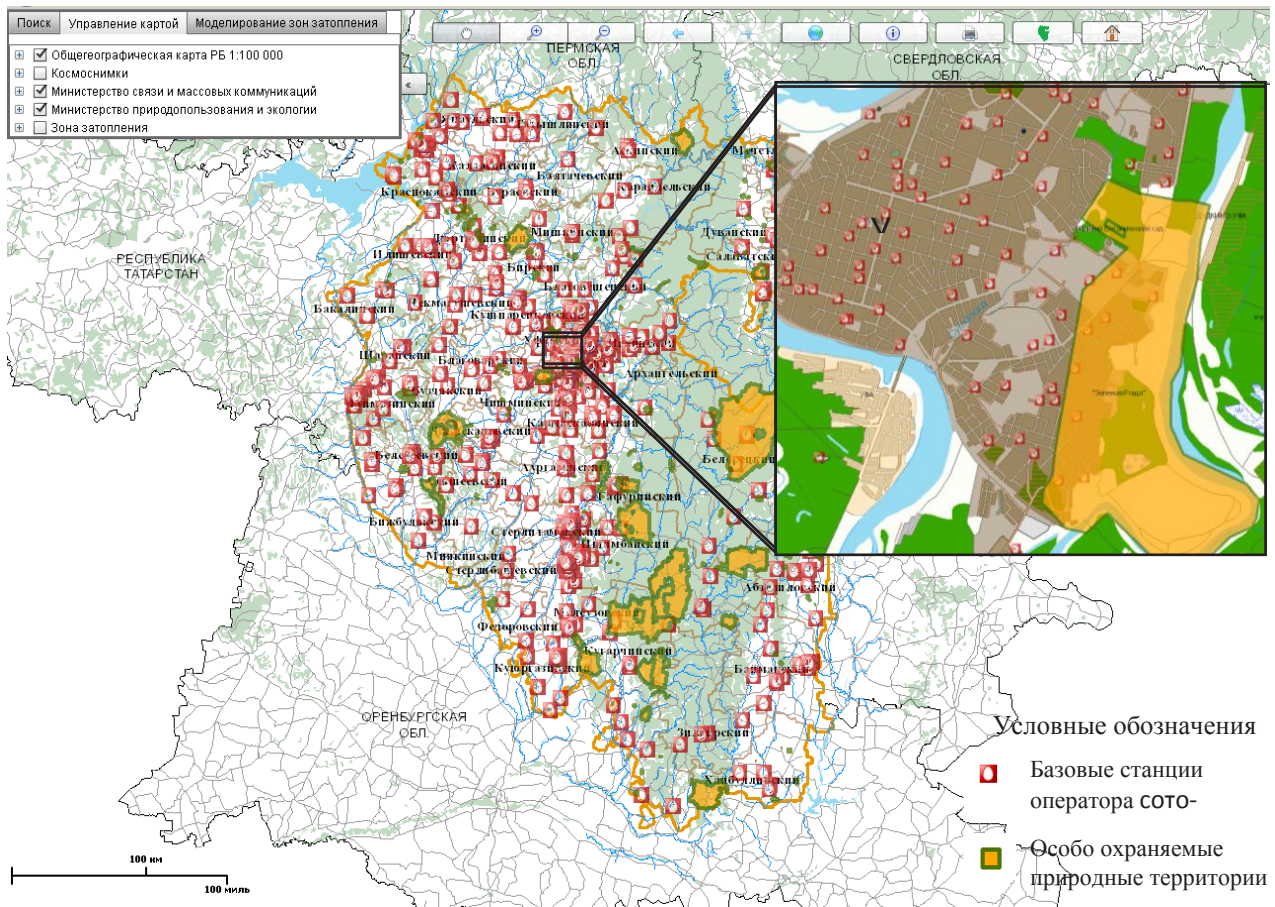


Рисунок 7 – Пример интеграции пространственных баз данных двух органов исполнительной власти

