КАЗАХСТАНСКИЙ ИНЖЕНЕРНО-ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

ТЕМИРБЕКОВ Н. М., ТЕМИРБЕКОВ А.Н., БАЙГЕРЕЕВ Д. Р., ОМИРЖАНОВА Б.Б., ЖАКСЫЛЫКОВА Ж.Р.

Распределенная информационная система научно-исследовательского института Академгородка города Алматы

Алматы 2020

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 3
1 ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОЙ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ФУНКЦИОНАЛЬНОСТИ И НОМЕНКЛАТУРЫ ВСЕХ ЕЕ КОМПОНЕНТ 7
1.1 Анализ возможностей участников научно-образовательного кластера 7
1.2 Построение радиосети Академгородка на базе телекоммуникационной технологии WiMAX 9
1.3 Программная часть распределенной информационной системы 16
1.4 Модель функционирования распределенной информационной системы 17
1.5 Заключение по разделу 18
2 ОТБОР И АДАПТАЦИЯ К ВЫБРАННОЙ АРХИТЕКТУРЕ НЕОБХОДИМОГО СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ 19
2.1 Выбор, установка и адаптация репозитория цифровых объектов 19
2.2 Выбор, установка и адаптация системы управления научными исследованиями 22
2.3 Выбор интегрирующего программного обеспечения 23
2.4 Заключение по разделу 24
3 СОЗДАНИЕ УНИКАЛЬНЫХ ПРОГРАММНЫХ КОМПОНЕНТ, НЕОБХОДИМЫХ ДЛЯ ФУНКЦИОНИРОВАНИЯ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ В ВЫБРАННОЙ АРХИТЕКТУРЕ 25
3.1 Разработка портала Академгородка 25
3.1.1 Архитектура портала Академгородка 26
3.2 Заключение по разделу 27
4 СОЗДАНИЕ НЕОБХОДИМЫХ УНИКАЛЬНЫХ ПРОГРАММНЫХ КОМПОНЕНТ, НЕОБХОДИМЫХ ДЛЯ ФУНКЦИОНИРОВАНИЯ РИС В ВЫБРАННОЙ АРХИТЕКТУРЕ 28
4.1 Интеграция DSpace и портала Академгородка 28
4.2 Административная часть портала 34
4.3 Модуль конференций 35
5 СОЗДАНИЕ ПОЛНОФУНКЦИОНАЛЬНОГО ПРОТОТИПА РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ НА БАЗЕ ОГРАНИЧЕННОГО ЧИСЛА УЧАСТНИКОВ 38
6 ОЦИФРОВКА НАИБОЛЕЕ ВАЖНЫХ НАУЧНЫХ ТРУДОВ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ ИНСТИТУТОВ 39
ЗАКЛЮЧЕНИЕ 41
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 43
ПРИЛОЖЕНИЕ А Схема данных и административная часть модуля конференции 46

ВВЕДЕНИЕ

На сегоднящний день технологический прогресс требует стремительного темпа роста вычислительно-информационных мощностей постпетафлопсных масштабов, предоставляющих получения новых фундаментальных знаний и внедрения инноваций в производственные, сельскохозяйственные, экономические и социальные сферы. В связи с этим стремительно развиваются информационные технологии, включающее теоретические и прикладные исследования, облачные вычисления и крупномасштабные компьютерные эксперименты, операции со сверхбольшими данными и искусственный интеллект. Решение подобных задач требует больших объёмов программного обеспечения нового поколения, архитектурных решений и конструктивных технологий. Мировые тренды ведут к созданию интегрированных вычислительно-информационных окружений, образующих инструментальную среду для автоматизации построения моделей и алгоритмов, их отображения на платформы супер ЭВМ и создания комфортных интерфейсов для пользователей с разной профессиональной подготовкой. Такая экосистема рассчитана на длительный жизненный цикл с непрерывным развитием и согласованным участием различных групп разработчиков. Фактически она приводит к образованию новой отрасли с массовыми профессиями высокой суперкомпьютерной грамотности. Обозначенные научно-технологические вызовы требуют компетентных организационных и инфраструктурных решений, включающих межведомственную координацию академических, образовательных и производственных коллективов [1]. При всей гармоничности развития различных социально-экономических направлении в XXI веке очевидным оказывается значение информационных технологий. Роль этих направлении кардинально возросла после массового появления персональных компьютеров, а затем в связи с рождением многопроцессорных ЭВМ терафлопсного и петафлопсного уровня 10^{12}–10^{15} операций в секунду. Скорость увеличения компьютерных мощностей подчиняется закону Г. Мура, одного из основателей компании «Интел», – экспоненциальный рост в 1000 раз за 11 лет. Хотя по физическим законам уже должно было наступить насыщение, аномальная динамика сохраняется. На сегодняшний день мощный суперкомпьютер имеет около 200 петафлопс, а появление экзафлопсников ожидается к концу 2020 года, которая может выполнять порядка 10^{18} арифметических действий в секунду. Такая многопроцессорная вычислительная система будет иметь сотни миллионов расчётных устройств, пользователи получат доступ к супервычислениям принципиально новых масштабов. Появится необходимость хранения и обработки огромных объёмов числовых данных [1]. Поэтому создание интегрированной распределенной библиотечной информационной системы долговременного хранения, накопления информации и обеспечения надежного защищенного доступа к цифровым объектам, в которой сбор и первичная обработка информации производится на местах, а хранение, накопление на централизованных облачных ресурсах является актуальной проблемой. В своем Послании от 1

