

УДК 004.9

РАЗРАБОТКА ИНТЕРАКТИВНОЙ СПРАВОЧНОЙ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ «РЕСПУБЛИКА ХАКАСИЯ»

Голубничий А.А., Замулин И.С., Коптева Д.Ю.

ФГБОУ ВО «Хакасский государственный университет им. Н.Ф. Катанова»,
Абакан, e-mail: zamulin_ivan@mail.ru

В статье дается краткий обзор ГИС, рассматриваются средства для создания геоинформационных систем средствами языков программирования. Дается краткое сравнение языков Python и R, на основании сравнения обосновывается выбор языка программирования R и дополнительных пакетов: leaflet, shiny и htmlwidgets. Обосновывается выбор структуры и средств реализации геоинформационной базы данных. В качестве основы для формирования базы данных используются плоские файлы, а именно формат CSV, как самый распространенный формат, применяемый в анализе данных. Геоинформационная база данных строится из трех файлов формата CSV, содержащих сведения о границах республики Хакасия, границах ее отдельных административных единиц и подробное описание каждой административной единицы в виде атрибутивных данных. Конкретизируются атрибутивные данные, включенные в геоинформационную базу данных. Описывается логика реализации программного продукта с использованием пакета leaflet. Перечисляются основные функции создания карты, описывается их логика и порядок вызова в скрипте программного кода. Обосновывается выбор в качестве подложки системы, данных открытого ресурса OpenStreetMap. Указывается потенциальная возможность масштабирования программного продукта посредством базовых возможностей пакета leaflet и других пакетов языка R. Представлены первичный и итоговый рендеринги интерфейса программного продукта, построенного с помощью пакета shiny.

Ключевые слова: ГИС, интерактивные веб-приложения, язык R, leaflet, shiny

DEVELOPMENT OF THE INTERACTIVE REFERENCE GEOINFORMATION SYSTEM «REPUBLIC OF KHAKASSIA»

Golubnichiy A.A., Zamulin I.S., Kopteva D.Yu.

Katanov Khakass State University, Abakan, e-mail: zamulin_ivan@mail.ru

The article gives a brief overview of GIS, considers the means for creating geoinformation systems using programming languages. A brief comparison of the Python and R languages is given, based on the comparison, the choice of the R programming language and additional packages: leaflet, shiny and htmlwidgets is justified. The choice of structure and means of implementation of the geoinformation database is substantiated. As a basis for the formation of the database, flat files are used, namely the CSV format, as the most common format used in data analysis. The geographic information database is built from three CSV files containing information about the borders of the Republic of Khakassia, the boundaries of its individual administrative units and a detailed description of each administrative unit in the form of attributive data. The attributive data included in the geoinformation database are specified. The logic of software product implementation using leaflet package is described. The main functions of creating a map are listed, their logic and the order of calling in the script of the program code are described. The choice of the data of the open resource OpenStreetMap as the system substrate is substantiated. Potential scalability of the software product using the basic capabilities of the leaflet package and other R language packages is indicated. The primary and final renderings of the interface of the software product built with the shiny package are presented.

Keywords: GIS, interactive web applications, R language, leaflet, shiny

Современные ГИС позволяют решать различные задачи экологии, управления, картографии и многие другие смежные задачи, а также применяются в сферах управления хозяйством, туризме, здравоохранении. За счет интеграции с мировыми ИТ-трендами – системами искусственного интеллекта, технологиями больших данных, дронами и роботами – геоинформационные технологии отходят от классического понимания карты и превращаются в многоцелевые решения, охватывая почти все сферы человеческой деятельности: от сферы недвижимости до индустрии развлечений. Широкое применение геоинформационные системы получили в авиации, наличие мощных инструментальных средств по интеграции данных разных типов, моделирования, 2D

и 3D карт позволяет решать задачи управления такими сложными комплексами с развитой пространственной инфраструктурой, как аэропорты, выполняя функции планирования маршрутов полетов, анализа данных и управления инфраструктурой, контроля окружающей среды, получения, обработки и анализа пространственных данных [1].

