

## АНАЛИЗ ИСТОЧНИКОВ ПРОСТРАНСТВЕННЫХ ДАННЫХ В РАЗВИТИИ ТЕХНОЛОГИИ НЕДРОПОЛЬЗОВАНИЯ

Г.А. Некрасов<sup>1</sup>, Д.Э. Поливода<sup>1</sup>, Е.Н. Прокофьева<sup>1</sup>

<sup>1</sup> Национальный исследовательский университет «Высшая школа экономики», Москва, Россия,  
e-mail: eprokofyeva@hse.ru

**Аннотация:** Проведено исследование разницы между растровыми и векторными данными; определены основные типы данных, которые используются для представления геопространственных данных; выявлены сильные и слабые стороны хранения данных в растровом формате; дано описание атрибутов, некоторых основных функций и типов растровых данных. Существует множество источников пространственных и атрибутивных данных. Источники пространственных данных: печатные карты, аэрофотоснимки, дистанционно воспринимаемые образы, выборочные данные из опросов, существующие файлы цифровых данных. Пространственные данные могут существовать во множестве форматов и содержат больше, чем просто информацию о конкретном местоположении. Различают два типа в соответствии с технологией хранения: растровые и векторные данные. Пространственные данные могут быть классифицированы как необработанные и производные. По мере того как собирают данные из окружающей среды, картографы используют свое восприятие для обнаружения шаблонов и последующей подготовки данных для создания карты. Статистические данные являются атрибутами пространственных объектов. Все источники пространственных данных формируются разными классами систем. Основное различие между пространственными и всеми другими типами данных, когда речь идет о статистическом анализе, заключается в необходимости учитывать такие факторы, как высота, расстояние и площадь в аналитическом процессе.

**Ключевые слова:** веб-скрейпинг, пространственные данные, геопространственное моделирование, автоматическое управление, аналитическая экосистема.

**Благодарность:** Публикация подготовлена в ходе проведения работы (№ 20-04-033 проекта) в рамках Программы «Научный фонд Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)» в 2020 – 2021 гг. и в рамках государственной поддержки ведущих университетов Российской Федерации «5-100».

**Для цитирования:** Некрасов Г.А., Поливода Д.Э., Прокофьева Е.Н. Анализ источников пространственных данных в развитии технологии недропользования // Горный информационно-аналитический бюллетень. – 2020. – № 5. – С. 164–176. DOI: 10.25018/0236-1493-2020-5-0-164-176.

### Sources of spatial data in advancing technologies of subsoil use

G.A. Nekrasov<sup>1</sup>, D.E. Polivoda<sup>1</sup>, E.N. Prokofeva<sup>1</sup>

<sup>1</sup> Higher School of Economics. National Research University, Moscow, Russia,  
e-mail: eprokofyeva@hse.ru

---

**Abstract:** The difference between the raster and vector data is studied. The basic types of data used in geo-spatial representations are determined. Advantages and disadvantages of the raster-type data storage are identified. The attributes, some key functions and types of the raster data are described. There are many sources of data for describing both space and attributes. The most popular sources of spatial data are: maps, air photos, remote sensing images, sample data of surveys and digitized information. Spatial data can exist in multiple formats and contain more various information rather than simply facts on a specific mineral deposit. There are two types of data in accordance with the storage technology, namely, raster and vector data. Spatial data can be classified as crude data and derived data. As data are being collected from the environment, map-makers use their perception to find patterns and to prepare the data for further mapping. Statistical data are the attributes of spatial objects. All sources of spatial data are formed by different classes of systems. The main difference between the spatial data and all other types of data, when we speak about the statistical analysis, is the requirement to take into account such factors as height, distance and area in the analysis.

**Key words:** web-scraping, spatial data, geo-spatial modeling, automated control, analytical ecosystem.

**Acknowledgements:** The publication has been prepared during implementation of Project No. 20-04-033 in the framework of the Program of the Scientific Foundation of the National Research University—Higher School of Economics in 2020–2021 and under governmental support of the leading universities of the Russian Federation, Project 5–100.

**For citation:** Nekrasov G. A., Polivoda D. E., Prokofeva E. N. Sources of spatial data in advancing technologies of subsoil use. *MIAB. Mining Inf. Anal. Bull.* 2020;(5):164-176. [In Russ]. DOI: 10.25018/0236-1493-2020-5-0-164-176.

---

## Введение

Геопространственные данные, то есть данные с информацией о местоположении, ежедневно генерируются в огромных объемах миллиардами мобильных телефонов, датчиков и других источников. Данные порождают данные, постоянно нарастающая неограниченные потоки геопространственных данных (для краткости «геоданные») в ожидании статистического анализа. Куда идут люди и их машины, что обнаруживает наше дистанционное зондирование и как наши устройства располагаются в пространстве и работают во времени, — все это очень важно для динамичности нашей экономики, здоровья нашей планеты, нашего общего счастья и благополучия, бытия. Геопространственная аналитика может предоставить инструменты и методы, которые необходимы, чтобы разобраться во всех этих данных и использовать их для решения проб-

лем, с которыми мы сталкиваемся на всех уровнях.