сентября 2020 года «Казахстан в новой реальности: время действий» Президент страны К.К. Токаев отметил, что конкурентноспособную экономику невозможно создать без развития сельского хозяйства. Системными проблемами в АПК являются дефицит профессиональных кадров, а также низкий уровень развития аграрной науки. Поэтому разработка и внедрение интегрированной распределенной информационной системы для обеспечения аграрной науки является актуальной задачей.

Во-первых, перспективными становятся технологии для дистанционного зондирования земли, точного земледелия, технологии визуализации космических данных на карте, оперативного доступа к высокодетальным снимкам для точного картографирования границ полей и севооборотов, зон плодородия, а также мониторинга состояния вегетации [2]. При этом используются данные со спутников, датчиков и других систем.

Во-вторых, большую прикладную значимость имеет внедрение цифровых технологий в библиотечные фонды и архивы научно-исследовательских институтов агропромышленного направления. Создание цифрового представления библиотечных ресурсов позволяет не только обеспечить их сохранность на многие столетия, но и предоставить их широкому кругу лиц. Следует признать, что большая часть работ, опубликованных до середины XX века учеными, так и не получила широкого распространения вследствие отсутствия единого хранилища или наукометрических баз данных.

В-третьих, актуальны системы онлайн-контроля элементов производственного цикла и управления агро-бизнесом: техники, полей и складов. Данные системы обрабатывают данные о погоде, индексе вегетации, состоянии почвы и севообороте. Например, российский проект ”Агросигнал” [3], к системе которого подключено свыше 150 хозяйств, обрабатывающих более 2 млн гектаров земли, сосредоточен в основном на контроле за посевами.

В-четвертых, немаловажную роль играют специализированные информационные сервисы, содержащие каталоги и базы знаний, а также данные наблюдений. Данные сервисы нередко используются для публикации новостей и аналитических статей, позволяют работникам агропромышленного направления обсудить острые вопросы и поделиться опытом. Академгородке города Алматы функционируют несколько научно-исследовательских институтов МОН РК и МСХ РК, которые проводят научные исследования в различных областях агропромышленного комплекса, перерабатывающей промышленности, по микробиологии и вирусологии, генетики и цитологии и другим направлениям. Значительные объемы информации, полученной в результате исследований данных институтов, их непрерывное увеличение и разнородный по многим признакам характер хранения и распространения, отсутствие унифицированного доступа к ним создают существенные проблемы их эффективного использования. Данные проблемы приводят к необходимости поиска новых подходов и решений проблем создания хранилищ информационных ресурсов, их организации, средств и способов доступа к ним пользователей. Такие подходы сегодня называют «цифровыми» или «электронными» библиотеками [4, 5]. Одним из актуальных направлений