В зависимости от цели разработки продукта, задач, которые он должен выполнять, для создания геоинформационных систем используются различные языки программирования. Так, например, для разработки ГИС часто применяется язык Python, который является языком программирования общего назначения с открытым кодом, а также языком программирования для скриптов в геообработке ArcGIS.

Еще одним языком, подходящим для разработки геоинформационных систем, является R. Язык R – это мощная и широко распространённая программная среда для статистических вычислений и графики, которая отлично подходит для анализа и обработки геопространственных данных [2]. Возможности геопространственного анализа с помощью R доступны через так называемые пакеты-модули, которые расширяют возможности R. Различные пакеты обеспечивают доступ как к большому числу традиционных алгоритмов, так и к самым новым из них, причём часто до того, как они появятся в другом открытом или проприетарном ПО. С помощью R можно обрабатывать точечные, линейные и полигональные векторные данные, а также растры. Благодаря этим возможностям, пользователи могут решать широкий спектр задач: статистический анализ для выявления пространственных связей и структуры свойств, классификация изображений и многое другое. Функции, доступные в R, дополняют и развивают возможности ГИС-пакетов и программ для работы с данными дистанционного зондирования Земли (ДЗЗ). ДЗЗ – незаменимый инструмент изучения и постоянного мониторинга нашей планеты, помогающий эффективно использовать и управлять ее ресурсами [3]. К тому же R обладает возможностями по созданию интерактивных карт. В качестве языка для разработки программного продукта решено выбрать язык R, так как он удовлетворяет потребностям для создания программы, а также полностью и понятно документирован [4].

Цель исследования: используя средства языка программирования R, разработать интерактивную справочную геоинформационную систему «Республика Хакасия».

Средства реализации проекта

Практически любой проект можно написать на «чистом» языке программирования, однако для ускорения процесса в реальной разработке используются специальные наборы функций – пакеты. Для реализации рассматриваемого программного продукта было принято решение использовать несколько сторонних пакетов, входящих в официальный репозиторий языка программирования R – CRAN.

Leaflet – пакет для создания интерактивных карт на языке R, основанный на одноименной библиотеке с открытым исходным кодом, написанной на языке JavaScript, являющейся одной из самых популярных библиотек для создания интерактивных карт. Данный пакет позволяет легко интегрировать и управлять картами в R. Основными

функциональными возможностями Leaflet по работе с картами являются:

- интерактивное масштабирование карт;
- добавление на карты слоев, маркеров, полигонов, линий, всплывающих окон, а также использование GeoJSON;
- создание карт из консоли R или с применением среды RStudio;
- использование карт совместно с Shiny, R Markdown и knitr;
- использование границ карты и событий мыши для управления картами;
- расширение функциональных возможностей при помощи различных плагинов [5].

Shiny – пакет, позволяющий с легкостью создавать многофункциональные интерактивные веб-приложения на языке R [6]. Разрабатываемое shiny-приложение может быть дополнено CSS-темами, html-виджетами и JavaScript. Любое shiny-приложение состоит из серверной части – server и пользовательского интерфейса – ui. В больших проектах можно разделять эти части на разные файлы с одноименными названиями и расширением .R, а также допускается объединять их в один файл [7].

Htmlwidgets – пакет, позволяющий создавать HTML-виджеты, используя JavaScript. Виджеты, созданные с помощью пакета htmlwidgets, можно использовать для анализа данных в консоли R, встраивать в документы R Markdown и приложения Shiny, а также сохранять как отдельные веб-страницы для обмена по почте [8].

Разрабатываемый программный продукт должен обладать следующими функциями:

- отображение территориальных границ республики Хакасия;
- отображение границ административно-территориального деления республики;
- отображение географического положения всех населённых пунктов региона;
- предоставление справочной информации по каждому населенному пункту, а именно: название на двух языках – русском и хакасском, административно-территориальная принадлежность, численность населения и почтовый индекс;
- предоставление инструмента для измерения расстояния между объектами;
- поиск населенных пунктов по названию;
- отображение координат текущего местоположения курсора.

