Федеральное государственное бюджетное образовательное учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»

Каукин А.С., Миллер Е.М.

Разработка пространственного и технологического атласа энергетической системы Российской Федерации

Москва 2020

Аннотация. Структура работы представлена четырьмя разделами. В первом разделе представлен обзор подходов и примеров реализации в мировой практике атласа энергетической системы. Во втором анализируется современное состояние и основные проблемы развития энергетической системы России. В третьем разделе рассматривается методология сбора и структуризации данных, а также описывается архитектура собранной базы данных по энергетической системе России. В четвертом разделе в качестве демонстрации возможностей собранной базы данных рассчитываются показатели удельного расхода топлива для генерации тепла и электроэнергии для каждого агрегата электростанций России, средний показатель по всем станциям России, а также средний показатель в региональном разрезе.

Ключевые слова: энергетика, инфраструктурные отрасли, экология, база данных, атлас энергетической системы, взаимосвязь энергетических объектов.

Abstract.The structure of the work is presented in four sections. The first section provides an overview of approaches and examples of the implementation of the atlas of the energy system in world practice. The second analyzes the current state and the main problems of the development of the energy system of Russia. The third section discusses the methodology for collecting and structuring data, as well as describes the architecture of the collected database on the energy system of Russia. In the fourth section, as a demonstration of the capabilities of the collected database, specific fuel consumption indicators for generating heat and electricity for each unit of Russian power plants are calculated, an average indicator for all stations in Russia, as well as an average indicator in a regional context.

Key words: energy, infrastructure industries, ecology, database, atlas of the energy system, the relationship of energy objects.

Каукин А.С. заведующий научно-исследовательской лабораторией системного анализа отраслевых рынков ИОРИ Российской академии народного хозяйства и государственной службы при Президенте РФ;

Миллер Е.М. старший научный сотрудник лаборатории системного анализа отраслевых рынков Российской академии народного хозяйства и государственной службы при Президенте РФ.

Данная работа подготовлена на основе материалов научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2019 год

СОДЕРЖАНИЕ

ВВЕДЕНИЕ....4
1 Обзор подходов и примеры реализации в мировой практике атласа энергетической системы....5
2 Анализ современного состояния, основных проблем и планов развития энергетической системы России....17
2.1 Анализ современного состояния энергетической системы России в сфере генерации энергии....17
2.2 Анализ современного состояния энергетической системы России в сфере потребления энергии....23
3 Создание базы данных географического расположения. технологического оснащения и взаимосвязей объектов энергетической системы России (Атлас энергетической системы России)....25
3.1 Методология сбора и структуризации данных....25
3.2 Описание архитектуры собранной базы данных по энергетической системе России....30
4 Определение удельного расхода топлива на выработку электрической энергии и теплоты на тепловых электростанциях....47
ЗАКЛЮЧЕНИЕ....53
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ....57

ВВЕДЕНИЕ

На сегодняшний день не существует единой базы данных, карты или атласа, которые содержали бы полную информацию о географическом расположении, технологическом оснащении (характеристики производственных функций) и взаимосвязях объектов энергосистемы РФ. Составление такой базы данных позволит кардинальным образом увеличить точность экономического моделирования в энергетической сфере, в частности, моделей полного или частичного равновесия для отдельных отраслевых и региональных рынков. Кроме того, учет детализированных данных позволит повысить качество соответствующих прогнозов и уточнить расчет последствий экономической политики в сфере энергетики и промышленной политики в целом.

Энергетическая система Российской Федерации характеризуется сложной географической структурой, на которую оказали влияние как расположение месторождений топливно-энергетических полезных ископаемых и других источников энергии, так и исторически сложившиеся центры концентрации населения, производства и потребления продукции, при изготовлении которой используются как отдельные специфические для конкретной отрасли виды энергоресурсов, так и электроэнергия и тепло. Элементы системы также являются весьма технологически неоднородными в силу разнообразия перерабатываемых источников энергии и длительной истории существования и эволюции энергосистемы.

1 Обзор подходов и примеры реализации в мировой практике атласа энергетической системы

Появление технологических возможностей обработки и хранения больших массивов данных создает спрос на их удобное представление, позволяющее проводить быстрый анализ исследуемой системы. Энергетическая отрасль является сложной системой, состоящей из большого числа объектов: добывающих первичные источники энергии месторождений, мощностей по их переработке в топливо, мощностей по генерации электрической и тепловой энергии, а также инфраструктуры, связывающей их. Визуализация на карте перечисленных выше объектов позволяет за короткий период времени оценить географическую концентрацию тех или иных полезных ископаемых, технологическое оснащение (характеристики производственных функций) и инфраструктурные ограничения регионов по их торговым потокам и т.п. Однако, на текущий момент нет подобной карты или атласа для энергетической системы России.

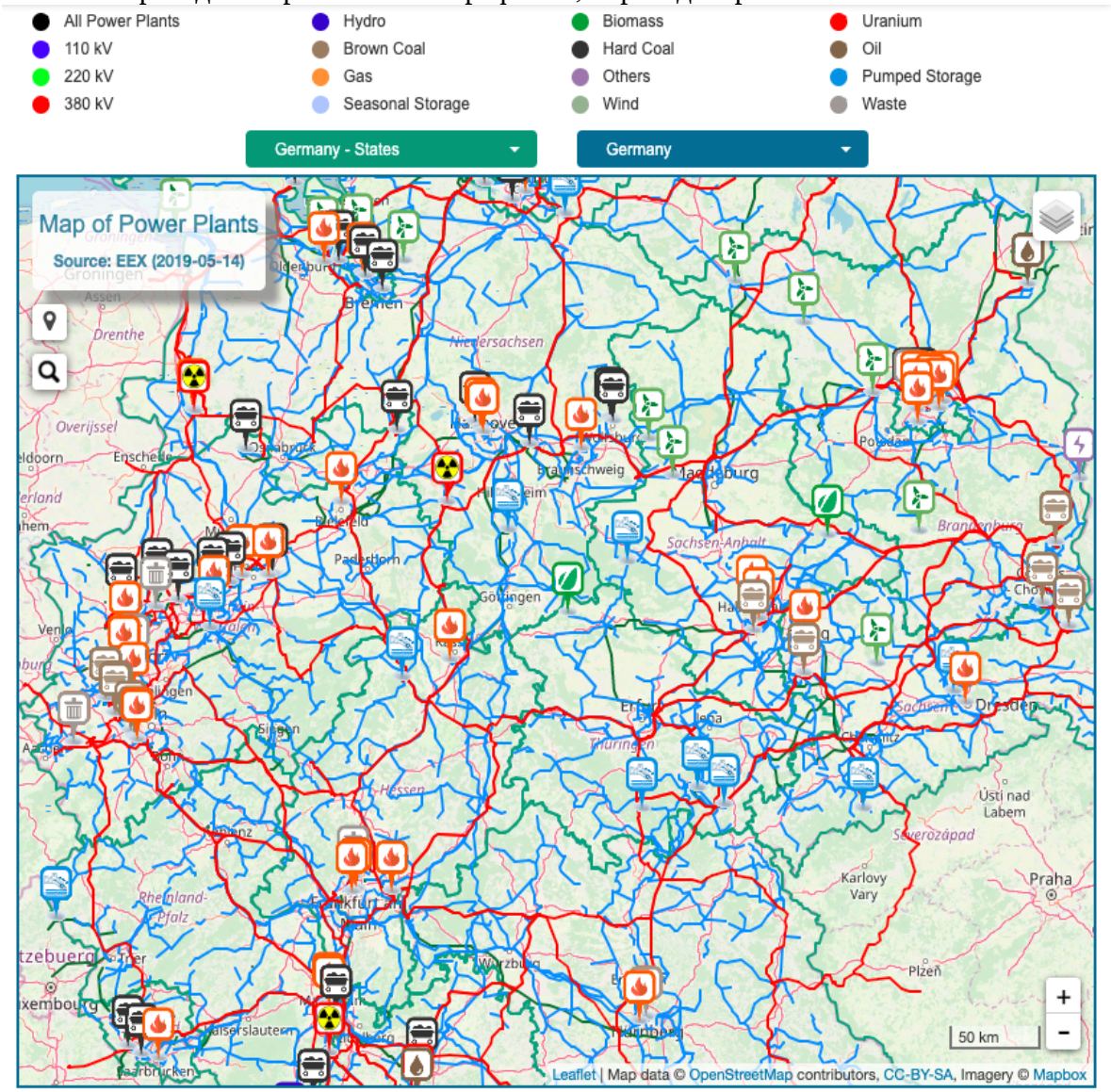
В разделе рассмотрены имеющиеся мировые аналоги Атласа энергетической системы: описывается их визуализация, основные плюсы и минусы, имеющийся функционал. В качестве аналогов к реализуемому в данной работе Атласу энергетической системы Российской Федерации, можно выделить:

- Map of Power Plants, интерактивную карту энергетических станций Европы (детализировано по Германии, остальные в разработке),
- Pan-European Thermal Atlas, атлас тепловых ресурсов Европы (представлено 14 европейских стран, что покрывает 90% рынка тепловой энергии Европы),
- EIA Maps, интерактивные карты Управления энергетической информации США (в разрезе штатов и мира),
- North American Energy Maps, Интернет-ресурс, содержащий статические и интерактивные карты по производственным мощностям нефтегазовой отрасли, отрасли производства электрической энергии и ВИЭ,
- Energy Web Atlas, атлас мирового рынка природного газа и СПГ.

Интерактивная карта энергетических станций Европы (Map of Power Plants)

На рисунке представлен снимок карты энергетических станций Германии, аналогичные объекты рисуются в границах всех территориальных делений Германии, а также для некоторых европейских стран: Австрии, Бельгии, Швейцарии, Чехии, Англии, Венгрии, Нидерландов и Словакии (без учета территориальных делений, только в целом по стране). На карте представлены станции в разрезе по используемому виду топлива, а также инфраструктура (линии передач), разделённая по напряжению. При поднесении курсора на тот или иной объект всплывает окно с краткой информацией по объекту (полное название, установленная мощность, основное топливо, страна). При нажатии на

объект появляется окно с дополнительными идентификаторами в базе данных ресурса и кнопками: переход к энергетическим графикам, переход к сравнению.



Примечание – Источник: Официальный сайт Fraunhofer ISE [CITATION Офи195 \l 1049].

Рисунок 1 – Интерактивная карта энергетических станций Германии

Источниками данных для часовых значений являются:

– EEX Transparency Platform – производство электроэнергии. Данные по всем видам топлива публикуются ежечасно с максимальной задержкой в один час. Прогноз на следующий день публикуется ежедневно в 18:00. Производство электроэнергии по виду топлива также публикуется ежечасно, однако, фактическое производство на единицу продукции отображается с задержкой в 5 дней из-за нормативных причин [CITATION EEX19 \l 1049],

– ENTSOE Transparency Platform – импорт и экспорт электроэнергии. Данные по всем попарным потокам между европейскими странами, предоставляемыми системными операторами [CITATION ENT19 \l 1049],

– EPEX SPOT – спотовые цены на электроэнергию. Данные по всем сделкам рынка Германии, Австрии, Швейцарии и Франции (включая 15-минутный контракт для Германии, Австрии и Швейцарии), данные аукциона на сутки вперед для Германии / Австрии, Франции и Швейцарии [CITATION Epe19 \l 1033].

Источником ежемесячной статистики является Федеральное статистическое управления Германии [CITATION Sta19 \l 1049], годовых – Федеральное министерство экономики и энергетики Германии [CITATION Fed19 \l 1049].

На официальном сайте ресурса, помимо интерактивной карты энергетических станций Европы, можно построить несколько интерактивных графиков, отображающих производство электроэнергии и цены на спотовом рынке в Германии. Сайт позволяет настраивать графики в соответствии с потребностями: можно выбрать один или несколько источников энергии или переключаться между графиками с абсолютными или процентными значениями. Числовые значения, отображаемые на графиках, можно просмотреть во всплывающем окне. Кроме того, можно выбрать период времени для просмотра и частоту данных (месяц, неделя). Легенда в верхней части каждого графика показывает доступные параметры, которые активируются нажатием. Открытые кружки в легенде указывают, что параметр в настоящее время не отображается, но его можно вставить, щелкнув по нему.

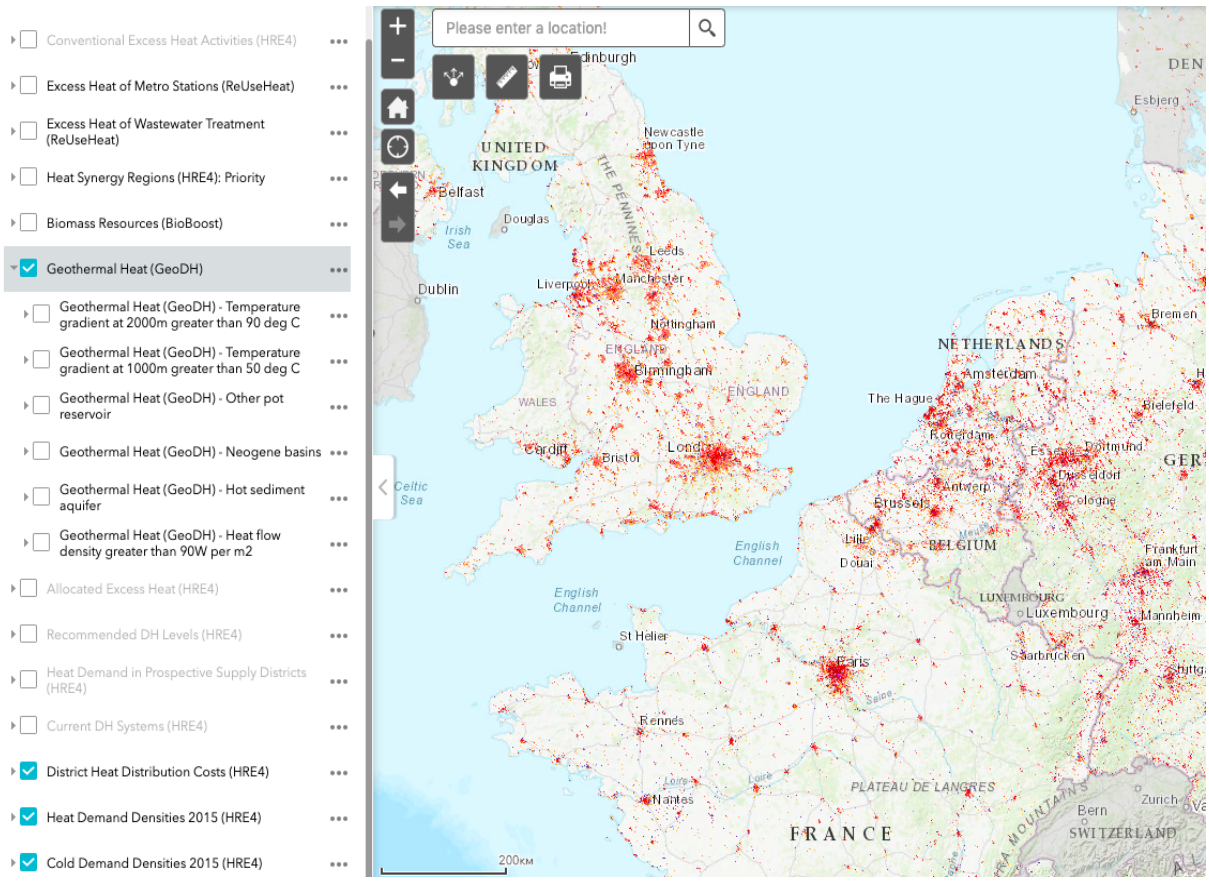
Для визуализации интерактивной карты использованы библиотеки javascript: D3.js управление данными документа, NVD3 повторно используемые графики для D3.js.

На ресурсе представлена новостная лента, которая сообщает о всех происходящих изменениях, также публикуются ежегодные отчеты о производстве электроэнергии в Германии.

Если говорить о минусах проекта, то можно выделить отсутствие возможности в выгрузке данных по которым осуществляется построение всех интерактивных графиков и нанесение на карту производственных объектов.

Атлас тепловых ресурсов Европы (The Pan-European Thermal Atlas – Peta)

Атлас тепловых ресурсов Европы (Peta) – это интерактивная карта, включающая около 90% всего европейского рынка тепла. Ресурс представляет собой первую европейскую цифровую географическую модель потребления централизованного отопления и кондиционирования воздуха (содержит объемы спроса), также ресурс содержит данные по инвестиционным затратам на инфраструктуру и перспективным площадкам поставок (размер сетки 100 м). Модель использует открытые данные Европейской комиссии, Европейского агентства по окружающей среде и Евростата. На рисунке представлен скриншот карты и ее возможных слоев.



Примечание – Источник: Официальный сайт The Pan-European Thermal Atlas – Peta [CITATION Офи197 \l 1049].

Рисунок 2 – Интерактивная карта тепловых ресурсов Европы

Атлас содержит карту по 14 странам Европейского союза[1], которая имеет следующие слои:

- Объемы избыточного тепла, разделенные по категориям промышленность, электростанции, выработка энергии с использованием отходов,
- Объемы избыточного тепла от станций метро,
- Объемы избыточного тепла при переработке мусора,
- Карта показателя синергии региона,
- Ресурсы биомассы,
- Геотермальные источники тепловой энергии,
- Избыточное предложение на тепловую энергию,
- Рекомендованный уровень потребления тепловой энергии,
- Объемы спроса на тепловую энергию,
- Инфраструктура передачи тепловой энергии,
- Региональные расходы на распределение тепловой энергии,
- Плотность спроса на тепловую энергию в 2015 году,
- Плотность спроса на охлаждение в 2015 году.

Как уже было отмечено выше, каждый квадрат слоя 100х100 м содержит: смоделированную потребность в тепле и охлаждении; локальную плотность обоих

1 Австрии, Бельгии, Чехии, Финляндии, Франции, Германии, Венгрии, Италии, Нидерландам, Польше, Румынии, Испании, Швеции, Великобритании

потребностей; базовую геометрию сети сбытовой компании тепловой энергии; установленную мощность, имеющихся объектов; и потенциал для возобновляемых источников энергии. Также на карте представлен потенциал централизованного теплоснабжения, который был количественно оценен по сравнению с затратами на инфраструктуру, наличием избыточного тепла и возобновляемыми источниками энергии, а также географическим положением.

Источником данных берется открытая статистика, доступная на сайтах Европейской Комиссии, Европейского агентства по защите окружающей среды и Евростата.