Сегодня карты могут быть легко созданы с помощью широкого спектра онлайн-инструментов любым человеком, имеющим доступ к Интернету. Однако карты, используемые в большинстве видов деятельности (от городского планирования, геологического изучения или управления окружающей средой до планирования поездок и навигации), по-прежнему, как правило, создаются специалистами, обладающими знаниями в области картирования или явлений, отображаемых на картах. Картографический процесс начинается с реальной или воображаемой среды, и картограф должен ответить на четыре вопроса: где? когда? какая? зачем? Пространственные данные обычно хранятся в виде координат, топологии и представляют собой данные, которые можно сопоставить. Пространственные данные часто доступ-

ны, обрабатываются или анализируются с помощью географических информационных систем (далее ГИС).

Пространственные данные также известны как географическая информация, которая определяет географическое местоположение объектов, таких как естественные или построенные объекты, океаны и многое другое, и границ на Земле. Пара координат широты и долготы определяет конкретное местоположение на земле.

В этой статье рассматриваются различные типы пространственных данных, используемых для геоморфологического картографирования как в аналоговом, так и в цифровом форматах. Как и в других областях науки и техники, переход от аналоговых к цифровым данным исторически важен в геоморфологии. Поэтому актуально рассмотреть аналоговые данные, затем цифровые данные, описать основные характеристики и привести некоторые примеры геоморфологического картирования с использованием данных.

### **Постановка задачи**

Дистанционное зондирование приобретает все большее значение и применение в горнодобывающей промышленности и экологических исследованиях благодаря широкому спектру быстро развивающихся технологий и методов, которые генерируют большие данные за меньшее время. Фотограмметрия, СНС, радар и набор бортовых датчиков сканирования и мониторинга выступают ключевыми системами, методами и инструментами.

Мини-, нано- и микроспутники относятся к числу современных систем наблюдения Земли, которые привлекают все больший интерес в мировом исследовательском сообществе, нацеленном на получение точных пространственных данных для поддержки интегрирован-

ных геопространственных процессов и устойчивых решений.

Пространственные данные являются основополагающими для любых картографических операций и могут быть классифицированы на два типа: необработанные и производные. Для геоморфологического картирования необработанные данные включают информацию о распределении высоты, такую как контурные линии и высоты точек на топографической карте, и растровую цифровую модель рельефа (далее ЦМР). В некотором смысле получение таких данных высот можно назвать геоморфологическим картированием. Кроме того, тематические карты, показывающие пространственное распределение единиц рельефа, являются типичными продуктами геоморфологической карты, для которых используются как необработанные, так и производные данные. Производные данные включают производные матрицы высот, такие как угол наклона, кривизна и аспект. Результаты визуальной интерпретации топографических карт и аэрофотоснимков также являются производными данными, полезными для прикладного геоморфологического картирования [1 – 6].

Другая распространенная двоичная классификация пространственных данных — аналоговая или цифровая. Классические пространственные данные представлены в аналоговом формате, таком как печатные карты и рукописные иллюстрации в полевых заметках. Хотя аналоговые данные внесли вклад в развитие геоморфологии, качественный и субъективный характер этих данных затрудняет их непосредственный анализ с помощью компьютеров. С 1990-х годов цифровые карты в значительной степени вытеснили традиционные аналоговые источники данных, топографические карты были заменены ЦМР, а их анализ облегчен соответствующи-

ми технологиями, такими как быстрые персональные компьютеры и географические информационные системы (далее ГИС).

Существуют два основных типа геопространственных данных: растровые и векторные данные. Растровые данные состоят из ячеек сетки, определенных строкой и столбцом. Весь географический район делится на группы отдельных ячеек, которые представляют изображение. Растровые данные хранятся в виде сетки значений, которые отображаются на карте в виде пикселей. Каждое значение пикселя представляет область на поверхности Земли, где каждый пиксель связан с конкретным географическим местоположением. Значение пикселя может быть непрерывным (например, высота) или категориальным (например, землепользование). Эта структура данных очень распространена, и именно так представляют любое цифровое изображение. Примерами растровых данных являются спутниковые изображения, фотографии, отсканированные изображения. Векторные данные состоят из точек, полилиний и многоугольников. Колодцы, дома представлены точками. Дороги, реки, ручьи представлены полилиниями. Деревни и города представлены полигонами. Данные атрибутов со-

держат соответствующую информацию о пространственных данных. Функция запроса работает на основе атрибутивных данных, то есть она привязана к геопространственным данным.

Типы данных атрибута:

- номинальные данные;
- порядковые данные;
- интервальные данные;
- соотношение данных.