Реализация программного продукта

Геоинформационная система тесно связана с понятием геоинформационной базы данных. Геоинформационная база данных (ГБД) – это организованная совокупность пространственных и атрибутивных данных, которые описывают некоторую территорию

и расположенные на ней объекты. Атрибутивные данные представляют собой таблицы, в которых все строки имеют одинаковые столбцы. Основные операции, которые выполняются с атрибутивными данными, – это поиск, сортировка данных, добавление и удаление записей, построение отчетов. Классы пространственных объектов на самом деле являются просто таблицами со специальными полями, содержащими информацию о геометрии объектов.

Для разрабатываемой геоинформационной системы необходимы следующие данные:

- координаты границ республики Хакасия (каждая точка представляет собой пару значений долгота/широта);
- координаты границ административно-территориального деления республики;
- информация обо всех населенных пунктах республики, а именно: названия на двух языках – русском и хакасском, координаты местоположения, численность населения, территориальная принадлежность, тип населенного пункта и почтовый индекс.

Данные для разрабатываемой системы решено хранить в формате CSV. CSV – текстовый формат, предназначенный для представления табличных данных. Строка таблицы соответствует строке текста, которая содержит одно или несколько полей, разделенных специальным символом, к примеру запятой или точкой с запятой. Для работы с этим форматом данных можно легко использовать любой табличный процессор.

Для удобства использования данные было решено разделить на три файла:

1. Файл «Khakasia_border_full.csv» содержит в себе координаты границы республики Хакасия. Она представлена многоугольником, состоящим из 6663 точек. Каждая точка хранит в себе значения latitude и longitude – широта и долгота в десятичном формате, а также id – порядковый номер, необходимый для проверки и корректировки данных.

2. Файл «adm_terit_del.csv» содержит значения координат границ административно-территориального деления республики. Республика Хакасия подразделяется на восемь районов: Алтайский, Аскизский, Бейский, Богградский, Орджоникидзевский, Таштыпский, Усть-Абаканский, Ширинский.

Также на территории РХ располагаются пять городских округов: городской округ Абаза, городской округ Абакан, городской округ Саяногорск, городской округ Сорск, городской округ Черногорск.

Каждый из районов и городских округов представлен одним или несколькими по-

лигонами, каждая точка которых содержит значения latitude, longitude и id. Первоисточником данных о границах субъектов Российской Федерации и границах административно-территориального деления субъектов является OpenStreetMap. Данные взяты с сайта mydata.biz, в архиве предоставляются данные о государственной границе, границах федеральных округов, субъектов федерации, городских округов, районов [9].

3. Файл «cities_full_csv.csv» хранит в себе названия всех населенных пунктов республики Хакасия на двух языках, а именно русском и хакасском. Также в файле содержится информация о местоположении каждого населенного пункта, представленная точкой со значениями широты и долготы, в которой будет стоять маркер. Помимо этого, здесь предоставлена информация о типе населенного пункта, административно-территориальной принадлежности, численности населения, а также почтовый индекс. Данные взяты из сети интернет, а также из закона РХ «О внесении изменений в отдельные законодательные акты республики Хакасия в части написания наименований административно-территориальных единиц и населенных пунктов республики Хакасия на государственных языках республики Хакасия» [10].

Основой геоинформационной системы является карта. Картографической основой данного проекта будут карты OpenStreetMap (OSM). Для отображения карты используется функция «leaflet()». Далее устанавливаем основной слой, а также дополнительные слои при необходимости, используя функции «addTiles()» и «addProviderTiles()». Разработчику доступно более 100 различных слоев карт. Чтобы установить первоначальную позицию отображения карты, используем «setView()» и передаем в нее значения lat и lng, отображающие центральную точку, а также zoom для начального масштаба карты. На рис. 1 показано начальное отображение карты при запуске системы.

