



НАУЧНАЯ СТАТЬЯ

УДК 004.9

DOI: <https://doi.org/10.30932/1992-3252-2024-22-2-1>

О формировании геопространственных знаний на основе преобразования геоописаний



Игорь РОЗЕНБЕРГ



Сергей ДУЛИН

Игорь Наумович Розенберг¹,
Сергей Константинович Дулин²

¹ Российский университет транспорта, Москва, Россия.

² Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия.

² Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), Москва, Россия.

¹ ORCID 0000-0001-9589-6783; Scopus Author ID: 15136568900; РИНЦ SPIN-код: 2958-4186; РИНЦ Author ID: 652172.

² РИНЦ SPIN-код: 7190-4891; РИНЦ Author ID: 3320.

✉ ² skdulin@mail.ru.

АННОТАЦИЯ

В рамках проблемы формализации и представления геоинформационных структур в информационной среде ГИС представлен подход, позволяющий упростить методы работы с атрибутивными характеристиками геоанных, основанный на использовании возможностей функционального программирования. В работе рассматриваются вопросы, связанные с формированием, сопоставлением и преобразованием геоописаний для актуализации геоинформации в информаци-

онной среде ГИС. Для осуществления преобразования геоописаний предложено упрощение формирования инструмента для представления геоинформации и набора геотекстов. Описанный в работе подход основывается на выборе парадигмы функционального программирования, реализованной в виде аналитических машин, что позволяет осуществлять преобразование геоописаний на различных геоязыках и поддерживать репозиторий геоанных в актуальном состоянии.

Ключевые слова: ГИС; формирование, сопоставление и преобразование геоописаний; средства функционального программирования.

Для цитирования: Розенберг И. Н., Дулин С. К. О формировании геопространственных знаний на основе преобразования геоописаний // Мир транспорта. 2024. Т. 22. № 2 (111). С. 6–11. DOI: <https://doi.org/10.30932/1992-3252-2024-22-2-1>.

Полный текст статьи в переводе на английский язык публикуется во второй части данного выпуска.
English translation of the full text of the article is published in the second part of the issue.

ВВЕДЕНИЕ

Обработка геоинформации с подключением знаковой системы обозначений вызывает ограничения при реализации семантической обработки георесурсов, которые обусловлены слабой стандартизацией атрибутивной и визуальной геоинформации и некачественными структурами ее описания, которые необходимы для создания взаимосвязанной совокупности георесурсов. Описания геоданных ввиду их гетерогенности невозможно представить исключительно в лингвистической форме, поэтому может оказаться продуктивной идея разработки подхода на основе функционального программирования, что позволило бы поставить в соответствие геоданным инкапсулированные функции их обработки, в какой-то степени отображающие их семантику [1; 2]. Предлагаемый подход позволит расширить систему используемых геоописаний различной геоязыковой принадлежности, включая представление визуальных геознаний.

РЕЗУЛЬТАТЫ

Два научных направления развития картографических систем

Синтез картографии и семиотики обеспечил зарождение новой научной дисциплины – картосемиотики, которая включала в себя анализ, интерпретацию и преобразование знаковых картографических систем.

Картосемиотика использует лингвистические описания в геоинформатике и дает возможность представлять атрибутивные характеристики карт, такие как условные обозначения и легенды, с помощью специальных языковых конструкций карт с модельной и содержательной составляющими.

С течением времени в картосемиотике были зафиксированы пять разделов, так или иначе связанных с созданием геоописаний: картосемиотический, картолингвистический, формально-логический, кибернетический и субязыковый. Тем не менее, формально-логический и субязыковый разделы не претерпели существенного развития и носят маргинальный характер.

Картосемиотика, будучи, по сути, прикладной семиотикой, включает в себя аналогичные семиотике три составные части: 1) картографический синтаксис знаковых совокупностей; 2) картографическая семантика, связанная с анализом смысла используемых знаков, корреспондирующих с описываемыми георесур-

сами; 3) картографическая прагматика, которая позволяет измерить утилитарность используемых знаков в зависимости от степени осведомленности и профессиональной подготовки потребителей георесурсов. Центральным звеном картосемиотики является разработка языка карты – корректной и непротиворечивой знаковой системы, предназначенной для отображения картографической информации. Качество языка карты характеризует уровень развития средств представления геознаний и возможности визуализации картографической информации. Наиболее значимый вклад в становление языка карты внес А. А. Лютый [3]. Созданный им язык карты включает два подязыка: один позволяет представить состав геообъектов, а другой представляет геометрию и топологию геообъектов. Повышение требований к организации георесурсов привело к бурному росту предложений и дополнений в области разработки геостандартов, позволивших более полно и точно формировать представления геознаний на основе вводимых геоописаний.

