



«Система персонифицированных предложений»

Содержание

Назначение документа	3
Общие сведения.....	3
Перечень и назначение компонентов	9
Общие сведения о проекте разработки подсистемы «CRM ИМ»	12
Задача.....	12
Сроки	12
Требования к технологическому стеку	13
Процесс разработки.....	13
Входные данные для разработки	14
Показатели назначения	14
Общее описание процесса функционирования.....	15
Перечень функциональных блоков модуля «CRM»	16
Перечень функций.....	17
Клиентский модуль идентификации пользователя	17
Загрузчик фактов Портала ИМ	18
Push-интерфейс загрузки фактов	20
Хранилище RAW.....	20
Модуль предпроцессинга RAW фактов	20
Модуль идентификации пользователей	21
Модуль кластеризации пользователей.....	21
Модуль управления поведенческими профилями	22
Модуль управления SpO (специальными предложениями)	22
Модуль публикации контента.....	23
Content API.....	24
Модуль управления пользователями и доступом	24
Требования к квалификации персонала.....	25
Требования к документированию.....	25

Назначение документа

Данный документ представляет собой свод требований к информационно-аналитической подсистеме «CRM», которая создается в рамках проекта «Инновационная мобильность», реализуемого в интересах группы РЖД.

Требования приведены в том виде, как это понимает Заказчик на момент издания документа и не представляют собой официальную позицию структур ОАО «РЖД».

Требования могут изменяться и дополняться заинтересованными лицами по ходу реализации проекта.

Документ предназначен для предоставления техническим специалистам организаций-претендентов на роль непосредственного разработчика программного продукта общих сведений о проекте, требованиях, способах реализации и ограничениях. Сведения раскрываются в объеме, достаточном для оценки достаточности своих ресурсов и квалификации, а также формирования коммерческого предложения.

Изложенная в данном документе информация является коммерческой тайной ООО «ОНЭЛИЯ» и не подлежит передаче третьим лицам или неограниченному кругу лиц полностью или частично без письменного согласия ООО «ОНЭЛИЯ».

Общие сведения

В данном разделе приведены общие сведения обо всей экосистеме проекта «Инновационная мобильность», включая информационную систему «CRM», но не ограничиваясь ею.

АСУ «Инновационная мобильность» (АСУ ИМ) (состоящая из ряда подсистем, включая CRM) создается в целях построения современной инфраструктуры дистрибуции билетов и комплексного мультиканального обслуживания пассажиров.

Она приходит на смену разрозненным и морально устаревшим системам, являющихся источниками недовольства пассажиров и партнеров.

Ожидается, что инновационные подходы, применяемые при создании АСУ ИМ позволят увеличить объем билетов и сопутствующих услуг, повысят

лояльность и удовлетворенность пассажиров, позволит выстроить мощную агентскую сеть дистрибьюции, а также лучше интегрировать ОАО «РЖД» / АО «ФПК» в международную систему кооперации относительно пассажирский перевозок.

АСУ ИМ содержит компоненты, которые используются несколькими группами пользователей:

- Пассажиры (прямые покупатели, b2c-сегмент);
- Агенты (дистрибутивная сеть) и корпоративные клиенты (b2b-сегмент);
- Сотрудники проекта «Инновационная мобильность».

Пассажиры используют АСУ ИМ посредством мобильного приложения в своих смартфонах и посредством b2c портала rzd.ru в сети Интернет. Целевой объем доступных пассажирам функций идентичен. Пассажир получает доступ к ресурсу и имеет возможность просмотреть доступный билетный контент и специальные предложения. Он также может осуществить поиск билетов, либо на конкретный вид транспорта, либо, воспользовавшись конструктором мультимодальных путешествий, подобрать себе пакетное предложение.

Для предоставления подбора билетов и услуг в АСУ ИМ используется подсистема «Агрегированные расписания», которая пользуется собственной базой данных (формируется за счет кэширования ранее выполненных запросов к инверторным системам), с запросами к инвенторным системам перевозчиков по необходимости, чтобы исключить излишнюю нагрузку на них.

Покупатель может оформить покупку через b2c-портал. Для этого он аутентифицируется на портале используя удобный для себя способ, в том числе механизмы авторизации существующих социальных сетей.

В случае, если пассажир ранее совершал покупки и пожелал сохранить свои персональные данные для будущих заказов, он может выбрать данные пассажиров из числа сохраненных, без повторного ввода персональных данных.

По мере оформления заказа АСУ ИМ предлагает сопутствующие услуги и специальные предложения, которые базируются на автоматизированном анализе предпочтений покупателя по данным АСУ ИМ и данных партнерских систем.

Встроенный шлюз позволяет оплатить покупку удобным образом, с незамедлительным подтверждением осуществления платежа.

Приобретенные билеты и услуги доставляются по выбранному каналу, в том числе в личный кабинет, в мобильный бумажник, по электронной почте, через инфокиоски и т.п.

