



**Автоматизированная информационная система
сбора и предварительной обработки данных
«Предбиллинг»
(АИС Предбиллинг)**

ФУНКЦИОНАЛЬНОЕ ОПИСАНИЕ СИСТЕМЫ

**Санкт-Петербург
2023**

Оглавление

1	Принятые термины и сокращения	3
2	Функциональное описание системы	4
2.1	Введение. Краткое описание системы.....	4
2.2	Возможности АИС Предбиллинг.....	4

1 Принятые термины и сокращения

Таблица 1 – Термины и сокращения

Термин	Определение
АИС Предбиллинг/ Система	Автоматизированная информационная система сбора и предварительной обработки данных «Предбиллинг»
Предприятие	Организация, в локальной сети которой развернута АИС «Предбиллинг»
ПО	Программное обеспечение
Биллинг	(англ. billing -- составление счёта) - это комплекс бизнес-процессов сбора информации об использовании услуг, тарификации, обработки платежей и выставления счетов потребителям.
Биллинговая система	Прикладное программное обеспечение для поддержки бизнес-процессов биллинга Предприятия.
АСУСЭ	Автоматизированная Система Управления Сбытом Энергии (АСУСЭ) – биллинговая система, собственная разработка ООО «АЛЬФА-ИНТЕГРАТОР ИНФОЭНЕРГО»
ГП	Гарантирующий поставщик - гарантирующий поставщик электрической энергии - коммерческая организация, которой в соответствии с законодательством Российской Федерации присвоен статус гарантирующего поставщика, которая осуществляет энергосбытовую деятельность и обязана в соответствии с Федеральным законом заключить договор энергоснабжения, договор купли-продажи (поставки) электрической энергии (мощности) с любым обратившимся к ней потребителем электрической энергии либо с лицом, действующим от своего имени или от имени потребителя электрической энергии и в интересах указанного потребителя электрической энергии и желающим приобрести электрическую энергию
АО АТС	Администратор торговой системы оптового рынка электроэнергии
ГТП	Группа точек поставки – на рынке электроэнергии – это одна или несколько точек поставки, относящихся к одному узлу расчетной модели и/или к единому технологически неделимому энергетическому объекту и ограничивающих территорию, в отношении которой купля-продажа электрической энергии (мощности) на оптовом рынке осуществляется только данным участником, и используемых для определения и исполнения участником оптового рынка обязательств, связанных с поставкой и оплатой электрической энергии (мощности). Группа точек поставки определяется Системным и Коммерческим операторами в соответствии с Положением о порядке получения статуса субъекта оптового рынка и ведения реестра субъектов оптового рынка электрической энергии и мощности. Для ГТП на сайте АО АТС ежемесячно публикуются цены и другие параметры для расчетов за электроэнергию на розничном рынке.

2 Функциональное описание системы

Система предназначена для автоматизации процессов сбора первичной информации для расчета за электроэнергию, теплоэнергию и прочие виды услуг. Система может использоваться как в связке с АСУСЭ, так и в качестве самостоятельной системы, интегрированной с Биллинговой системой Предприятия. Рассмотрим функционирование АИС Предбиллинг на примере получения данных электрической энергии.

2.1 Введение. Краткое описание системы.

Для проведения расчетов с потребителями и выставления счетов Биллинговой системе необходимо значительное количество первичной информации, включающей в себя (для расчетов за электроэнергию): интегральные и почасовые показания приборов учета, часы пиковой нагрузки в текущем отчетном периоде, почасовые и свободные тарифы, их составляющие и т.д. Эти данные обычно ежемесячно вводятся в Биллинговую систему в разных форматах и в основном в ручном режиме. При работе с этими данными приходится решать вопросы, связанные с:

- Множеством источников поступления показаний, из которых сохраняются только последние поступившие;
- Различными форматами поступления данных о показаниях от контрагентов;
- Ручной загрузкой показаний из файлов;
- Ручным запуском интеграции с системами АСКУЭ;
- Ежемесячной ручной загрузкой файлов цен и часов, публикуемых на сайте регулятора.

С целью решения этих вопросов, была разработана *Автоматизированная информационная система сбора и предварительной обработки данных «Предбиллинг»* (далее - АИС Предбиллинг).

2.2 Возможности АИС Предбиллинг

- Получение информации в виде файлов с данными по электронной почте, с сетевых папок и ftp.
- Автоматическая загрузка файлов по расписанию с web-сайтов с настройкой правил построения ежемесячной ссылки на файл.
- Хранение полной истории всех поступающих файлов и результатов их разбора.
- Автоматическое определение вида и формата файла с использованием нейросети.
- Возможность обработки файлов в произвольном формате – использование нейросети для определения расположения полей с информацией в файле.
- Полностью автоматическая работа.
- Автоматический запуск интеграции с системами-источниками данных по расписанию.
- Web-интерфейс администратора.