А. А. Лютый рассматривал три формы представления геознаний: «невербальные знания, вербальные знания и знания, представимые и в вербальной, и в невербальной формах» [3]. Из этого следует, что карты, аэрофотоснимки, литологические и стратиграфические колонки, геодиаграммы, палеотектонические схемы, геохимические диаграммы и схемы корреляции представляют собой не что иное, как разнообразные геотексты.

В качестве одного из наиболее значимых вкладов А. А. Лютого в геоинформатику можно указать разработку семиотики геоинформации, названной им «геосемиотикой» [3]. В геосемиотике географические концепции подразделяются на пять категорий (рис. 1).

А. А. Лютый изучал проблему соотношения образного и вербального в рамках задачи геосемантического описания геопространства для

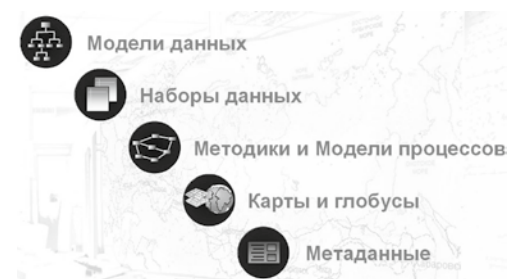


Рис. 1. Категории географических концепций.





Рис. 2. Задачи сопоставления геоописаний.

разработки электронного образа Земли. Его исследования о расположении и взаимовлиянии элементов геопространства позволили решать задачи геоинформатики на основе геоописаний.

Задачи сопоставления геоописаний

Гетерогенность геоданных обуславливает большие затраты на создание, хранение, обработку и поддержку геоописаний. Дополнительную нагрузку вызывает их интероперабельность из-за не коррелируемых форматов и различия поставщиков геоданных и программных сервисов. Задача сопоставления геоописаний для их совместного согласованного использования является важным мероприятием, требующим программных, технологических и лингвистических решений [4; 5]. Для совместного использования геоданных на основе сопоставления геоописаний нужен семантический уровень интероперабельности, но автономность разрабатываемых программных систем приводит к локализации специфики геоописаний. Попытки интеграции разнородных геоописаний с не коррелируемыми форматами приводят к затратам на согласование противоречивых моделей геоданных, превышающим собственно разработку геоописаний.

Проектируемые базы геоданных чаще всего имеют персонифицированный или ведомственный характер, что требует дополнительных мероприятий для их совместного использования. Если при этом базы геоданных имеют тенденцию к перманентному пополнению, то возникают дополнительные требования к масштабируемости ГИС и разработке моделей сопоставления геоописаний, обеспечивающих непротиворечивость и целостность геоданных.

Сопоставление геоописаний включает процедуры обработки ссылок и установления связей между разными уровнями детализации

для реализации интероперабельности между сопоставляемыми геообъектами [6]. Интероперабельность необходима для функциональной совместимости и корректного сопоставления геоданных разной детализации.

Дополнительные сложности вызывает наличие мультилингвальности геоописаний, которая является не таким уж редким явлением. Совокупное геоописание строится на основе синтеза геоописаний, представленных на разных естественных языках. Задачи сопоставления геоописаний на разных языках представлены на рис. 2.

Успешное сопоставление геоописаний обеспечивает:

- улучшение информативности;
- повышение семантики геоописания;
- выявление качества геоописаний;
- повышение интероперабельности экспертов-аналитиков;
- апробацию методов сопоставления геообъектов;
- разработку методов преобразования геоописаний.