Пассажир в любой момент может получить доступ к своим заказам, осуществить возвраты или приобрести доступные дополнительные услуги, если это предусмотрено его тарифом.

По мере приближения поездки и во время поездки АСУ ИМ формирует уведомления и напоминания, чтобы сделать поездку удобной и комфортной.

Все указанные этапы осуществляются в полностью автоматизированном режиме. Однако, пассажир имеет возможность в любой момент обратиться в контактный центр, чтобы получить поддержку от «живого» оператора проекта «Инновационная мобильность». Работа контакт-центра поддерживается по телефону, через популярные сервисы обмена сообщениями, через социальные сети и формы на сайте. Также предоставляется возможность получения помощи через имитацию общения с «живым» оператором с задействованием чат-ботов, программ, которые имитируют общение с использованием естественного языка.

В то же время, организации-Агенты используют АСУ ИМ посредством подсистемы b2b-шлюза/портала. С помощью него они осуществляют:

- Подбор (поиск) билетов и услуг в интересах конечных клиентов;
- Оформление билетов и услуг в интересах конечных клиентов;
- Осуществление взаиморасчетов по результатам клиринга.

Для подбора (поиска) билетов и услуг в интересах конечных клиентов так же, как и в случае поиска билета пассажиром, для поиска подходящих вариантов используется система «Агрегированных расписаний».

Организации-агенты задействуют b2b-шлюз/портал в трех вариантах:

- Интерактивно. Через веб-приложение «Личный кабинет агента» сотрудники организаций-агентов могут выполнять весь комплекс операций с билетами, без необходимости приобретения собственного программного обеспечения;

- Через программный интерфейс (XML). Путем интеграции своих программных продуктов в b2b шлюз АСУ ИМ, используя открытый SDK (комплект разработчика);

- Через программный интерфейс (NDC). Путем интеграции своих программных продуктов в b2b шлюз АСУ ИМ, используя инновационный протокол NDC.

В случае интерактивного взаимодействия, сотрудник организации-агента аутентифицируется в своем «личном кабинете». Личный кабинет всегда доступен через зашифрованный канал связи посредством интернет-браузера сети Интернет.

В своем личном кабинете сотрудник компании-агента имеет возможность осуществлять подбор билетов и услуг в интересах конечных клиентов и проводить их оформление в рамках своих лимитов.

Оплата приобретаемых билетов может осуществляться различными способами, такими, например, как списанием с лицевого счета организации-агента или перенаправлением конечного клиента на платежную страницу платежного шлюза, где он может совершить оплату банковской картой.

Приобретенные билеты доставляются конечным пользователям по каналам, указанным агентом.

В личном кабинете агента используются экранные формы и отчетные документы, которые стилизованы в фирменном стиле организации-агента.

Вне зависимости от места приобретения билетов пассажир имеет возможность видеть свои заказы на b2c портале и в мобильном приложении ОАО «РЖД» /АО «ФПК».

В случае использования организацией-агентом интеграции через XML/NDC, обслуживание конечного покупателя осуществляется с помощью программного обеспечения организации-агента, которое взаимодействует с b2b-шлюзом АСУ ИМ посредством сообщений в открытом XML/NDC формате. В этом случае b2b шлюз обеспечивает выполнение информационных запросов, запросов на оформление билетов и дополнительных услуг, процессинг платежных операций и клиринговых сообщений.

Аналогичным образом функционирует программный комплекс, который поддерживает продажи корпоративным клиентам, для нужд их сотрудников.

Отличие состоит в способе расчета по проданным билетам, который в этом случае осуществляется напрямую с перевозчиком, без взимания комиссий и участия в расчетах «ИМ».

Сотрудники проекта «Инновационная мобильность» используют АСУ ИМ для следующих групп операций:

- Проведение сверок и расчетов с перевозчиками и поставщиками дополнительных услуг;
- Обеспечение непрерывности бизнеса организаций-агентов, проведение сверок и расчетов;
- Содействие пассажирам в комфортном получении услуг, разрешение сложных ситуаций.

В целом, деятельность сотрудников «ИМ» связана с регистрацией новых участников, поддержанием функционирования инфраструктуры, мониторингом безопасности операций, осуществлением клиринга и взаиморасчетов.

Архитектура решения

Укрупненная компонентная диаграмма проекта "Инновационная мобильность"

На рисунке приведен состав и взаимосвязи основных подсистем организационных участников процесса обслуживания пассажиров.