В рамках проекта были разработаны 4 различных сценария развития энергетической системы для каждой из 14 стран. Первый сценарий является базовым, вся имеющаяся инфраструктура и мощности находятся на уровне 2015 г. Второй сценарий учитывает политику ЕС в энергопотреблении принятую в марте 2018 г. Третий сценарий – сценарий декарбонизации к 2050 г., который поощряет развитие сектора возобновляемой энергетики, но не фокусирует действие данных программ в секторах отопления и кондиционирования. Четвертый сценарий повторяет третий, за исключением того, что программы направлены и на сектор отопления и кондиционирования. Результаты представлены в открытом доступе, как детализированные (расчеты в разрезе 12 стран по всем имеющимся моделям: Peta, FORECAST, JRC-EU-TIMES, EnergyPLAN), так и агрегированные в одну общую таблицу [CITATION Hea19 \l 1049].

Основным минусом проекта является сложность анализа рассчитанных данных, так как результаты представлены не интерактивно, также их нельзя скачать, только посмотреть в отчетах и презентациях. Это увеличивает время на сбор и обработку данных у исследователя, планирующего воспроизвести результаты или посчитать что-то свое.

Интерактивные карты Управления энергетической информации США (EIA Map)

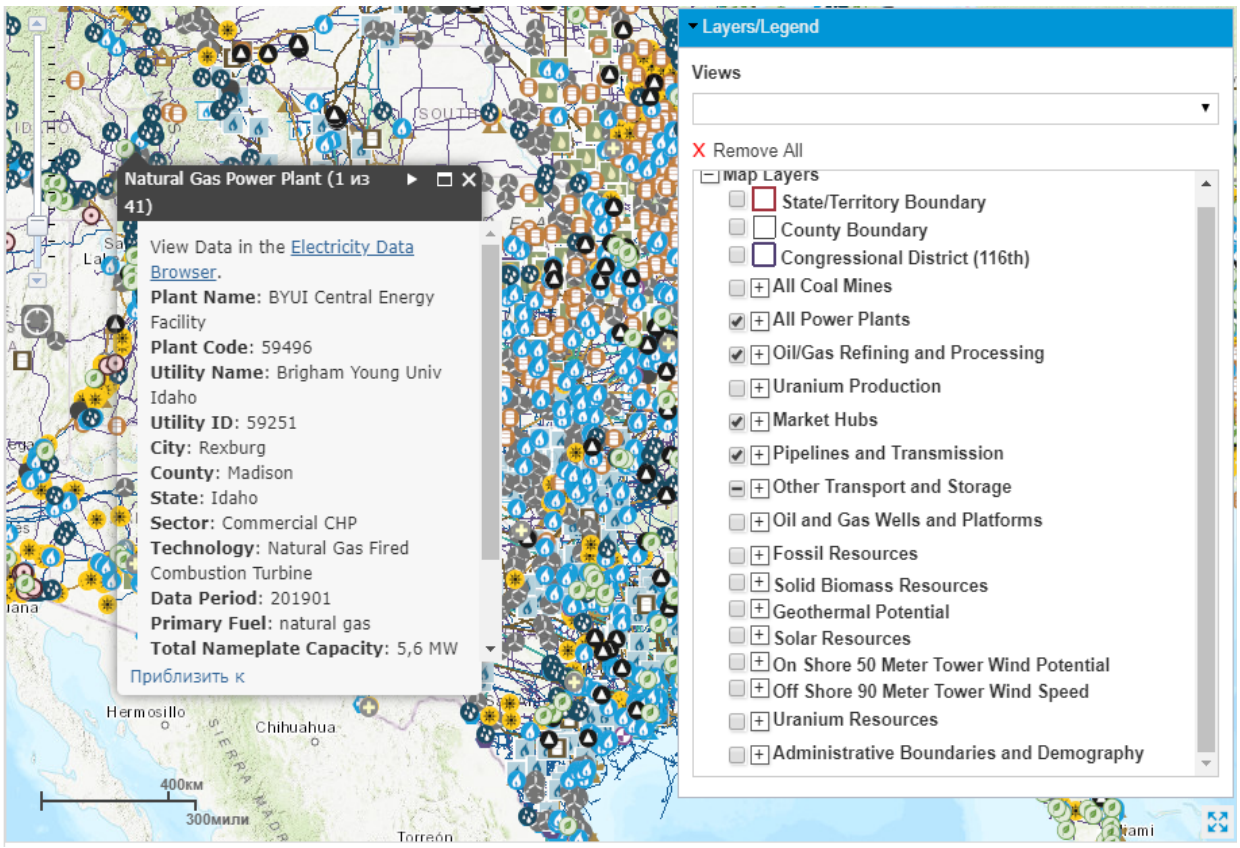
Официальный портал Управления энергетической информации США (EIA) содержит несколько интерактивных карт, в том числе с региональной разбивкой по 48 штатам:

- Карту энергетической системы США (U.S. States Portal), которая включает статистику по энергетической инфраструктуре, объемам производства энергии, объемам ее потребления и ценам,
- Карту потенциальных угроз от штормов, наводнений и повышений уровня моря для энергетической инфраструктуры (Flood Vulnerability Assessment Map). Информация транслируется в реальном времени,
- Карту залегания топливно-энергетических полезных ископаемых (Fossil Fuel Resource Map),
- Карту угольной инфраструктуры (Coal Infrastructure Map),

- Карту нефтяной инфраструктуры (Petroleum Infrastructure Map),
- Карту газовой инфраструктуры (Natural Gas Infrastructure Map),
- Карту возобновляемых источников энергии (Renewable Infrastructure Map),
- Карту электрической инфраструктуры (U.S. Electricity Infrastructure Map),
- Карту перебоев (сбоев) в энергетической инфраструктуре США (Energy Disruptions) от наводнений, штормов, лесных пожаров и землетрясений,
- Карту низко проницаемых нефтяных и газовых бассейнов (Low Permeability Oil and Gas Plays). Карта содержит статистику границ нефтегазоносных бассейнов, глубину залегания в них топливно-энергетических полезных ископаемых.

Для наиболее стратегически важных районов в плане развития энергетической системы ресурс имеет индивидуальные карты со слоями, функционал схож с региональными. В частности, подобную страницу имеет карта по Мексиканскому заливу, так как на долю добычи нефти на шельфе приходится 17% от общего объема добычи нефти в США, газа – 5%, также вдоль побережья расположено более 45% от общего объема нефтеперерабатывающих и 51% газоперерабатывающих мощностей США.

Портал EIA предоставляет достаточно широкий перечень данных по энергетической системе США, в том числе и на уровне штата: на сайте представлено более 150 интерактивных графиков, диаграмм и таблиц. Также имеется сравнение штатов по этим показателям (ранжирование). Интерактивная карта с нанесение слоев границ штатов, месторождений (угольных, нефтегазовых, урановой руды), перерабатывающих нефть заводов, трубопроводной транспортной инфраструктуры, линий железных дорог, рыночных хабов, терминалов хранения и транспортировки нефти и газа другими видами транспорта, геотермального потенциала, ВИЭ. Пример снимка экрана интерактивной карты по энергетической системе США представлен на рисунке .



Примечание – Источник: Официальный сайт U.S. Energy Information Administration (U.S. EIA) [CITATION USE19 \l 1049].

Рисунок 3 – Интерактивная карта Управления энергетической информации США

Каждая карта имеет несколько слоев. Общими слоями являются границы территориальных делений (город, штат, район конгресса). Остальная группа слоев зависят от назначения карты. Функционал аналогичен вышеописанному: при поднесении курсора на тот или иной объект выбранного слоя всплывает окно с краткой информацией по объекту (полное название, установленная мощность, основное топливо, страна, штат и т.п.). При нажатии на объект происходит переход на страницу с дополнительной информацией по объекту (браузер данных), например, если это объект генерирующий электроэнергию: объемы чистой генерации; объемы потребления топливо для генерации электроэнергии, тепловой энергии; объемы розничных продаж электроэнергии; выручка; средняя розничная цена; запасы ископаемого топлива для производства электроэнергии; средняя стоимость сырья (топлива). Все перечисленные данные представлены в ретроспективе с частотой месяц. Также на странице имеется возможность посмотреть статистику агрегировано в целом по стране в годовой, квартальной и месячной разбивке. Статистику можно скачать для дальнейшей обработки и проведения собственных исследований. Кроме того, браузер данных представлен для:

– Почасовых объемов выработки электроэнергии, спросе в регионах США, как фактические значения, так и прогнозные, т.е. представление оперативной

статистики по электроэнергетической системе США (U.S. Electric System Operating Data),

- Краткосрочных прогнозов объемов энергопоставок США (спрос, импорт/ экспорт, цены) на ближайшие 18-24 месяца,
- Прогнозных значений объемов добычи и потребления различных первичных источников энергии (AEO Table Browser). Прогнозы представлены по нескольким сценариям,
- Ретроспективной и оперативной статистики по угольной промышленности (объемы добычи угля, количество часов работы шахт, численность работников на шахтах, цены по типу залегания и виду в региональном разрезе и т.п.),
- Статистики по использованию альтернативного вида топлива (Alternative Fuel Vehicle Data) по типу и категориям собственности транспортных средств.

Аналогичные браузеры данных представлены для статистики по другим странам, однако не с такой точной детализацией – агрегация доступна на уровне всей страны.

Интерактивный Интернет-ресурс, созданный Управлением энергетической информации США, состоит из достаточно подробной базы данных по энергетической инфраструктуре Америки в разрезе регионов. В ней представлено местоположение, основные характеристики (цены и объемы), и прогнозные показатели по основным добывающим первичные источники энергии, перерабатывающим, и генерирующим электрическую и тепловую энергию мощностям, а также транспортная система ЛЭП, трубопроводы, дороги. Помимо интерактивной составляющей: наглядного представления и проведения быстрого сравнительного анализа, выбранных характеристик, на сайте имеется возможность по выгрузке данных и проведения собственных исследований по имеющимся в базе данным. Большая часть статистики собирается Министерством с помощью форм, заполняемых производителями, а также с учетом статистики системных администраторов и т.п.

В качестве недостатка можно выделить лишь неудобное представление: нарушена логика идентичности. На странице указан перечень «различных» интерактивных карт, хотя по сути некоторые из них дублируются – при переходе на страницу интерактивного представления карты в качестве автоматически выбранного слоя подсвечивается тот, что указан в название карты. Например, карта энергетической системы США имеет несколько слоев, показывающих инфраструктуру нефтяной, газовой, угольной и других отраслей. При выборе карты нефтяной инфраструктуры в результате будет показана интерактивная карта энергетической системы США с автоматически выбранным слоем – инфраструктура нефтяной отрасли, аналогично для газовой и угольной отрасли.

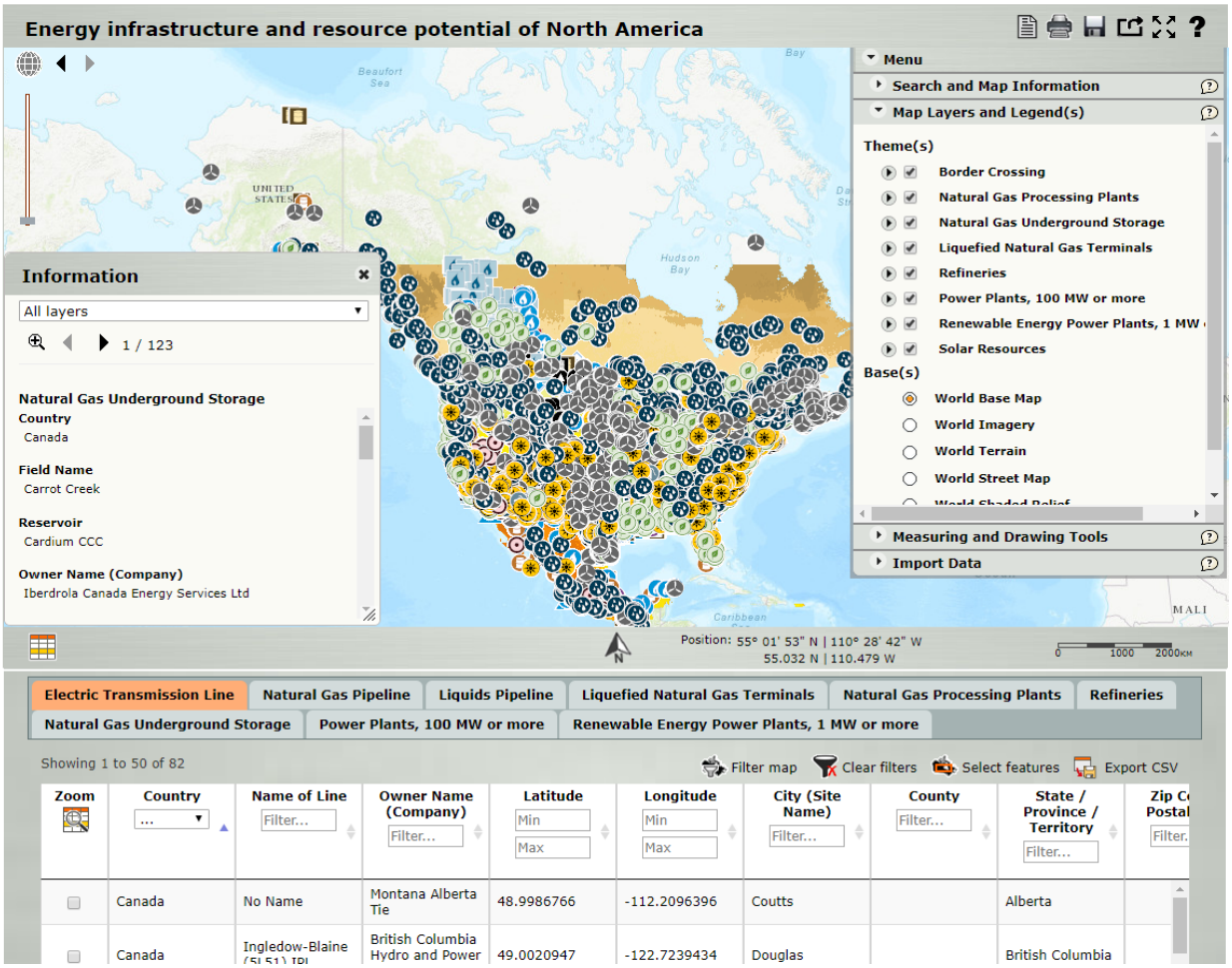
Данный аналог является очень близким представлением желаемого интерактивного ресурса (результата данной работы).

Интернет-ресурс статических и интерактивных карт энергетической системы Северной Америки (North American Energy Maps)

На сайте Министерства природных ресурсов Канады представлено несколько видов карт Североамериканской энергетической системы. Статические, которые включают в себя карты: мощностей по переработке природного газа (Natural Gas Processing Plants); терминалов по импорту и экспорту СПГ (Liquefied Natural Gas Import and Export Terminals); нефтеперерабатывающих заводов (Refineries and Upgraders); электростанций с деление по виду используемого топлива и установленной мощности (Electric Power Plants); источников возобновляемой энергетики (Renewable Power Plants); солнечной активности (Solar resources, Global Horizontal Irradiance and solar power plants); линий электропередач (Electric Transmission Lines); трубопроводов (Natural Gas Pipelines и Liquids Pipelines).

В 2014 году между США, Канадой и Мексикой был подписан трёхсторонний меморандум о сотрудничестве в области предоставления энергетической информации, учитывая взаимодействие EIA с Канадским Министерством энергетики, интерактивная карта энергетической системы Северной Америки напоминает интерактивную карту энергетической системы США с добавление инфраструктуры, мощностей и месторождений на территории Канады и Мексики (рисунок). Данный ресурс можно по большей части назвать дополнением к интерактивной карте Энергетической системы США, так как страна граничит на севере с Канадой и на юге с Мексикой, имеет соединения энергетической инфраструктуры.

К сожалению, дополнение (информация по энергетическим системам Канады и Мексики) представлено не в аналогичной детализации, как по США. Минусом является и различие интерфейса представления данных по сравнению с энергетической системой США, в частности структура данных при выгрузке и сам процесс фильтрации необходимой статистики.



Примечание – Источник: Официальный сайт the Department of Natural Resources Canada [CITATION Офи198 \l 1049].