Векторные структуры данных представляют определенные особенности на поверхности Земли и присваивают атрибуты этим объектам. Чтобы работать с растровыми и векторными наборами данных, необходимо понимать основные структуры этих типов данных и сами типы данных, которые могут быть использованы для дальнейшей визуализации и представления [7 – 10]. Геопространственный растр отличается от цифровой фотографии только тем, что он сопровождается пространственной информацией, которая связывает данные с определенным местоположением. Это включает в себя экстенд растра и размер ячейки, количество строк и столбцов и его систему координат. Некоторые примеры непрерывных растров включают в себя: карты осадков, карты высоты деревьев, полученные из данных LiDAR, значения высот для региона. На рис. 1

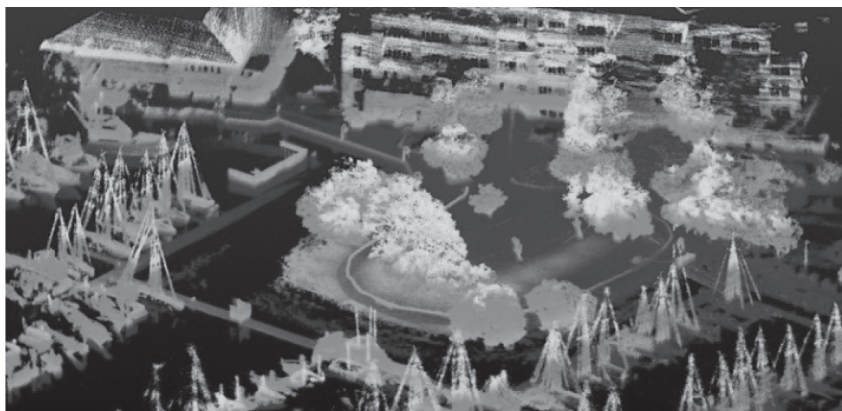


Рис. 1. Карты сканирования высот с помощью датчика NEON AOP LiDAR

Fig. 1. NEON AOP LiDAR scanned elevation maps

приведено изображение карты высот, полученное с помощью датчика NEON AOP LiDAR.

Некоторые растры содержат категориальные данные, где каждый пиксель представляет отдельный класс, такой как тип земного покрова (например, «лес» или «луг»), а не непрерывное значение, такое как высота или температура.

Некоторые примеры карт могут включать в себя:

- карты земельного покрова/землепользования (рис. 2);
- карты высоты деревьев классифицируются как короткие, средние и высокие деревья;
- карты высот классифицируются как низкие, средние и высокие.

Растровые данные имеют ряд важных преимуществ:

- представление непрерывных поверхностей;
- потенциально очень высокий уровень детализации;
- данные «невзвешенные» по всей своей протяженности — геометрия неявно выделяет объекты;
- клеточные вычисления могут быть очень быстрыми и эффективными.

Недостатки растровых данных:

- очень большой размер файла, так как размер ячейки становится меньше;
- в настоящее время популярные

форматы не очень хорошо встраивают метаданные;

- сложно представить сложную информацию;

Атрибутами растровых данных являются протяженность (экстент) и разрешение. Пространственный экстент — это географическая область, охватываемая растровыми данными.

Пространственная протяженность пространственного объекта представляет собой географический край или местоположение, которое находится дальше всего на севере, юге, востоке и западе. Другими словами, экстент представляет собой общий географический охват пространственного объекта. Разрешение растра представляет собой область на земле, которую охватывает каждый пиксель растра.

Растровые данные могут быть разных форматов. Например, формат GeoTIFF, который имеет расширение .tif. Метаданные файлы хранят и атрибуты о файле, вложенные tif tags. Например, камера может хранить тег, который описывает марку и модель камеры или дату, когда была сделана фотография при сохранении .tif. GeoTIFF является стандартным .tif форматом изображения с дополнительной пространственной (географической) информацией, встроенной в файл в виде тегов. Эти теги должны включать следующие растровые метаданные:

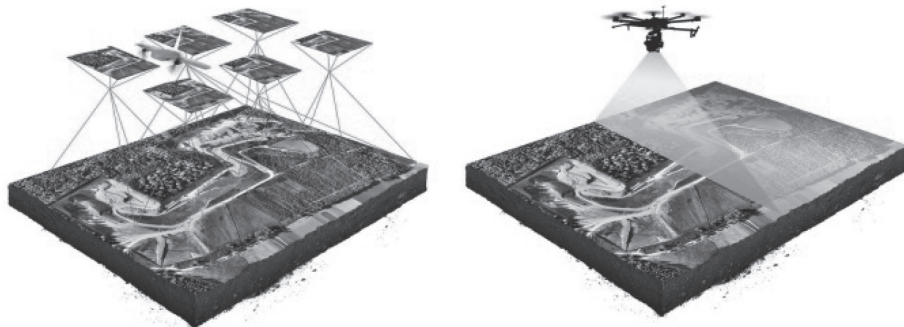


Рис. 2. Технология Фотограмметрии (а) и сканирования с помощью датчика LiDAR (б)

Fig. 2. Technology of photogrammetry (a) and LiDAR scanning (b)



- экстент (степень);
- разрешение;
- система координат (далее CRS — coordinate reference system);
- значения, которые представляют отсутствующие данные (NoDataValue).