является создание и применение распределенных компьютерных систем для интенсивной работы с данными, что актуально как для решения новых научных задач, использование больших объемов научных данных, генерируемых современными измерительными средствами, так и экономических и социальных проблем на основе больших данных и технологии их обработки и анализа. Одним из основных результатов научной деятельности является создание и накопление опыта предыдущих поколений. Существует достаточно много технологических разработок информационных систем для электронных библиотек. Наиболее известными в России являются, например, Euro CRIS (eurocris.org), eLibrary (elibrary.ru), Информика (informika.ru), MathNET (mathnet.ru). В Казахстане функционирует Республиканская межвузовская электронная библиотека (РМЭБ) на базе университетских образовательных ресурсов по единой технологии, предложенной Ассоциацией вузов Республики Казахстан и научно-образовательной компьютерной сетью Казахстана KAZRENA. Одним из пионеров, и как следствие, одним их мировых лидеров в развитии научно-образовательной информационной системы является «Электронная библиотека Сибирского отделения РАН», который сегодня называется Порталом СО РАН (sbras.ru). Данный ресурс по данным Webometrics занимает первое место в России среди научных организаций и 45-е место в мире. Однако основной недостаток большинства систем – ограниченность интеграции ресурсов как внутри каждой из них, так и с внешними системами. Серьезной проблемой является идентификация информационных ресурсов, позволяющая получать библиографические сведения, а также устанавливать связи определенного ресурса с другими фактами и объектами. В настоящей элетронной книге поставлена цель создать интегрированную распределенную библиотечную информационную систему Академгородка города Алматы, позволяющую сохранить результат интеллектуальной деятельности научно-исследовательских институтов МОН РК, МСХ РК и КазИТУ, расположенных в Академгородке г. Алматы, объединяемых с помощью телекоммуникационной технологии WiMAX. В книге описан процесс использования телекоммуникационной технологии WiMAX, разработанная с целью предоставления универсальной беспроводной связи между НИИ на больших расстояниях для широкого спектра устройств. На основе предложенной технологии разработана и реализована распределенная библиотечная информационная система научно-образовательного кластера НИИ и КазИТУ, расположенных в Академгородке г. Алматы. С помощью технологии WiMAX оцифрованные наиболее важные научные труды НИИ и КазИТУ, расположенных в Академгородке г. Алматы сразу поступают на сервер баз данных. Для этого организовано прямое соединение НИИ и ЦОД ТОО «Академсеть». За счет него обеспечивается более высокий уровень информационной безопасности. Информационная система позволяет хранить колоссальные объемы информации при минимизации объема хранения, осуществлять их выбор, обеспечивая быстроту и точность поиска. Созданная информационная система ориентирована на обеспечение потребностей участников научно-образовательного кластера, основана на передовых

информационных технологиях и реализована на основе свободно распространяемого программного обеспечения и уникальных программных модулях, разработанных в процессе выполнения.

Для достижения поставленной цели предполагается решение следующих задач:

1. Анализ возможностей и потребностей всех участников научно-образовательного кластера для построения оптимальной архитектуры распределенной информационной системы (РИС).
2. Определение оптимальной архитектуры РИС, функциональности и номенклатуры всех ее компонентов.
3. Отбор и адаптация к выбранной архитектуре необходимого свободно распространяемого программного обеспечения.
4. Создание необходимых уникальных программных компонент, необходимых для функционирования РИС в выбранной архитектуре.
5. Создание полнофункционального прототипа РИС на базе ограниченного числа участников.
6. Оцифровка наиболее важных научных трудов указанных научно-исследовательских институтов, накопленных десятилетиями (за всю историю существования).
7. Наполнение РИС актуальной информацией и информацией, созданной в процессе оцифровки документальных источников.
8. Опытная эксплуатация прототипа РИС.
9. Создание полного комплекта документации по РИС и отдельным ее компонентам.
10. Обеспечение научных исследований современными компьютерными ресурсами через Центр обработки данных, доступного участникам проекта с помощью радиотехнологий WiMAX.

Научная новизна исследования состоит в следующем:

1. Будет предложена методика проектирования распределенных информационных систем (РИС), основанных на распределенных базах данных, что позволяет своевременный и беспрепятственный доступ к информации.
2. Будет предложена технология реализации распределенных информационных систем с распределенной базой данных на основе сервис-ориентированной архитектуры (SOA).
3. На основе предложенной методики будет разработана и реализована распределенная библиотечная информационная система научно-образовательного кластера НИИ и КазИТУ, расположенных в Академгородке г. Алматы. С помощью технологии WiMAX оцифрованные наиболее важные научные труды НИИ и КазИТУ, расположенных в Академгородке г. Алматы сразу поступают на сервер баз данных. Для этого организовано прямое соединение НИИ и ЦОД ТОО «Академсеть».