Формирование геоописаний

Содержимое геоинформационного портала не исчерпывается визуальной картографической информацией. В базе геоданных содержится и обрабатывается большое количество атрибутивных (семантических) данных в виде разного характера текстов и документов. Геоинформационный портал, как отмечал А. А. Лютый, представляет собой «электронную библиотеку с геотекстами на разных естественных языках». Гетерогенность геоданных приводит к семиотической неоднородности геотекстов, хранящихся на геоинформационном портале. С позиции запросно-ответного режима базы данных обработка существующих геотекстов должна обеспечивать получение запрашиваемой информации. На рис. 3 изображены блоки, обеспечивающие анализ и обработку геоданных, представленных геоописаниями в виде геотекстов.

Геоинформационное пространство для создания и сопровождения геоописаний представляет собой синергию информационного наполнения ГИС, среды хранения репозитория и программ, ориентированных на манипулирование необходимыми геоданными. Следует отметить специфику репозитория геоданных, который выполняет роль коорди-

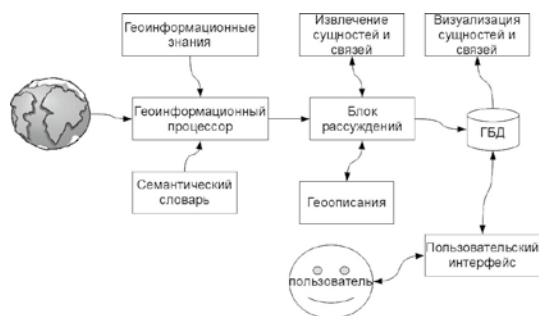


Рис. 3. Анализ и обработка геоописаний в виде геотекстов.

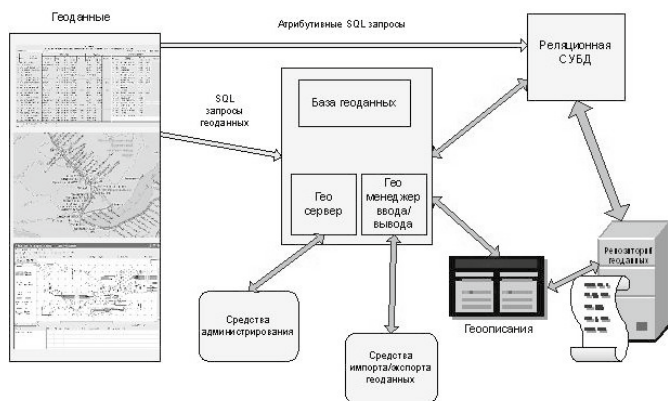


Рис. 4. Геоинформационное взаимодействие геоописаний.

натора реализации сопоставления геоописаний с другими геоописаниями, обеспечивая релевантную интероперабельность геоописаний (рис. 4).

Преобразование геоописаний

Необходимость преобразования геоописаний возникает, если требуется проанализировать и сопоставить мультилингвальные геоописания, синтезируя взаимодополненные релевантные геоописания. Возможно, такое преобразование имеет наибольшее значение для формирования репозитория геоданных ввиду меньшего количества ограничений на хранимые структуры по сравнению с базой геоданных.

Лингвистический характер задачи преобразования геоописаний делает ее похожей на задачу перевода естественно-языковых текстов с помощью специальных программ, осуществляющих межязыковое преобразование [7]. Если выбрать сильно формализованный геоязык, на котором выполнено геоописание, то можно существующие методы межязыкового преобразования использовать при преобразовании геоописаний.

Ключевой идеей предлагаемого подхода к анализу геоязыковых структур является

простое предложение: представлять грамматические правила в виде функций в математическом смысле, то есть в виде отображения вида $A \rightarrow B$, где A и B – некоторые конечные множества. Для каждого из геоязыков существует соответствующая аналитическая машина, благодаря которой грамматический разбор можно выполнить в автоматическом режиме, то есть посредством вычислительных процедур. Использование аналитической машины, основанной на парадигме функционального подхода, предполагает, что будет проводиться анализ нормализованных текстов на естественном языке. Нормализованный текст – это текст, обработанный синтаксическим и/или морфологическим анализатором. В качестве еще одного ограничения следует отметить необходимость задания предметной области анализируемого геотекста.