Растр может содержать одну или несколько полос (многополосные растровые данные). Одним из типов многоканальных наборов растровых данных, который знаком многим, является цветное изображение. Основное цветное изображение состоит из трех полос: красного, зеленого и синего. Каждая полоса представляет свет, отраженный от красной, зеленой или синей части электромагнитного спектра. Яркость пикселей для каждой полосы при компоновке создает цвета, которые мы видим на изображении. Можно построить каждую полосу многоканального изображения индивидуально или объединить все полосы вместе, чтобы сделать цветное изображение. В многоканальном наборе данных растры всегда будут иметь одинаковую протяженность, разрешение и CRS. Многополосные растровые данные могут также содержать:

- Временные ряды: одна и та же переменная, в той же области, во времени.
- Мульти- или гиперспектральные изображения: растровые изображения, которые имеют 4 или более (мульти-спектральные) или более 10—15 (гиперспектральных) полос.

Как определялось ранее, векторные структуры данных представляют определенные особенности на поверхности Земли и присваивают атрибуты этим объектам. Векторы состоят из дискретных геометрических положений (значений  $x$ ,  $y$ ), известных как вершины, которые определяют форму пространственного объекта. Организация вершин определяет тип вектора: точка, линия или многоугольник. Каждая точка определяется одной координатой  $x$ ,  $y$ . В файле

векторных точек может быть много точек. Примеры точечных данных включают: места отбора проб, местоположение отдельных деревьев или местоположение участков съемки.

Линии состоят из множества (как минимум 2) точек, которые связаны между собой. Например, дорога или поток могут быть представлен линией. Эта линия состоит из серии сегментов, каждый изгиб дороги или потока представляет собой вершину, которая определила местоположение  $x$ ,  $y$ .

Полигон состоит из 3 или более вершин, которые соединены и замкнуты. Контуры границ съемочных площадок, озер, океанов и штатов или стран часто представлены полигонами.

Иногда пограничные слои, такие как государства и страны, хранятся в виде линий, а не полигонов. Однако эти границы, представленные в виде линии, не будут создавать замкнутый объект с определенной областью, которая может быть заполнена.

Векторные данные имеют ряд важных преимуществ:

- сама геометрия содержит информацию о том, что создатель набора данных считал важным;
- геометрические структуры содержат информацию сами по себе — например, зачем выбирать точку над многоугольником?
- каждый геометрический объект может содержать несколько атрибутов вместо одного, например, база данных городов может иметь атрибуты для имени, страны, населения и прочее;
- хранение данных может быть очень эффективным по сравнению с растрами.

Недостатки векторных данных включают в себя:

- потенциальная потеря детализации по сравнению с растром;
- потенциальный уклон в наборах данных — что не было зарегистрировано;

- расчеты, включающие несколько векторных слоев, должны выполнять математику как для геометрии, так и для атрибутов, поэтому могут быть медленными по сравнению с растровой математикой.

Наборы векторных данных используются во многих отраслях промышленности, кроме геопространственных полей. Например, компьютерная графика в значительной степени основана на векторах, хотя используемые структуры данных имеют тенденцию соединять точки, используя дуги и сложные кривые, а не прямые линии. Компьютерное проектирование (САПР) также основано на векторах. Разница в том, что наборы геопространственных данных сопровождаются информацией, привязывающей их особенности к местам реального мира [11 – 18].

Как и растровые данные, векторные данные также могут иметь различные форматы. Например, формат Shapefile, который имеет расширение .shp. А .shp файл хранит географические координаты каждой вершины в векторе, а также метаданные, включая:

- экстенд — пространственный экстенд шейп-файла (то есть географическая область, которую охватывает шейп-файл). Пространственный экстенд для шейп-файла представляет собой объединенный экстенд для всех пространственных объектов в шейп-файле;

- тип объекта — содержит ли шейп-файл точки, линии или многоугольники;

- система координат (CRS);

- другие атрибуты — например, шейп-файл линии, которые содержат местоположения потоков и могут содержать имя каждого потока.

Поскольку структура точек, линий и многоугольников различна, каждый отдельный шейп-файл может содержать только один векторный тип (все точки, все линии или все многоугольники). Как

следствие, не найдется смесь точечных, линейных и многоугольных объектов в одном шейп-файле.

Очень немногие форматы могут содержать как растровые, так и векторные данные. Наборы векторных данных обычно привязаны к одному типу геометрии, например, только к точкам. Наборы растровых данных обычно могут кодировать только один тип данных, например, не может быть многоканального GeoTIFF, где один слой представляет собой целочисленные данные, а другой — с плавающей запятой. Для этого есть веские причины, стандарты формата легче определять и поддерживать, как и метаданные. Эффект от конкретных манипуляций с данными более предсказуем, если все входные данные имеют одинаковые характеристики.

Таким образом, векторные структуры данных представляют специфические особенности на поверхности Земли вместе с атрибутами этих особенностей. И векторные объекты — это либо точки, линии или многоугольники.

Структура данных не может считаться геопространственной, если она не сопровождается информацией системы координат (CRS) в формате, который геопространственные приложения могут использовать для правильного отображения и манипулирования данными. Информация CRS связывает данные с поверхностью Земли с использованием математической модели. CRS (система координат) и SRS (система пространственных координат) являются синонимами и обычно взаимозаменяемы. CRS, связанный с набором данных, сообщает картографической программе, где растр расположен в географическом пространстве. Он также сообщает картографическому программному обеспечению, какой метод следует использовать для выравнивания или проецирования растра в географическом пространстве.