В качестве загружаемой информации может выступать:

- Почасовые и получасовые расходы по приборам учета.
- Профили потерь по точкам поставки.
- Профили потребления по ГТП оптового рынка.
- Интегральные показания приборов учета.

- Составляющие предельных уровней нерегулируемых цен.
- Часы пиковой нагрузки.
- Объемы планового почасового потребления.
- Цены для потребителей 1 ЦК, публикуемых на сайтах ГП.
- Объемы плановой покупки у розничной генерации.
- Другие виды данных.

Перечень обрабатываемых сущностей может расширяться без изменения кода программы путем настройки дополнительных типов сущностей и описания выделяемых параметров. Процесс поступления и обработки данных схематично можно представить в следующем виде (Рисунок 1):

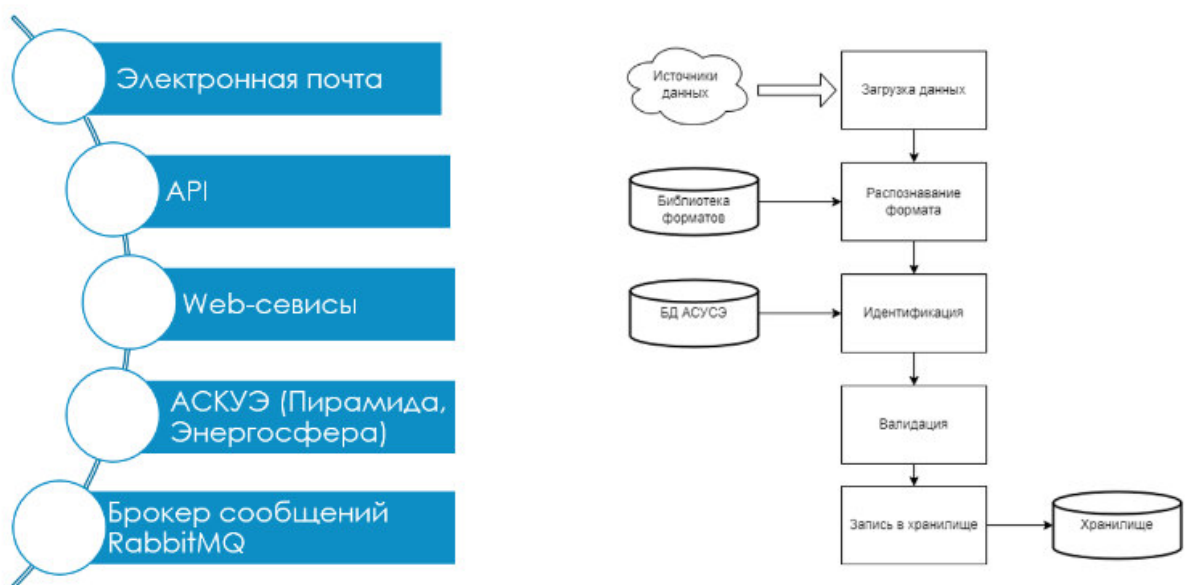


Рисунок 1. Схема поступления данных в АИС Предбиллинг

2.3 Структура системы и алгоритм работы сервисов.

Система состоит из функциональных микросервисов, работающих как отдельные процессы на сервере. Между собой микросервисы общаются путем передачи сообщений через собственный брокер сообщений RabbitMQ. При необходимости, возможна интеграция со сторонними системами с использованием сервиса очередей сообщений – входящая информация может поступать прямо в очередь сообщений Системы. Структурная схема системы представлена на Рисунок 2.