- системы внешних поставщиков
- готовые системы ИИМ
- перспективные системы ИИМ
- разрабатываемые системы

Актуально по состоянию на 19.01.2018

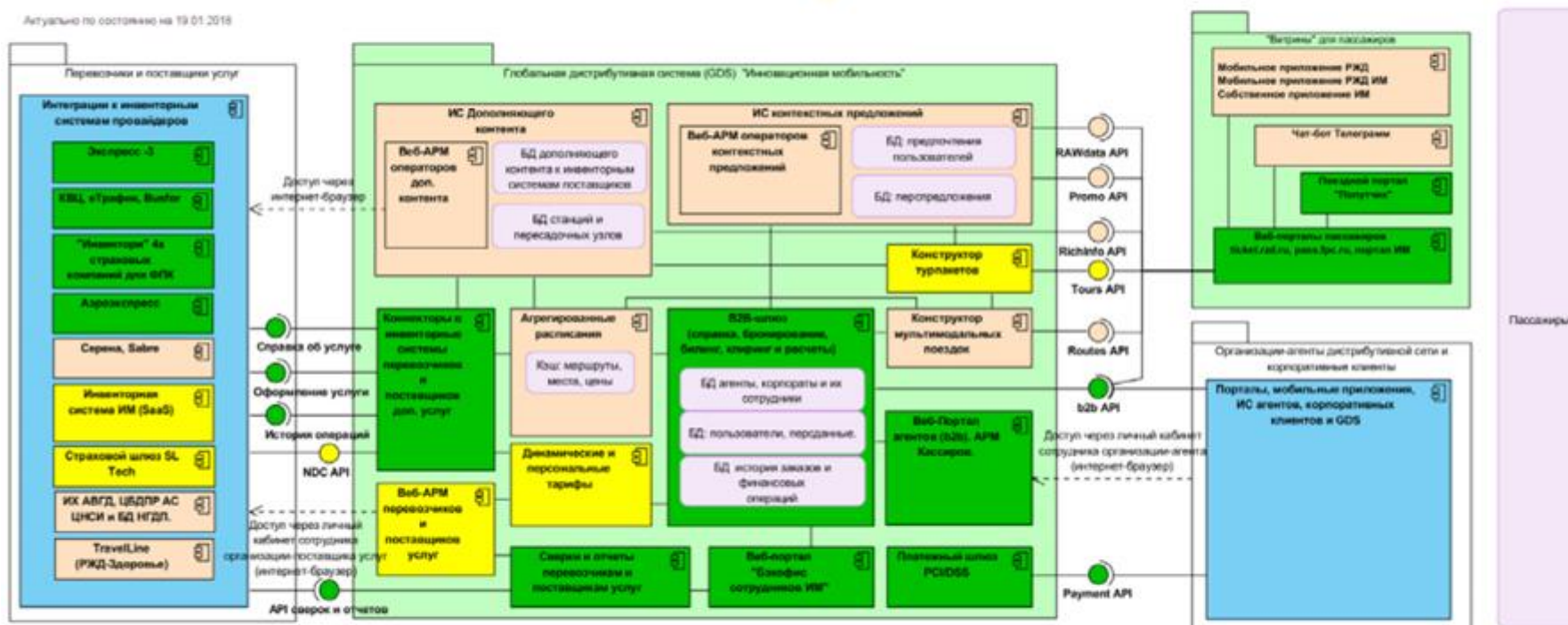


Рисунок 1. Укрупненная компонентная диаграмма

Перечень и назначение компонентов

На рисунке 1 приведены следующие компоненты:

Наименование	Назначение
Инвенторные системы перевозчиков	Являются источником билетного контента, дистрибьюция которого осуществляется посредством проекта «Инновационная мобильность».
Коннекторы в инвенторные системы перевозчиков и поставщиков доп. услуг	Интеграционная подсистема, которая позволяет взаимодействовать с унаследованными системами перевозчиков в принятых у них форматах обмена для получения информации об услугах, билетах и расписаниях, а также оформления покупки.
Агрегированные расписания (ИС АР)	Подсистема, обеспечивающая интеллектуальное «кэширование» запросов и ответов справки по расписаниям, по количеству и стоимости мест от организаций-агентов дистрибутивной сети, конструктора мультимодальных поездок.
Конструктор мультимодальных поездок	Подсистема, которая по запросам пользователей формирует составные маршруты (цепочки из сегментов различных видов транспорта с дополнительными услугами) на базе критериев запроса, имеющейся справочной информации от поставщиков и пользовательских предпочтений.
ИС Дополняющего контента (ИС ДК)	Подсистема, которая позволяет осуществлять ведение справочников станций, населенных пунктов и других общесистемных (в рамках всего проекта) справочников. Позволяет осуществлять сбор, упорядочивание и поддержку в актуальном состоянии описательной информации по продуктам перевозчиков и поставщиков сопутствующих услуг. Предоставляет потребителям актуальную описательную информацию по продуктам перевозчиков и поставщиков сопутствующих услуг.
ИС Контекстных предложений (CRM)	Подсистема, осуществляющая идентификацию зарегистрированных и анонимных пользователей,