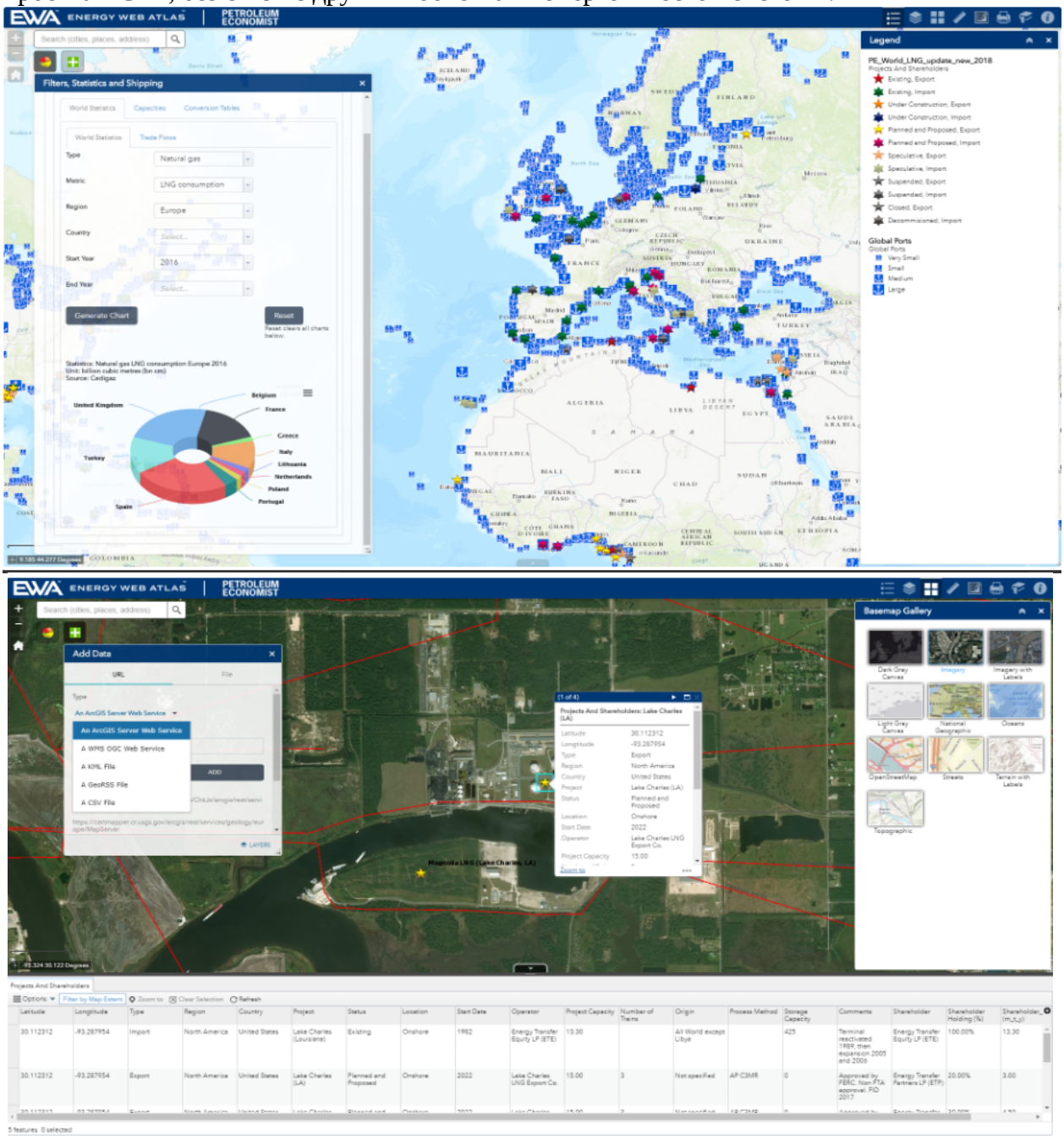
Рисунок 4 – Интерактивная карта Североамериканской энергетической системы

Атлас мирового рынка природного газа и СПГ (Energy Web Atlas)

Интерактивная карта Energy Web Atlas содержит подробную информацию о каждом проекте СПГ в мире. Карта основана на ESRI ГИС-платформе, что позволяет запрашивать информацию о проектах в режиме реального времени (статус, установленную мощность, используемую технологию, основных акционеров). Также система содержит дополнительные карты, содержащие информацию по газоперерабатывающим мощностям США и по их основным потребителям, а также информацию по проектам, которые используют возобновляемые источники энергии. Продукт доступен по подписке: расширенный выбор инструментов взаимодействия со слоями и объектами карты, выгрузка данных, представление данных в виде графиков и диаграмм.

Разработкой платформы занимается компания Gulf Energy Information совместно с ведущими отраслевыми журналами: Hydrocarbon Processing, Gas Processing, Petroleum Economist, Pipeline & Gas Journal и World Oil. На рисунке представлены снимки экрана фрагментов интерактивной карты, которые были представлены на сайте. Из него можно

сделать вывод, что ресурс является узконаправленным: статистика представлена только по проектам СПГ, без связи с другими объектами энергетической системы.



Примечание – Источник: Официальный сайт Energy Web Atlas [CITATION Офи199 \l 1049]

Рисунок 5 – Интерактивная карта Атласа мирового рынка природного газа и СПГ

В разделе был рассмотрен мировой опыт визуализации базы данных по производственным мощностям и инфраструктурным объектам энергетической системы, перечислены их основные сильные и слабые стороны. Анализ показал, что на данный момент многие подобные системы используют статические карты, а не интерактивное представление; выгрузка дезагрегированной оперативной статистики по производственным мощностям в основном осуществляется по платному доступу;

ведущим в подобном представлении данных являются развитые страны, в частности США. Краткие выводы анализа представлены в таблице .

Таблица 1 – Сравнение аналогов Атласа энергетической системы России между собой по функциональным возможностям

Название аналога	Map of Power Plants		The Pan-European Thermal Atlas – Peta	U.S. Energy Mapping System	IEA Atlas of Energy	North American Energy Maps	Energy Web Atlas
Данные в открытом доступе	-	+	+	+	+	+	-
Данные доступны по подписке	+	-	-	-	-	-	+
Частота статистики:	+	-	+	+	+	-	-
– год							
– месяц	+	-	+	+	-	-	-
– неделя	-	-	-	-	-	-	-
– день	-	-	-	-	-	-	-
– час	-	+	-	-	-	-	-
– в реальном времени	-	-	-	+	-	+	+
Региональная разбивка	+	+	+	+	+	+/-	-
Интерактивное представление данных	+		+	+	+	+	+
Возможность выгрузки данных	-		+	+	+	+	+
Наличие аналитических показателей	-		+	+	+	-	-
Наличие периодических публикаций	+		+	+	+	-	-

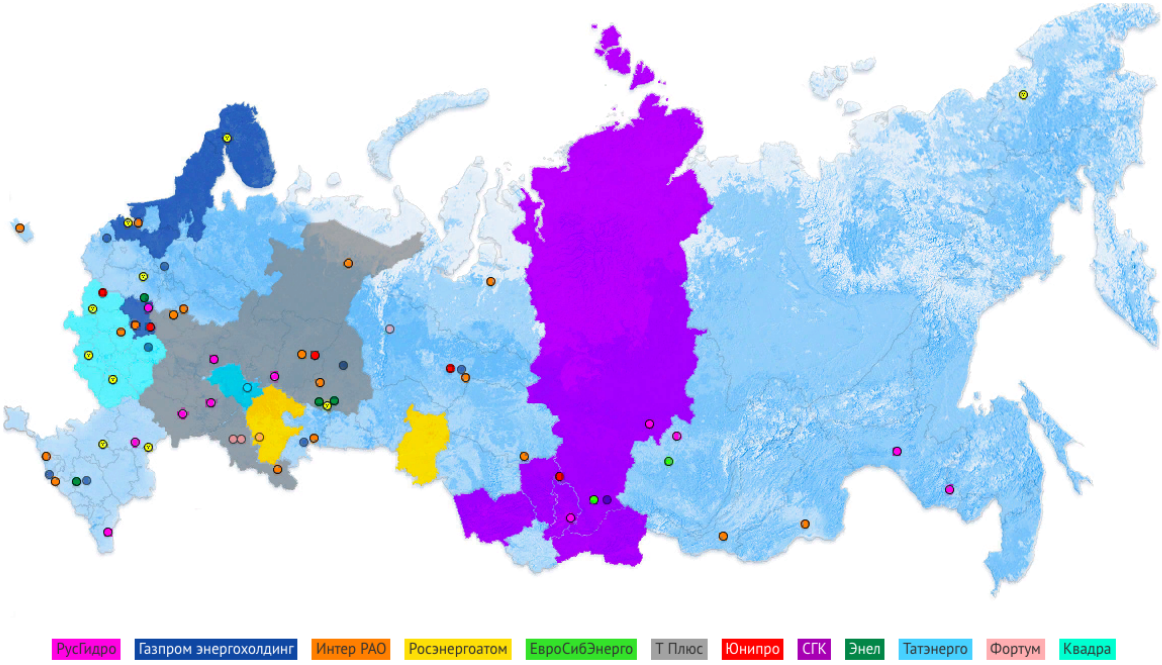
Примечание – Источник: составлено авторами

В качестве образца наиболее подходящим по описанию и функционалу к желаемому результату данного исследования является интерактивная карта энергетической системы США от EIA. Во-первых, потому что она содержит, как хорошую дезагрегированную по отдельным производственным мощностям, так и агрегированную на уровне региона и страны, статистику. Во-вторых, между представленными данными есть взаимосвязь: например, связность инфраструктуры между регионами и соседними странами, связность объектов различных отраслей через аналитические показатели и с помощью одновременного нанесения на карту и т.п. В-третьих, возможность выгрузки данных.

2 Анализ современного состояния, основных проблем и планов развития энергетической системы России

2.1 Анализ современного состояния энергетической системы России в сфере генерации энергии

Российский рынок электроэнергии состоит их четырех сегментов: генерация, передача и распределение энергии, сбыт, потребление. Генерация представлена, как государственными, так и частными компаниями (смотри рисунки и). В число государственных входят: ПАО «Интер РАО», АО «Концерн Росэнергоатом», ПАО «РусГидро», ООО «Газпром энергохолдинг». Среди частных помимо отечественных (ПАО «Квадра», АО «ЕвроСибЭнерго», ООО «Сибирская генерирующая компания», другие), представлены и иностранные компании (ПАО «Юнипро», ПАО «Энел Россия», ПАО «Фортум»). Данный сегмент является конкурентным, а генерирующие компании могут поставлять электроэнергию в сеть или напрямую крупным промышленным предприятиям.



Примечание – Источник: Сайт проекта «Моя энергия»

Рисунок 6 – Структура сегмента генерации отрасли Российской Федерации

Сетевые компании являются естественной монополией, и ее деятельность регулируется государством, что подразумевает установление тарифов на передачу электроэнергии и предоставление потребителям равноправного, недискриминационного доступа к услугам сетей (структура сетевых компаний в региональном разрезе представлена на рисунке).

Примечание – Источник: Сайт проекта «Моя энергия»

Рисунок 7 – Установленная мощность в разрезе крупных генерирующих компаний и объемы производства электроэнергии в год

Сбыт осуществляется независимыми энергосбытовыми компаниями, гарантией того, что конечный потребитель не окажется в ситуации, когда с ним отказались заключать договор все сбытовые организации, служит наличие в электроэнергетической системе гарантирующих поставщиков.

Примечание – Источник: Сайт проекта «Моя энергия»

Рисунок 8 – Региональная структура сетей с крупными подстанциями

В качестве гарантирующего поставщика выступает компания, которая обязана заключить с любым обратившимся к ней физическим или юридическим лицом, находящемся в зоне ее деятельности, договор предоставления электроэнергии.

Рынок электрической энергии и мощности (рассматриваются, как два отдельных товара) в России представляет собой двухуровневую систему: оптовый и розничный рынки. Все генерирующие станции с установленной мощностью выше 25 МВт осуществляют продажу электроэнергии и мощности на оптовом рынке. Станции, у

которых мощность меньше 25 МВт могут выбрать на каком рынке им работать: оптов или розничном рынке электроэнергии и мощности.

На оптовом рынке взаимодействуют между собой продавцы, в лице генерирующих компаний, и покупатели – крупные промышленные предприятия, сбытовые компании и гарантирующие поставщики. На розничном рынке продавцами являются сбытовые компании и гарантирующие поставки, компании малой генерации и ВИЭ с установленной мощностью менее 25 МВт; покупателями – население и приравненные к нему категории потребителей и небольшие предприятия.

Мощность, как товар, представляет собой покупку права на поддержание в готовности генерирующего оборудования для выработки электроэнергии установленного качества в объеме, необходимом для удовлетворения потребности в электроэнергии указанного участника.

Ценообразование на российском рынке электроэнергии зависит от того, на каком рынке осуществляются деятельность поставщики и покупатели, и от категории, к которой относятся поставщики и покупатели. В таблице представлена классификация рынка электроэнергии и мощности в Российской Федерации, а также структура объемов реализации электроэнергии и мощности в секторах оптового рынка электроэнергии.

Тариф для конечного потребителя на электроэнергию и мощность формируется на основе пяти составляющих: цена электроэнергии (цена покупки электроэнергии на оптовом рынке или у розничного генератора), цена мощности (цена покупки мощности энергосбытовой компанией на оптовом рынке или у розничного генератора), цена передачи по сети с дифференциацией по уровню напряжения: тарифы ФСК на передачу по магистральным сетям, тарифы МРСК на передачу по сетям среднего напряжения и тариф ТСО на передачу по сетям низкого напряжения, инфраструктурные платежи: плата за услуги СО ЕЭС, АТС, ЦФР (размер платы регулируется ФАС России и Ассоциацией «НП Совет рынка»), сбытовая надбавка. В структуре тарифа на электроэнергию и мощность для конечного потребителя на розничном рынке в среднем около 68% занимает стоимость электроэнергии и мощности, 28.4% составляют услуги по передаче, 3.7% — сбытовая надбавка, 0.1% — инфраструктурные платежи.

Теплоэнергетика чрезвычайно важна для России, так как теплоснабжение по обороту сопоставимо с 2.5% ВВП страны и составляет более 50% в платеже гражданина за ЖКУ [CITATION Мин15 \l 1049].

Таблица 2 – Классификация рынков электроэнергии и мощности в России

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Тип рынка</td></tr>
<tr><td colspan="2">Оптовый рынок
Продавцы: генерирующие компании
Покупатели: крупные промышленные предприятия, сбытовые компании и гарантирующие поставщики</td><td>Розничный рынок
Продавцы: сбытовые компании, гарантирующие поставщики, компании малой генерации, ВИЭ с установленной мощностью <25 МВт
Покупатели: население, приравненные к нему категории потребителей, небольшие предприятия</td></tr>
<tr><td rowspan="5">Тип товара</td><td rowspan="2">Э/э – 63%</td><td>Рынок на сутки вперед 52%</td><td>Балансирующий рынок 1%</td><td>Рынок нерегулируемых цен
(для потребителей, входящих в 1 и 2 ценовые зоны)</td></tr>
<tr><td>Рынок регулируемых договоров 6%
(для населения, гарантирующих поставщиков)</td><td>Рынок свободных договоров 2%
(для крупных платежеспособных потребителей)</td><td>Рынок регулируемых цен
(для населения и приравненных к нему групп потребителей и потребителей, не входящих в 1 и 2 ценовые зоны)</td></tr>
<tr><td rowspan="3">Мощность – 37%</td><td>Рынок регулируемых договоров 5%
(для населения)</td><td>Договор о предоставлении мощности (ДПМ) 16%</td><td rowspan="2">Рынок нерегулируемых цен
(для потребителей, входящих в 1 и 2 ценовые зоны)</td></tr>
<tr><td>Рынок свободных договоров 4%</td><td>Конкурентный отбор мощности 8%</td></tr>
<tr><td colspan="2">Реализация в статусе вынужденного генератора 2%</td><td>Рынок регулируемых цен
(для населения и приравненных к нему групп потребителей и потребителей, не входящих в 1 и 2 ценовые зоны)</td></tr>
</table>

Примечание – Источник: Обзор электроэнергетической отрасли России EY [CITATION Ару18 \l 1049]

Ценовые зоны: 1 – Европейская часть России и Урала. ТЭС (природный газ), АЭС; 2 – Сибирь. ТЭС (уголь), ГЭС;

Неценовая зона: Архангельская, Калининградская области, Республика Коми, регионы Дальнего Востока реализация э/э и мощности осуществляется по особым правилам и по регулируемым ценам (тарифам).

Связано это, во-первых, с тем, что большая часть страны находится в климатических зонах низкого температурного фона – средняя температура в осенне-зимний период в Европейской части России, в Сибири и на Дальнем Востоке составляет -5.5°С [CITATION Гру18 \l 1049], а продолжительность отопительного сезона длится 9 месяцев. А, во-вторых, с тем, что тепловая энергетика является очень энергоемкой отраслью, занимающей ведущее место в совокупном потреблении топливно-энергетических ресурсов в стране.

По результатам на начало 2017 г., в отрасли действовало около 21 тыс предприятий, 67% из них находилось в государственной и муниципальной собственности, 33% в частной, данное соотношение в течение 5 лет существенно не менялось [CITATION Рос181 \l 1049]. За 2016 г. было введено 6 637 новых котельных, в основном прирост обеспечило строительство малых котельных в сельской местности.

Суммарная установленная мощность источников теплоснабжения общего пользования в начале 2017 г. составила 844.7 тыс Гкал/ч. Доля отопительных котельных составляет около 70%, ТЭЦ – 30%. По итогам 2016 г. средняя загрузка источников тепловой энергии в отопительный период составила 30.1%, среднегодовая – 17.8%.

Общая протяженность теплопровода в России в двухтрубном исчислении составила 171.5 тыс км. Согласно статистике Росстата, наиболее развитые тепловые сети находятся в Центральном федеральном округе, их общая длина составляет 44 158 км, это почти 26 % от общей протяженности. Наименее развитые тепловые сети находятся в Северокавказском федеральном округе (3 421 км или 2% от общей протяженности). Протяженность тепловых сетей большего диаметра характеризует уровень централизации отрасли. Наибольшую длину теплопровода диаметром свыше 400 мм имеет Центральный федеральный округ, наименьшую – Северо-Кавказский федеральный округ.

По данным Росстата 28.8% теплопроводов систем теплоснабжения нуждаются в замене. Согласно Генеральной схеме размещения объектов электроэнергетики до 2035 г., около 46% установленных мощностей в России имеют возраст более 36 лет. Доля вновь введенного энергетического оборудования остается на достаточно низком уровне. Дефицит инвестиций в строительство новых, ремонт и модернизацию имеющихся мощностей может привести к техническим ограничениям и снизить надежность энергоснабжения.

Основными регионами-производителями тепловой энергии являются: Центральный федеральный округ из-за достаточно большой плотности населения; и Уральский федеральный округ из-за высокой плотности населения, большого количества крупных промышленных предприятий и низкого показателя среднегодовой температуры [CITATION Рос164 \l 1049].

Производство тепловой энергии зависит от погодных условий и продолжительности отопительного сезона. Объем производства тепловой энергии по данным за 2016 г. составил 1 284 млн Гкал: 46.5% было произведено централизованными котельными, которые производят только тепло и имеют низкую топливную эффективность по сравнению с системами ТЭЦ; 45.9% с помощью когенерационных установок (тепловыми электростанциями или ТЭС), работающие на основе теплового и электрического режима (ТЭЦ). При этом потери при передачи тепловой энергии конечному пользователю оцениваются в районе 10%. В настоящее время около 80% тепла производится с помощью систем теплоснабжения, которые обеспечивают его передачу в виде горячей воды из централизованной котельной или когенерационной установки по трубопроводной сети [CITATION Гру18 \l 1049]. Основной вид топлива для электростанций и котельных России является природный газ (более 70%). Доля сжигаемых в котельной нефтепродуктов высока, по сравнению с американским рынком – 7%, в США 1% [CITATION Гру18 \l 1049].

С 2013 г. выделяют в отчетностях отпуск тепла от пиковых водогрейных котельных (ПВК) и редукционноохладительных установок (РОУ). Многие ТЭС в последние годы злоупотребляют отпуском тепла от РОУ, что ранее применялось исключительно в особых случаях и крайне нежелательно, так как существенно снижает тепловую экономичность тепловых электростанций.

В таблице представлена статистика по инвестициям в основной капитал в сфере централизованного теплоснабжения. Из таблицы видно, что финансирование отрасли осталось почти на том же уровне в 2016 г. по сравнению с 2012 г. Поменялась внутренняя структура, больше средств стало идти на производство тепловой энергии, сократились инвестиции, направляемые на передачу тепловой энергии.