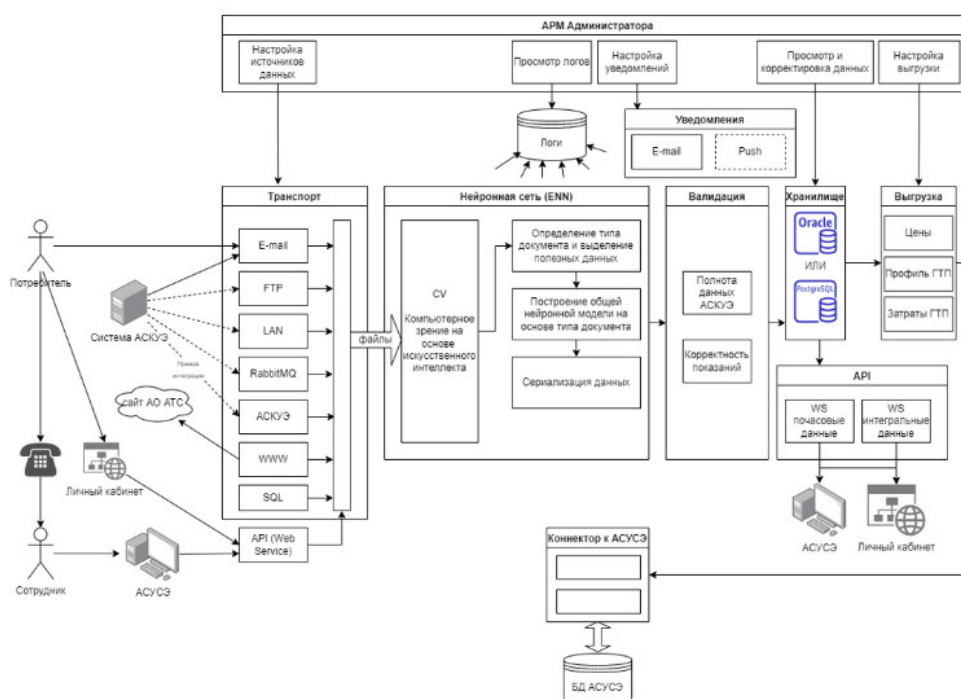


Рисунок 2. Структурная схема АИС Предбиллинг

«**Транспорт**»- обеспечивает поступление информации. Состоит из процедур опроса сетевых ресурсов, ftp, почты, CRQ-интерфейса «Энергосфера», сервера «Пирамида», сайта АТС, сайтов ГП и др. Полученная информация в виде файлов сохраняется в файловое хранилище и передается на распознавание

«**Нейронная сеть (ENN)**» - Специально разработанная нейронная сеть, задача которой определять типы документов и распознавать внутреннее содержимое документов.

«**Компьютерное зрение (CV)**» - Вспомогательный блок нейронной сети, который позволяет преобразовать любые файл в данные для обработки нейронной сетью. Наделяет сервис способностью «видеть» и извлекать информацию из увиденного.

«**Сохранение**» - сохранение данных в локальное хранилище АИС Предбиллинг, а также передача данных в Биллинговую систему с использованием коннектора.

«**Валидация**» - Набор методов для проверки правильности полученных данных.

«**API**» - интерфейс для подключения сторонних поставщиков и потребителей данных.

Наиболее сложными с технической точки зрения являются процессы определения формата, идентификации и валидации для XLS-файлов. Разбор файлов происходит в 3 этапа: первичный анализ документа компьютерным зрением, построение модели при помощи нейронной сети и вариация собранной модели. К модели Excel могут быть предварительно приведены файлы форматов PDF (кроме сканированных) и DOC.

Первичный анализ документа компьютерным зрением: Самый сложный алгоритмически шаг, так как для сервиса документ представляет собой просто набор букв и цифр в матрице. Нужно проанализировать документ и «увидеть» таблицы и другие данные.

Сервис умеет считывать из ячеек видимые границы и строить из них модели таблиц, разделяя на заголовки и табличные данные. Так же это позволяет разделить данные в документе на табличные и остальные ячейки. Предусмотрен набор алгоритмов, которые помогают достраивать таблицы, если они были плохо построены изначально или построить таблицы, не имеющих видимых границ.

Построение модели при помощи нейронной сети: После нахождения таблиц и отделения от других ячеек происходит нейроструктурное определение документа. Нейронная сеть считывает все заголовки таблиц ячейки вокруг таблиц и заметки в документе. Далее она формирует результат в виде типа документа: почасовые данные, интегральные показание или другие, основываясь на модели, полученной в результате обучения на примерах. Дополнительно накладываются некоторые структурные особенности документов, которые корректируют результат нейронной сети.

Валидация: после построения модели и получения типа документа происходит проверка модели на правильность определения типа и полноту данных в ней. После успешного прохождения валидации документ проходит на этап сохранения, а при ошибке валидации формируется отчет и направляется в виде ответного сообщения источнику. Если документ пришел по электронной почте, ответ будет выслан ответным сообщением.

Обработка PDF: PDF документ проходит конвертирование в Excel формат и подаётся на обработку как Excel.

Обработка TXT, CSV, HTML: Так как эти форматы являются примитивными, то они обрабатываются линейной логикой с небольшой поддержкой нейронной сети для определения типов документов.

Обработка XML: В процессе загрузки происходит проверка формата файла на соответствие xsd-схемам. Файлы, которые не соответствуют ни одной из схем из библиотеки, не загружаются, отправителю возвращается ответ о неуспешной загрузке.

Система нормализации документов: Примерно в 80% отправляемых документах присутствует ошибки в заполнении данных и/или неправильное оформление документов. Например, «сломанное» оформление таблицы, много пустых ячеек с табличным оформлением и т.д. и т.п. Это усложняет анализ документа и является основным местом появления ошибок при анализе документа. Перед анализом компьютерным зрением системы пытается нормализовать документ исправляя ошибки, допущенные человеком при заполнении.