

ПРИМЕНЕНИЕ МЕТОДА ФАСЕТНОЙ КЛАССИФИКАЦИИ ДЛЯ ОРГАНИЗАЦИИ КАТАЛОГА МЕТАДАННЫХ В СОСТАВЕ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ ОРГАНОВ ИСПОЛНИТЕЛЬНОЙ ВЛАСТИ РЕСПУБЛИКИ БАШКОРТОСТАН

Ефремова О.А., Кравченко Р.А.

ФГБОУ ВПО «Уфимский государственный авиационный технический университет», Уфа, Россия (450000, Уфа, ул. Карла Маркса, 12), e-mail: efremova-oa@yandex.ru, rom.kr@mail.ru

В статье проведен анализ использования различных подходов к организации каталога метаданных в составе Геоинформационной системы органов исполнительной власти Республики Башкортостан. Предложена структура каталога метаданных на базе фасетного метода классификации. В качестве примера рассмотрена задача поиска метаданных в составе каталога метаданных Геоинформационной системы органов исполнительной власти Республики Башкортостан с использованием метода фасетной классификации. Работу дерева каталога метаданных предлагается реализовать, используя принципы навигации в форме фасетного запроса объектов. На основе предложенного подхода разработана блок-схема и осуществлена программная реализация алгоритма поиска пространственных данных в составе каталога метаданных с использованием фасетной формулы.

Ключевые слова: метаданные, геоинформационная система, классификация, фасетный метод, иерархический метод, органы исполнительной власти, каталог метаданных.

APPLICATION OF FACETED CLASSIFICATION FOR ORGANIZING METADATA CATALOG AS A PART OF THE GEOGRAPHIC INFORMATION SYSTEM OF THE EXECUTIVE AGENCIES OF THE REPUBLIC OF BASHKORTOSTAN

Efremova O.A., Kravchenko R.A.

Ufa State Aviation Technical University, K. Marx Street 12, Ufa, The Republic of Bashkortostan, Russian Federation, 450000, e-mail: efremova-oa@yandex.ru, rom.kr@mail.ru

The article analyzes the use of different approaches to construction metadata catalog as part of the Geographic information System of the executive agencies of the Republic of Bashkortostan. The structure of the metadata catalog based on faceted classification method. As an example, consider the problem of searching for metadata as part of the directory using faceted classification and developed a flowchart search using faceted formula. Directory metadata is proposed to implement using the principles of navigation in the form of facet query objects. On based this approach was developed by a block diagram and software implementation of a search algorithm for spatial data as part of the metadata catalog using faceted formula.

Key words: metadata, geographic information system, classifications, faceted method, hierarchical method, executive agencies, metadata catalog.

В соответствии с Концепцией программы «Развитие информационного общества в Республике Башкортостан на 2012–2017 годы», одобренной распоряжением Правительства Республики Башкортостан от 1 августа 2011 года № 962-р, а также принятой Правительством Республики Башкортостан Концепцией Геоинформационной системы органов исполнительной власти Республики Башкортостан, в республике поэтапно создается Геоинформационная система органов исполнительной власти – (ГИС ОИВ РБ).

К основным функциям ГИС ОИВ РБ относятся:

- 1) централизованное хранение и управление всеми базовыми пространственными информационными ресурсами;
- 2) совместное использование пространственных информационных ресурсов;

3) организация распределенной системы сбора, хранения и обработки специализированной пространственной информации, в том числе с использованием облачных моделей вычислений;

4) информационное обеспечение формирования и принятия управленческих решений, а именно:

- поиск и предоставление пользователям всей имеющейся информации в виде, удобном для анализа и использования;

- тематический анализ данных по направлениям деятельности органов исполнительной власти Республики Башкортостан;

- моделирование и прогнозирование процессов и явлений, происходящих на территории республики;

5) обеспечение на основе тематических пространственных информационных ресурсов группы государственных услуг, предоставляемых в рамках электронного Правительства Республики Башкортостан органам местного самоуправления Республики Башкортостан, организациям и гражданам;

б) обеспечение целостности, непротиворечивости, защищенности и безопасности пространственных информационных ресурсов.

Создание такого рода масштабной геоинформационной системы подразумевает обработку больших массивов пространственной информации из различных неоднородных источников.