На рис. 3 отображены различия в форме, связанные с каждой проекцией. Эти различия являются прямым результатом расчетов, использованных для сглаживания данных на 2-мерной карте [6]. Существует множество полезных ресурсов, которые описывают системы координат и проекции более подробно. Важно отметить, что данные из одного и того же местоположения, но сохраненные в разных проекциях, не будут выстроены ни в одной ГИС или другой программе. Таким образом, при работе с пространственными данными важно идентифицировать систему координат, примененную к данным, и сохранять ее в процессе обработки и анализа данных.

Информация CRS состоит из трех компонентов:

- Элемент данных: модель формы земли. Он имеет угловые единицы, то есть градусы, и определяет начальную точку, поэтому углы указывают на значимое место на земле. Общими глобальными данными являются WGS84 и NAD83. Базовые данные также могут быть локальными — соответствовать определенным областям земного шара, но плохо подходят за пределами области предполагаемого использования.
- Проекция — математическое преобразование угловых измерений на круглой земле в плоскую поверхность (например, на бумаге или экране компьютера). Единицы, связанные с данной

проекцией, обычно линейны (футы, метры и т. д.).

- Дополнительные параметры часто необходимы для создания полной системы координат. Один общий дополнительный параметр — это определение центра карты. Количество необходимых дополнительных параметров зависит от того, что требуется для каждой конкретной проекции.

Если представить аллегорию, что земля аналог «апельсина», то, как его очищают, и затем выравнивают, похоже на то, как проецируется. То есть проекция — это как очистить «апельсин», а затем выровнять его. Чтобы определить, подходит ли прогноз для данных, необходимо ответить на следующие вопросы: какова область минимального искажения и какой аспект данных он сохраняет.

Существует несколько общих систем, используемых для хранения и передачи информации CRS, а также для трансляции между различными CRS. Эти системы в целом соответствуют стандарту ISO 19111. Общие системы описания CRS включают EPSG, PROJ и OGC WKT. PROJ — это библиотека с открытым исходным кодом для хранения, представления и преобразования информации CRS. PROJ.5 был недавно выпущен, но PROJ.4 использовался в течение 25 лет. PROJ представляет информацию CRS в виде текстовой строки



Рис. 3. Карта Гавайи в нескольких проекциях

Fig. 3. Different Hawaii map projections



пар ключ-значение, что облегчает ее настройку и с небольшой практикой, легко читается и интерпретируется. Строка PROJ4 содержит следующую информацию:

- proj =: проекция данных;
- зона =: зона данных (это специфично для проекции UTM);
- Datum =: использование данных;
- unit =: единицы для координат данных;
- эллипс =: эллипсоид (как рассчитывается округлость Земли) для данных.

Другие общеизвестные базы данных CRS — это система EPSG и Стандарт OGC WKT. В системе EPSG, поддерживаемой Международной ассоциацией производителей нефти и газа, набор данных содержит определения CRS и информацию о том, как безопасно преобразовать данные из одного CRS в другой. Использовать EPSG легко, поскольку каждый CRS имеет целочисленный идентификатор, например, WGS84 — это EPSG: 4326. Недостатком является возможность использования только тех CRS, которые EPSG определяет, и не предусмотрено самостоятельного их настраивания их.

Стандарт OGC WKT использует ряд важных геопространственных приложений и программных библиотек. WKT — это вложенный список геодезических параметров. Структура информации определяется на их сайте. WKT интереснее тем, что информация CRS более прозрачна, чем в EPSG, но ее сложнее читать и сравнивать, чем PROJ. Кроме того, стандарт WKT непоследовательно реализован на различных программных платформах, а сама спецификация имеет недостатки в использовании.

Геопространственные данные библиотеки GDAL имеют открытый исходный код. Это набор программных инструментов, которые преобразуют практически любой геопространственный

формат, широко используемый сегодня (и некоторые не очень распространенные). GDAL также содержит инструменты для редактирования и обработки как растровых, так и векторных файлов, включая перепроектирование данных в различные CRS. GDAL может использоваться, как автономный инструмент командной строки, или встраиваться в другое программное обеспечение ГИС. Несколько ГИС-программ с открытым исходным кодом используют GDAL для всех операций импорта/экспорта файлов.

Пространственные данные бесполезны без метаданных. Основные метаданные включают в себя информацию CRS, но правильные пространственные метаданные охватывают больше, чем только это. Варианты использования должны быть документированы в метаданных. Эта информация должна сопровождать набор пространственных данных, куда бы она ни направлялась. На практике это может быть сложно, так как многие форматы пространственных данных не имеют встроенного места для хранения такого рода информации. Метаданные часто должны храниться в сопутствующем файле, создаваться и поддерживаться вручную.

Большинство традиционных ГИС-работ выполняется в автономных приложениях, целью которых является предоставление комплексных геопространственных решений.