Таблица 3 – Инвестиции в основной капитал в сфере централизованного теплоснабжения в 2012-2016 гг.. млрд руб.

	2012	2013	2014	2015	2016	2016/2012
Производство. передача и распределение тепловой энергии. всего	98.73	111.49	86.85	95.69	100.07	101%
Производство тепловой энергии. в т.ч.:	38.58	52.09	42.59	45.92	51.23	133%
ТЭС	15.29	25.05	16.18	19.62	21.85	143%
прочими электростанциями	0.06	0.12	0.12	0.35	0.42	700%
котельными	23.2	26.92	26.29	25.88	28.96	125%
Передача тепловой энергии	39.07	19.74	21.32	18.08	19.35	50%
Распределение тепловой энергии	14.61	33.22	18.38	28.25	25.84	177%
Деятельность по обеспечению работоспособности котельных	1.88	1.94	1.5	0.86	1.17	62%
Деятельность по обеспечению работоспособности тепловых сетей	4.58	4.49	3.02	2.29	2.04	44%
Торговля тепловой энергией	0.01	0.01	0.04	0.30	0.44	4439%

Примечание – Источник: Министерство энергетики России «Теплоэнергетика и централизованное теплоснабжение России в 2015-2016 гг.». 2018 [CITATION Док18 \l 1049]

Цены (тарифы) на тепловую энергию в России устанавливаются в соответствии с Федеральным законом от 27.07.2010 №190-ФЗ «О теплоснабжении». постановлением Правительства Российской Федерации от 22.10.2012 №1075 «О ценообразовании в сфере теплоснабжения». приказом Федеральной службы по тарифам от 13.06.2013 №760-э «Об утверждении Методических указаний по расчету регулируемых цен (тарифов) в сфере теплоснабжения». Для населения и приравненных к ним категорий тарифы регулируются государством. для остальных нет. Цена, отпущенная котельными выше, чем электростанциями, что связанно с более эффективным и менее затратным способом производства тепловой энергии на когенерирующих установках.

Исходя из проведенного анализа, можно заключить, что основная нагрузка по выработке электрической энергии лежит на ТЭС: более половины электрической энергии вырабатывается именно данными станциями. Растет доля выработки электричества из газа и атомными станциями, сокращая при этом угольную генерацию. Основными особенностями отрасли тепловой энергетики являются: высокая изношенность основных фондов (большая часть инфраструктуры была построена еще в СССР); высокая централизация и невозможность размещения теплогенерирующих станций на большом расстоянии от непосредственных потребителей, из-за высоких потерь при передаче энергии; исторически сложившееся совместное введение бизнеса генерации электроэнергии и продажи тепла, в результате прибыль от первого перекрывает убыток от второго (неэффективное перекрестное субсидирование между электричеством и теплом, между паром и теплом, между промышленностью и населением, между территориями); невозможность отключения подачи тепла у должников, так как данная услуга является социально значимой.

2.2 Анализ современного состояния энергетической системы России в сфере потребления энергии

Крупнейшим потребителем тепловой энергии в России в 2018 году является промышленность, его доля в структуре потребления составляет 42.6% [CITATION Рос164 \l 1049]. Около 31% использует население, остальное идет на коммунально-бытовой и непромышленный сектора [CITATION Рос181 \l 1049]. Учитывая существенную долю населения в структуре производства тепловой энергии и ее социальную значимость (невозможно отключить за неуплату), отрасль терпит серьезные убытки: по данным Росстата, долги за отопление на конец октября 2017 г. составили 184

млрд руб. [CITATION Офи179 \l 1049]. Примечательно то, что потери тепла в сети составляют долю, сопоставимую с потреблением тепла химической отрасли и большую, чем потребление тепла, например, металлургического производства. Это уже было отмечено выше связано с большой изношенностью основных фондов отрасли генерации электроэнергии и тепла.

По данным Министерства энергетики России в 2018 году потребление электроэнергии по всей стране составило 1076.2 млрд кВт·ч (по ЕЭС России 1055.6 — млрд кВт·ч), что превышает показатели 2017 года на 1.6% (по данным ЕЭС России — на 1.5%).

В качестве основного фактора увеличения потребления является температура (снижение среднегодовой температуры на 0.6°C относительно прошлого года), оказывающая значительное влияния в марте, октябре и декабре 2018 года. В перечисленные месяца среднемесячная температура достигала максимальных значений. Помимо этого, произошло значительное увеличение потребления электроэнергии промышленными предприятиями (в большей степени металлургическими, деревообрабатывающими, и объектами газопроводного и железнодорожного транспорта) [CITATION Рос181 \l 1049].

Таким образом, можно сделать вывод о том, что крупнейшим потребителем тепловой и электрической энергии в Российской Федерации последние годы является промышленность. Доля во всем потреблении 46%. Далее следует населения, занимая чуть более 30% от всего потребления. Основными факторами, влияющими на изменение динамики потребления тепловой и электрической энергии, являются появление новых крупных промышленных объектов и температура. Если первый фактор можно хорошо спрогнозировать, учитывая инвестиционные программы крупных промышленных программ, то по второму возникают вопросы. Например, в работе Лозинской А., Шенкман Е., Тетериной Д. и Гордиенко А. [CITATION Гор19 \l 1049] был предложен способ кластеризировать регионы России по схожести климатических условий (по температуре перехода от эффекта подогрева на эффект охлаждения). Данный метод и его развитие (усложнение признаков для кластеризации регионов) может помочь в моделировании.

3 Создание базы данных географического расположения. технологического оснащения и взаимосвязей объектов энергетической системы России (Атлас энергетической системы России)

3.1 Методология сбора и структуризации данных

Сбор данных предполагает использование открытых источников с применением инструментов автоматического извлечения данных. К ним относятся: программные интерфейсы сайтов (API – application programming interface), которые позволяют напрямую взаимодействовать с базами данных сервисов и выгружать большие объемы информации; библиотеки для извлечения данных из HTML и XML структур в случае отсутствия API. Процесс сбора данных был проведен с учетом следования следующим принципам:

- Первичный источник публикуемой информации (там, где это возможно),
- Открытый доступ к данным,
- Сравнение информации по нескольким источникам.

Сбор информации в автоматическом режиме осуществляется исключительно с сайтов. не ограничивающих право на использование в коммерческих целях информации, хранящейся на ресурсе. К примеру. одним из источников для сбора информации о геолокации энергообъектов являлся OpenStreetMap[CITATION Ope19 \l 1033]. Ресурс имеет программный интерфейс – API Overpass. Этот достаточно гибкий инструмент позволяет получать в любом удобном формате примерно до 10 миллионов элементов карты за несколько минут. Отбор элементов осуществляется с помощью критериев поиска. таких как: местоположение. тип объектов. свойства тега. близость или их комбинации.

Для сбора и автоматического обновления данных из HTML и XML структур используется библиотека «BeautifulSoup4» для Python3, что относится к свободно распространяемому программному обеспечению. Важной особенностью библиотеки является поддержка обнаружения кодирования, которая может обеспечить лучшие результаты для HTML-страниц, которые неправильно декларируют свою кодировку. Предполагается, что написанные скреперы будут использоваться с некоторой периодичностью для поиска обновлений в информационных системах. Рассмотрим более подробно работу скрепера на примере извлечения описательных характеристик объектов электроэнергетики с Wikipedia.org. Свободная энциклопедия служит источником структурированной информации по неизменяемым характеристикам объектов электроэнергетики. К таковым относятся: год постройки, установленная мощность.

количество силовых агрегатов. характеристики силовых агрегатов. ссылка на официальный сайт. географический координаты и прочее.

Загруженная страница сайта, при помощи инструментов библиотеки, в среду Python, представляет собой html-дерево, для извлечения информации из которого используется анализатор текста. Встроенный анализатор текста позволяет перемещаться по тегам html-дерева и извлекать текст или целые блоки текста, находящиеся между открывающим и закрывающим тегами. Таким образом, для получения географических координат и прочей информации для объектов электроэнергетики, был составлен список индивидуальных страниц энергообъектов. При этом перечень страниц объектов электроэнергетики составлялся не вручную, а использовались списки открытой энциклопедии, что позволило собрать все имеющиеся данные. После чего, по циклу по отобранным страницам из тегов, хранящих определенною информацию, были извлечены данные для всех объектов.

Результатом работы скреперов является массив данных, который зачастую представляет из себя лишь частично пригодную информацию для ее дальнейшего анализа. Это происходит по ряду причин: численные показатели могут содержать нечисловые символы (указывать на единицы измерения. отступ между разрядами чисел и прочее), в результате чего формат данных будет определен, как текстовый; в свою очередь текстовые характеристики могут содержать лишние символы, являющиеся артефактами html-разметки или не относящиеся к сути слова; возможны несовпадения единиц измерения для однотипных характеристик; ошибки при заполнении и многое другое. Строгие ограничения на формат хранения и порядок занесения данных в базу, о которой пойдет речь в следующем разделе. практически исключают возможность использования некорректной информации, т.е. не соответствующей ожидаемым характеристикам. Для этого до занесения в базу данные проходят процедуру предобработки и очистки. Для минимизации риска занесения в базу несоответствующей действительности информации применяется сравнение по различным источникам и используются первоисточники в том случае, если это возможно.

Предобработка данных предполагает очистку первичных данных, сравнение числовых и иных характеристик объектов из различных источников и формирование конечных массивов данных для их размещения в базе данных.

Результатом работы скреперов, могут быть массивы данных в различной степени, содержащие ошибки и искажения. Ошибки, как правило, вызваны неправильным наполнением численных и текстовых блоков сайтов (человеческий фактор), с которых в автоматическом режиме считываются и сохраняются данные. Так же возможны изменения структуры html разметки сайта после написания скрепера, что приводит к его

неправильной работе (технический фактор) и частичной или полной загрузке ненужных данных, если не будет вызвана ошибка при чтении кода. Искажения могут быть вызваны также человеческим фактором, когда одна категория данных для различных объектов на едином ресурсе представлена не в единой шкале. К примеру, для числовых данных искажением будет считаться случай, когда часть данных однотипных показателей измеряется в тысячах некоторых едини, а основная масса в млн. некоторых единиц. Выше также упоминалось и о распространенной проблеме, связанной с извлечением блоков из xml-структур, когда интересуемые данные содержат лишние символы. Основной задачей очистки данных является поиск и устранение таких ошибок и несоответствий. Очитка загруженных данных производится в Python с использованием различных пакетов обработки данных, таких как: «Pandas», «Numpy, «Regular expression operation».

Основная часть работы по обработке первичных данных осуществляется автоматически. Для этого используется Python-скрипт с инструкциями проверки и исправления данных с учетом их предполагаемого типа. Поскольку база данных на настоящий момент включает использование семи типов данных, инструкция содержит семь функциональных блоков. Далее приведена краткая характеристика таких блоков.

В силу того, что проект «энергетический атлас» прежде всего предполагает предоставление пользователю аналитики, основанной на численном подходе, важной задачей является обработка численной информации. Наибольшее количество численной информации удобно представлять в формате вещественных чисел (тип данных real для sql), поскольку показатели объектов энергосистемы часто записываются в виде десятичной дроби (к примеру: 19.23 кг/Гкал – удельный расход условного топлива для работы котельной. 9155.9 млн. кВт*ч – годовая выработка электростанции и т.д.). Однако первичные данные, которые потенциально относятся к вещественному типу, часто имеют текстовый тип данных, поскольку содержат приписку в виде указания размерности и величины измерения, а также различные символы разметки. такие как пробелы между разрядами чисел (например, 1 000 вместо 1000) и прочие артефакты xml разметки. Основная задача функционального блока – это поиск в текстовой строке численного выражения, которое может быть в виде десятичной дроби, с разделением целой и дробной части запятой или точкой; натурального числа, с пробелами между разрядами или без. Для поиска таких последовательностей используются операции с регулярными выражениями. Найденная числовая последовательность преобразуется в вещественный тип данных. а оставшийся текст. указывающий на величину измерения. проверяется на соответствие принятой по молчанию для обрабатываемого показателя. После чего в рамках этого же функционального блока создается запись, которая может быть инкорпорирована в

соответствующую типу данных таблицу в базе данных. Кроме того, в возможности функционально блока входит поиск и сохранение объектов, с незаполненными данными по обрабатываемому показателю. Это необходимо для ведения статистики относительно полноты информации по объектам определенной отрасли энергосистемы.

Данный функциональный блок технически выполняет все операции, что и описанный блок выше с единственной разницей в том, что найденная числовая последовательность преобразуется в натуральное число. Блок применяется к таким данным, которые нет необходимости преобразовывать в вещественный тип. В первую очередь – это год строительства. введения в эксплуатацию. реконструкции; численность сотрудников; количество энергоблоков или других агрегатов объекта и т.д.

Числовые последовательности из вещественных чисел в данных, как правило. содержат информацию о диапазонах изменчивости показателей. К примеру, важнейшее технологической свойство нефти – вязкость. измеряется в мПа*с и часто представляется в виде некоторого диапазона в связи с неоднородностью нефти в одном месторождении. Задачей функционального блока является поиск числовых последовательностей и приведение их в вид, в котором может быть осуществлена машинная запись в базу данных в поле с типом «array» для вещественных чисел.

Результат применения функции данного блока аналогичен, описанному выше с разницей в типе числовых данных, используемом внутри «array». После обработки текстовых строк и получения диапазона чисел. последним назначается целочисленный тип данных. К примеру, газопровод в своем составе может иметь различные по диметру трубы. И одной из характеристик газопровода будет диапазон диаметров труб. Диаметр трубы газопровода измеряется в мм и не предполагает измерение в микрометрах. Следовательно, для хранения данных разумно использовать целочисленный тип данных, дающий возможность эти данные векторизовать. Другим примером может служить глубина залегания нефти на месторождении. Глубина залегания оценивается в метрах и имеет некоторый оценочный диапазон, уточнение которого вплоть до см не производится.

Обработка текстовых данных главным образом предполагает очистку текста от артефактов синтаксиса xml, к которым могут относиться символы для переноса каретки на следующую строку. различные открывающие и закрывающие теги и прочее. Регулярные выражения, используемые в функциональном блоке. позволяют найти и исключить подобные ненужные символы. после чего на выходе из блока может быть сформирован массив из очищенных текстовых атрибутов, соответствующих им объектов для записи в таблицу базы данных.

Данный блок использует аналогичные функции для очистки текста, как и вышеописанный блок, но при этом. в случае наличия перечислений. будет записывать

элементы в последовательность array. Подобные перечисления часто встречаются в описательных характеристиках объектов энергетической инфраструктуры. К примеру, перечисления типов генераторов на электрических станциях; видов нефтехимической продукции на нефтеперабатывающих заводах и прочее.

Хранение в базе данных географических координат предполагает использование специального геометрического типа данных. В случае СУБД PostgreSQL это тип данных «point», который более подробно будет описан в следующей части работы с объяснением выбора модели и архитектуры базы данных. Задачей на данном шаге является приведение структуры координат к форме, которая принимается для записи в такой тип данных. В явном виде это пара координат x и y. перечисленных через запятую. где x – широта. а y – долгота. Функция обработки учитывает то, что на различных ресурсах широта и долгота могут быть записаны, как в различных ячейках, так и в одной, переставлены местами (долгота-широта) и содержать текстовые дополнения. Для учета последнего к координатам применяется функция обработки вещественных чисел.

Описанные функциональные блоки, в свою очередь являются функциями, к которым обращается основная функция, подготавливающая данные перед их записью в базу данных. Основная функция получает на вход необработанный вектор однотипных характеристик и указание на то, к какому типу данных его необходимо привести. После применения функции очистки и исключения пропусков формируется таблица с полученными вектором характеристик и прочими атрибутами, определенными архитектурой базы данных для каждого типа данных. Подробнее о построении таблиц для каждого типа дынных и механизмах, связывающих все воедино под управлением СУБД (система управления базами данных) будет написано в следующем разделе. Кратко и структурировано выводы по обработке собранных данных представлены в таблице .

Таблица 4 – Функциональные блоки обработки первичных данных в соответствии с их типами

Функциональный блок	Задача
Обработка вещественных чисел	Поиск в текстовой строке численного выражения, которое может быть представлено в виде десятичной дроби. с разделением целой и дробной части запятой или точкой. натурального числа. с пробелами между разрядами или без и преобразование в вещественный тип данных.
Обработка натуральных чисел	Поиск в текстовой строке численного выражения, которое может быть натуральным числом. с пробелами между разрядами или без и преобразование в целочисленный тип данных.
Обработка числовых последовательностей из вещественных чисел	Поиск числовых последовательностей и приведение их в вид. в котором может быть осуществлена машинная запись в базу данных в поле с типом «array» для вещественных чисел.
Обработка числовых последовательностей из натуральных чисел	Поиск числовых последовательностей и приведение их в вид. в котором может быть осуществлена машинная запись в базу данных в поле с типом «array» для натуральных чисел.
Обработка текстовых данных	Очистка текста от артефактов синтаксиса xml. к которым может относиться перенос каретки на следующую строку. различные открывающие и закрывающие теги и прочее.

Продолжение таблицы Ошибка: источник перекрёстной ссылки не найден

Функциональный блок	Задача
Обработка последовательностей из текстовых данных	Очистка текста от артефактов синтаксиса xml и разбиение на блоки в виде последовательности array для списков с перечислениями.
Обработка координат	Преобразование первоначальной структуры записи координат к форме, воспринимаемой СУБД для записи в тип данных географических координат. Функция обработки учитывает то. что на различных ресурсах широта и долгота могут быть записаны. как в различных ячейках. так и в одной. переставлены местами (долгота-широта). могут содержать текстовые дополнения.

Примечание – Источник: составлен автором.

3.2 Описание архитектуры собранной базы данных по энергетической системе России

В данном разделе будет описана выбранная модель и архитектура базы данных. Архитектура базы данных разрабатывалась под PostgreSQL 9.6, что связано с версией СУБД, установленной на сервере академии. где физически размещается база данных. PostgreSQL – объектно-реляционная система управления базами данных. Имеет свободную лицензию[2], что позволяет бесплатно использовать СУБД в том числе для коммерческих целей.

