УДК 004. 553.7

ОБРАБОТКА ГЕОЛОГИЧЕСКИХ ДАННЫХ ГОРНОЙ ПРОМЫШЛЕННОСТИ КЫРГЫЗСТАНА И ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ BIG DATA

Кабаева Гульнара Джамалбековна, *д.ф.-м.н, профессор, декан факультета ИТ, КГТУ им. И.Раззакова, Кыргызстан, 720044, г.Бишкек пр.Мира 20, kabgd@mail.ru*
Аманкулова Нургуль Асимкановна, *к.ф.-м.н, доцент, ИГД и ГТ им. У. Асаналиева, Кыргызстан, 720010, г.Бишкек пр.Чуй 215, a_nur4@mail.ru*

Аннотация. В данной работе на примере создания web -приложения по геотермальным источникам Кыргызстана обсуждаются этапы разработки и принципы построения web-ориентированной геоинформационной системы, а также возможности современных технологий для сбора, накопления, систематизации и представления данных горнодобывающей промышленности Кыргызстана. Постоянное накопление такого рода

данных требует внедрения новейших технологий анализа и хранения больших объемов данных, т.е. технологий BIG DATA.

Ключевые слова: *ГИС, веб-сервис, геопространственные данные, web - ориентированная ИС, веб-приложение, веб-картография, Big Data, Hadoop. база данных.*

PROCESSING OF GEOLOGICAL DATA OF MINING INDUSTRY OF KYRGYZSTAN AND PROSPECTS OF APPLICATION OF TECHNOLOGIES BIG DATA

Kabaeva Gulnara Jamalbekovna, *Doctor of Physics and Mathematics, Professor, Dean of the Faculty of IT, KSTU. I.Razzakova, Kyrgyzstan, 720020, Bishkek, Prospect Mira 20, kabgd@mail.ru*
Amankulova Nurgul Asimkanovna, Ph.D., *Associate Professor, IGD and GT them. U. Asanalieva, Kyrgyzstan, 720017, Bishkek Ave. Chuy 215, a_nur4@mail.ru*

Annotation. In this paper are discussed the development stages and principles for building a web-based geoinformation system, as well as the capabilities of modern technologies for collecting, accumulating, systematizing and reporting mining industry data in Kyrgyzstan using the example of creating a web-application on geothermal sources in Kyrgyzstan. Constant accumulation of such data requires the introduction of new technologies for analyzing and storing large amounts of data, i.e. BIG DATA technologies.

Keywords: *GIS, web service, geospatial data, web-oriented IS, web application, web cartography, Big Data, Hadoop. database.*

Введение. В настоящее время объемы данных горнодобывающей промышленности постоянно растут и внедрение новейших технологий анализа и обработки больших данных, остается для нашей Республики делом времени и возможностей. К тому же, согласно Национальной долгосрочной стратегией развития экономики нашей Республики, повышению качества и внедрению современных информационных технологий будет уделяться большое внимание. Постоянное накопление данных приводит к созданию баз и банков данных такого объема, когда существующие технологии экспертных систем оказываются неприменимы для их обработки, в силу различных объективных причин. И наиболее большие перспективы в этом направлении определились для технологий обработки больших данных (BIG DATA) и машинного обучения (Machine Learning), которые получают широкое распространение в развитых странах, как свидетельствуют интернет публикации на различных бизнес сайтах.

Однако внедрение технологий работы с большими данными требует создания дорогостоящего Центра обработки данных (Дата-центра), по крайней мере, малого размера (80-100 серверов). Например, широко используемый проект с открытым программным обеспечением Hadoop представляет собой структуру с библиотекой программного обеспечения Apache Hadoop, позволяющей распределенную обработку больших наборов данных в кластерах компьютеров и предназначен для масштабирования от отдельных серверов до тысяч компьютеров, каждый из которых предлагает локальные вычисления и хранение. Hadoop состоит из компонентов:

1) распределенной файловой системы Hadoop (англ. HDFS), которая отвечает за хранение данных на кластере Hadoop;

2) Hadoop YARN : структуры планирования рабочих мест и управления ресурсами кластера;

3) системы MapReduce, предназначенной для вычислений и обработки больших объемов данных на кластере (в этом пакете система на основе YARN для параллельной обработки больших наборов данных);

4) Hadoop Ozone: хранилище объектов для Hadoop; а также общие утилиты для поддержки модулей Hadoop. Hadoop реализует вычислительную парадигму с именем Map / Reduce , где приложение разделено на множество небольших фрагментов работы, каждый из которых может быть выполнен или повторно выполнен на любом узле кластера [3,8].