Самые ранние компьютерные системы работали без графического интерфейса пользователя (GUI), полагаясь только на интерфейс командной строки (CLI). Поскольку картирование и пространственный анализ являются строго визуальными задачами, приложения ГИС получили большую выгоду от появления графических интерфейсов и быстро стали полагаться на них. Большинство современных ГИС-приложений имеют очень сложные графические интерфейсы со

всеми общими инструментами и процедурами, доступ к которым осуществляется с помощью меню [19–24].

Существует множество мощных библиотек геопространственной обработки для языков программирования общего назначения, таких как Java и C++.

Языки сценариев более высокого уровня, такие как R и Python, легче изучать и использовать. Как следствие, они имеют свои собственные пакеты, которые объединяют эти библиотеки геопространственной обработки и облегчают доступ к ним и их безопасное использование. Ключевым примером является Java Topology Suite (JTS), который реализован в C++ как GEOS. GEOS доступен в R через sf пакет и в Python через shapely. R и Python также имеют пакеты интерфейсов для GDAL и для конкретных приложений ГИС. Интерфейсные пакеты дают возможность доступа к функциям, уникальным для конкретных программ.

### **Заключение**

В рамках данного исследования установлено, что все наборы геопростран-

ственных данных (растровые и векторные) связаны с конкретной системой координат. Система координат включает в себя данные, проекцию и дополнительные параметры, специфичные для набора данных. При работе с пространственными данными важно идентифицировать систему координат, примененную к данным, и сохранять ее в процессе обработки и анализа данных. Многие существующие форматы файлов были изобретены разработчиками программного обеспечения ГИС, часто в среде с закрытым исходным кодом. Это привело к большому количеству предлагаемых сегодня форматов передачи данных между программными средами. Новые инструменты и развитие алгоритмов искусственного интеллекта и машинного обучения расширяют возможности для изучения сложных данных о добыче полезных ископаемых и окружающей среде, а также служат основой для моделирования сценариев управления опасными явлениями и геопространственными рисками, определения оптимальных вариантов использования технологий недоропользования.

### **СПИСОК ЛИТЕРАТУРЫ**

1. Chadalawada J., Espinoza-Molina D., Datcu M. Assessment of earth observation data content based on data compression – application to settlements understanding // IEEE International Geoscience and Remote Sensing Symposium. 2012, pp. 6130–6133. DOI: 10.1109/igarss.2012.6352207.
2. Lionel Gueguen Classifying compound structures in satellite images: a compressed representation for fast queries // IEEE Transactions on Geoscience and Remote Sensing, 2014, Vol. 53, pp. 1803–1818. DOI: 10.1109/tgrs.2014.2348864.
3. Hansaem Park, Kwangseob Kim, Kiwon Lee Geo-data visualization on online and offline mode of mobile web using HTML5 / 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA). 2016. pp. 237–240. DOI: 10.1109/eorsa.2016.7552804.
4. Scrapy 1.4 documentation. 2019. Available at: <https://docs.scrapy.org/en/latest> (accessed 21 May 2019).
5. Olston C., Najork M. Web crawling // Foundations and Trends in Information Retrieval, 2010, vol. 4, no 3, pp. 175–246.
6. Hsieh J. M., Gribble S. D., Levy H. M. The architecture and implementation of an extensible web crawler / Proceedings of the 7th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2010, April 28–30, 2010, San Jose, CA, USA.