Для получения данных из файлов можно использовать функции «read.csv» и «read.csv2», принципиальное различие этих функций заключается в значении разделителя. В первом случае используется запятая, а во втором точка с запятой. Однако более удобным вариантом будет функция «read_delim», позволяющая вручную настроить типы считываемых данных, разделители, использование строки-заголовка. Данная функция доступна из пакета «readr», предназначенного для чтения матричных данных из таких файлов, как CSV, TSV, FWF [11]. Для файла «cities_full_csv.csv» вручную задаем столбцу index тип значений

character, а lat и lng тип значений number. Аналогично для других файлов в столбцах, хранящих значения широты и долготы, вручную задаем тип значений number.

Для отображения границ республики и административно-территориального деления применяется функция «addPolygons()». В качестве параметров lat и lng передаются векторы значений широты и долготы из переменных «khakasia_border» и «adm_terit_del». Также задаются параметры отображения подсказок, при наведении на конкретную область, цвет, прозрачность выделения, прорисовка контура и другие параметры создаваемых полигонов.

Для отображения географического положения населенных пунктов путем установки маркеров используется функция «addMarkers()». В качестве параметров передаем значения lat и lng из файла «cities_full_csv.csv». Для отображения справочной информации о населенных пунктах используем всплывающие окна роруп, текст которых сформирован функцией «unific_label_text». Кластеризация маркеров реализована путем добавления «clusterOptions». Функции создания полигонов, маркеров, слоев имеют параметр «group» для дальнейше-

го управления их отображением. Используя функцию «addLayersControl», на карте в правом верхнем углу появляется выпадающий список с кнопками переключения и флажками, где пользователь может вручную выбрать, какие элементы необходимо отображать, а какие скрывать.

При создании пользовательского интерфейса shiny-приложения разработчику доступны такие элементы, как заголовки, поля ввода и вывода текстовой информации, графиков и гистограмм, карт, изображений, таблиц, кнопки, чекбоксы, переключатели, панели навигации и многое другое [12]. За создание заголовка страницы отвечает функция «titlePanel()». Основная информация страницы будет отражена в mainPanel(). Функцией «leafletOutput()» создаем область для рендеринга карты, здесь же можно установить размеры области, а также id для обращения к ней в серверной части. Для отображения координат курсора используется поле вывода «verbatimTextOutput», позволяющее отображать постоянно меняющиеся значения, в отличие от «textOutput», предназначенного для вывода постоянных значений. На рис. 2 отображен конечный вид разрабатываемой программы.