Одним из ключевых моментов является сбор данных, и в данном случае, мы можем решать эти задачи с помощью современных геоинформационных систем (ГИС), которые подходят практически для всех видов деятельности горнодобывающей промышленности.
В этом случае решается проблема обработки, оценивания и визуализации информации, представленной статистическими данными [6].

В геологии Кыргызстана имеется огромное количество разнородной информации полученной из геологических изысканий, при этом не все данные переведены в электронный вид. В связи с этим внедрение современных информационных технологий для систематизации, хранения больших массивов геологических данных, а также представления геологической информации в современном и удобном виде является актуальной проблемой, особенно в связи с развитием цифровых технологий, которые имеют возможность передать информации различного формата в реальном времени.

Цель работы состоит в разработке информационной системы на основе технологий ГИС для накопления, систематизации, представления и обработки геотермальных данных Кыргызстана, предоставляющих возможности для дальнейшего анализа каких-либо закономерностей, моделирования и описания природы объекта.

Описание задачи. Создание геоинформационной системы опирается на сбор, структуирование и анализ геотермальных ресурсов в виде карт по территории Кыргызской Республики, которые послужат основой для оценки и прогноза состояния и развития использования данного вида источника энергии. При разработке геоинформационных систем совокупность данных рассматриваемого объекта составляют такие данные как: количество точек, состав, количество статических данных и др., они имеют довольно большие по объему и разнообразные по составу характеристики.

Ресурсы геотермальных вод в Кыргызстане относятся к нетрадиционным возобновляемым источникам энергии (НВИЭ), что очевидно составляет немаловажную причину для сбора данных о них и исследования альтернативных источников энергии.
В качестве входных данных использовались данные из реесстра месторождений и проявлений полезных ископаемых КР. В области геологии традиционным представлением информации являются карты, таблицы или диаграммы.

Геотермальные источники распространены во многих горных районах республики. Их известно около 25 групп, но наиболее известны геотермальные источники в Жалалабаде, Ысыкате, Аксуу, Жетыогузе, Теплоключенке и других местах [1]. Большинство геотермальных источников расположены в курортной зоне При Иссык- кулья, где к настоящему времени пробурено около 40 скважин глубиной от 1300 до 2500 м с максимальным дебитом от 1—3 до 37 л/с и температурой воды от 40 до 60°С. Минерализация этих вод составляет от 0,3—1,2 до 40 г/л [4]. Среди этих месторождений Карабулунский, Джыргаланский, Кошкөлский, Чаекский, Курский, Чолпонатинский, Алтынарашанский, Барбулакский и другие источники, их суммарные естественные ресурсы составляют 35,3 тыс.м3 [5].

Обсуждение результатов. В основе разрабатываемой информационной системы по геотермальным водам лежат СУБД MySQL и ГИС- технологии. На данный момент разработана база данных и web -интерфейс которой были написаны на двух языках программирования: PHP и JavaScript. А также использован язык прикладного интерфейса программирования API (Application Programming Interface) для подключения к web-картографической информации. Процесс разработки веб-ориентированной информационной системы в общем случае состоит из следующих этапов, приведенных на рис.1.