Наименование	Назначение
	<p>пользователей веб-порталов и мобильных приложений, их кластеризацию по поведенческим и маркетинговым параметрам, профилирование кластеров в поведенческие профили, генерацию специальных предложений в привязке к профилям пользователей, а также предрасчёт и публикацию окончательного контента СПО для интеграции в веб-порталы и мобильные приложения с последующим сбором сведений о пользовательском интересе к ним (конверсии). Предоставляет пользователям персонализированные рекомендательные предложения на основе выявления их предпочтений, в т.ч. персональные тарифы на перевозки.</p>
<p>Сверки и отчеты перевозчикам и поставщикам услуг интерфейс и личный кабинет веб-портала перевозчиков.</p>	<p>Автоматическая система, которая по данным о фактических операциях из подсистемы бэкофиса регулярно формирует реестры и отчеты по взаиморасчетам для поставщиков услуг в требуемых форматах. Результаты доступны через программный интерфейс и личный кабинет веб-портала перевозчиков.</p>
<p>B2b-шлюз</p>	<p>Подсистема, которая обеспечивает непосредственный процессинг информационных запросов и заказов на приобретение билетов и дополнительных услуг со стороны порталов, мобильных приложений и прочих информационных систем агентов и собственных «витрин» проекта «Инновационная мобильность», таких как «Мобильное приложение РЖД» и веб-портал самообслуживания пассажиров pass.rzd.ru. Предоставляет API для простого подключения систем агентской сети дистрибуции.</p>
<p>Веб-портал агентов (b2b). АРМ Кассиров</p>	<p>Веб-портал самообслуживания агентов, который доступен через браузер в сети интернет. Позволяет уполномоченным сотрудникам компаний-агентов и корпоративных клиентов через свой «личный кабинет»</p>

Наименование	Назначение
	осуществлять полный спектр операций по обслуживанию своих клиентов в части организации поездок и приобретения билетов.
Клиринг, взаиморасчеты, бэкофис	Подсистема, которая на основе совершенных операций ведет непрерывный расчет финансовых позиций агентов и перевозчиков в системе, контролирует лимиты, нивелирует финансовые риски сторон, организует оперативные взаиморасчеты.
Платежный шлюз	Подсистема в составе проекта «Инновационная Мобильность», которая обеспечивает «онлайн» расчеты по банковским картам в соответствии с правилами национальных и международных платежных систем.
Порталы, мобильные приложения, информационные системы агентов и глобальных дистрибутивных систем	Внешние по отношению к проекту «ИМ» системы организаций-агентов, которые через программный интерфейс получают билетный и информационный контент с целью его дальнейшей дистрибуции конечным пользователям через свои собственные «торговые витрины».
Веб-порталы пассажиров ticket.rzd.ru, pass.fpc.ru, портал ИМ	Веб-порталы самообслуживания пассажиров. Предоставляет функционал полного цикла обслуживания: информирование, выбор, продажа, оказание услуги, поддержка, обратная связь, защита персональных данных, учет предпочтений. Являются «витриной» к основному набору компонентов проекта «ИМ».
Мобильные приложения РЖД, РЖД ИМ, ИМ	Приложение, предоставляющее функционал веб-портала пассажиров pass.rzd.ru с использованием специфических возможностей смартфонов и планшетов.
Чат-бот Телеграмм	Позволяет выполнять поиск и подбор билетов с использованием популярного мессенджера.
Поездной портал «Попутчик»	Поездной портал «Попутчик» предоставляет информационный и развлекательный контент для пассажиров (новости, расписание, меню вагона-ресторана, фильмы, музыку, книги, информацию о городах на пути следования поезда и аудиогид по достопримечательностям маршрута, а также предоставляет услуги покупки билета, заказа такси, и чат для общения пассажиров).

Наименование	Назначение
Вспомогательные системы сторонних поставщиков	Унаследованные информационные системы, контент и функционал которых используется на портале pass.rzd.ru и в мобильном приложении РЖД для обеспечения полноты, предоставляемой пассажирам информации и услуг.
Пассажир	Ключевое заинтересованное лицо проекта, на нужды которого сфокусирована реализация требований в первую очередь.

Общие сведения о проекте разработки подсистемы «CRM ИМ»

Задача

Требуется:

- Детализировать требования и согласовать результат с ООО «ОНЭЛИЯ»;
- В соответствии с детализированными требованиями разработать программный продукт «CRM ИМ»;
- Документировать (автодоки html);
- Развернуть программный продукт на мощностях компании ООО «ОНЭЛИЯ»;
- Провести испытания и устранить замечания по их результату;
- Запустить продукт в опытную эксплуатацию и разработать комплект рабочей документации;
- Обеспечить реакцию на инциденты в ходе опытной эксплуатации;
- Провести доработки по результатам опытной эксплуатации;
- Сдать систему в промышленную эксплуатацию.

Сроки

5 мес.

Требования к технологическому стеку

Технологические компоненты системы должны:

- быть лицензированы как OpenSource: MIT, Apache, GPL, Freeware;
- функционировать под управлением ОС Linux Ubuntu или Centos Linux;
- иметь прецеденты успешного применения для решения аналогичных задач в известных проектах со сходными показателями назначения.