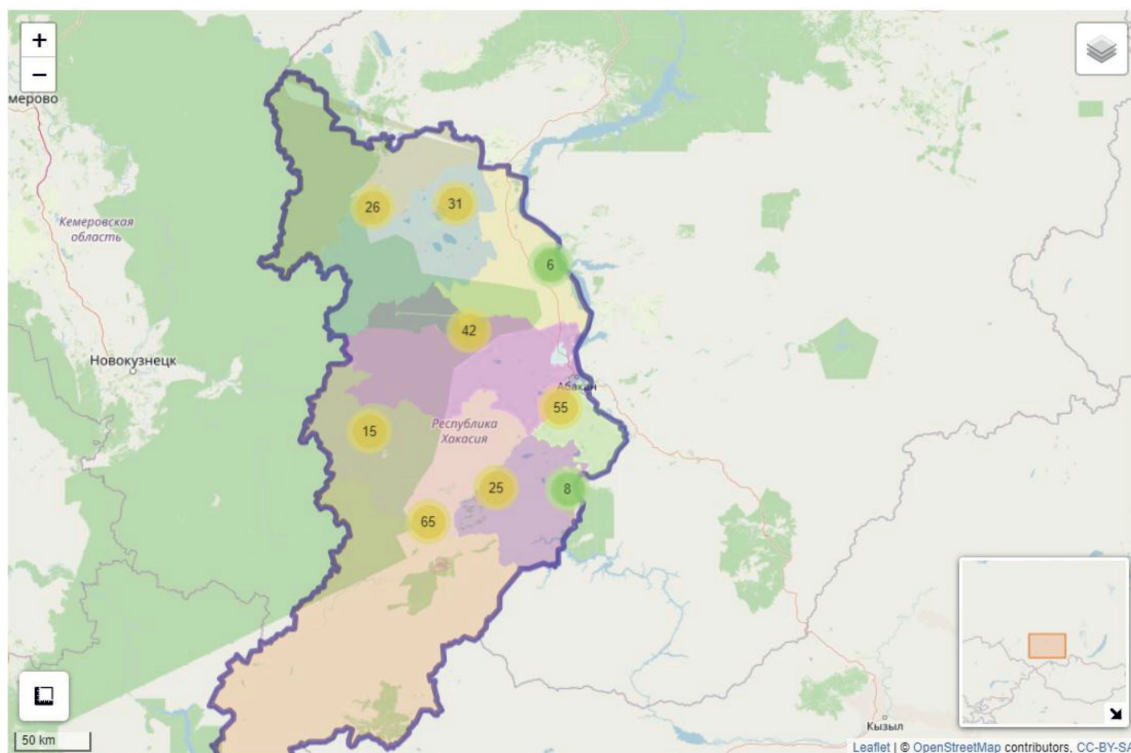


Рис. 1. Рендеринг карты при запуске системы

Интерактивная справочная геоинформационная система Республика Хакасия



Широта: 53.654767399467
Долгота: 90.9118652291363

Введите название населенного пункта

Тигей

location	population	lat	lng	index	name_kh	type	Region
Тигей	91.00	53.42	91.12	655118	Тигей	Поселок	Усть-Абаканский район

Рис. 2. Рендеринг разработанной системы

Выводы

Разработанное приложение позволяет получить основную справочную информацию обо всех населенных пунктах республики Хакасия, наглядно отобразить границы административно-территориального деления и всего региона в целом.

Список литературы

1. Зачем людям нужны ГИС-сервисы. [Электронный ресурс]. URL: <https://hightech.fm/2018/12/25/gis-services> (дата обращения: 26.08.2020).
2. Хакимова Т.В., Голубничий А.А. Популярность и перспективы языка программирования R // E-Scio. 2019. № 6 (33). С. 817–821.
3. Дистанционное зондирование Земли. [Электронный ресурс]. URL: <https://www.roskosmos.ru/24707/> (дата обращения: 26.08.2020).
4. Robin Lovelace, Jakub Nowosad, Jannes Muenchow. Geocomputation with R: R. Lovelace, J. Nowosad, J. Muenchow. M.: Taylor & Francis Ltd, 2019. 353 с.

5. CRAN – Package leaflet. [Электронный ресурс]. URL: <https://CRAN.R-project.org/package=leaflet> (дата обращения: 26.08.2020).

6. Голубничий А.А., Чернявская К.А. Создание web-приложения средствами shiny для организации интерактивного обучения в вузе // Вестник Хакасского государственного университета им. Н.Ф. Катанова. 2017. № 20. С. 95–98.

7. Chris Beeley. Web Application Development with R Using Shiny. M.: Packt Publishing Ltd, 2013. 110 с.

8. CRAN – Package htmlwidgets. [Электронный ресурс]. URL: <https://CRAN.R-project.org/package=htmlwidgets> (дата обращения: 26.08.2020).

9. Данные об административно-территориальном делении субъектов РФ [Электронный ресурс]. URL: https://mydata.biz/ru/catalog/databases/borders_ru (дата обращения: 26.08.2020).

10. РЕЕСТР административно-территориальных единиц Республики Хакасия. [Электронный ресурс]. URL: <http://docs.cntd.ru/document/453362324> (дата обращения: 26.08.2020).

11. CRAN – Package readr. [Электронный ресурс]. URL: <https://CRAN.R-project.org/package=readr> (дата обращения: 26.08.2020).

12. Build a user interface. [Электронный ресурс]. URL: <https://shiny.rstudio.com/tutorial/written-tutorial/lesson2/> (дата обращения: 26.08.2020).