Основываясь на вышесказанном, для обеспечения выполнения такого рода функций, как централизованное хранение, сбор, обработка и поиск разрозненной и неоднородной пространственной информации в составе Геоинформационной системы органов исполнительной власти, необходимо решить проблему каталогизации данных, для решения которой предлагается создать каталог метаданных и организовать последовательное подключение к нему источников пространственных данных всех органов исполнительной власти республики [1].

Основные подходы к организации каталога метаданных

Как и любое автоматизированное хранилище информации, каталог метаданных требует разработки подхода к организации, хранению, обработке и поиску данных. Как правило, для использования автоматизированных средств используются методы классификации, представляющие собой систему распределения объектов по классам в соответствии с перечнем наименований и кодов классификационных группировок. Для создания подобных хранилищ наиболее популярны иерархический и фасетный методы классификации [2,6].

Иерархический метод классификации заключается в последовательном разделении множества объектов на подчиненные классификационные группировки. Деление

производится по определенному выбранному признаку на крупные группировки, затем каждая из них делится по другому признаку на ряд последующих группировок, детализируя объект классификации, и в итоге между группировками устанавливается четкая иерархия классов.

В иерархической системе классификации каждый объект на любом уровне должен быть отнесен к одному классу, который характеризуется конкретным значением выбранного классификационного признака. Для последующей группировки в каждом новом классе необходимо задать свои классификационные признаки и их значения, при этом выбор признаков будет зависеть от семантического содержания того класса, для которого необходима группировка на последующем уровне иерархии [2]. Таким образом, основными преимуществами иерархической системы классификации являются:

- 1) четкость построения связей между уровнями иерархии;
- 2) легкость навигации по каталогу;
- 3) простота отображения каталога в виде дерева;
- 4) четкое отслеживание расположения метаданных на каждом узле.

Применительно к Геоинформационной системе органов исполнительной власти фрагмент структуры каталога метаданных с использованием иерархической системы классификации представлен на рисунке 1.

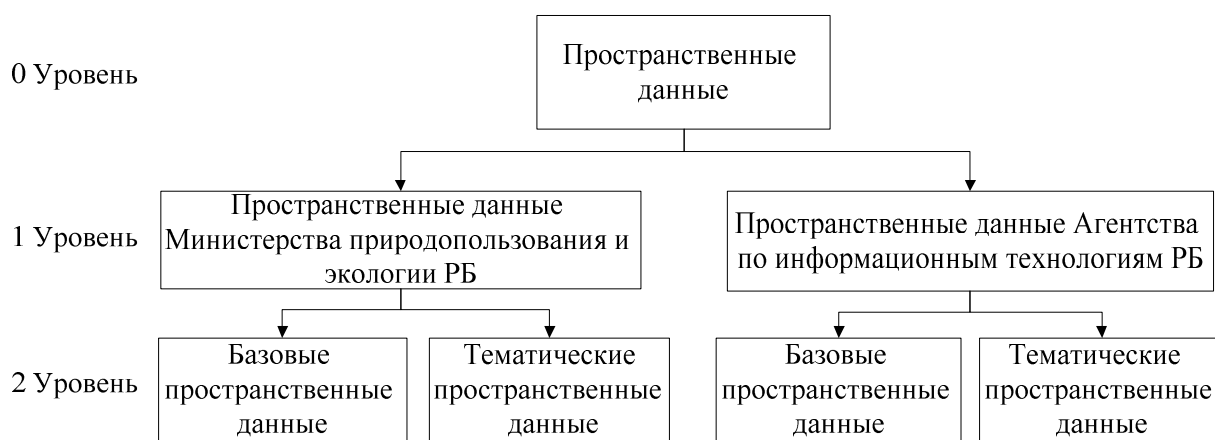


Рис. 1. Фрагмент структуры каталога метаданных с использованием иерархической системы классификации

Однако в ходе функционирования каталога метаданных с использованием иерархического метода выявился ряд проблем при работе с метаданными. Первая проблема – это сложность доработки классификационной схемы из-за жесткости структуры каждого уровня иерархии, т.е. необходимо заранее планировать, какие пространственные данные, и каких органов исполнительной власти будут включены в состав ГИС ОИВ РБ.