Структура данных, изложенная в предыдущем разделе, предопределила модель базы данных. Большое количество параметров при относительно небольшом количестве экземпляров является главной характеристикой накопленной статистики. Другими словами, количество объектов энергетической инфраструктуры может исчисляться десятками, в то время как. количество описывающих его параметров – сотнями. Подобное наблюдается не всегда, но систематически встречается. К тому же энергетическая система страны потенциально может быть описана в тысячах параметрах в различных отраслях с множеством уровней детализации. Следовательно, в будущем ожидается систематическое пополнение и расширение имеющихся данных, что предполагает удобство масштабирования базы данных. Такие особенности хорошо могут быть учтены в модели Объект-Атрибут-Значение[CITATION Din \l 1049]. Такого рода модель, как средство представления данных общего назначения. возникла с понятием «списки ассоциаций» (пары атрибут-значение). Обычно используемые сегодня. они были впервые представлены на языке LISP[CITATION Fre \l 1049]. Название модели отражает ее сущность. В дальнейшем будет использоваться ее общепринятое краткое наименование – EAV, являющееся сокращением от «Entity-Attribute-Value». В самом простом виде архитектура модели сводится к трем таблицам. Основанная таблица имеет три столбца: в первом перечисляются все имеющиеся объекты. во втором – их атрибуты (то есть

2Интернет ресурс: https://postgrespro.ru/docs/postgresql/9.6/intro-whatis

характеристики), и в третьем хранятся непосредственно значения атрибутов. Одно из преимуществ такого представления состоит в отсутствии необходимости хранить пустые значения для атрибутов напротив объектов с отсутствующей информацией, что неминуемо при развернутой широкой таблице. Возвращаясь к особенностям статистики по энергетическому комплексу. где для относительно небольшого количества объектов определенной сферы может существовать множество атрибутов. важно сказать, что далеко не по всем объектам будут иметься значения. Преимущественно это связано с труднодоступностью информации. Таким образом, модель позволяет развернуть множество широких таблиц, в которых одна запись или строка соотносит объект с множеством его характеристик. в единственную вертикальную таблицу: «тощую», но длинную. При этом специальным образом сконструированные запросы к такой базе дынных позволяют получить привычные табличные представления. Пример привычной таблицы. с которой удобно работать в любом редакторе данных с интерфейсом. в общем виде приведен в таблице .

Таблица 5 – Общий вид таблицы в широком представлении

Объект	Атрибут 1	Атрибут 2	Атрибут 3	Атрибут 4	…	Атрибут m
Объект 1	X11	X12	NA	NA	…	X1m
Объект 2	X21	X22	NA	NA	…	X2.m
…	…	…	…	…	…	…
Объект n	NA	NA	Xn.k-2	Xn.k-1	…	Xn.k

Примечание – Источник: составлен автором.

Физическое хранение подобной таблицы. которая представляет собой разряженную матрицу. в PostgreSQL не эффективно. В данном случае количество атрибутов (m) относительно количества объектов (n) достаточно велико. а пропусков в данных много. Под пропуски в записях (строках) таблицы СУБД обычно резервирует ячейки памяти. несмотря на то. что информация для них может никогда и не появиться. Кроме того, количество таких таблиц может быть множество, но они не имеют между собой отношений, т.е. существует множество групп объектов, каждую из которых характеризует свой определенный набор атрибутов. В таком случае. с точки зрения эффективности хранения данных. таблицу лучше преобразовать в вертикальный вид следующим образом (таблица).

В такой таблице отсутствуют пропуски, но при этом многократно дублируются названия атрибутов и объектов. Конечно. базе данных совершенно не нужно хранить одну и ту же повторяющуюся информацию в различных ячейках памяти. Поэтому в архитектуру EAV базы данных обязательно входит еще две таблицы, о которых было упомянуто выше. Во второй хранится список объектов (таблица). а в 3 хранится весь перечень атрибутов (таблица).

Таблица 6 – Таблица Объект-Атрибут-Значение в общем виде

Объект	Атрибут	Значение

Объект 1	Атрибут 1	X11
Объект 1	Атрибут 2	X12
…	…	…
Объект 1	Атрибут m	X1m
Объект 2	Атрибут 1	X21
Объект 2	Атрибут 2	X22
…	…	…
Объект 2	Атрибут m-1	X2.m-1
Объект 2	Атрибут m	X2.m
…	…	…
Объект n	Атрибут 4	Xn.k-1
Объект n	Атрибут m	Xn.k

Примечание – Источник: составлен автором.

Таблица 7 – Список объектов для модели Объект-Атрибут-Значение в общем виде

№	Объект
1	Объект 1
2	Объект 3
3	Объект 4
4	Объект 2
5	Объект 5
6	Объект 6
… n	… Объект n

Примечание – Источник: составлен автором.

Таблица 8 – Список атрибутов для модели Объект-Атрибут-Значение в общем виде

№	Атрибут
1	Атрибут 1
2	Атрибут 2
3	Атрибут 3
4	Атрибут 4
… m-2	… Атрибут m-2
m-1	Атрибут m-1
m	Атрибут m

Примечание – Источник: составлен автором.

В основной таблице поля с объектами и атрибутами содержат лишь ссылки на их собственные таблицы, т.е. являются внешними ключами и не хранят непосредственно значения в своих ячейках. Стоит заметить, что структура данных в главной таблице может не удовлетворять некоторым нормальным формам БД.

В теории реляционных баз данных выделяют 6 нормальных форм, условия которых рекомендуется выполнять при проектировании[CITATION Хом \l 1049]. Каждая нормальная форма способствует минимизации избыточности отношений и самих данных в таблицах БД. Соответствие нормальным формам позволяет считать базу данных эффективной в плане занимаемого пространства и производительности пользовательских запросов. Однако строгое соответствие всем нормальным формам уместно не всегда. Все зависит от структуры данных, требований к производительности и удобству написания пользовательских запросов.

Считается, что таблица находится в первой нормальной форме, если все ее атрибуты простые и не содержат повторений. Примером поля, не удовлетворяющим первой нормальной форе из-за наличия повторений, может быть столбец с типами

электростанций: АЭС, ГЭС, ТЭЦ, когда объекты - электростанции. В данном случае поле атрибута «Тип электростанции» может быть заполнено для каждой электростанции только одним из трех возможных значений. Следовательно, наличие в первой и второй таблицах такого атрибута приведет к многократному дублированию хранимой информации. Таблицы могут быть приведены к нормальной первой форме путем выделения поля с типами электростанций в отдельную таблицу, наподобие таблицы с перечислением объектов или атрибутов.

Однако, на практике выделять отдельную таблицу, индексировать ее, а в основной таблице создавать поле со ссылками на таблицу типов не обязательно, поскольку существует тип данных «перечисления». При помощи него можно определить статический набор значений, который будет использоваться для заполнения поля атрибута. Использование такого типа данных вместо текстового, по сути, аналогично операции с созданием таблицы. Требование относительно простоты атрибутов означает атомарность или неделимость записываемых значений в одну ячейку. К примеру, числовые характеристики объекта должны быть записаны в разные ячейки с соответствующими им атрибутами, а не перечисляться в одной ячейке. Однако это необходимо в случае, если это логически неделимый вектор, такой как географические координаты[3].

Таблица находится во второй нормальной форме, когда каждый неключевой атрибут зависит от всего ключа, а не только от его части. Для широкой таблицы вторая нормальная форма может быть нарушена, в том случае, когда ключом таблицы является составной ключ. К примеру, состоящий из полей: название электростанций и ID электростанции (т.е. тип) объекта, а один из атрибутов таблицы – мощность генератора. Можно сказать, что атрибут зависит от названия электростанции, но не зависит от ID, а скорее от того, какая марка генератора установлена (при этом на электростанции, как правило, несколько генераторов). Наличие подобных отношений приводит снова к излишнему дублированию хранящейся информации, поскольку в действительности марок генераторов существует конечное число, и оно на порядок меньше количества электростанций. В результате таблица базы данных хранит для одних и тех же генераторов их технические характеристики в различных ячейках памяти. Этого можно избежать, если выделить в отдельную таблицу список всех имеющихся марок генераторов с их техническими характеристиками. Тогда, в основной таблице можно будет только хранить ссылки на марки генераторов, который установлены на определенной электростанции, а при необходимости подгружать необходимую техническую информацию В вертикальной таблице, присутствующий одни неключевой атрибут – это поле значений, может также

3 Для географических координат существует свой тип данных point, о котором будет сказано в части описания собственной архитектуры базы данных.

неоднозначно зависеть от своего составного ключа (Объект-Атрибут). В случае с примером про генераторы, EAV таблица значений будет соответствовать второй нормальной форме, если поле объекты в таблице будет содержать, как названия электростанций, так и марки генераторов. Тогда одна запись таблицы EAV будет иметь вид: «Электростанция1, Марка генератора – ТВФ-110-2ЕУ3», а другая запись: «ТВФ-110-2ЕУ3 – мощность генератора – 60 МВт». Конечно, такой пример является искусственным и в действительности необходимо бы было создание отдельной таблицы, но не для выделения отдельно всех генераторов, а потому как одна таблица значений в EAV модели предполагает один тип данных. В нашем примере марка генератора – текстовый тип, а мощность – числовой. Для примера примем это допущение. Важно заметить, что, удовлетворяя принципу второй нормальной формы в данном примере, усложняется написание запроса, в результате которого пользователь мог бы видеть: «электростанция – тип генератора – мощность генератора». В случае простой таблицы достаточно было бы вывести одно поле, которому принадлежит интересуемый объект, отображая только необходимые атрибуты. Для двух таблиц не EAV представления, можно использовать объединение таблиц при условии равенства двух атрибутов: марки генератора в поле таблицы электростанций и марки генератора в ключевом поле таблицы генераторов. В случае одной EAV таблицы необходимо создавать запрос, где по аналогичному условию, только исходя из поиска соответствия внутри одной таблицы, будет добавляться значение мощности генератора к основной строке. К тому же не факт, что запрос, содержащий подобное условие, будет выполняться с необходимой скоростью из-за длины таблицы. Для ускорения подобных запросов можно воспользоваться представлениями – создавать заранее таблицы с наборами данных, к которым могут быть частые обращения со стороны пользователей, на основе базовых таблиц. Это позволит при пользовательском запросе не запускать процедуру поиска соответствий по всей длине таблицы, а выдавать уже практически готовый результат. Однако, предусмотреть все такие случаи, по крайней мере на этапе проектирования базы данных сложно, поэтому можно оставить непрямое соответствие составному ключу EAV таблицы и вместо записи «ТВФ-110-2ЕУ3 – мощность генератора – 60 МВт» хранить запись: «Электростанция1 – мощность генератора – 60 МВт», Это в случае, если у электростанции один электрогенератор и: «Электростанция2 – мощность генератора – 60 МВт; 50 МВт; 70 Мвт», если генераторов несколько. Атрибут «марка генератора» для данной электростанции будет хранить вектор из трех марок генераторов, которым соответствуют установленные мощности. В таком случае поиск по частичному индексу, индексирующему только объекты – названия электростанций, даст возможность получить результат значительно быстрее, когда

необходим вывод различных атрибутов. Именно поэтому изначально утверждалось, что строгое соответствие требованиям нормальных форм уместно не всегда. Скорее это всегда компромисс между многими факторами: сохранением ресурса памяти, производительностью пользовательских запросов, сложностью их написания и прочее.

Соответствие третьей нормальной форме означает отсутствие транзитивных зависимостей. В случае вида таблицы , третья нормальная форма может быть нарушена, если атрибут 3 зависит от атрибута 2, а не от объекта, который является ключом. К примеру, это может произойти, если атрибут 2 содержит информацию о марке генератора, а атрибут 3 от мощности генератора. В таком случае, для удовлетворения требованиям третьей нормальной формы, необходимо выделить в отдельную таблицу все марки генераторов, где они будут ключами, а атрибутами – всевозможные их характеристики. Вертикальная таблица находится по умолчанию в третьей нормальной форме, если она находится во второй и первой нормальной форме, поскольку содержит только одно поле с атрибутами, не входящими в составной ключ, следовательно, от других не ключевых атрибутов зависимости быть не может.

Соответствие четвертой нормальной форме предполагает, что составной ключ однозначно определяет атрибут таблицы. Соответствие этой форме особенно важно для вертикальных таблиц, поскольку они имеют единственный неключевой атрибут, а нарушение формы фактически означает возможность повторения комбинации составного ключа для различных атрибутов. Тогда при пользовательском запросе СУБД будет извлечен первый попавшийся атрибут. Рассмотрим пример с таблицей , где в качестве составного ключа используются: название электростанции (Объект) и название атрибута (Атрибут), а значение – это марка генератора.

Таблица 9 – Нарушение четвертой нормальной формы для вертикальной таблицы

Объект	Атрибут	Значение
Электростанция 1	Генератор	S21P515
Электростанция 1	Генератор	ВСГ 325/49-32
Электростанция 2	Генератор	ВСГ 325/49-32

Примечание – Источник: составлен автором.

Из таблицы очевидно, что Электростанция 1 имеет 2 генератора, однако ключ их однозначно не определяет и результатом запроса, сформированного по ключу, пользователь увидит только марку генератора, которая первой встретиться СУБД в процессе сквозного просмотра таблицы по заданной комбинации ключа. Чтоб избежать подобной проблемы, но при этом сохранить вертикальную структуру, на практике можно хранить все марки генераторов, имеющиеся у одной электростанции в виде вектора, используя тип данных «array».

Другой пример, встречающийся на практике, связан с неоднозначностью ключа для значений атрибутов, изменяющихся во времени. Для однозначного определения значения

атрибута можно добавить с составной ключ еще поле с датами. Для широких таблиц несоответствие четвертной нормальной форме состоит ровно в том же, а для исправления ситуации, используется, как правило, разделение таблиц.

Пятая нормальная форма предполагает отсутствие сложных зависимых соединений между атрибутами. На практике такая ситуация встречается редко и скорее является теоретической. Важно заметить, что каждая последующая нормальная форма предполагает нахождение таблиц базы данных в нормальных формах предыдущего уровня. Следовательно, если существуют сложные зависимости, то существуют и простые, которые устраняются разделением таблиц. Таким образом, подходя к пятой нормальной форме имеющиеся данные должны находиться в нескольких таблицах. Нарушение пятой нормальной формы происходит, когда при попытке создания представления (соединения) исходных таблиц в одну происходит дублирование строк. Это возможно, когда в исходных данных атрибут 1 зависит от атрибута 2, атрибут 2 зависит от атрибута 3, который в свою очередь зависит от атрибута 1. Как правило СУБД снабжены специальными механизмами, которые позволяют не допускать дублирования в объединенных таблицах.

Шестая нормальная форма предполагает, что таблица содержит первичный ключ и не более одного неключевого атрибута. В общем смысле нормализация отношений в таблице предполагает ее дробление на множество подтаблиц в целях устранения дублирования информации и упрощения связей. Шестая форма является конечной и содержательно представляет собой вертикальную таблицу с ограничениями, состоящими в обязательном исполнении требований предыдущих нормальных форм.

Таким образом, полностью нормализованная база данных представляет собой множество вертикальных таблиц. При наличии такого представления размещение новых типов данных становится нетрудоемкой задачей (при условии относительно небольшого количества самих таблиц). Идея нормальных форм и удобства масштабирования базы данных лежит в основе теории вертикальных баз данных «сущность-атрибут-значение». При этом сами таблицы могут не соответствовать или частично не соответствовать нормальным формам, но, как следует из приведенного выше анализа, это соответствие не всегда означает лучшие характеристики базы данных: производительность, объем занимаемой памяти, удобство написания запросов к базе данных.

В начале раздела утверждалось, что построение базы данных для энергетического атласа России – задача многоэтапная, и, следовательно, расширение базы данных будет происходить регулярно. Для этого необходима гибкость модели «сущность-атрибут-значение». Нормализация отношений в такой модели будет произведена с учетом приоритетности простоты архитектуры, позволяющей без трудоемких запросов извлекать необходимую информацию. С другой стороны, тема оптимизации баз данных далеко не

ограничивается теорией о нормальных формах, а представляет весьма обширный перечень инструментов, позволяющих оптимизировать целевые характеристики базы данных, и включает в себя работу с индексами, представлениями и прочее. В следующем подразделе будут продемонстрированы основные шаги по созданию архитектуры базы данных на основе модели «сущность-атрибут-значение» для целей проекта, а также описаны методы оптимизации, которые были применены.

Архитектура базы данных проектировалась в логике модели «сущность-атрибут-значение». Концепция простой модели предполагает создание как минимум трех таблиц: таблицы с перечислением сущностей, таблицы с перечислением всех используемых атрибутов, таблицы со списком значений. Собственная база данных включает лишь два усложнения простой модели, что связано с поиском оптимального выбора между производительностью и удобством написания пользовательских запросов.

Первое отличие от простой модели состоит в разделении таблицы с сущностями на две: первая содержит агрегированные сущности – отрасли энергетического рынка страны, а вторая – объекты соответствующих отраслей. Это необходимо для быстрого извлечения всей информации, содержащейся в базе данных по определенной отрасли. Как в дальнейшем будет продемонстрировано, индикатор принадлежности к классу входит в каждую таблицу со значениями, что и позволяет быстро совершать подобную операцию. Поскольку одной из основных задач проекта является возможность отображения на интерактивной карте всех объектов энергетики, то желательно, чтоб скорость выполнения запроса к базе данных являлась минимальной. К тому же, можно предположить, что потенциальных пользователей «Энергетического атласа» будут интересовать в первую очередь массивы информации по отраслям, а не характеристики единичных объектов. В связи с этим важно заложить в логику модели возможность быстрого поиска и извлечения большого количества данных. Преимущества использованного разбиения и включения поля с отнесением объекта к классу (id) в каждую таблицу со значениями связаны с удобством индексирования записей таблиц, и будут изложены позже.

Вторым отличием от простой модели «сущность-атрибут-значение» является разбиение таблицы со значениями на семь таблиц в соответствии с их типами данных. Такое решение приводит к увеличению производительности СУБД, и при этом незначительно усложняет структуру базы данных и написание пользовательских запросов.