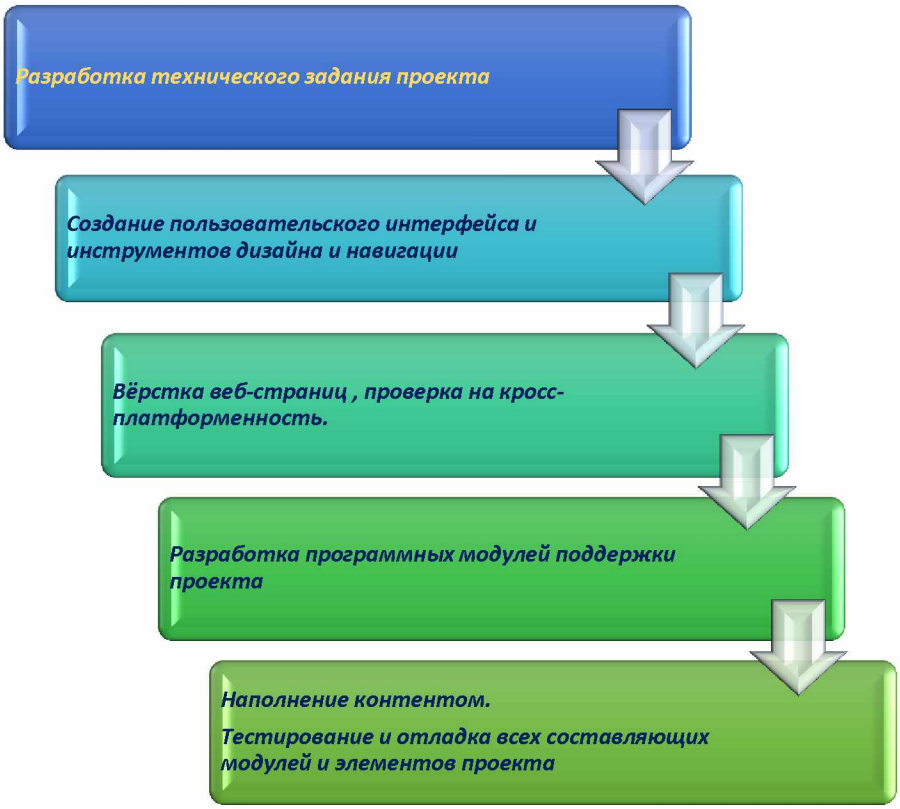

Рис.1. Этапы разработки web-приложений.

На каждом этапе требуется проведение подробного комплексного анализа, определяющего критерии, которым должен соответствовать будущий web-проект. Интерфейс и общая информация по исследованию геотермальных мест Кыргызстана представлены на геопортале как веб-страница (рис. 2).

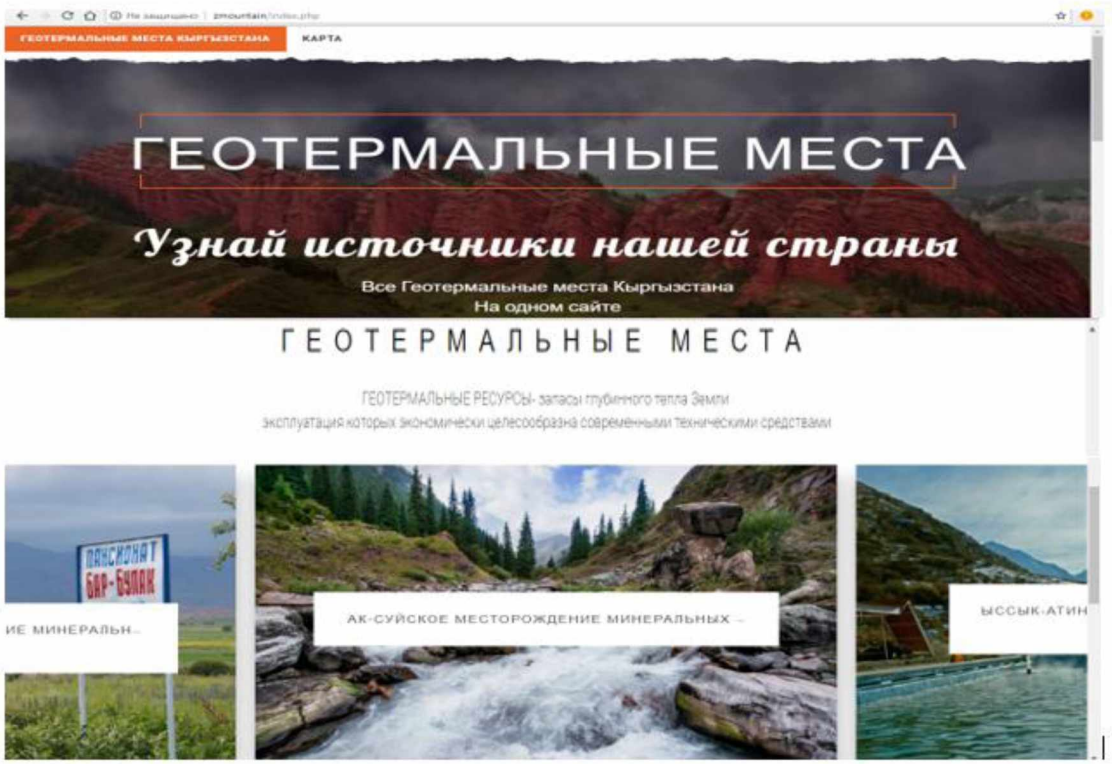

Рис.2. Web-интерфейс геотермальных источников Кыргызстана.