Так как порядок уровней иерархии задается жестко при создании каталога метаданных, возникает вторая проблема – изменение его структуры в случае необходимости. В-третьих, наблюдается сложность добавления объектов с новым набором характеристик в каталог

метаданных. Все это, в конечном итоге, приводит к тому, что даже небольшие изменения классификационных задач требуют переработки всей классификационной схемы [3].

Рядом преимуществ, по сравнению с рассмотренным выше иерархическим методом классификации, обладает фасетный метод классификации, который, в свою очередь, основан на разделении всего множества объектов на независимые подмножества по различным признакам классификации.

Фасет ϕ – это набор значений отдельного признака классификации, все признаки взаимно независимы. Данный метод предполагает, что исходное множество объектов A разбивается на подмножества группировок по независимым между собой признакам классификации – фасетам (1), где n – количество фасетов.

$$A = \{\phi_1, \phi_2, \dots, \phi_n\} \quad (1)$$

Каждый объект $Obj \in A$ одновременно имеет классификационные признаки (K) из различных фасетов, а классификационные группировки создаются путем задания фасетной формулы (2) из последовательности значений (X) классификационных признаков выбранных фасетов.

$$K = \{X_1, X_2, \dots, X_n\} \quad (2)$$

В свою очередь, каждый фасет может быть разделен на ряд дополнительных фасетов (признаков). Емкость фасетной классификации при числе фасетов N и для фасета имеющего K_n признаков равна величине H :

$$H = \sum_{i=1}^N C_N^i * \prod_{n=1}^i K_n \quad (3)$$

Использование данного подхода при разработке структуры каталога метаданных Геоинформационной системы органов исполнительной власти Республики Башкортостан (рис. 2) позволяет обеспечить:

- 1) возможность расширения множества классифицируемых объектов и внесение необходимых изменений в структуры классификации;
- 2) возможность сопряжения с другими классификациями объектов;
- 3) возможность решения комплексных задач различного уровня;
- 4) хорошую приспособленность каталога к меняющимся характеристикам задач.

Исходя из вышесказанного, разработку структуры каталога метаданных целесообразно осуществлять с использованием фасетного метода классификации.