Первая таблица с сущностями (таблица) содержит шесть классов энергетических объектов. Это означает, что база данных включает различные сведения по следующим категориям: нефтеперерабатывающие заводы (NPZ), газовые компрессорные станции (Compressor), нефтяные и газовые месторождениям (il_Gas_fields), нефтяные и газовые трубопроводы (Piple), нефтяные терминалы (Terminal), электрические станции

(Power_plants). Номер по порядку в таблице свидетельствует о том, что база данных хранит записи не в упорядоченном виде и не согласно последовательности внесения данных, а в некотором случайном порядке. Упорядочивание по значениям столбца «id» было произведено после выгрузки из базы данных в редакторе Python. Во всех приведенных далее таблицах первый столбец также не имеет прямого отношения к базе данных и был добавлен позднее, для нумерации записей в порядке их извлечения. Стоить заметить, что особенностью языка python является нумерация, начинающаяся с нуля. Поэтому первой извлеченной записи присваивается нулевой номер. Однако упорядоченное хранение записей в базе данных возможно и даже необходимо для ускорения выполнения запросов для чего необходимо применять индексирование. Добавление столбца с номером необходимо для демонстрации важности некоторых решений, использованных при проектировании таблиц, связанных с индексированием, что будет изложено позднее.

Таблица 10 – Таблица с сущностями, агрегирующая объекты до уровня отрасли

№	id	type
2	1	NPZ
1	2	Compressor
0	3	Oil_Gas_fields
3	4	Piple
4	5	Terminal
5	6	Power_plants

Примечание – Источник: составлен автором.

Вторая таблица с атрибутами-объектами содержит полный перечень наименований объектов по всем классам. Таблица является фрагментом аналогичной таблицы из базы данных и содержит по три объекта из всего множества для каждого класса. Поле «id» таблицы не заполняется числовыми значениями как таковыми. СУБД в соответствующих ячейках хранит ссылки на родительскую таблицу, где каждому номеру сопоставлено название класса. Это позволяет при необходимости быстро восстановить смысл числового значения идентификационного номера или совершить обратную операцию: для определенного класса извлечь все имеющиеся объекты. Поле «objects» имеет текстовый тип данных «varchar» с максимально допустимым размером в 200 символов. Если использовать текстовый тип «charter» с ограничением на длину строки, то неиспользованное пространство символов будет заполнено пробелами, что означает нерациональное использование памяти. PostgreSQL так же предлагает использовать тип «text», на который не налагается ограничение по длине строки, при этом он в худшую сторону отличается по быстродействию (сортировка по полю с указанным типом данных выполняется медленнее[CITATION Obe \l 1049]).

Таблица 11 – Частичный список наименований объектов энергетики

№	id	objects
0	1	Александровский НПЗ
1	1	Ангарская НХК

2	1	Анжерская НГК
89	2	ЛПДС Салым
90	2	ПХГ Дамборжице
91	2	КС-32 Ромны
563	3	Западно-Салымское месторождение
564	3	Снежное месторождение
565	3	Майское месторождение
889	4	Газопровод Касимовское ПХГ - КС Воскресенск
890	4	Газопровод Голубой поток
891	4	Газопровод Уренгой - Помары - Ужгород
983	5	Нефтяной терминал «Шесхарис»
984	5	Морской терминал Каспийского Трубопроводного Консорциума
985	5	Нефтяной терминал «Таманьнефтегаз»
699	6	Омская ТЭЦ-4
1006	6	Саяно-Шушенская ГЭС
1007	6	Красноярская ГЭС

Примечание – Источник: база данных проекта «Энергетический атлас»

Наименования объектов, как правило, лаконичны и не требуют большого объема пространства памяти и, тем более, безразмерного типа текстовых данных. Поле таблицы «objects» содержит ограничение на уникальность. Это предотвращает дублирование, которое может произойти при объединении данных из разных источников, если в процессе сравнения один и тот же объект был воспринят, как два отличных. Для поля «objects» создается уникальный индекс, налагающий ограничения на уникальность записи с учетом регистра. Подобное ограничение позволит избежать создания дублирующих друг друга объектов, даже в том случае, когда на стадии предобработки данных, собранных из различных источников и объединенных в одну таблицу, не была проведена проверка на наличие совпадающих наименований, записанных в разных регистрах.

Таблица представляет собой классическую таблицу атрибутов в EAV модели. В данном случае представлен фрагмент со случайными атрибутами. Поле таблицы «key» содержит транслитерацию сокращенных названий всех имеющихся показателей в базе данных. Это позволяет таблицам с данными не хранить повторяющиеся текстовые строки названий показателей, а использовать ссылки с ключами на данную таблицу, по которой можно восстановить полное описание показателя.

Далее будут представлены фрагменты таблиц с данными, в соответствии с разделением данных на семь различных типов и описаны основные механизмы их создания и взаимодействия в среде PostgreSQL.

Таблица 12 – Таблица с определениями атрибутов

№	key	description
9	Glubina_pererabotki	Глубина переработки, %
40	God_vvoda	Год ввода в эксплуатацию
60	Zapasy_MPR1	Запасы нефти по классификации МПР (Приказ 477), извлекаемые запасы A+B1+B2+C1 (млн тонн)
83	Diametr_trub	Диаметр труб (мм)
89	Maks_part	Максимальное количество отгружаемой партии нефти, тыс. тонн
92	Emkost_rezervuarnogo_parka	Емкость резервуарного парка, м. куб.
97	dedvejt	Максимальное водоизмещение (дедвейт) принимаемого судна, тыс.

№	key	description
		тонн
100	Napravlenia_postavki	Направления поставки
103	Godovaa_vyrabotka_elektricestva	Годовая выработка электричества, КВт*ч
110	URUT_e_e	Удельный расход условного топлива на э/э (УРУТ э.э.). г/кВт*ч

Примечание – Источник: база данных проекта «Энергетический атлас»

Наибольший объем данных на момент написания настоящего отчета содержался в таблице с численными характеристиками, где в каждом показатели находились или потенциально могли находиться дробные значения. Не учитывая добавленный столбец с нумерацией извлечения записей[4] из базы данных, таблица состоит из пяти полей: класс объекта «id», наименование объекта «objects», название показателя «key», дата внесения показателя «date» и численное значение показателя «value_n». Четыре первых поля являются составным индексом таблицы. Это означает, что только комбинация из четырех значений однозначно определяет числовое значение, хранящееся в последнем поле. При этом первые три поля фактически не содержат значений, а хранят лишь ссылки на соответствующие значения в таблицах, описанных выше. Каждое поле составного ключа проиндексировано автоматически самой СУБД. Наличие индексов ускоряет выборку данных. Поскольку СУБД изначально хранит записи каждой таблицы в неупорядоченном виде, то при запросе к определенной строке, происходит перебор всех строк, пока не выполниться условие поиска. Индексирование предполагает хранение информации в таблице уже в упорядоченном виде, что ускоряет перебор строк.

Тип данных для последнего поля таблицы «numeric». В PostgrSQL этот тип данных обозначает вещественное число с заданной точностью. Задание точности предполагает определение максимального масштаба и точности (precision). Масштаб указывает максимальное количество цифр, стоящих после запятой, разделяющей целую и дробную часть. Точность – общее количество цифр, стоящих слева и справа от запятой. В данном случае тип «numeric» используется без ограничения на точность и масштаб. Это вызвано тем, что существует вероятность необходимости приведения некоторых показателей к одному масштабу в результате перенесения десятичной запятой на несколько порядков. Для того, чтоб избежать потери точности при этом действии, а также при возможных математических операциях внутри PostgreSQL, с задействованием значений такого типа, ограничения не использовались.

Целочисленный тип данных выделен в отдельную таблицу. Фрагмент таблицы. хранящейся в базе данных, представлен в таблице . Столбец с нумерацией, как и в предыдущем случае, был добавлен после выгрузки и представляет сквозную нумерацию для каждого типа объектов. В качестве примера приведены по три первых объекта

4 В терминах информационных систем записью таблицы называется одна ее строка.

соответствующих своему классу. Выделение целочисленного типа данных в отдельную таблицу, как и в целом разделение по типам данных, приводит к повышению производительности СУБД. Процесс извлечения данных включает в себя сортировку и перебор значений. Осуществлять сортировку данных можно намного эффективнее, если все данные имеют один и тот же тип[CITATION ЕПМ17 \l 1049]. Таблица хранит значения для таких атрибутов, как: год основания; год ввода в эксплуатация; год начала и окончания модернизации; количество газоперекачивающих агрегатов; стадия разработки и прочее.

Таблица 13 – Пример данных из таблицы значений с целочисленным типом

№	id	objects	key	date	value_int
0	1	Александровский НПЗ	God_osnovania	2019-03-03	2001
1	1	Ангарская НХК	God_osnovania	2019-03-03	1953
2	1	Анжерская НГК	God_osnovania	2019-03-03	2010
96	2	КС-32 Ромны	Kolicestvo_gazoperekacivausih_agregatov	2019-03-03	3
97	2	КС-33 Гребенковская	Kolicestvo_gazoperekacivausih_agregatov	2019-03-03	3
98	2	КС-35 Ставишенская	Kolicestvo_gazoperekacivausih_agregatov	2019-03-03	3
1118	3	Холмогорское месторождение	Stadia_razrabotki	2019-03-03	4
1119	3	Пограничное месторождение	Stadia_razrabotki	2019-03-03	4
1120	3	Воргенское месторождение	Stadia_razrabotki	2019-03-03	1
1149	4	Газопровод Касимовское ПХГ - КС Воскресенск	God_vvoda_v_ekspluataziu	2019-03-03	2009
1150	4	Газопровод Голубой поток	God_vvoda_v_ekspluataziu	2019-03-03	2002
1151	4	Газопровод Уренгой - Помары - Ужгород	God_vvoda_v_ekspluataziu	2019-03-03	1983
1293	5	Нефтяной терминал «Шесхарис»	God_vvoda_v_ekspluataziu	2019-03-03	1964
1294	5	Морской терминал Каспийского Трубопроводного Консорциума	God_vvoda_v_ekspluataziu	2019-03-03	2001
1295	5	Нефтяной терминал «Таманьнефтегаз»	God_vvoda_v_ekspluataziu	2019-03-03	2012
1101	6	Пермская ГРЭС	God_vvoda_v_ekspluataziu	2019-03-03	1986
1166	6	Рязанская ГРЭС	God_vvoda_v_ekspluataziu	2019-03-03	1973

Примечание – Источник: база данных проекта «Энергетический атлас»

Ряд показателей в базе данных имеет некоторый диапазон значений. При этом непосредственно значения могут быть, как целыми, так и дробными. Для подобных числовых последовательностей была создана отдельная таблица, в которой первые четыре поля стандартные: класс объекта, название объекта, название показателя (ключ), дата внесения данных.

Как и в таблицах со значениями выше, четыре поля в совокупности являются составным ключом таблицы и хранят ссылки на значения в соответствующих таблицах. кроме поля с датами. Последнее поле имеет тип данных «array numeric». Это означает, что ячейки такого поля могут хранить одномерные массивы из вещественных чисел. Таблица содержит случайную выборку показателей с интервалами значений. Среди них: плотность нефти при 15 градусах, длина причала и прочие.

Таблица 14 – Пример данных из таблицы значений, хранящей диапазоны значений с вещественным типом данных

№	id	objects	key	date	value_fl
0	3	Талинское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.67. 0.7]
1	3	Самотлорское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.85]
2	3	Варьеганское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.77. 0.86]
3	3	Кулешовское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.821. 0.828]
4	3	Бариновско-Лебяжинское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.856. 0.883]
5	3	Мухановское месторождение	Plotnost__nefti_pri_15_grad	2019-03-03	[0.807. 0.832]
49	5	Нефтяной терминал «Шесхарис»	Dlina_pricala	2019-03-03	[490]
50	5	Терминал АО "Усть-Луга Ойл"	Dlina_pricala	2019-03-03	[300]
51	5	Терминал ООО «Новороссийский мазутный терминал»	Dlina_pricala	2019-03-03	[220]
52	5	АО "Туапсинский морской торговый порт"	Dlina_pricala	2019-03-03	[300]
53	5	Терминал ООО «Курганнефтепродукт»	Dlina_pricala	2019-03-03	[163]
54	5	Распределительный перевалочный комплекс (РПК) — Высоцк «ЛУКОЙЛ-II»	Dlina_pricala	2019-03-03	[260]

Примечание – Источник: база данных проекта «Энергетический атлас»

В отдельную таблицу также выделены данные, которые представляют собой диапазон значений из натуральных чисел. Случайная выборка показателей из аналогичной таблицы в базе данных представлена в таблице .

В качестве примеров приведены такие атрибуты, как: глубина залегания нефти (измеряется в метрах и из-за неравномерности распределения пласта по глубине залегания, как привило предоставляется в виде интервальной оценки); диаметр труб газопровода. Диаметр труб в данном случае имеет диапазон значений поскольку один трубопровод, как правило, имеет ветвления, и в него входят трубы с разным диаметром.

Для целей проекта «Энергетический атлас» желательно располагать информацией относительно протяженности и географических координат для каждой трубы газопровода, однако на данный момент подобная информация отсутствует и база данных хранит описательную статистику по каждому газопроводу.

Таблица 15 – Пример данных из таблицы значений, хранящей диапазоны значений из натуральных чисел

№	id	objects	key	date	value_arn
0	3	Западно-Мессояхское месторождение	Glubina_zalegania	2019-03-03	[800]
1	3	Восточно-Мессояхское месторождение	Glubina_zalegania	2019-03-03	[800]
2	3	Холмогорское месторождение	Glubina_zalegania	2019-03-03	[2850]
3	3	Вынгапуровское месторождение	Glubina_zalegania	2019-03-03	[2519. 2622]
4	3	Средне-Итурское месторождение	Glubina_zalegania	2019-03-03	[3025]
5	3	Новогоднее месторождение	Glubina_zalegania	2019-03-03	[2825]
76	4	Газопровод Касимовское ПХГ - КС Воскресенск	Diametr_trub	2019-03-03	[1220]
77	4	Газопровод Голубой поток	Diametr_trub	2019-03-03	[610. 1400]
78	4	Газопровод Уренгой - Помары - Ужгород	Diametr_trub	2019-03-03	[1420]
79	4	Газопровод СРТО - Торжок	Diametr_trub	2019-03-03	[1420]

№	id	objects	key	date	value_arn
80	4	Газопровод Сахалин - Хабаровск - Владивосток	Diametr_trub	2019-03-03	[700. 1220]
81	4	Газопровод Ямал - Европа	Diametr_trub	2019-03-03	[1420]

Примечание – Источник: база данных проекта «Энергетический атлас»

В таблице представлены примеры значений атрибутов с текстовым типом данных. Текстовые характеристики объектов занимают значительную часть в общем объеме имеющихся данных. Среди таких характеристик: официальное название объекта; названии компании собственника; источник поставок; тип объекта; состояние; названия отдельных агрегатов объекта и прочее. Тип данных для поля со значениями был использован «Varchar» с ограничением на количество символов в 1000 значений.

Таблица 16 – Пример данных из таблицы значений с текстовыми данными

№	id	objects	key	date	value
312	1	Ангарская НХК	Ofizial_noe_nazvanie	2019-03-03	ОАО «Ангарская нефтехимическая компания»
313	1	Анжерская НГК	Ofizial_noe_nazvanie	2019-03-03	ООО "АНГК"
314	1	Антипинский НПЗ	Ofizial_noe_nazvanie	2019-03-03	АО «Антипинский НПЗ»
0	2	КС-32 Ромны	Istocniki_postavok	2019-03-03	КС Курская
1	2	КС-33 Гребенковская	Istocniki_postavok	2019-03-03	КС-32 Ромны
2	2	КС-35 Ставишенская	Istocniki_postavok	2019-03-03	КС-34 Софиевка
3464	3	Западно-Салымское месторождение	Kompania	2019-03-03	«Салым Петролеум Девелопмент Н.В.»
3465	3	Снежное месторождение	Kompania	2019-03-03	Imperial Energy
3466	3	Майское месторождение	Kompania	2019-03-03	Imperial Energy
5349	4	Газопровод Касимовское ПХГ - КС Воскресенск	Kompania	2019-03-03	ПАО «Газпром»
5350	4	Газопровод Голубой поток	Kompania	2019-03-03	ПАО «Газпром»
5733	5	Нефтяной терминал «Шесхарис»	Tip	2019-03-03	нефтяной
5735	5	Нефтяной терминал «Таманьнефтегаз»	Tip	2019-03-03	нефтяной
3044	6	ТЭЦ-22	Sostoanie	2019-03-03	в эксплуатации
3235	6	ВЭС "Каневская"	Sostoanie	2019-03-03	строится
3446	6	ВЭС-Щербиновский	Sostoanie	2019-03-03	строится

Примечание – Источник: база данных проекта «Энергетический атлас»

Отдельная таблица предусмотрена для данных, представляющих собой, массив с текстовой информацией. В таблице приведены случайные примеры атрибутов и соответствующих им значений для определенных объектов из базы данных проекта. К примеру, среди таких данных встречаются: виды продукции, выпускаемой нефтеперерабатывающими заводами; энергоблоки электростанций и их характеристика, и прочее.

Таблица 17 – Пример данных из таблицы значений с массивами текстовых данных

№	id	objects	key	date	value_art
0	1	Александровский НПЗ	Produkzia	2019-03-03	['Бензин газовый стабильный СТО 1-2009г.'. 'Бензин прямогонный экспортный. марки N2. ТУ 38.001256-76'. 'Дизельное топливо Л-0.2-40. ГОСТ 305-82'. 'Топочный мазут марки М-100 ГОСТ 10585-9']
1	1	Ангарская НХК	Produkzia	2019-03-03	['Бензин неэтилированный Премиум Евро-95 вид II (АИ-95-К4)'. 'Бензин неэтилированный

Продолжение таблицы

№	id	objects	key	date	value_art
					Премиум Евро-95 вид III (АИ-95-К5)'. 'Бензин неэтилированный Супер Евро-98 вид II (АИ-98-К4)'. 'Бензин неэтилированный Супер Евро-98 вид III (АИ-98-К5)'. 'Неэтилированный бензин Регуляр-92 (АИ-92-К4)'. 'Мазут топочный 100. 1.50%. малозольный. 25°С'. 'Топливо маловязкое судовое. вид I. II. III'. 'Топливо судовое остаточное RME 180 вид II'. 'Топливо дизельное ЕВРО класс 2 вид III (ДТ-3-К5)'. 'Топливо дизельное ЕВРО класс 4 вид III (ДТ-3-К5)'. 'Топливо дизельное ЕВРО Сорт С вид III (ДТ-Л-К5)'. 'Топливо дизельное ЕВРО сорт Е вид II (ДТ-Е-К4)'. 'Топливо ТС-1. в/с']
2	1	Анжерская НГК	Produkzia	2019-03-03	['первичная переработка нефти']
3	1	Антипинский НПЗ	Produkzia	2019-03-03	['Бензин газовый стабильный;'. 'Топливо дизельное ЕВРО класса 5;'. 'Кокс нефтяной (анодный. электродный);'. 'Сжиженные углеводородные газы (ПА. ПБА. ПБТ);'. 'Сера техническая гранулированная газовая.']
4	1	Арктикнефть	Produkzia	2019-03-03	['первичная переработка нефти'. 'Дизельное топливо. мазут. прямой бензин для собственных нужд и местных точек продаж']
30	6	Няганская ГРЭС	Energobloki	2019-03-03	['Энергоблок 1 в эксплуатации Мощность: 421 МВт Год ввода в эксплуатацию: 2013'. ' Энергоблок 2 в эксплуатации Мощность: 424 МВт Год ввода в эксплуатацию: 2013'. ' Энергоблок 3 строится Мощность: 420 МВт']
41	6	Уренгойская ГРЭС	Energobloki	2019-03-03	['Энергоблок 1 в эксплуатации Мощность: 24 МВт'. ' Энергоблок 2 в эксплуатации Мощность: 450 МВт Год ввода в эксплуатацию: 2012']
53	6	Псковская ГРЭС	Energobloki	2019-03-03	['Энергоблок 1 в эксплуатации Мощность: 215 МВт'. ' Энергоблок 2 в эксплуатации Мощность: 215 МВт']
75	6	Пермская ГРЭС	Energobloki	2019-03-03	['Энергоблок 1 в эксплуатации'. ' Энергоблок 2 не установлено Мощность: 800 МВт'. ' Энергоблок 4 не установлено Мощность: 800 МВт']

Примечание – Источник: база данных проекта «Энергетический атлас»

Последняя вертикальная таблица в базе данных содержит географические координаты объектов, по которым имеется подобная информация. В качестве примера, смотри таблицу .