Подход к описанию правил грамматики, как к заданию математической функции (с ее областью определения и областью значения) дает возможность при программной реализации аналитической машины применить парадигму функционального программирования, в основе которого лежат чистые функции и неизменяемые структуры данных. Следует также отметить, что использование среды



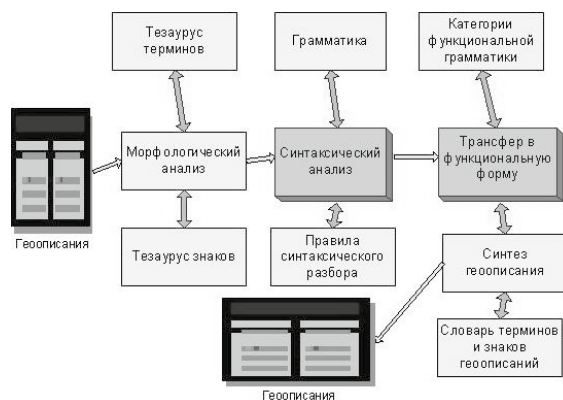


Рис. 5. Блоки преобразования геоописаний.

функционального программирования как аналитической машины дает возможность проектировщику использовать кортежи и «функции-защитники».

«Атомом» назовем любой однозначно определенный идентификатор. Атомы предназначены для задания атрибутивных характеристик атрибутов геоязыка. Таким образом, задаются все возможные атрибутивные характеристики геоязыковых структур, и формализм функционального программирования оказывается достаточным формализмом для анализа геоязыковых структур, так что нет необходимости использовать какие-либо другие математические средства.

Использование в качестве инструмента аналитической машины на основе языка функционального программирования позволит спроектировать аналитическую машину, реализующую формальные правила преобразования геоописаний в виде математических функций.

Основываясь на грамматике функционально заданных геотекстов и выбрав среду функционального программирования [8; 9], можно получить аналитическую машину для осуществления анализа геотекстов, релевантных выбранной грамматике. Кроме того, выбор среды функционального программирования открывает разработчикам доступ к управлению на уровне таких структур, как атомы языка и кортежи, при обработке атрибутивных характеристик геоописаний.

Определение свойств атрибутов приобретает здесь другое значение. При переходе на уровень кортежей и атомов в функциональном языке работа с характеристиками атрибутов уже не является прерогативой базы геоданных, эту нагрузку берет на себя язык функционального программирования. Соот-

ветственно, поддержка целостности и обработки характеристик атрибутов осуществляется с помощью функционального языка, а не геоинформационной системы. Это обстоятельство сильно упрощает создание аналитической машины для преобразования геоописаний, ввиду того, что работа с характеристиками атрибутов происходит вне геоинформационной модели.

Основным преимуществом использования функционально-вычислимых грамматик [9] при работе с атрибутивными характеристиками геоописаний является то, что правила грамматики уже являются элементами функционального языка программирования, что упрощает решение ряда вопросов при проектировании геоинформационной системы, таких как разрешение неоднозначностей трактовки правил или задание самих правил. При этом при задании правил существует возможность использования как средств самого функционального языка, так и его расширений, таких как «функции-защитники», которые позволяют задавать правила в гибкой форме, что в свою очередь позволит сделать работу с атрибутивными характеристиками геоописаний более адаптируемой для решаемой задачи.

Таким образом, имея формализованные атрибутивные характеристики геоописаний и набор правил грамматики, заданный в функциональном виде, можно для их обработки использовать язык функционального программирования и не проектировать отдельно аналитическую машину, а также вынести обработку правил из геоинформационной модели.

Следует заметить, что описанный выше подход позволяет не только анализировать атрибутивные характеристики геоописаний,

но и генерировать их в необходимом виде для дальнейших задач системы.

Система преобразования геоописаний представлена на рис. 5.

Нетрудно заметить, что этапы обработки при преобразовании геоописаний аналогичны этапам обработки при межъязыковой трансформации [10; 11]. Морфологический и синтаксический анализаторы, как было отмечено выше, необходимы для получения нормализованного геотекста, который передается на блок трансфера в функциональную форму. Средства функционального программирования используют грамматические правила, синтезируя функциональный вид геоописания. Завершающая стадия полученного геоописания может обеспечить перевод его на другой геоязык.

Предложенная схема преобразования геоописаний ориентирована на обеспечение возможности при работе с ГИС обрабатывать геоописания, представленные на разных геоязыках.