Стандартный веб-браузер обеспечивает возможность интерактивной навигации по картографическому изображению, представленному на рис. 2. с изменением масштаба, построение запросов по объектам карты щелчком мыши, управление видимостью слоев карты в легенде, выбор картографической основы – карты различных поставщиков, мозаики спутниковых снимков, цифровой рельеф, и проч.

Предварительную схему работы картографического сервиса можно представить в виде стандартного алгоритма, приведенного на рис.3 [2].

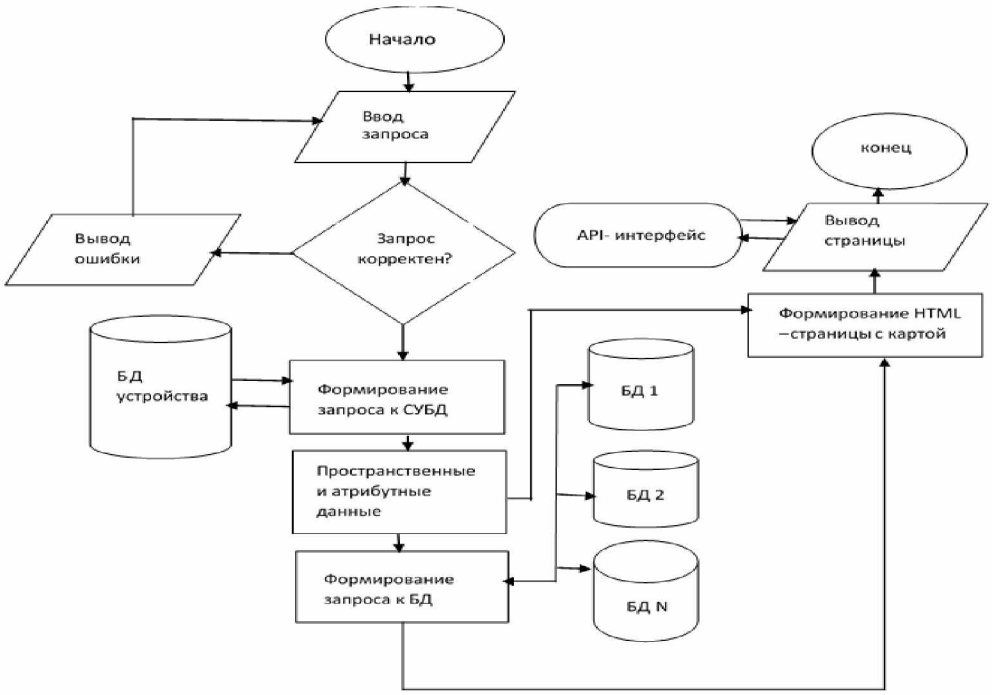

Рис.3. Алгоритм работы веб-гис сервера при вводе запроса.

Результат картографического представления геотермальных данных приведено на рис.4.