Список источников

- [1] Павлов, С.В. Интеграция пространственной информации в геоинформационной системе органов исполнительной власти на основе сервис-ориентированной архитектуры / С.В. Павлов, О.А. Ефремова, И.У. Ямапов // Вестник УГАТУ - 2013. Т. 17, №5(58). - С. 129-139.
- [2] Смирнов, А.В. Онтологии в системах искусственного интеллекта: способы построения и организации (часть 1) / А.В. Смирнов, М.П. Пашкин, Н.Г. Шилов, Т.В. Левашова // Новости искусственного интеллекта, № 1 (49) 2002. - С.3-13.
- [3] Кучуганов, М.В. Синтез схем баз данных на основе онтологии / М.В. Кучуганов // Онтология проектирования. - 2016. - Т. 6, №4(22). - С. 475-484. - DOI: 10.18287/2223-9537-2016-6-4-475-484.
- [4] Ефремова, О.А. Применение системного подхода к анализу проблемы использования пространственной информации для поддержки принятия решений региональными органами исполнительной власти. Электронный научный журнал «Инженерный вестник Дона». №2, 2014. <http://www.ivdon.ru/magazine>
- [5] Ефремова, О.А. Применение метода фасетной классификации для организации каталога метаданных в составе Геоинформационной системы органов исполнительной власти Республики Башкортостан / О.А. Ефремова, Р.А. Кравченко // Современные проблемы науки и образования. - 2014. - № 5; - <http://www.science-education.ru/119-14524>.
- [6] *Ontology Summit 2013 Communique. Towards Ontology Evaluation across the Life Cycle. Current Version is: v1.0.4 - 2013.05.31* http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2013_Communique. Коммюнике Онтологического саммита 2013. Оценка онтологий в течение всего жизненного цикла. Перевод с англ. Онтология проектирования. - 2013. - №2. - С. 66-74.
- [7] Виттих, В.А. Организация сложных систем. - Самара: Самарский научный центр РАН, 2010. - 66 с.
- [8] Лапшин, В. А. Онтологии в компьютерных системах. — М.: Научный мир, 2010.

- [9] **Боргест, Н.М.** Онтологии: современное состояние, краткий обзор / Н.М. Боргест, М.Д. Коровин. // Онтология проектирования. - №2(8), 2013. - С.49-55.
- [10] **Гвоздев, В.Е.** Предупреждение дефектов на ранних стадиях проектирования аппаратно-программных комплексов на основе положений теории интересубъективного управления / В.Е. Гвоздев, Д.В. Блинова, Л.Р. Черняховская // Онтология проектирования. - 2016. - Т.6, №4(22). - С.452-464.- DOI: 10.18287/2223-9537-2016-6-4-452-464.
- [11] **Ризванов, Д.А.** Интеллектуальная поддержка принятия решений при управлении ресурсами сложных систем на основе многоагентного подхода / Д.А. Ризванов, Н.И. Юсупова // Онтология проектирования. - 2015.- Т.5, №3(17). - С.297-312.- DOI: 10.18287/2223-9537-2015-5-3-297-312.
- [12] **Смирнов, С.В.** Онтологическое моделирование в ситуационном управлении / С.В. Смирнов // Онтология проектирования. - 2012.- №2. - С.16-24.

ONTOLOGICAL MODEL FOR INTEGRATION OF STRUCTURALLY HETEROGENEOUS SPATIAL DATABASES OF VARIOUS SUBJECT AREAS INTO A UNIFORM REGIONAL DATABASE

S.V. Pavlov¹, O. A. Efremova²

Ufa State Aviation Technical University, Ufa, Russia

¹psvgis@mail.ru, ²efremova-Oa@yandex.ru

Abstract

Article is devoted to a solution of the problem of integration of spatial databases into the uniform regional database for the organization of information support of decision-making for management of a large industrial region. The authors propose an ontological model for semantic-level integration that would allow description of structure of heterogeneous spatial databases within a single methodological framework highlighting in their composition the ontologies of spatial, attributive and metadata, and thereby ensuring the possibility of their further integration into a single regional database. Within the framework of the proposed ontological model, an algorithm for constructing and integrating ontologies into the software for integrating spatial and spatial databases that are heterogeneous in structure and subject has been developed. Approbation of the developed algorithm is carried out on the example of creating a single regional database of spatial data in the Geoinformation System of the Republic of Bashkortostan's administration.

Key words: ontology, spatial data, databases, integration, industrial region, decision-making.

Citation: Pavlov SV, Efremova OA. Ontological model for integration of structurally heterogeneous spatial databases of various subject areas into a uniform regional database. *Ontology of designing*. 2017; 7(3): 323-333. DOI: 10.18287/2223-9537-2017-7-3-323-333.