ЗАКЛЮЧЕНИЕ

В статье представлен подход, позволяющий учитывать семантику геообъектов за счет анализа и преобразования геоописаний их атрибутивных характеристик. Возможность упрощения обработки атрибутивных характеристик геоданных основана на использовании средств функциональных языков программирования. Рассмотрены вопросы, связанные с формированием, сопоставлением и преобразованием геоописаний для актуализации геоинформации в информационной среде ГИС. Для осуществления преобразования геоописаний предложено упрощение формирования инструмента на основе парадигмы функционального программирования для представления геоинформации и набора геотекстов на различных геоязыках.

Реализация системы преобразования геоописаний на основе среды функционального программирования позволит решить задачу

преобразования геоописаний аналогично преобразованию межъязыковых текстов.

В качестве дальнейшего исследования в области геоописаний предполагается разработка специализированной среды функционального программирования, предназначенной для решения задач обработки и анализа геоописаний.

СПИСОК ИСТОЧНИКОВ

1. Дулин С. К., Дулина Н. Г., Никишин Д. А. Проблемы обеспечения семантической геоинтероперабельности и согласования понимания семантики геоданных // Системы и средства информатики. – 2016. – Т. 26. – № 1. – С. 86–108. DOI: 10.14357/08696527160107.
2. Тарасов В. Б. Проблема понимания: настоящее и будущее искусственного интеллекта // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2015): Материалы V Международной научно-техн. конференции (Минск, 19–21 февраля 2015). – 2015. – № 5. – С. 25–42. EDN: ZIAXUN.
3. Лютий А. А. Язык карты: сущность, система, функции – М.: ГЕОС, 2002. – 327 с. ISBN 5-89658-018-5.
4. Shihong Du, Qimin Qin, Qiao Wang, Haijian Ma. Evaluating structural and topological consistency of complex regions with broad boundaries in multi-resolution spatial databases. Information Sciences, 2008, Vol. 178, Iss. 1, pp. 52–68. DOI: <https://doi.org/10.1016/j.ins.2007.07.023>.
5. Schreiber, G. Knowledge acquisition and the web. International Journal of Human-Computer Studies, 2013, Vol. 71, pp. 206–210. [Электронный ресурс]: <https://www.cs.vu.nl/~guus/papers/Schreiber13a.pdf>. Доступ 17.04.2024.
6. Макаренко С. И., Олейников А. Я., Черницкая Т. Е. Модели интероперабельности информационных систем // Системы управления, связи и безопасности. – 2019. – № 4. – С. 215–245. DOI: 10.24411/2410-9916-2019-10408.
7. Козеренко Е. Б. Лингвистическое моделирование для систем машинного перевода и обработки знаний // Информатика и ее применения. – 2007. – Т. 1. – № 1. – С. 54–65. EDN: KVICUX.
8. Душкин Р. В. Функциональное программирование на языке Haskell. – М.: ДМК Пресс, 2006. – 608 с. ISBN 5-94074-335-8.
9. Дулин С. К., Ермаков П. В. Функционально-вычислимые грамматики в задачах анализа текстов на естественном языке – М.: ВЦ РАН, 2011. – 22 с.
10. Dulina, N., Kozhunova, O. Information monitoring system: a problem of linguistic resources consistency and verification // Третья международная конференция «Проблемы кибернетики и информатики» (PCI'2010) – Баку: Elm, 2010. – С. 56–58.
11. Козеренко Е. Б. Лингвистические аспекты информатики // Системы и средства информатики // Специальный выпуск «Научно-методические проблемы информатики». – 2006. – № 16. – С. 88–111. [Электронный ресурс]: http://www.ipiran.ru/journal/collected/2006_16a/01-03.pdf. Доступ 15.02.2024. ●

Информация об авторах:

Розенберг Игорь Наумович – доктор технических наук, профессор, член-корреспондент Российской академии наук, научный руководитель Российского университета транспорта, Москва, Россия, info@science-rut.ru.

Дулин Сергей Константинович – доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; главный научный сотрудник Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), Москва, Россия, skdulin@mail.ru.

Статья поступила в редакцию 16.02.2024, одобрена после рецензирования 29.04.2024, принята к публикации 15.05.2024.