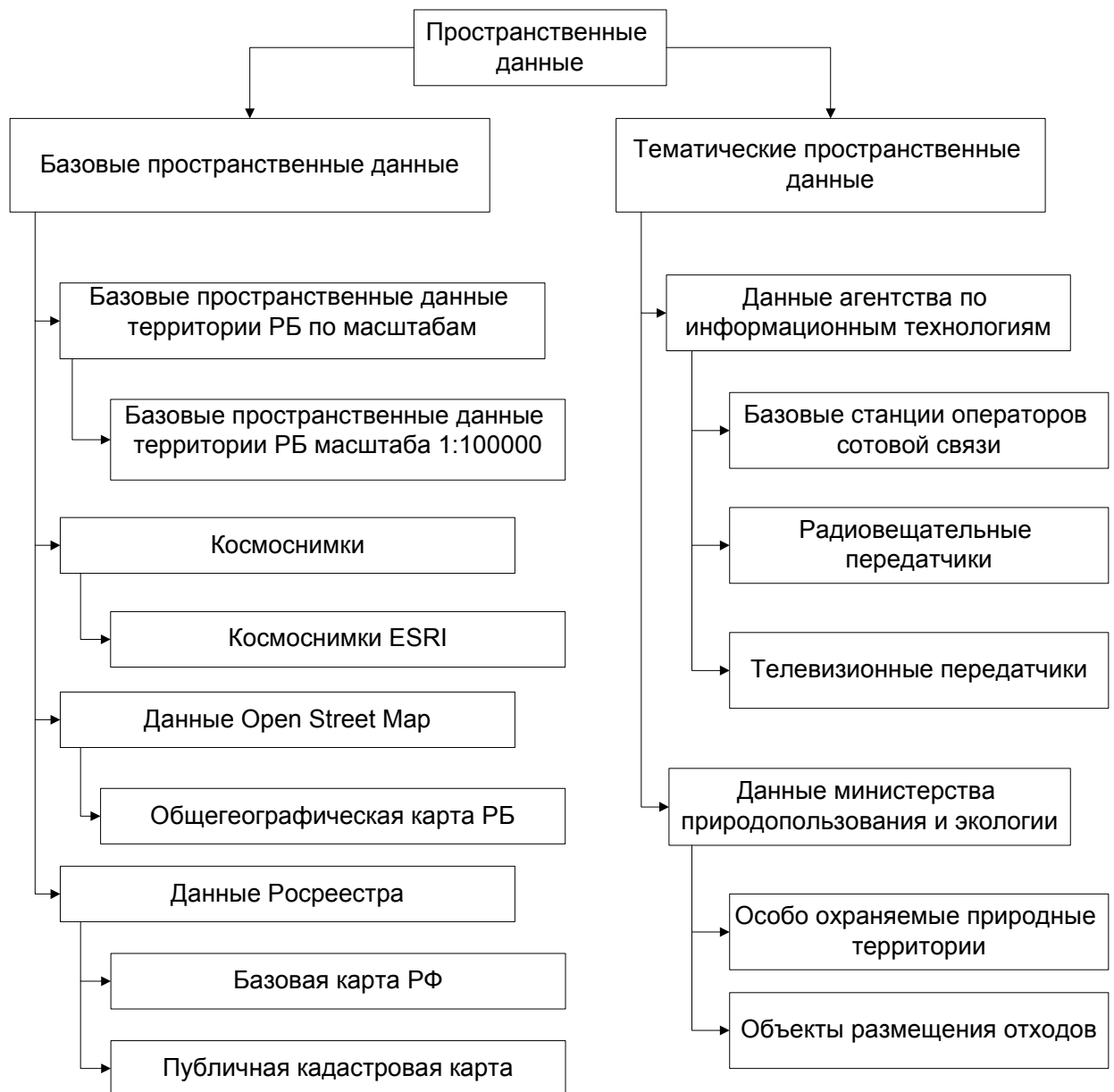


Рис. 2. Фрагмент структуры каталога метаданных с использованием фасетной системы классификации

Пример решения задачи поиска метаданных с использованием метода фасетной классификации

Работу дерева каталога метаданных предлагается реализовать, используя принципы навигации в форме фасетного запроса объектов, представленного в виде (4). Фасетный признак – список значений фасетного признака для данного объекта, список объектов, которые содержат данные значения фасетного признака.

$$K = \{ \phi_i | X_1 \dots X_n \} \quad (4)$$

Данный подход сводит задачу поиска к тому, что каждый объект по каждому фасетному признаку характеризуется массивом значений фасетного признака, характеризующих объект. Набор фасетных признаков диктуется направлением объекта, также возможен анализ

сочетаний фасетных признаков и их заданных значений, что позволяет в процессе работы системы выстраивать логические связи между метаданными, которые можно модифицировать в зависимости от их актуального состояния [4].

Метаданные каталога удобно сгруппировать по следующему фасетному признаку: имеется специально выделенный фасет базовых пространственных данных и фасет тематических пространственных данных. При поиске базовой информации пользователь опирается на её удобство использования в группе с другими данными, а именно такие показатели, как масштаб отображения, тип данных (космоснимки, кадастровые карты, границы, города и т.д.). При поиске тематических данных пользователь опирается на предназначение данных, поэтому информацию целесообразно группировать указанным на рисунке 2 образом.

Общий обзор всех фасетных признаков и полного списка их значений обеспечивает пользователю представление о направлениях каталога метаданных и как итог, все поисковые задачи могут быть сведены к построению запроса на фасетную выборку объектов.

Формирования запроса с целью получения выборки объектов будет выглядеть как задание классификационной группировки путем фасетной формулы вида (5), где Z – конкретное значение классификационного признака, которое представляет собой множество объектов фасетов, удовлетворяющих значениям признака.

$$K = \{\phi_1 | Z_1 \in X_1, \phi_2 | Z_2 \in X_2, \dots, \phi_n | Z_n \in X_n\} \quad (5)$$

Блок-схема предложенного в соответствии с (5) модуля формирования поискового запроса с использованием фасетной формулы представлена на рисунке 3.

Задача пользовательского интерфейса, реализующего предложенный подход, сводится к просмотру всех основных фасетов, выбору значений фасета, формированию выборки объектов путем фасетной формулы и достройки фасетной формулы до получения удовлетворительного результата [6]. Пример интерфейса пользователя каталога метаданных, реализованного по вышеописанному алгоритму, представлена на рисунке 4.