Предпочтительный технологический стек:

- ОС: Ubuntu Linux;
- Языки реализации: Ruby, Python, Javascript;
- Client web: React или Angular2;
- Административный бэкенд: Python, Ruby + соответствующие фреймворки;
- BigData: Spark, Cassandra, Kafka;
- Веб-фронтенд: Nginx;
- Веб-серверы: nodejs, puma;
- Devops: Jenkins / TeamCity, Ansible, Chef, Docker.

Предлагаемые к использованию технологии должны быть письменно согласованы с заказчиком.

Процесс разработки

- Continuous integration pipeline;
- Итеративная разработка, длина итерации 1-2 недели по Agile-методологии;
- Демонстрационный стенд с публикацией релизов и соответствующей документации по завершению итераций;
- Менеджер продукта со стороны ООО «ОНЭЛИЯ», включен в процесс планирования каждой итерации;
- Предпочтительные системы управления процессом разработки - Youtrack, Jira; - с гостевым доступом для кураторов проекта со стороны компании ООО «ОНЭЛИЯ»;

- Ежедневный мониторинг прогресса (потраченные/оставшиеся часы) по задачам каждого разработчика;
- Все участники команды должны присутствовать на демонстрациях по окончанию каждой итерации (удаленно, посредством средств конференцсвязи);
- Гибкое управление требованиями - в соответствии с актуальными приоритетами;
- Каждый участник команды должен быть доступен для коммуникации.

Входные данные для разработки

- Перечень функциональных требований;
- Описание API систем, с которыми взаимодействует модуль «CRM»;
- Требования к стеку используемых технологий;
- Требования к срокам;
- Требования к документированию;
- Требования к организации процесса разработки.

Показатели назначения

- Система должна обеспечивать сбор и хранение исторических данных о поведенческих фактах (10 миллионов фактов в сутки), получаемых с расширяемого количества площадок (не менее 2);
- Система должна обеспечивать постоянное автоматическое обновление кластеров пользователей на основании вновь поступающей информации. При этом задержка обработки новых данных не должна составлять более 5 минут;
- Система должна обеспечивать обработку 10 миллионов запросов в сутки на возврат персонифицированного контента по запросу с идентификатором пользователя;
- Не допускается сохранение и обработка ПДн.

Общее описание процесса функционирования

CRM ИМ

Эскизный процесс использования CRM ИМ представлен рисунке 2, ниже:

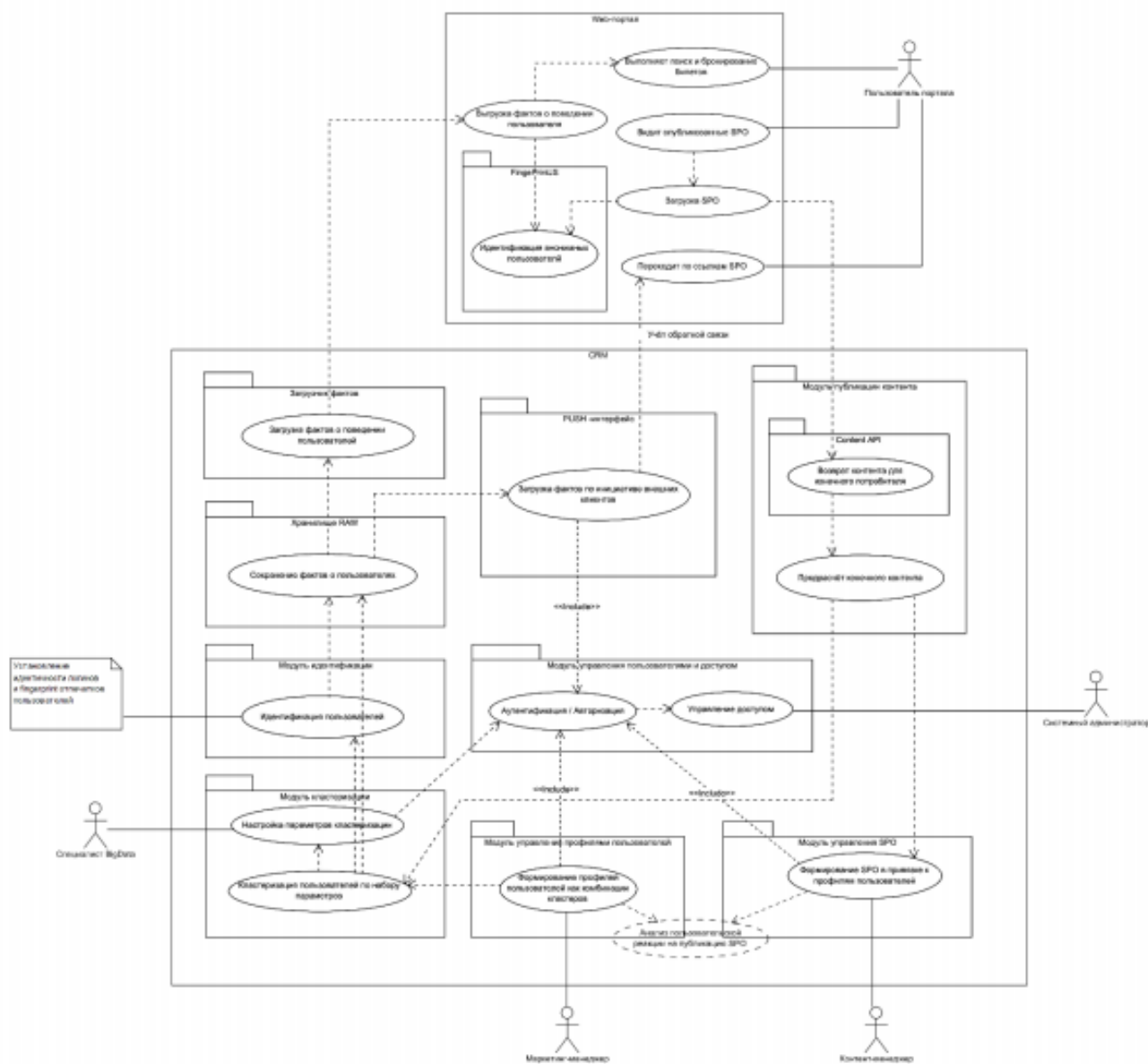


Рисунок 2 – Эскизный процесс использования CRM ИМ

Перечень функциональных блоков модуля «CRM»

Перечень функциональных блоков и их краткое описание приведены в таблице:

Наименование	Назначение
Клиентский модуль идентификации пользователя	Предназначен для идентификации пользователей на веб-порталах АСУ ИМ.
Загрузчик фактов Портала ИМ	Осуществляет загрузку фактов о поведении пользователей на веб-порталах с обезличенной идентификацией по Fingerprintjs
Собственный механизм сбора фактов	Осуществляет сбор и выявление фактов о поведении пользователей, ценных с точки зрения CRM.
PUSH-интерфейс загрузки фактов	API-Интерфейс для загрузки фактов в систему по инициативе внешних аутентифицированных пользователей (мобильные приложения, B2B-шлюз).
Хранилище RAW	Пополняемое хранилище исторических «сырых» фактов о пользователях.
Модуль идентификации пользователей	Программно-алгоритмический модуль, обеспечивающий идентификацию анонимных и зарегистрированных пользователей веб-порталов ИМ на основании построения поведенческого профиля.
Модуль кластеризации пользователей	Инженерно-программный модуль, обеспечивающий кластеризацию идентифицированных и слабо идентифицированных пользователей на основе массива «сырых» данных и конфигурируемых параметров кластеризации.
Модуль управления профилями пользователей	Программный модуль, позволяющий оператору CRM-системы формировать и управлять профилями пользователей, как приоритезированными комбинациями кластеров.
Модуль управления SpO	Программный модуль, позволяющий оператору CRM-системы формировать и управлять SpO (специальными предложениями (от англ. – Special Offer)) и их контентом в привязке к профилям пользователей, а также анализировать статистику

Наименование	Назначение
	реакции пользователей на опубликованные SpO.
Модуль публикации контента	Инженерно-программный модуль, обеспечивающий предварительную генерацию конечного контента для показа SpO на целевых площадках.
Content API	API осуществляющий возврат конечного контента SpO по параметрам, идентифицирующим анонимного или зарегистрированного пользователя, подсчёт и статистику количества запросов для различных профайлов и площадок, а также сбор обратной связи - фактов о кликах пользователей по контенту SpO.
Модуль управления пользователями и доступом	Программный модуль, позволяющий: 1) администратору CRM ИМ управлять доступом операторов к различным функциональным блокам системы; 2) Управлять учетными записями для доступа к PUSH-интерфейсу.
Модуль анонимной идентификации	Программная библиотека (предположительно FingerprintJS), осуществляющая идентификацию пользователя по техническим параметрам его компьютера.

Перечень функций

Ниже приведен сгруппированный по функциональным блокам перечень функций, которые необходимо реализовать в рамках текущего этапа проекта (границы проекта).

Клиентский модуль идентификации пользователя

Идентификация пользователей решения «Инновационная мобильность» осуществляется несколькими способами:

- Логин в ЛК / мобильном приложении;
- Снимок портрета анонимного пользователя на основании характеристик компьютера – с использованием библиотеки fingerprintjs (fingerprintjs2).

Клиентский модуль идентификации предназначен идентификации пользователей на веб-порталах АСУ ИМ при сборе и обработке данных о поведении пользователя на сайте.

Загрузчик фактов Портала ИМ

Модуль предназначен для загрузки фактов о поведении пользователей на вебпорталах АСУ ИМ.