1 ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОЙ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ФУНКЦИОНАЛЬНОСТИ И НОМЕНКЛАТУРЫ ВСЕХ ЕЕ КОМПОНЕНТ

1.1 Анализ возможностей участников научно-образовательного кластера

В соответствии с полученной от руководств научно-исследовательских институтов информацией, подготовлен реестр информационных ресурсов участников, линий связи и вычислительных мощностей, на основании которого сделано следующее заключение:

– четыре из шести научно-исследовательских институтов и КазИТУ имеют веб-сайты в сети Интернет, содержащие общую информацию об организации, новости о последних достижениях и др.;

– отсутствует единое хранилище публикаций сотрудников научно-исследовательских институтов и других типов информационных ресурсов; большая часть результатов научных исследований не оцифрована;

– инфраструктура ИКТ большинства участников научно-образовательного кластера изношена;

– сети внутренних коммуникаций и шлюз выхода в Интернет не защищены в соответствии с требованиями к сетям государственных органов;

– в некоторых из названных НИИ серверная инфраструктура отсутствует, что делает невозможным покупку и эксплуатацию серверного оборудования, а там где она есть, ее характеристики не соответствуют отраслевым стандартам;

– компьютерный парк, несмотря на относительную работоспособность и возможность выполнения простейших задач, требует модернизации или замены;

– во всех научно-исследовательских институтах используется разнородное программное обеспечение и системы управления базами данных, что затрудняет их унификацию, необходимую для оптимизации затрат;

– современное программное обеспечение не используется в работе научно-исследовательских институтов по причине неподготовленности собственного, либо привлекаемого ИТ-персонала. Привлечение специалистов более высокой квалификации на данный момент представляется затруднительным в силу ограниченности бюджета. По этой же причине затруднена покупка и эксплуатация более современных аппаратных ИКТ комплексов.Проанализировав приведенные факты сделан единственно возможный вывод о целесообразности и необходимости выноса систем хранения данных за пределы научно-исследовательских институтов (участников научно-образовательного кластера) в Центр обработки данных ТОО «Академсеть», оставив рабочие места для первичной обработки информации непосредственно в научно-исследовательских институтах. Такая распределенная централизованная архитектура систем хранения данных позволяет:

– унифицировать инструментарий хранения и обработки данных;

– использовать современное, в достаточной мере производительное

аппаратное обеспечение, в том числе серверы и системы хранения данных, обеспечив, при необходимости, их модернизацию и масштабирование;

– привлечь к разработке создаваемых систем IT-персонал достаточно высокой квалификации, разделив в итоге затраты по его содержанию на всех участников проекта. По суммарным затратам это, возможно, будет не дешевле существующих затрат, но такому персоналу будут доступны задачи более высокого современного уровня;

– обеспечить создаваемую информационную систему необходимыми средствами защиты от атак и взлома. В том числе и в соответствии с требованиями к сетям государственных органов. Такое требование, кажущееся на первый взгляд избыточным к научно-исследовательским организациям, в результате может оказаться необходимым, поскольку информация, размещаемая в информационной системе может попадать под различные уровни секретности.В дальнейшем эта информация будет передаваться внешним пользователям только при наличии соответствующей авторизации через единый шлюз доступа в Интернет. Шлюз также должен получить развитие с общим развитием проекта, как по увеличению функционала защиты от атак и взлома, так и по разграничению прав доступа и защите подключаемых каналов связи. Потенциальная возможность присутствия в рамках проекта информации с ограниченными категориями доступа требует соответствующей защищенности каналов связи от рабочих мест НИИ до Центра обработки данных (ЦОД). На этапе отладки допустимо использование общих Интернет-каналов, но с развитием проекта весь комплекс должен опираться только на защищенные каналы связи НИИ-ЦОД. Более надежными могут считаться закрытые каналы связи типа "точка-точка", которые намного надежнее открытых каналов даже при использовании в последних различных алгоритмов шифрования. Каналы связи от рабочих мест НИИ до ЦОД должны иметь возможность существенного увеличения пропускной способности. В этом случае у НИИ появится возможность использования других сервисов ЦОД. В частности, технологий виртуализации, сервисов IaaS, DaaS, SaaS, PaaS, CCTVaaS, SecaaS и других. Это позволит существенно расширить возможности IT-парка всех НИИ на базе облачных мощностей используемого ЦОД и объединить их в итоге в единый научно-образовательный кластер (НОК).