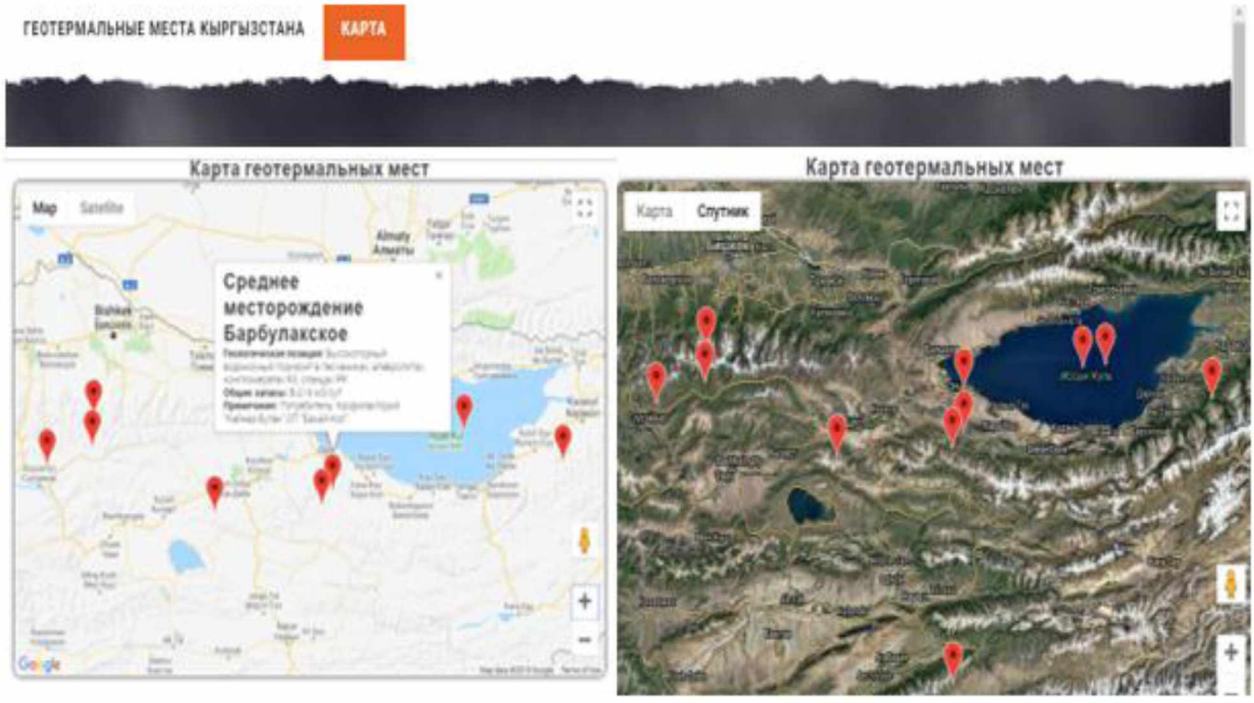

Рис.4. Картографический web-сервис по источникам геотермальных мест.

База данных разработана на СУБД MySQL, для создания модуля управления базой данных был использован "phpMyAdmin" , вид окна которого показан на рис.5 .