References

- [1] **Pavlov SV., Efremova OA., Yamalov IU.** Integration of spatial information in a geographic information system of executive agencies on a basis service - the oriented architecture [In Russian]. UGATU bulletin: scientific log of the Ufa State Aviation Technical University. 2013. V. 17, No.5(58): 129-139.
- [2] **Smirnov AV., Pashkin MP., Shilov NG., Levashova TV.** Ontologies in systems of an artificial intelligence: methods of creation and organization (part 1) [In Russian]. News of an artificial intelligence. 2002; No.1(49): 3-13.
- [3] **Kuchuganov MV.** Synthesis database schemas based on ontology [In Russian]. *Ontology of designing*. 2016; 6(4): 475-484. – DOI: 10.18287/2223-9537-2016-6-4-475-484.
- [4] **Efremova OA.** The Application of a systematic approach to the analysis of the problem of use of spatial information to support decision-making by executive agencies [In Russian]. The electronic scientific journal "Engineering journal of Don". No. 2, 2014. - <http://www.ivdon.EN/ru/magazine>.
- [5] **Efremova OA., Kravchenko RA.** The application of the method of faceted classification for the organization of the metadata catalog in the composition of the geographic information system of the executive agencies of the Republic of Bashkortostan [In Russian]. The Modern problems of science and education. 2014; No.5; - <http://www.science-education.ru/119-14524>.
- [6] **Ontology Summit 2013 Communique. Towards Ontology Evaluation across the Life Cycle.** - http://ontology.cim3.net/cgi-bin/wiki.pl?OntologySummit2013_Communique.

- [7] *Vitikh VA.* Organization of complex-systems [In Russian]. - Samara: Scientific Center of RAS, 2010. – 66 p.
- [8] *Lapshin VA.* Ontologies in computer systems [In Russian]. - M.: Scientific world, 2010.
- [9] *Borgest NM., Korovin MD.* Ontologies: current status, an overview [In Russian]. *Ontology of designing.* 2013; No.2(8): 49-55.
- [10] *Gvozdev VE., Blinova DV., Chernyakhovsk LR.* Hardware-software complexes on the basis of the positions of the theory of intersubjective management [In Russian]. *Ontology of Designing.* 2016; 6(4): 452-464. - DOI: 10.18287/2223-9537-2016-6-4-452-464.
- [11] *Rizvanov DA., Yusupova NI.* Intelligent decision support for resource management of complex systems based on multi-agent approach [In Russian]. *Ontology of designing.* 2015; 5(3): 297-312. - DOI: 10.18287/2223-9537-2015-5-3-297-312.
- [12] *Smirnov SV.* Ontological modeling in situational management [In Russian]. *Ontology of designing.* 2012; No.2(4): 16-24.

Сведения об авторах



Павлов Сергей Владимирович, 1955 г.р. Окончил Башкирский государственный университет по специальности вычислительная математика (1977). Заведующий кафедрой геоинформационных систем Уфимского государственного авиационного технического университета. Доктор технических наук, профессор. Заслуженный деятель науки Республики Башкортостан. Проводит исследования в области системного анализа, информационных систем, разработки многомерных моделей данных, интеграции разрозненных баз данных, геоинформационных систем.

Sergey Vladimirovich Pavlov (b. 1955) graduated from the Bashkir state University majoring in Computational Mathematics (1977). Head of Department of geoinformation systems of Ufa State Aviation Technical University. Doctor of Technical Sciences, professor. Conducts research in the areas of systems analysis, information systems design, multidimensional data

models, integrating disparate databases, geographic information systems.



Ефремова Оксана Александровна, 1977 г.р. Окончила Уфимский государственный авиационный технический университет (УГАТУ) по специальности автоматизированные системы обработки информации и управления (1999). Заместитель зав. кафедрой геоинформационных систем по научной и инновационной деятельности УГАТУ. Доцент, кандидат технических наук. Область научных интересов: системный анализ, разработка информационных систем в сфере организации общественной безопасности и защиты в чрезвычайных ситуациях, разработка многомерных моделей данных, геоинформационные системы.

Oksana Alexandrovna Efremova (b. 1977) graduated from the Ufa state aviation technical University on a speciality "the engineer of automated systems of information processing and management" (1999). At the present time — Deputy head the Department of geo-information

systems for research and innovation activities USATU. Associate Professor, candidate of technical Sciences. Research interests: systems analysis, information systems development in the sphere of public security and protection in emergency situations, the development of multidimensional data models, geographic information systems.