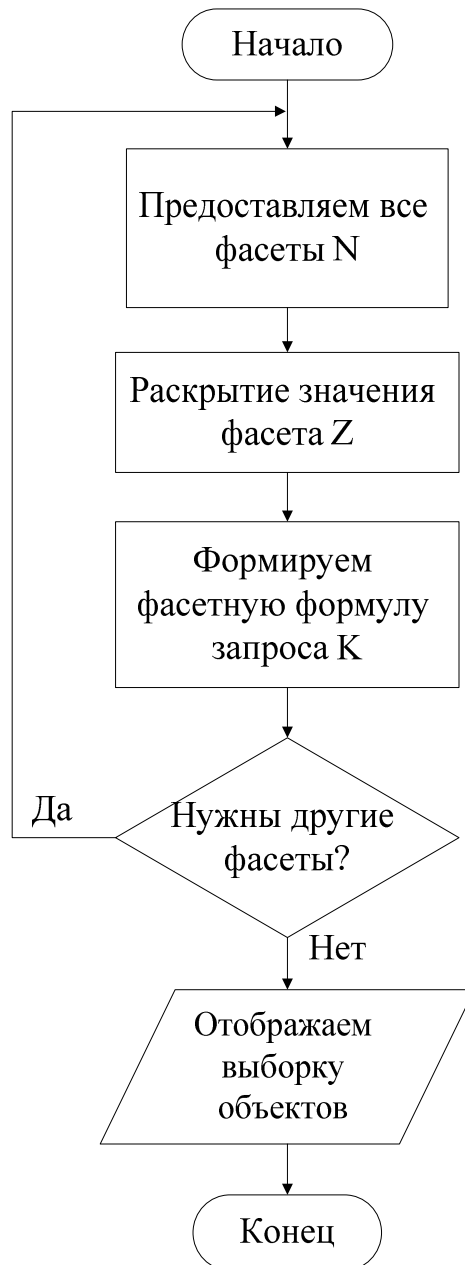


Рис. 3. Блок-схема алгоритма поиска пространственных данных в составе каталога метаданных с использованием фасытной формулы