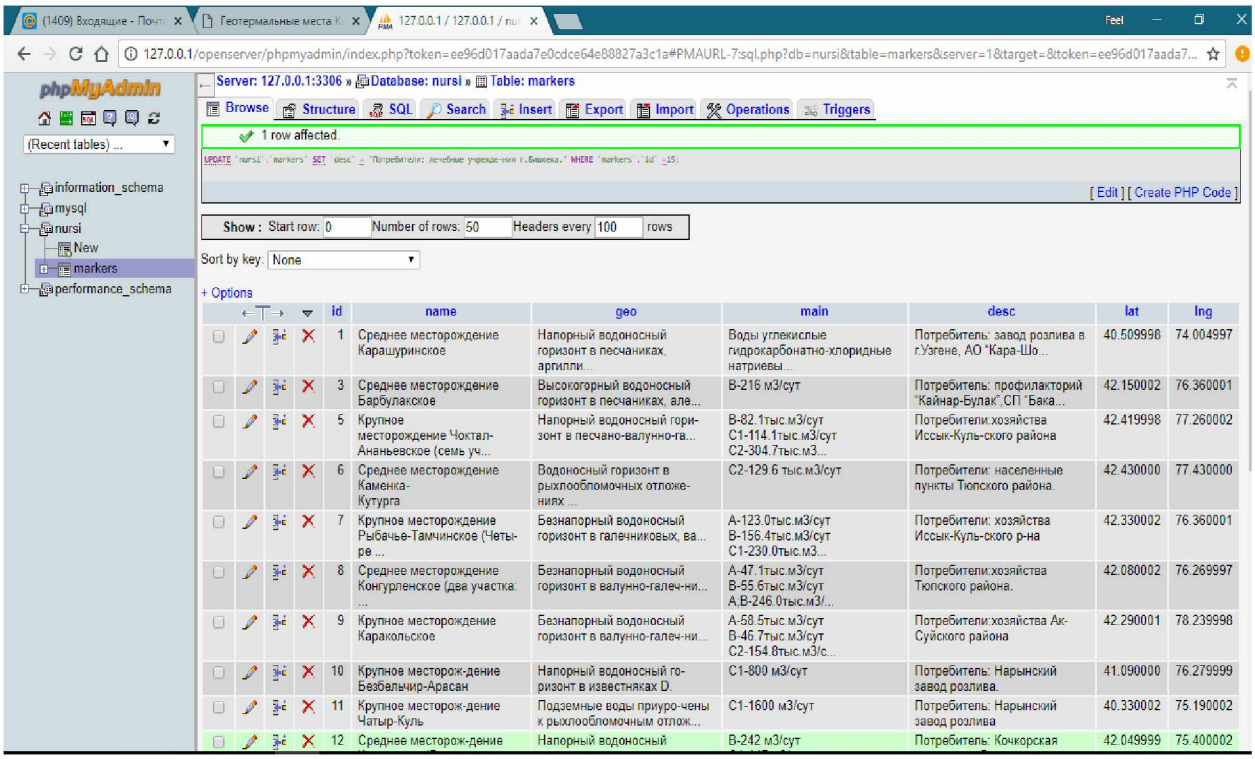

Рис.5. Вид окна создания базы данных.

В процессе разработки геоинформационной системы по геотермальным источникам Кыргызстана проведен обзор по данным исследований [7], в которых отмечено, что геотермальные источники в быту используются недостаточно в связи с некоторыми объективными причинами, связанными с отсутствием возможности приобретения соответствующего оборудования в зависимости от состава термальных вод и др.

В связи с этим введение базы данных и картографического сервиса на основе геологических исследований геотермальных мест с отображением состава и зоны охвата могут быть полезными для проведения статистического анализа и мониторинга использования соответствующего оборудования для геотермальных вод Кыргызстана.

Организация различного рода информационных систем, баз и банков данных, развитие и совершенстование методов обработки и представления создают основу для проведения исследований на качественно новом уровне. Во многих случаях современные информационные технологии представления данных, прежде всего развитые базы данных, позволяют получать новые научные результаты и решать такие проблемы, которые в отсутствие этих баз данных не только не могли бы быть решены. Внедрение современных информационных технологий в различных направлениях горно-геологических наук, повышает эффективность научных исследований.

Выводы. Таким образом, на примере создания web -приложения по геотермальным источникам Кыргызстана рассмотрены этапы разработки и принципы построения web-ориентированной геоинформационной системы, а также возможности современных технологий для сбора, накопления, систематизации и представления данных горнодобывающей промышленности Кыргызстана. Поскольку данные геологических исследований Кыргызстана огромны, но недостаточно переведены на электронный вид, таким же образом можно внедрение возможностей современных информационных технологий для их обработки.

Дальнейшая работа в этом направлении предусматривает изучение и применение программ обработки больших данных Hadoop и ГИС. А также, необходимо развивать инструментальную часть пользовательского интерфейса системы, для расширения его функциональных возможностей.

Разработанную систему можно использовать для стратегического анализа и обработки статистических данных, для создания базы данных исходной и расчетной информации; и для оценки состояние использования геотермальных источников для нужд народного хозяйства. Описание технологии разработки ГИС и саму БД можно использовать как методические указания в учебном процессе при изучении специальных дисциплин: Проектирование ИС, Технология обработки данных, ИС в минералогии, геологии для студентов ИГДиГТ.

Литература

1. Беляков Ю. П., Рахимов К.Р. Энергетические ресурсы Кыргызстана и их использование. – Бишкек: Илим. 1993.
2. Вагизов М.Р. Разработка интерактивного картографического сервиса для определения лесотаксационных показателей насаждений программно-техническим методом.// Диссертация на соискание ученой степени к.т.н. Санкт-Петербург – 2016.
3. BIG DATA – путь к умному производству// http://www.umpro.ru Выпуск 41 от 03.18.
4. Использование возобновляемых источников энергии в практике народного хозяйства республики. /Тезисы докладов республиканской научно-практической конференции. – Фрунзе: Илим. 1988, С.180.
5. Иссык-Куль. Нарын. Энциклопедия. – Фрунзе: гл. ред. М. Борбугулов, КСЭ, 1991,С. 512.
6. Очертание местности: роль геопространственных данных в горнодобывающей отрасли //ArcReview, № 3 (74), 2015.
7. Эркимбаев М.Т. Экономическая сущность нетрадиционных возобновляемых источников энергии.// https://docplayer.ru С.1-6.
8. http://hadoop.apache.org/.