В качестве архитектуры централизованной системы хранения данных на начальном этапе проекта при запланированном относительно небольшом объеме данных, выбрана архитектура типа NAS. Этот вывод сделан в силу того, что типовая процедура обработки данных в реализуемом проекте строится по схеме, когда клиентский запрос серверу обрабатывается одним или несколькими серверами обработки данных, которые получая необходимые данные, производит их обработку и возвращает результат этой обработки пользователю. При архитектуре системы хранения данных типа NAS обработка запросов ведется в рамках одной сети. При этом по этой сети идут как сами запросы, так и все необходимые для обработки этого запроса данные. Поэтому для того, чтобы сеть, объединяющая НИИ и ЦОД, обладала необходимой пропускной способностью, она должна обеспечивать необходимую полосу

пропускания для больших объемов данных практически по всей сети. Это не всегда возможно и всегда достаточно дорого в реализации. Обработка и предварительная подготовка данных по определенным стандартам будет производиться на рабочих местах в НИИ, без загрузки сети, и только впоследствии она будет передаваться на хранение в библиотеки на централизованную систему хранения данных, размещаемую в ЦОД.

1.2 Построение радиосети Академгородка на базе телекоммуникационной технологии WiMAX

В рамках первых двух этапов работы была разработана интегрированная распределенная библиотечная информационная система (РБИС) долговременного хранения, накопления информации и обеспечения надежного защищенного доступа к цифровым объектам, в которой сбор и первичная обработка информации производится на местах, а хранение, накопление на централизованных облачных ресурсах в специализированном дата-центре. Внешний интерфейс для стороннего, регламентированного, множественного доступа к накопленным ресурсам размещен на портале https://acagor.kz/, представляющий собой авторскую разработку участников проекта в части построения архитектуры и написания программного кода. Получено авторское свидетельство на созданную распределенную информационную систему научно-исследовательских институтов Академгородка города Алматы выданное РГП «Национальный институт интеллектуальной собственности» МЮ РК. В рамках проекта для нескольких НИИ – участников проекта, в число которых входят НИИ почвоведения и агрохимии им. Успанова, НИИ перерабатывающей и пищевой промышленности, НИИ плодоовощеводства, НИИ микробиологии и вирусологии, НИИ физиологии человека и животных, НИИ генетики и цитологии, образовательного учреждения КазИТУ, за счет гранта были оборудованы выделенные компьютеризированные рабочие места, предназначенные для сканирования и первичной цифровой обработки печатного материала, отображающего результаты интеллектуальной деятельности , накопленные в этих научных учреждениях за десятилетия. Значительная часть указанных печатных научных трудов представляет собой безусловную ценность для отечественной науки. Но, ввиду ведомственных пертурбаций, перманентного бюджетного секвестирования и сложностей хранения, существует очевидная вероятность их безвозвратной утери для отечественной науки. Тогда как, перевод этих результатов работ нескольких поколений ученых в цифровой формат позволяет сохранить их для более длительного и менее затратного хранения в рамках РБИС. Кроме того, цифровая информация может быть использована в рамках учебного процесса образовательной системы РК. Все указанные компьютеризированные терминалы были распределены непосредственно по вышеназванным НИИ. Следует отметить, что в настоящее время все НИИ относятся к различным министерствам и ведомствам, что априори предполагает различную стратегию их развития и как следствие самостоятельную политику информатизации. То