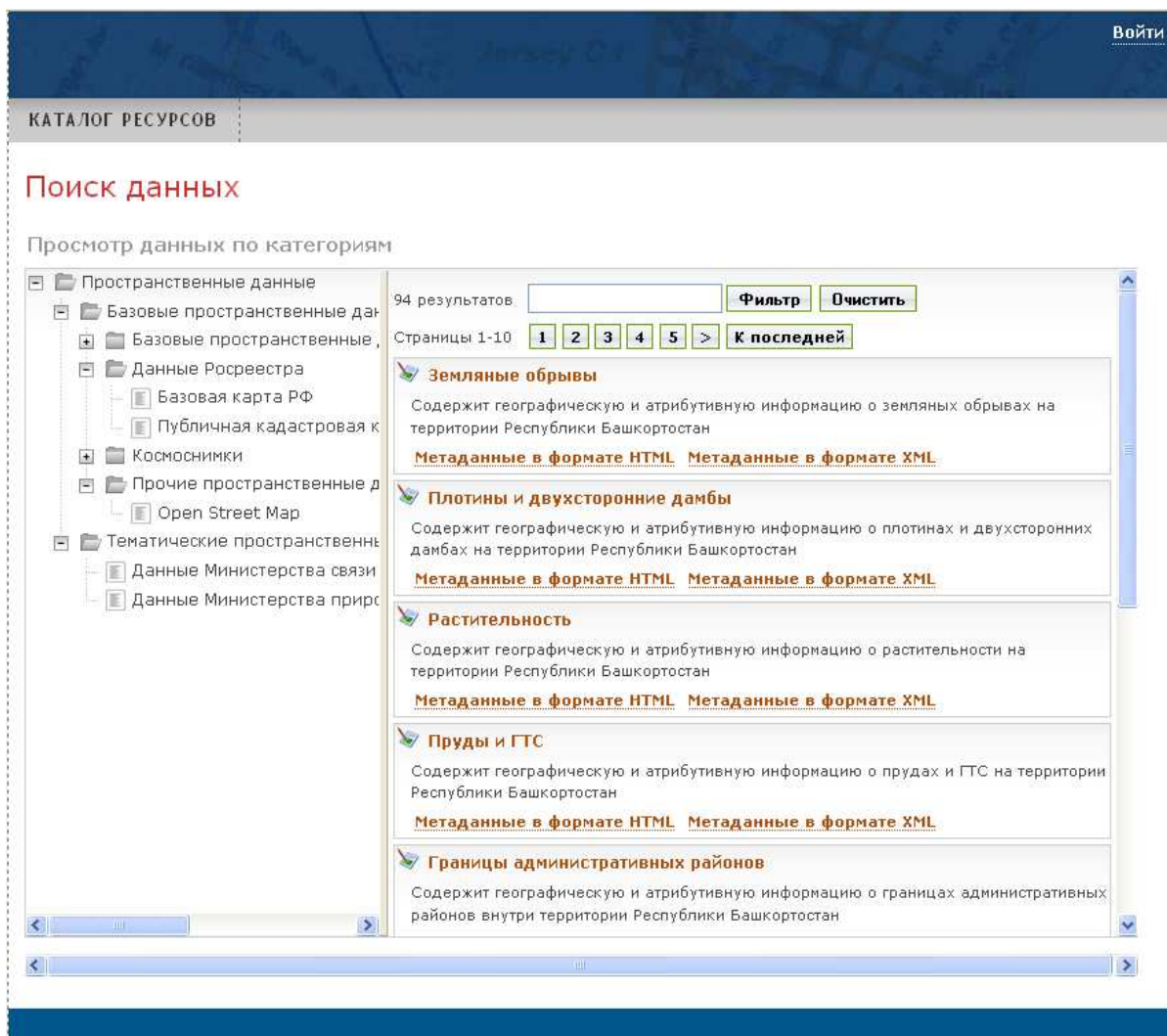


Рис. 4. Интерфейс пользователя каталога метаданных

Заключение

В рамках разработки каталога метаданных Геоинформационной системы органов исполнительной власти Республики Башкортостан проведен анализ подходов к организации структуры каталога, выявлены основные проблемы функционирования при использовании иерархической классификации, разработана структура каталога и алгоритм поиска с использованием фасетного метода классификации, а также программная реализация каталога метаданных с использованием предложенного подхода.

Список литературы

1. Ефремова О.А., Бахтизин Р.Н., Павлов С.В., Павлов А.С. Архитектура Геоинформационной системы Республики Башкортостан: межвузовский научный сборник: Геоинформационные технологии в проектировании и создании корпоративных информационных систем. – Уфа: УГАТУ, 2011.

2. Кузьминский А.Н. Хозяйственный учет на базе микроЭВМ / А.Н. Кузьминский, Ж.Б. Бонев, В.И. Смолянинов. – М.: Финансы и статистика, 1989. – 255 с.
3. Павлов С.В., Ефремова О.А., Павлов А.С., Информационная вычислительная система для обеспечения органов исполнительной власти региона пространственными данными // Научный журнал «Электротехнические и информационные комплексы и системы». – Уфа, 2013. – № 2, Т. 9. – С. 98-108.
4. Павлов С.В., Ефремова О.А., Ямалов И.У. Интеграция пространственной информации в Геоинформационной системе органов исполнительной власти на основе сервис-ориентированной архитектуры // Вестник УГАТУ: научный журнал Уфимского государственного авиационного технического университета. – Уфа, 2014. – Т.17, № 5. – С.129-139.
5. Павлов С.В., Ефремова О.А., Соколова А.В. Формализованное описание пространственной информации в составе трехмерных моделей потенциально опасных объектов на основе теоретико-множественного подхода // Научный журнал «Электротехнические и информационные комплексы и системы». – Уфа, 2014. – Т.10, № 1. – С. 66-72.
6. Ранганатан Ш. Р. Классификация двоеточием. Основная классификация: пер.с англ. / под. ред. Т.С. Гомолицкой. – ГПНТБ СССР, 1970. – 422 с.

Рецензенты:

Гвоздев В.Е., д.т.н., профессор, зав. каф. АПрИС, ФГБОУ ВПО УГАТУ, г. Уфа;

Павлов С.В., д.т.н., профессор каф. ГИС, ФГБОУ ВПО УГАТУ, г. Уфа.