Система CRM ИМ предусматривает сбор данных о поведении пользователей осуществлялся с учетом их идентификаторов – логина, идентификатора сессии и fingerprint хэша, также доступных идентификаторов иных аналитических сервисов.

Фиксация фактов о поведении пользователей должна осуществляться асинхронно, с минимальной временной задержкой.

Примеры фактов (расширяемый перечень):

1. Поведенческие:

- Вход/выход на страницу;
- Наведение на html-элемент по его id (управляемый набор отслеживаемых ID);
- Клик на html-элемент по его id (управляемый набор отслеживаемых ID);
- Отображение в видимой области html-элемента по его id (управляемый набор отслеживаемых ID);
- Увеличение/уменьшение окна браузера на странице;
- Закрытие окна/вкладки браузера;
- Постановка закладки на страницу;
- Использование вертикальной/горизонтальной прокрутки;
- Поиск варианта по направлению А->В;
- Выбор класса;
- Выбор способа сортировки;
- Выбор критерия фильтрации;
- Размещение заказа;

- Оплата заказа;
- Отмена заказа до оплаты;
- Отмена оплаты заказа;
- Возврат заказа после оплаты;
- Переоформление заказа (повышение класса, изменение дат);
- Отправка информации знакомым;
- Выбор способа получения заказа (мейл, телефон, wallet, печать);
- Переключение формата вывода результатов подбора вариантов поездки/размещения;
- Смена языка локализации;
- Увеличение/уменьшение шрифта средствами браузера;
- Пол/возраст.

2. Общестатистические:

- страна
- браузер
- операционная система
- разрешение экрана
- размеры окна браузера
- наличие возможности сенсорного ввода
- поддержка Java/javascript
- глубина цвета
- браузерный агент
- примерная локация

Интеграционное решение должно предусматривать pub-sub архитектуру с использованием решения, обеспечивающего отказоустойчивость и гарантию сохранности данных при разрывах связи или отключении системы.

Push-интерфейс загрузки фактов

API-интерфейс, предназначенный для загрузки в систему фактов по инициативе внешних компонентов, таких как (но не ограничиваясь):

- мобильное приложение ИМ;
- B2C-портала
- B2B-шлюз
- Чат-бот Telegram
- Веб-ресурсы организаций-агентов.

Интерфейс должен обеспечивать аутентифицированный доступ по протоколу REST JSON API и загрузку фактов о поведении пользователей, идентифицированных по fingerprint, логину в ЛК ИМ, идентификатору сессии на веб-портале или «внешнему идентификатору», доступному для чтения с пользовательского устройства.

Хранилище RAW

- Обеспечивает хранение, резервное копирование и восстановление исторических данных о всех фактах поведения пользователей на всех отслеживаемых площадках.

- Хранилище должно обеспечивать возможность произвольного доступа ко всем записям в соответствии с требованиями модулей идентификации и кластеризации пользователей.

Модуль предпроцессинга RAW фактов

Предназначен для выполнения предварительной обработки RAW-фактов для их подготовки к кластеризации.

Например, по информации о входе и выходе со станции по определенному fingerprint рассчитывает факт о нахождении на странице больше установленного порога времени или «кратко», «средне», «долго».

Другой пример, расчет возрастной категории по абсолютной величине возраста.

Модуль идентификации пользователей

Модуль идентификации пользователей реализует в системе CRM функцию уточнения идентичности анонимных и неанонимных пользователей. В эту функцию входят:

- Автоматическое связывание fingerprint-идентификатора анонимного пользователя с логином, идентификатором сессии b2c портала и «внешним» идентификатором сторонних сервисов;
- Автоматическое отвязывание fingerprint-идентификатора пользователя от логина, в случае, если поведенческие характеристики носителя fingerprint отпечатка позволяют сделать вывод о том, что данный отпечаток принадлежит большому количеству пользователей.

В рамках работ требуется разработать и согласовать перечень поведенческих критериев, на основании которых будет реализована данная функция.

Модуль должен функционировать постоянно в фоновом режиме и обновлять существующие кластеры в соответствии с вновь поступающими данными в RAW хранилище.

Модуль идентификации пользователей может быть архитектурно совмещен с модулем кластеризации.

Модуль кластеризации пользователей

Выполняет следующие функции:

- Автоматическая группировка идентифицированных пользователей в кластеры на основе схожих поведенческих шаблонов и предметных фактов о пользователе;
- Обновление существующих кластеров в соответствии с вновь поступающими данными в RAW-хранилище в фоновом режиме.

Задержка между вновь поступившими данными и обновлением кластеров с использованием этой информации не должна составлять более 5 минут.

В рамках работ требуется разработать и предложить перечень способов кластеризации на основе общестатистических и поведенческих фактов о пользователях.

При обработке каждого факта должна учитываться его:

- Достоверность;
- Ценность (вес, значимость).

Модуль управления поведенческими профилями

Предназначен для ручного введения перечня «поведенческих профилей», их именованья и сопоставления с автоматически выделенными кластерами пользователей, значимыми с точки зрения аналитика.