7. Бойко В. В., Савинков В. М. Проектирование баз данных информационных систем. 2-е изд. — М.: Финансы и статистика, 1989. — 350 с.
8. Тиори Т., Фрай Дж. Проектирование структур баз данных. — М.: Мир, 1985.
9. Jibo Xie, Guoqing Li Implementing next-generation national earth observation data infrastructure to integrate distributed big earth observation data // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2016. DOI: 10.1109/IGARSS.2016.7729042.
10. Mamatha Y. N., Ananth A. G. Content based image retrieval of satellite imageris using soft query based color composite techniques // International Journal of Computer Applications. 2010, Vol. 7, No 5. DOI: 10.5120/1156-1379.
11. Vostrikov A. V., Prokofeva E. N., Goncharenko S. N., Griбанov I. V. Analytical modeling for the modern mining industry // Eurasian Mining. 2019. No 2(32). Pp. 30—35. DOI: 10.17580/em.2019.02.07.
12. Гончаренко С. Н., Ле Б. З., Стоянова И. А., Петров И. В. Моделирование параметров инновационных водоохраных мероприятий на основе производственно-технологических показателей добычи угля на предприятиях Вьетнама // Горный журнал. — 2014. — № 9. — С. 143—146.
13. Гончаренко С. Н., Коростелев Д. Б. Системный анализ и прогноз показателей и индикаторов эффективности деятельности в сфере охраны окружающей среды и природопользования // Горный информационно-аналитический бюллетень. — 2018. — № 9. — С. 104—110. DOI: 10.25018/0236-1493-2018-9-0-104-110.
14. Гончаренко С. Н., Коростелев Д. Б. Методы и модели комплексной оценки системной связи эффективности природоохранной политики и управленческих решений в сфере использования природных ресурсов // Горный информационно-аналитический бюллетень. — 2018. — № 11. — С. 70—76. DOI: 10.25018/0236-1493-2018-11-0-70-76.
15. Умербеков Ж. Ж., Гончаренко С. Н. Обоснование эффективности внедрения целевой модели управления производственной безопасностью горнодобывающей компании // Горный информационно-аналитический бюллетень. — 2019. — № 8. — С. 225—234.
16. Prokofeva E. N., Vostrikov A. V., Shapovalenko G. N., Alvarez A. The development of effective geomonitoring for mining area with industrial review // Eurasian Mining. 2017. No 2. Pp. 61—63. DOI: 10.17580/em.2017.02.15.
17. Alvarez A., Fernandez E., Prokofeva E. N., Vostrikov A. V. The building of effective systems of training and development for mining engineers with the basis of digital technologies // Eurasian Mining. 2019. No 1(31). Pp. 49—52. DOI: 10.17580/em.2019.01.12.
18. Zotov L., Frolova N., Shum C. Gravity changes over russian river basins from GRACE / Planetary Exploration and Science: Recent Results and Advances. Berlin: Birkhauser/Springer, 2015, pp 45—59.
19. Shepel T., Grafe B., Hartlieb P., Drebenstedt C., Malovyk A. Evaluation of cutting forces in granite treated with microwaves on the basis of multiple linear regression analysis // International Journal of Rock Mechanics and Mining Sciences 2018, Vol. 107, Pp. 69—74. DOI:10.1016/j.ijrmms.2018.04.043.
20. Temkin I., Deryabin S., Konov I. Soft computing models in an intellectual open-pit mines transport control system // Procedia Computer Science, 2017. Vol. 120, pp. 411—416. DOI: 10.1016/j.procs.2017.11.257.
21. Темкин И. О., Дерябин С. А., Конов И. С., Клебанов Д. А. Метод определения состояния технологических дорог карьера при управлении взаимодействием роботизированных элементов горнотранспортного комплекса // Горный журнал. — 2018. — № 1. — С. 78—82.
22. Prokofeva E. N., Vostrikov A. V., Fernandez E., Borisov N. Navigation satellite systems as the audit foundation for mining companies // Eurasian Mining. 2017. No 1. Pp. 30—32. DOI: 10.17580/em.2017.01.08.

23. Рьльникова М. В., Галченко Ю. П. Возобновляемые источники энергии при комплексном освоении недр. — М.: ИПКОН РАН, 2015. — 122 с.

24. Rylnikova M., Ainbinder I., Radchenko D. Role of safety justification of mining development for the regulatory framework formation and mineral resources management // E3S Web of Conferences. 2018. 41, Article 01033. DOI: 10.1051/e3sconf/20184101033. **MIAB**

## REFERENCES

1. Chadalawada J., Espinoza-Molina D., Datcu M. Assessment of earth observation data content based on data compression — application to settlements understanding. *IEEE International Geoscience and Remote Sensing Symposium*. 2012, pp. 6130—6133. DOI: 10.1109/igars.2012.6352207.

2. Lionel Gueguen Classifying compound structures in satellite images: a compressed representation for fast queries. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, Vol. 53, pp. 1803—1818. DOI: 10.1109/tgrs.2014.2348864.

3. Hansaem Park, Kwangseob Kim, Kiwon Lee Geo-data visualization on online and offline mode of mobile web using HTML5 / *4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*. 2016. pp. 237—240. DOI: 10.1109/eorsa.2016.7552804.

4. *Scrapy 1.4 documentation*. 2019. Available at: <https://docs.scrapy.org/en/latest> (accessed 21 May 2019).

5. Olston C., Najork M. Web crawling. *Foundations and Trends in Information Retrieval*, 2010, vol. 4, no 3, pp. 175—246.

6. Hsieh J. M., Gribble S. D., Levy H. M. The architecture and implementation of an extensible web crawler / *Proceedings of the 7th USENIX Symposium on Networked Systems Design and Implementation*, NSDI 2010, April 28—30, 2010, San Jose, CA, USA.

7. Boyko V. V., Savinkov V. M. *Proektirovanie baz dannykh informatsionnykh sistem*. 2-e izd. [Designing databases of information systems. 2nd edition], Moscow, Finansy i statistika, 1989, 350 p.

8. Tiori T., Fry J. *Proektirovanie struktur baz dannykh* [Designing database structures], Moscow, Mir, 1985.

9. Jibo Xie, Guoqing Li Implementing next-generation national earth observation data infrastructure to integrate distributed big earth observation data. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2016. DOI: 10.1109/IGARSS.2016.7729042.

10. Mamatha Y. N., Ananth A. G. Content based image retrieval of satellite imageris using soft query based color composite techniques. *International Journal of Computer Applications*. 2010, Vol. 7, No 5. DOI: 10.5120/1156-1379.

11. Vostrikov A. V., Prokofeva E. N., Goncharenko S. N., Gribanov I. V. Analytical modeling for the modern mining industry. *Eurasian Mining*. 2019. No 2(32). Pp. 30—35. DOI: 10.17580/em.2019.02.07.