Таблица 18 – Пример части таблицы с координатами объектов

№	id	objects	location
0	1	Александровский НПЗ	(56.466504.84.983483)
1	1	Ангарская НХК	(52.557177.3.9154552)
2	1	Анжерская НГК	(55.326287.6.1167618)
79	2	ПХГ Дамборжице	(49.042798.16.916696)
80	2	КС-32 Ромны	(51.039722.34.630278)
81	2	КС-33 Гребенковская	(50.448333.33.3975)
552	3	Западно-Салымское месторождение	(60.288173.70.897522)
553	3	Снежное месторождение	(58.924854.79.694824)
554	3	Майское месторождение	(60.390791.72.907505)
870	5	Нефтяной терминал «Шесхарис»	(44.707981.37.845197)
871	5	Морской терминал Каспийского Трубопроводного Консорциума	(44.67085.37.65388)

№	id	objects	location
872	5	Нефтяной терминал «Таманьнефтегаз»	(45.152385.36.67614)
890	6	Нерюнгринская ГРЭС	(56.668889.124.853333)
891	6	ТЭЦ-8	(55.72425.37.696611)
895	6	Саяно-Шушенская ГЭС	(52.825842.91.370626)

Примечание – Источник: база данных проекта «Энергетический атлас»

Поскольку изменение локации объекта представляется маловероятным, таблица не содержит поля с датой. Однако, при возникновении прецедента, можно воспользоваться преимуществом EAV модели, состоящим в легкости ее изменения и добавить информацию о дате занесения данных и хранить историю изменения геолокации объекта в пространстве. Поле объект однозначно определяет атрибут и может использоваться в качестве ключа, однако в таблице используется составной ключ, включающий также и класс объекта. Это необходимо для индексирования, которое позволяет хранить СУБД данные в упорядоченном по классам и названиями объектов виде. Подобное упорядочивание может существенно ускорить загрузку координат из базы данных при необходимости отразить на карте все объекты определенного класса.

При необходимости вывода информации для пользователя в привычном представлении широкой таблицы, где по строкам перечислены объекты определенного типа, а по столбцам собраны всевозможные характеристики из различных таблиц баз данных, применяются различные подходы к объединению таблиц, в том числе левые соединения, внутренние соединения, рекурсивные запросы. В таблице обобщены типы данных PostgreSQL, использованные при создании таблиц с примерами их применения.

В данной главе были приведены основные преимущества выбранной модели базы дынных «Сущность-Атрибут-Значение». В первую очередь к ним отнесено удобство масштабирования, что является критически-важной особенностью для целей выполняемого проекта, а также удобство хранения информации для случаев, когда количество характеристик превышает количество объектов для определенного класса. Был описан процесс создания собственной архитектуры базы данных на основе выбранной модели; приведены фрагменты таблиц базы данных и их основные характеристики.

Таблица 19 – Перечень типов данных PostgreSQL, используемых при конструировании базы данных

Тип данных PostgresSQL	Описание	Применение
Integer	Знаковое четырехбайтное число	Числовые данные, предполагающие только целые числа (год основания, численность сотрудников, год модернизации и прочее)
Numeric	Вещественное число заданной точности	Числовые данные, которые могут быть представлены в виде десятичной дроби (годовая выработка электроэнергии, среднесуточная добыча нефти, коэффициент извлечения нефти, мощность газопровода и другие).
Point	Геометрическая точка на плоскости	Географические координаты объектов энергетики
Array numeric	Массив из вещественных чисел заданной точности	Числовые данные. которые имеют некоторый диапазон значений действительных чисел

Тип данных PostgresSQL	Описание	Применение
		(плотность нефти, вязкость нефти)
Array integer	Массив из знаковых четырехбайтных чисел	Числовые данные, которые имеют некоторый диапазон значений целых чисел (глубина залегания нефти, диаметр труб газопровода, длина причала)
Array text	Массив из символьных строк переменной длинны	Перечисления однородных характеристик объекта (виды производимой продукции, марки энергетических установок, характеристики энергоблоков и прочее)
Varchar	Символьная строка переменной длины	Текстовые характеристики объектов (основное топливо, тип нефти по вязкости, степень освоения месторождения, тип причалов и прочее)
Serial	Четырехбайтное целое с автоувеличением	Применяется для присвоения id записи
Date	Календарная дата (год, месяц, день)	Дата обновления данных или дата, на которую данные были актуальны.

Примечание – Источник: Официальная документация PostgreSQL[CITATION Pos18 \l 1033]

4 Определение удельного расхода топлива на выработку электрической энергии и теплоты на тепловых электростанциях

Для демонстрации возможности собранной базы данных по энергетической системе России были рассчитаны показатели удельных затрат топлива на генерацию тепловой и электрической энергии на теплоэлектростанциях. На данные объекты энергетической инфраструктуры приходится основная нагрузка по выработке электричества, по данным Минэнерго на 2018 г. до 60 %, а также значительный объем генерации тепла – до 46% от общей выработки. Тепловая электростанция (ТЭС) – это электростанция, которая вырабатывает электрическую энергию за счет преобразования химической энергии топлива в процессе сжигания в тепловую, а затем в механическую энергию вращения вала электрогенератора. Выделяют два типа тепловых электростанций: конденсационные электростанции и теплоэлектроцентраль.

Конденсационные электростанции (КЭС) производят только электрическую энергию, как правило располагаются вблизи мест добычи используемого топлива. Принцип работы, следующий: под большим давлением в котёл, подводится питательная вода, топливо и атмосферный воздух для горения. Топливо в котле сгорает, преобразуя химическую энергию в тепловую, нагревая стенки котла, экранированные трубами, по которым бежит питательная вода, преобразующаяся в пар высокого давления при нагревании. Полученный пар попадает в паровую турбину с помощью трубопроводов, заставляя ее вращаться (вырабатывать механическую энергию), при этом он расширяется до очень низкого давления и потенциальная энергия сжатого и нагретого до высокой температуры пара превращается в кинетическую энергию вращения ротора турбины. Турбина приводит в движение электрогенератор, преобразующий кинетическую энергию вращения ротора генератора в электрический ток. Конденсатор необходим для конденсации пара, который поступает из турбины (далее возвращается обратно в котел) и создания глубокого разрежения, из-за которого происходит расширение пара в турбине. Он создаёт вакуум на выходе из турбины, поэтому пар, поступив в турбину с высоким давлением, движется к конденсатору и расширяется, что обеспечивает превращение его потенциальной энергии в механическую работу [CITATION Бур07 \l 1049].

Теплоэлектроцентрали (ТЭЦ) могут одновременно вырабатывать как тепло, так и электрическую энергию. ТЭЦ различаются по устройству турбин и используемому топливу. Основное отличие в принципе работы ТЭЦ от КЭС в возможности отобрать часть тепловой энергии пара после того, как он выработает электрическую энергию, для нагрева воды, которая составляет горячее водоснабжение и теплофикацию.

Согласно методике расчета нормативов удельных расходов топлива по теплоэлектростанциям расчеты выполняются по каждому турбоагрегату и каждому котлоагрегату, входящему в состав группы оборудования. По группе в целом показатель определяется путем суммирования или взвешивания расчетов показателей входящих в состав агрегатов [CITATION Зак1 \l 1049]. Поэтому используя собранную статистику из открытых источников (базы данных Атласа энергетической системы России), касающуюся технологических характеристик основного оборудования теплоэлектростанций (турбинных установок и котлоагрегатов) и описанный выше принцип работы станции, рассчитаем удельные затраты топлива на генерацию тепловой энергии для каждого котлоагрегата (уравнение (3)) через формулу расчета его КПД, представленную уравнением (1), использовав преобразование (2):

$$\eta_к = \frac{D \cdot \left(h_{п.пара} - h_{п.воды}\right)}{B \cdot q_{нр}}, \quad (1)$$

$$Q_{полезное} = D \cdot \left(h_{п.пара} - h_{п.воды}\right), \quad (2)$$

$$b_т = \frac{B}{Q_{полезное}} = \frac{1}{\eta_к \cdot q_{нр}} \quad (3)$$

где

$\eta_к$ – КПД котла;

D – расход пара;

B – расход топлива;

$h_{п.пара}$ – удельная энтальпия перегретого пара;

$h_{п.воды}$ – удельная энтальпия полезной воды[5];

$q_{нр}$ – удельная теплота сгорания топлива[6];

$Q_{полезное}$ – полезная теплота сгорания;

$b_т$ – удельный расход топлива на выработку теплоты на электростанции;

При расчете удельных затрат топлива на генерацию электроэнергии предполагаем, что теплоэлектростанция работает в комбинированном режиме, получаем уравнение (4):

$$b_{ээ} = \frac{0{,}123}{\eta_к \cdot \eta_{турб}} \quad (4)$$

где

$b_{ээ}$ – удельный расход топлива на выработку электроэнергии на электростанции;

5 Для расчета удельной энтальпии перегретого пара и полезной воды используются таблицы или калькулятор перевода свойств перегретого пара в значение удельной энтальпии [https://www.tlv.com/global/RU/calculator/superheated-steam-table.html]

6 Сравнительная таблица теплотворности некоторых видов топлива [https://a-invest.com.ua/aktualno/tablitsa-teplotvornosti.html]

$\eta_к$ – КПД котла;

$\eta_{турб}$ – электромеханический КПД турбины.

Авторы работы понимают, что в действительности данный показатель зависит от количества, вырабатываемого станцией в конденсационном режиме, схемы соединения агрегатов между собой, а также от фактического спроса на тепло и электроэнергию, поэтому полученные результаты в разрезе агрегатов, теплоэлектростанций и регионов можно считать оценками сверху.

При расчете учитывается возраст оборудования по формуле (5):

$$\eta_{i,2019} = \eta_{i,\text{год ввода}} \cdot (1 - П_c) \quad (5)$$

где

$\eta_{i,\text{год ввода}}$ – КПД нового агрегата,

$\eta_{i,2019}$ – текущий КПД (в 2019 г.) с учетом износа,

$П_c$ – показатель старения (зноса), %.

Под износом понимается технологически невосстанавливаемый ремонтами естественный физический износ оборудования при длительной его эксплуатации, ухудшение технического состояния, снижение экономичности. Данный фактор увеличивает показатель удельных затрат топлива на генерацию энергии.

В таблице представлены рекомендуемые Минэнерго значения при отсутствие испытаний, проводимых на теплоэлектростанциях, о зависимости износа оборудования и удельного расхода топлива на генерацию энергии, которые были использованы при расчетах.

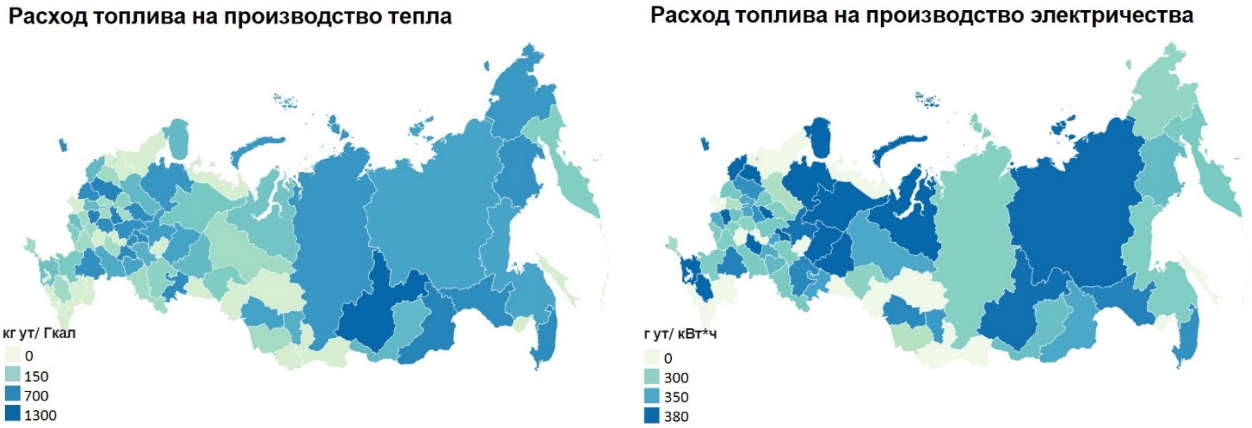
Полученные численные значения удельных затрат топлива на генерацию тепла и электроэнергии по агрегатам усреднялись на уровне станции. Далее полученные значения усреднялись по станциям, находящимся в одном и том же субъекте федерации. Результаты расчетов в региональном разрезе представлены на рисунке .

Таблица 20 – Рекомендуемая Минэнерго таблица учета износа оборудования при расчетах удельных затрат топлива на генерацию тепла

Котлоагрегаты	Показатель старения ($П_c$), % к КПД котлоагрегата			
	более 5 до 10 лет	более 10 до 15 лет	более 15 до 20 лет	более 20 до 25 лет
Котел водогрейный производительностью более 20 Гкал/ч	-	0,5	1,0	2,0
Котел водогрейный производительностью более 4 до 20 Гкал/ч	0,5	1,0	2,0	3,0
Котел водогрейный производительностью до 4 Гкал/ч	0,75	1,5	3,0	4,0
Котел паровой с рабочим давлением до 14 кгс/см2	-	1,0	1,5	2,0

Примечание – Источник: Приказ Министерства энергетики РФ от 30 декабря 2008 г. N 323 «Об утверждении порядка определения нормативов удельного расхода топлива при производстве электрической и тепловой энергии» (с изменениями и дополнениями)

Для содержательной интерпретации полученных показателей на региональном уровне необходимо взглянуть на используемую в расчетах статистику на аналогичном уровне представления (тоже в региональном разрезе): тип основного топлива, используемого для генерации тепловой и электрической энергии в регионе; средний возраст генерирующего оборудования в регионе.



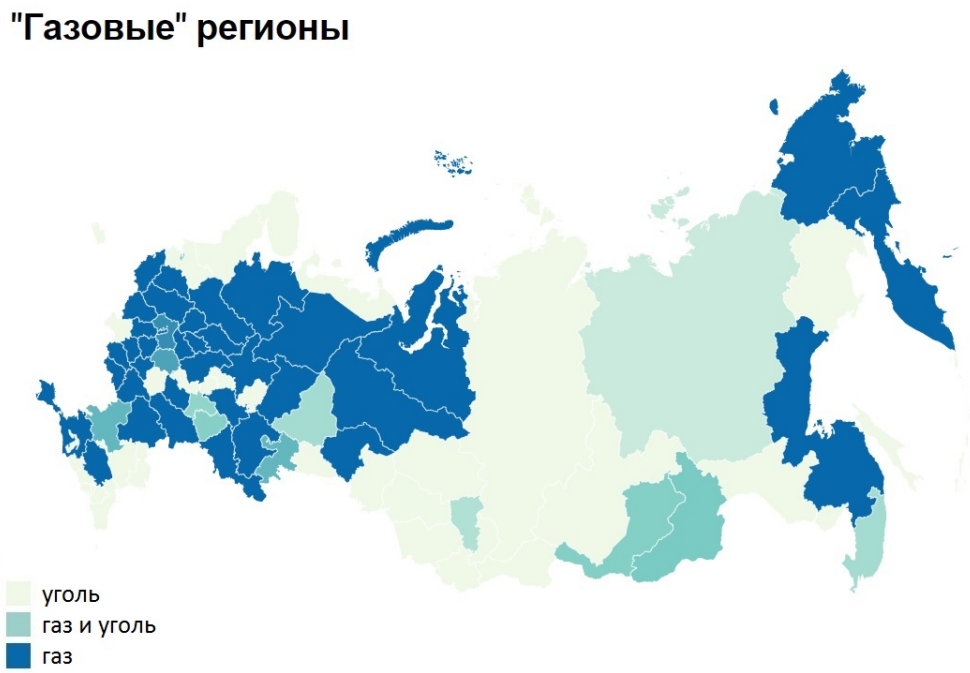
Примечание – Источник: расчеты авторов.

Рисунок 9 – Удельный расход топлива на производство тепловой энергии (кг ут/ Гкал) и электрической энергии (г ут/ кВТ*ч)

На рисунке представлены регионы, разделенные на три кластера: в кластер 1 входят регионы, которые для генерации тепла и электроэнергии используют в качестве основного топлива природный газ; в кластер 2 входят регионы, которые для генерации тепла и электроэнергии используют в качестве основного топлива уголь; в кластер 3 регионы, которые для генерации тепла и электроэнергии используют в качестве основного топлива и природный газ, и уголь. Разделение на кластеры производилось с учетом доли установленной мощности в общем объеме генерации.

Видно, что большинство газовых регионов сосредоточилось в западной части России, также в этот кластер (кластер 1) вошли: Хабаровский край, Чукотский АО и Камчатский край.