В рамках работы требуется разработать и предложить:

- Виды отношений между кластерами пользователей и поведенческими профилями;
- АРМ редактора поведенческих профилей, обеспечивающий возможность создавать, удалять и настраивать (уточнять параметры) поведенческие профили и сопоставлять их автоматически идентифицированным кластерам.

Пользовательские интерфейсы должны обеспечивать возможность анализа количества идентифицированных пользователей в рамках профиля (включая исторические данные - динамику).

Изменение параметров профиля пользователей должно приводить к запуску процедуры пересчета конечного контента для целевых идентифицированных пользователей в модуле публикации контента.

Модуль управления SpO (специальными предложениями)

- Формирование контента специальных предложений в привязке к сформированным профилям пользователей;
 - Привязка SpO к различным поведенческим профилям должна выполняться с заданием весового коэффициента, определяющего приоритезацию показа SpO для заданного профиля;
 - Контент SpO должен:
 - иметь возможность интеграции на сайт портала ИМ;
 - предусматривать возможность задания параметрического контента для каждого идентифицированного (включая анонимных) пользователя;

- последнее купленное направление;
- часто просматриваемое направление;
- прочие.

■ поддерживать возможность включения «статического» SVG, Flash, Javascript из CDN.

- Публикация и снятие с публикации SpO;
- Анализ конверсии опубликованных SpO:

- возможность оценить интерес (конверсию) целевого пользовательского профиля в отношении данного SpO, включая исторические данные (динамику);
- возможность оценить интерес (конверсию) кластеров пользователей, включая исторические данные (динамику).

Модуль управления SpO должен предоставлять возможность имперсонированного предварительного просмотра контента.

Модуль управления SpO должен предусматривать наличие различных статусов SpO:

- черновик;
- на модерации;
- опубликовано;
- снято с публикации.

Совокупность опубликованных SpO, привязанных к одному профилю пользователей формирует очередь ротации SpO в рамках одного профиля пользователей. Публикация SpO (изменение статуса) приводит к запуску предрасчета контента для соответствующих профилей.

В рамках работы должны быть разработаны и согласованы пользовательские интерфейсы модуля управления SpO.

Модуль публикации контента

Предназначен для предварительной генерации конечного контента и его индексирования таким образом, чтобы обрабатывать большое количество запросов

(10 млн показов в сутки) с учетом весовых коэффициентов SpO и очередей ротации контента для каждого пользователя.

Перегенерация и индексирование контента должно запускаться автоматически в случаях:

- создания, удаления и модификации профилей пользователей (составляющих их кластеров и их взаимоотношений);
- публикации и снятия с публикации SpO.

Перегенерация контента не должна отражаться на работоспособности существующего контента до завершения регенерации.

Время генерации контента не должно быть более, чем линейно зависимо от числа поведенческих профилей, для которых предназначается конечный контент.

Модуль публикации контента должен включать все необходимые метки в генерируемых в SpO ссылках для сбора данных о конверсии SpO посредством модуля «Загрузчик фактов Портала ИМ».

Content API

- Реализует функцию возврата контента SpO по параметрам идентификации пользователей (логин/fingerprint), сессии веб-портала или «внешнего» идентификатора, доступного для чтения с клиентского устройства;
- Реализует функцию ротации контента для каждого пользователя в рамках одной сессии, с учетом весовых коэффициентов SpO для поведенческих профилей, к которым отнесен конкретный посетитель на основании попадания в один и более автоматически идентифицированных кластеров.

В рамках работы должен быть разработан и согласован алгоритм ротации SpO.

Модуль управления пользователями и доступом

- Создание, правка, блокировка и удаление внутренних пользователей системы (маркетинг-менеджеров и контент-менеджеров) администратором, а также функции сброса пароля для зарегистрированных пользователей;
- Создание, правка, блокировка и удаление внешних пользователей системы (пользователи PUSH-интерфейса) и их ключей аутентификации;

Требования к квалификации персонала

- Опыт реализации проектов автоматического анализа больших данных;
- Опыт реализации нагруженных веб-систем.

Требования к документированию

Каждый релиз-кандидат системы должен сопровождаться документацией в электронном виде, включающую в себя:

- инструкции по установке и настройке программных компонентов;
- инструкции по обновлению;
- описания сценариев развертывания и обслуживания системы;
- документированный программный код: назначение и примеры использования программных компонентов: модулей, классов и их методов;
- спецификации API;
- описания и UML-диаграммы прикладной логики программных компонентов и специальных алгоритмов;
- обоснование и описание математической модели автоматической кластеризации фактов;
- диаграммы (wireframe) интерфейсов.

Для документирования всех составляющих должны использоваться соответствующие применяемым технологиям системы генерирования документации.

Документация актуальной версии должна публиковаться в демонстрационной среде наравне с релиз-версиями системы.