12. Goncharenko S. N., Le Binh Duong, Stoyanova I. A., Petrov M. V. Modeling of parameters of innovation water-protection measures on the basis of industrial-technological indices of coal mining at Vietnam enterprises. *Gornyi Zhurnal*, 2014, no 9, pp. 143—146.

13. Goncharenko S. N., Korostelev D. B. System analysis and prediction of performance efficiency figures and indicators in the area of environmental protection and nature management. *Gornyy informatsionno-analiticheskiy byulleten'*. 2018, no 9, pp. 104–110. [In Russ]. DOI: 10.25018/0236-1493-2018-9-0-104-110.

14. Goncharenko S. N., Korostelev D. B. Methods and models for integrated estimation of system connection between the nature protection policy efficiency and managerial solutions in the field of use of natural resources. *Gornyy informatsionno-analiticheskiy byulleten'*. 2018, no 11, pp. 70—76. [In Russ]. DOI: 10.25018/0236-1493-2018-11-0-70-76.

15. Umerbekov Zh. Zh., Goncharenko S. N. Validation of efficiency of the target production safety management model introduction in the mining industry. *MIAB. Mining Inf. Anal. Bull.* 2019;(8):225-234. [In Russ]. DOI: 10.25018/0236-1493-2019-08-0-225-234.

16. Prokofeva E. N., Vostrikov A. V., Shapovalenko G. N., Alvarez A. The development of effective geomonitoring for mining area with industrial review. *Eurasian Mining*. 2017. No 2. Pp. 61 – 63. DOI: 10.17580/em.2017.02.15.

17. Alvarez A., Fernandez E., Prokofeva E. N., Vostrikov A. V. The building of effective systems of training and development for mining engineers with the basis of digital technologies. *Eurasian Mining*. 2019. No 1(31). Pp. 49 – 52. DOI: 10.17580/em.2019.01.12.

18. Zotov L., Frolova N., Shum C. Gravity changes over russian river basins from GRACE. *Planetary Exploration and Science: Recent Results and Advances*. Berlin: Birkhauser/Springer, 2015, pp 45 – 59.

19. Shepel T., Grafe B., Hartlieb P., Drebenstedt C., Malovyk A. Evaluation of cutting forces in granite treated with microwaves on the basis of multiple linear regression analysis. *International Journal of Rock Mechanics and Mining Sciences*. 2018, Vol. 107, Pp. 69 – 74. DOI:10.1016/j.ijrmms.2018.04.043.

20. Temkin I., Deryabin S., Konov I. Soft computing models in an intellectual open-pit mines transport control system. *Procedia Computer Science*, 2017. Vol. 120, pp. 411 – 416. DOI: 10.1016/j.procs.2017.11.257.

21. Temkin I. O., Deryabin S. A., Konov I. S., Klebanov D. A. Method of determining the state of the haul road career in the management of the interaction between robotic elements of the mining transportation complex. *Gornyi Zhurnal*. 2018, no 1, pp. 78 – 82. [In Russ].

22. Prokofeva E. N., Vostrikov A. V., Fernandez E., Borisov N. Navigation satellite systems as the audit foundation for mining companies. *Eurasian Mining*. 2017. No 1. Pp. 30 – 32. DOI: 10.17580/em.2017.01.08.

23. Ryl'nikova M. V., Galchenko Yu. P. *Vozobnovlyaemye istochniki energii pri kompleksnom osvoenii nedr* [Renewable sources of energy in integrated subsoil development], Moscow, IPKON RAN, 2015, 122 p.

24. Ryl'nikova M., Ainbinder I., Radchenko D. Role of safety justification of mining development for the regulatory framework formation and mineral resources management. *E3S Web of Conferences*. 2018. 41, Article 01033. DOI: 10.1051/e3sconf/20184101033.

## ИНФОРМАЦИЯ ОБ АВТОРАХ

Некрасов Глеб Александрович<sup>1</sup> – студент, e-mail: alax-27@mail.ru,

Поливода Денис Эдуардович<sup>1</sup> – студент,  
e-mail: depolivoda@edu.hse.ru,

Прокофьева Екатерина Николаевна<sup>1</sup> – канд. техн. наук,  
профессор, e-mail: eprokofyeva@hse.ru,

<sup>1</sup> Национальный исследовательский университет  
«Высшая школа экономики».

**Для контактов:** Прокофьева Е.Н., e-mail: eprokofyeva@hse.ru.

## INFORMATION ABOUT THE AUTHORS

G.A. Nekrasov<sup>1</sup>, Student, e-mail: alax-27@mail.ru,

D.E. Polivoda<sup>1</sup>, Student, e-mail: depolivoda@edu.hse.ru,

E.N. Prokofyeva<sup>1</sup>, Cand. Sci. (Eng.), Professor,  
e-mail: eprokofyeva@hse.ru,

<sup>1</sup> Higher School of Economics. National Research University,  
143072, Moscow, Russia.

**Corresponding author:** E.N. Prokofyeva, e-mail: eprokofyeva@hse.ru.

Получена редакцией 08.02.2020; получена после рецензии 17.03.2020; принята к печати 20.04.2020.

Received by the editors 08.02.2020; received after the review 17.03.2020; accepted for printing 20.04.2020.