В кластер 2, использующий уголь в качестве основного топлива при генерации электричества и тепла вошли все восточные регионы и некоторые западные, в частности: Адыгея, Ингушетия, Кабардино-Балкария, Калмыкия, Дагестан, Северная Осетия, Чечня, Карелия, Мурманская область, Ненецкий АО.



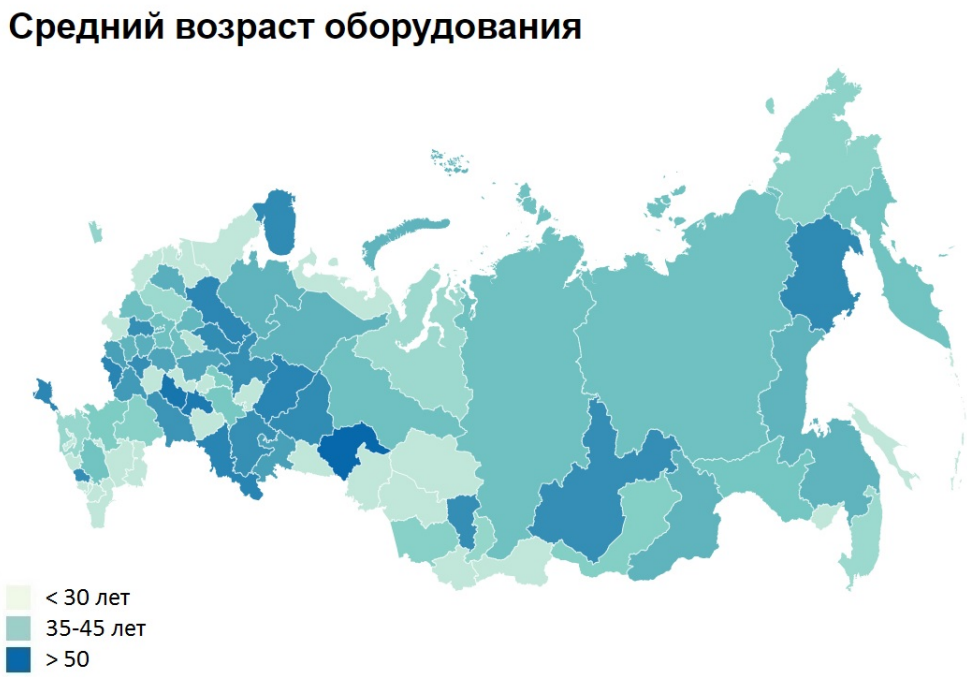
Примечание – Источник: расчеты авторов.

Рисунок 10 – Региональные кластеры по типу основного топлива, используемого при генерации тепла и электричества

В кластер 3 вошли: Москва, Московская область, Челябинская область, Свердловская область, Самарская область, Ульяновская область, Ростовская область, Кемеровская область, Забайкальский край, Приморский край, Бурятия, Якутия. При совместном рассмотрении статистики по основному топливу, необходимому для генерации тепловой и электрической энергии и средним удельным затратам топлива для генерации тепловой и электрической энергии можно отметить следующий вывод: расход топлива на производство тепловой энергии меньше в регионах использующих в качестве основного топлива природный газ, так как теплотворная способность угля ниже теплотворной способности газа почти в 2 раза (удельная теплота сгорания 1 кг угля 6 450 кКал, метана[7] – 11 950 кКал). Что касается типа топлива и генерации электроэнергии, то взаимосвязь неочевидна, необходимо использовать другую дополнительную информацию. Если высокие удельные затрат на генерацию электроэнергии в Мурманской области, Якутии, Иркутской области можно объяснить тем, что в качестве основного топлива используется уголь, то для ЯНАО, Коми, Архангельской области, Пермского края, Свердловской области данный вывод сделать нельзя (основным топливом является газ, тем не менее издержки генерации электроэнергии выше 380 г ут/ кВт*ч).

Рассмотрим еще один показатель в региональном разрезе, который учитывался при расчете удельного расхода топлива на генерацию тепла и электричества: средний возраст оборудования, представленный на рисунке .

7 По магистральным трубопроводам транспортируется именно метан.



Примечание – Источник: расчеты авторов.

Рисунок 11 – Средний возраст оборудования, генерирующего тепловую и электрическую энергию в региональном разрезе

Возраст оборудования рассчитывался на 2019 г., датой отчета служил ввод в эксплуатацию оборудования. Чем старше оборудование, тем выше его износ, а значит ниже эффективность (удельный расход топлива на генерацию энергии выше). Данное предположение выполняется для генерации тепловой энергии, видно, что регионы со средним возрастом службы оборудования, превышающего 30 лет: Магаданская область, Иркутская область, Кемеровская область, Тюменская область, Ульяновская область, большая часть регионов Центрального федерального округа тратят больше топлива на генерацию 1 Гкал тепла, чем регионы имеющие более новое оборудование. При генерации электрической энергии данное предположение выполняется для: Иркутской области, Кемеровской области, Мурманской области, Свердловская область, Пермский край, Северная Осетия.

Анализ литературы, показал, что на расход топлива при генерации электрической энергии помимо типа топлива, возраста оборудования, влияет степень загрузки турбины и температура сетевой воды. В статье Липатова Т.В. [CITATION Лип13 \l 1049], было экспериментально показано, что снижение степени электрической нагрузки от максимальной до минимальной вызывает рост удельного расхода топлива на 19-64%. Рост температуры сетевой воды вызывает рост удельного расхода на 17-47% .

Таким образом, можно говорить о неэффективном использование мощностей при высоком расходе топлива на производство электрической энергии в случае, когда в качестве основного топлива используют природный газ и генерирующее энергию

оборудование имеет короткий срок эксплуатации. Например, Краснодарский край, Ставропольский край, республика Коми.

Сравнение усредненных показателей по России в целом с значения, рассчитываемыми Минэнерго, показало расхождение: удельный расход топлива на производство тепла по данным Минэнерго на 2018 г. составляет 157,9 кг ут/ Гкал, по расчету – 163,1 кг ут/ Гкал (рассчитанный показатель находится на уровне 2016 г.); удельный расход топлива на генерацию электричества по данным Минэнерго на 2018 г. составил 309,8 г ут/ кВт*ч, по расчету – 341,9 г ут/ к Вт*ч (рассчитанный показатель находится на уровне начала 2000-х годов). Расхождение можно объяснить отсутствием статистики по году ввода каждого агрегата на рассматриваемых станциях (из открытых источников было собрано 38% информации по данной переменной, год ввода агрегата заменялся на год начала эксплуатации электростанции при отсутствии статистики), а также тем, что в расчетах Минэнерго принимают участие только станции, имеющие установленную электрическую мощность выше 25 МВт.

Таким образом, рассчитанные показатели удельных затрат топлива на генерацию тепла и электричества могут выступать в качестве индикаторов «проблемных» теплоэлектростанций, т.е. тех, что используют неэффективно имеющееся оборудование, требуют ремонта/ модернизации.

ЗАКЛЮЧЕНИЕ

Появление технологических возможностей обработки и хранения больших массивов данных создает спрос на их удобное представление, позволяющее проводить быстрый анализ исследуемого объекта. Энергетическая система Российской Федерации характеризуется сложной географической структурой, на которую оказали влияние как расположение месторождений топливно-энергетических полезных ископаемых и других источников энергии, так и исторически сложившиеся центры концентрации населения, производства и потребления продукции, при изготовлении которой используются как отдельные специфические для конкретной отрасли виды энергоресурсов, так и электроэнергия и тепло. Элементы системы также являются весьма технологически неоднородными в силу разнообразия перерабатываемых источников энергии и длительной истории существования и эволюции энергосистемы. Визуализация на карте перечисленных выше объектов позволяет за короткий период времени оценить географическую концентрацию тех или иных полезных ископаемых, технологическое оснащение (характеристики производственных функций) и инфраструктурные ограничения регионов по их торговым потокам и т.п. Однако, на текущий момент нет подобной карты или атласа для энергетической системы России, поэтому целью работы было его создание. Для этого была собрана статистика по следующим объектам:

- различным типам электростанций,
- разрабатываемым месторождениям топливно-энергетических полезных ископаемых,
- мощностям по их переработке,
- а также информация об используемых технологиях на этих объектах.

Анализ мирового опыта представления информации о географическом расположении, технологическом оснащении и взаимосвязях объектов энергосистемы показал, что на данный момент интерактивная визуализация достаточно дезагрегированной статистики (вплоть до статистики на уровне производственных установок) и возможность выгрузки данных представлены только для американской энергетической системы, разрабатываемой Управлением энергетической информации США (EIA). Поэтому данная реализация была взята за образец при создании макета разметки сайта Атласа энергетической системы России (предварительное техническое задание на визуализацию информационного ресурса), а также при сборе и структурирование данных, необходимых для заполнения базы данных ресурса.

Сбор данных в основном осуществлялся с использованием открытых источников:

– информационной платформы EnergyBase.ru, осуществляющей услуги по рекомендации поставщиков на энергетическом рынке, рекламные услуги. Статистика, представленная на сайте, по техническим характеристикам объектов предоставляется зарегистрированным на платформе собственниками, поэтому собранные данные можно считать актуальными и достоверными. В результате была собрана статистика по нефтегазовым месторождениям (название компании собственника лицензии, объемы запасов, объемы добычи, коэффициент извлечения, химические свойства ископаемого первичного источника энергии, координаты и другое), нефтегазовым магистральным трубопроводам (тип транспортируемого продукта, протяженность, диаметр трубы, направление поставки, точки входа, выхода, мощность), компрессорам (название компании собственника, источник поставок, направление поставок, мощность, годовой объем транспортировки, емкость резервуаров и хранилищ, координаты), перерабатывающим заводам (название компании собственника, мощность, глубина переработки, выход светлых нефтепродуктов, имеющиеся установки, ассортимент выпускаемой продукции), терминалам (название компании собственника, производительность налива, емкость резервуарного парка, вид входящего транспорта, направление поставки, объем отгрузки, координаты и другое), электростанциям (тип станции, название компании собственника, установленная мощность по выработке тепла, установленная мощность по выработке электроэнергии, КИУМ, координаты),

– интернет-энциклопедии Wikipedia.com. Статистика, представленная на персональных страницах объектов энергетической системы ресурса, была выборочно проверена на актуальность и достоверность по мощностям, генерирующим электрическую и тепловую энергию. Проверка осуществлялась с привлечением официальных интернет-ресурсов компаний собственников. В результате верификации расхождений обнаружено не было. С данного источника были собраны технологические характеристики тепло-, электрогенерирующих объектов (ТЭЦ, ТЭС, ГРЭС, СЭС, ГЭС, ВЭС, АЭС), в частности характеристики основного оборудования (тип установки, название агрегата, тип агрегата, основное топливо, резервное топливо, год ввода в эксплуатацию, характеристики пара, мощность и другое),

– технических паспортов агрегатов, входящих в основное оборудование, генерирующих тепло и электроэнергию объектов (КПД, мощность, производительность пара, давление пара, температура питательной воды), в том числе с официальных сайтов компаний-производителей агрегатов.

Для сбора и структуризации статистики использовались автоматические методы сбора, в частности парсер BeautifulSoup для синтаксического разбора файлов HTML/XML, написанный на языке программирования Python, который преобразовывает в том числе и неправильную разметку страницы в дерево синтаксического разбора. Он поддерживает простые и естественные способы навигации, поиска и модификации дерева синтаксического разбора, что позволяет экономить время на сбор данных по сравнению с обычным ручным методом. Ограничение сбора «полной» базы данных из открытых источников является разреженность статистики, поэтому для заполнения пробелов необходимо использовать платные ресурсы (например, статистика по объемам первичной переработки и объему производства нефтепродуктов в продуктовом разрезе и разрезе по НПЗ).

В качестве методологии формирования архитектуры базы данных собранной статистики была выбрана модель «Сущность-Атрибут-Значение» из-за удобства масштабирования, хранения статистики для случаев, когда количество характеристик превышает количество объектов для определенного класса.

Для демонстрации возможности собранной базы данных по энергетической системе России были рассчитаны показатели удельных затрат топлива на генерацию тепловой и электрической энергии на теплоэлектростанциях. Данные показатели могут выступать в качестве индикаторов «проблемных» теплоэлектростанций, т.е. тех, что используют неэффективно имеющееся оборудование, требуют ремонта/ модернизации.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Официальный сайт Fraunhofer ISE: [сайт]. [2019]. URL: https://energy-charts.de/osm.htm (дата обращения: 20.Май.2019).

2 Официальный сайт EEX Transparency Platform: [сайт]. [2019]. URL: https://www.eex-transparency.com/power/ (дата обращения: 20.Май.2019).

3 Официальный сайт ENTSOE Transparency Platform: [сайт]. [2019]. URL: https://transparency.entsoe.eu/transmission/r2/explicitAllocationsIntraday/show?name=&defaultValue=false&viewType=TABLE&areaType=BORDER_BZN&atch=false&dateTime.dateTime=01.01.2019+00:00|UTC|DAY&dateTime.endDateTime=22.05.2019+00:00|UTC|DAY&border.values=C (дата обращения: 20.Май.2019).

4 Официальный сайт Epexspot: [сайт]. [2019]. URL: https://www.epexspot.com/en/market-data/market_data_offers (дата обращения: 20.Май.2019).

5 Официальный сайт Statistisches Bundesamt (Destatis): [сайт]. [2019]. URL: https://www.destatis.de/EN/Home/_node.html (дата обращения: 20.Май.2019).

6 Официальный сайт Federal Ministry for Economic Affairs and Energy: [сайт]. [2019]. URL: https://www.bmwi.de (дата обращения: 20.Май.2019).

7 Официальный сайт The Pan-European Thermal Atlas – Peta: [сайт]. [2019]. URL: https://heatroadmap.eu/peta4/ (дата обращения: 20.Май.2019).

8 Официальный сайт Heat road map: [сайт]. [2019]. URL: https://heatroadmap.eu/energy-models/ (дата обращения: 20.Май.2019).

9 Официальный сайт U.S. Energy Information Administration: [сайт]. [2019]. URL: https://www.eia.gov/tools/ (дата обращения: 6.Июнь.2019).

10 Официальный сайт the Department of Natural Resources Canada: [сайт]. [2019]. URL: http://geoappext.nrcan.gc.ca/GeoCanViz/map/nacei-cnaie/en/index.html (дата обращения: 6.Июнь.2019).

11 Официальный сайт Атласа мирового рынка природного газа и СПГ (Energy Web Atlas): [сайт]. [2019]. URL: http://www.energywebatlas.com/ (дата обращения: 6.Июнь.2019).

12 Официальный сайт Министерства энергетики Российской Федерации: [сайт]. [2018]. URL: https://minenergo.gov.ru/ (дата обращения: 02.Апрель.2018).

13 Арутюнян Г et al., "Обзор электроэнергетическогой отрасли России," Эрнст энд Янг, 2018.

14 Министерство энергетики Российской Федерации, "О целевой модели рынка тепловой энергии," Министерство энергетики Российской Федерации, 2015.

15 Группа по оказанию услуг компаниям ТЭК и электроэнергетики в СНГ, "Обзор электро-энергетической отрасли России," ООО "Эрнст энд Янг — оценка и консультационные услуги", 2018. [Online]. https://www.ey.com/Publication/vwLUAssets/EY-power-market-russia-2018/$File/EY-power-market-russia-2018.pdf

16 Российское энергетическое агентство, "Теплоэнергетика и централизованное теплоснабжение России в 2015-2016 гг.," ФГБУ "РЭА" Минэнерго России, Информационно-аналитический доклад № АД-П9-107пр , 2018.

17 Доклад о состоянии сферы теплоэнергетики и теплоснабжения в Российской Федерации: [сайт]. [2018]. URL: https://minenergo.gov.ru/node/10850

18 Официальный сайт Москва24. Росстат посчитал долги россиян за отопление // Официальный сайт Москва24. 2017. URL: https://www.m24.ru/news/ZhKH/09112017/16086

19 Open Street Map: [сайт]. [2019]. URL: http://www.openstreetmap.ru

20 TechDifferences: [сайт]. [2018]. URL: https://techdifferences.com

21 Dinu V., Nadkarni P., Brandt C. Pivoting approaches for bulk extraction of Entity–Attribute–Value data //Computer methods and programs in biomedicine. – 2006. – Т. 82. – №. 1. – С. 38-43.

22 Free Software Foundation (10 June 2007), GNU Emacs Lisp Reference Manual, Boston, MA: Free Software Foundation, pp. Section 5.8.

23 Хомоненко А. Д., Цыганков В. М., Мальцев М. Г. Базы данных: Учебник для высших учебных заведений/Под ред. проф. АД Хомоненко //СПб.: КОРОНА принт. – 2009.

24 Obe R. O., Hsu L. S. PostgreSQL: Up and Running: a Practical Guide to the Advanced Open Source Database. – "O'Reilly Media, Inc.", 2017.

25 Моргунов Е.П. Язык SQL. Базовый курс Учебно-практическое пособие. Москва: Postgres Professional, 2017.

26 Group P.G.D. Документация к PostgreSQL 9.6.9. 2018.

27 Maymala J. PostgreSQL for data architects. – Packt Publishing Ltd, 2015.

28 Буров В.Д., Дорохов Е.В., and Елизаров Д.П. Тепловые электрические станции: Учебник для вузов. 2nd ed. Москва: Издательский дом МЭИ, 2007. 466 pp.

29 Минэнерго. Доклад о состоянии сферы теплоэнергетики и теплоснабжения в Российской Федерации // Официальный сайт Министерства энергетики России. 2018. URL: https://minenergo.gov.ru/node/10850

30 МакКинзи Россия, "Основы устойчивого развития генерирующего сектора российской энергетики," МакКинзи Россия, 2018.

31 Законопроект № 1086603−6 О внесении изменений в ФЗ «О теплоснабжении» и о внесении изменений в отдельные законодательные акты РФ по вопросам совершенствования системы отношений в сфере теплоснабжения (18 июля 2017г).

32 Липатов Т.В., "Выработка электроэнергии на базе теплового потребления: расчет относительного прироста топлива на прирост тепла," Тепловая энергетика и ЖКХ, Vol. 6, No. 3, Июнь 2013.

33 Пресс-центр Региональной энергетической комиссии Свердловской области. От чего зависит величина тарифа на тепловую энергию // Официальный сайт Региональной энергетической комиссии Свердловской области. 2012. URL: https://rek.midural.ru/news/show/id/332/news_category/35 (дата обращения: 19.Сентябрь.2018).