есть, ввиду ведомственной разобщенности, каждый НИИ представляет из себя некую автономную систему, с неопределенным уровнем общей квалификации персонала по части ИКТ. Как правило квалификация ИТ персонала в неспециализированных НИИ находится не на самом высоком уровне. На практике, это является осязаемой проблемой для всего искомого проекта в целом. Необходимо отметить, что оцифровка с последующим размещением в РБИС проводится в сквозном порядке, без предварительной экспертизы. В результате существует вероятность того, что часть размещаемого на портале РБИС материала, гипотетически может попадать в разряд государственных секретов, свободный, а тем более несанкционированный доступ к которой извне - недопустим. Дуализм данной проблемы очевиден. С одной стороны, во избежание утери необходимо оцифровать все подряд. Но с другой стороны, существует риски несанкционированного доступа. Авторами проекта было принято решение проводить сквозную оцифровку и хранить накопленные цифровые объекты в закрытом хранилище данных. Вопросы организации политики доступа к информации оставить на рассмотрении администрации научных учреждений. Тем не менее, еще на этапе подготовки проекта его инициаторами прорабатывались различные аспекты обеспечения информационной безопасности, как комплекса мероприятий проводимых во избежание утери и утечки информации в общей распределенной системе. Если на самом портале был задействован комплекс общеизвестных программных средств для защиты от несанкционированного доступа извне, то, со стороны мест первичной подготовки материала на рабочих местах в НИИ существует достаточно много рисков. Поскольку персонал и оборудование на этих местах полностью подчинены внутренней политике своего НИИ. Учитывая тот факт, что для этих рабочих мест в РБИС отведена важная роль с расширенными правами доступа на запись, в случае слабой защиты внутренней сети данного НИИ от внешнего взлома существует высокая вероятность атаки, а возможно и взлома всей РБИС со стороны, именно этих рабочих мест. Поэтому можно однозначно утверждать, что работа данных рабочих терминалов проекта с открытым доступом в Интернет, может представлять определенные риски. На практике можно производить передачу накопленной цифровой информации от рабочих терминалов к центральному депозитарию курьерами на внешних носителях, но тогда возникают другие риски. Появляются дополнительные затраты на транспортировку информации, учет, дополнительной меры защиты и т.д. В то же время при передаче информации в закрытых сетях на канальном уровне (классическое определение дано согласно классической сетевой модели OSI [6]), при передаче не используется внешний IP адрес, а лишь динамически распределяемый внутренний адрес. Нет внешнего IP адреса, соответственно сетевой объект невидим для прямых внешних угроз и атак из Интернет. Согласно общепринятых постулатов информационной безопасности абсолютной защиты от взлома и внешних атак не существует. Существует лишь вопрос цены взлома. В данном случае взлом закрытой сети с неизвестным алгоритмом построения внутренней адресации, будет стоить потенциальным злоумышленникам ощутимо дорого. Причем, не факт, что они смогут достичь

цели. С учетом вышесказанного, участниками проекта было предложено связать закрытыми каналами передачи данных все рабочие места с центральным депозитарием РБИС и объединить их в единую закрытую сеть без доступа в Интернет с самих рабочих терминалов. По сути, это - чисто техническая часть работы, построения корпоративной сети городского масштаба. В традиционной классификации компьютерных сетей LAN, MAN, WAN, создаваемую сеть можно считать сетью класса MAN (municipality area network - сеть городского масштаба). Особенности построения таких сетей описаны во многих учебниках. Например, в работе [6] относительную проблематику с технической и организационной точки зрения создает ведомственная разобщенность НИИ участников проекта, чьи разноплановые локальные сети были объединены в единую сеть городского масштаба.

В системе образования РК уже известны различные попытки объединения в сеть образовательных учреждений. Так сеть Schoolnet связывает 207 школ города Алматы, сеть KazRena связывала ВУЗы Алматы. В большинстве своем названные сети передачи данных создавались для организации централизованного доступа узлов сети (школ или ВУЗов) на каналах связи стороннего оператора, а не на собственных каналах связи. Такой подход предполагает дополнительные перманентные расходы государства на эксплуатацию сети. К тому же основной целью инициаторов этих проектов была организация доступа хостов (рабочих узлов) сети к внешним ресурсам. По сути, это было решением last mile доступа в Интернет для всех сегментов сети. Тогда как, описываемая сеть изначально строится вокруг своих ресурсов и вокруг ЦОД. Обеспечение доступа к сети Интернет научных учреждений, это автономная задача каждого из НИИ. Впрочем, преимущества централизованного доступа учреждений у сети Интернет с централизованными функциями по информационной безопасности, также были продемонстрированы участникам проекта в ходе его реализации. Что удивительно, опыта объединения сетью НИИ и Образовательных учреждений из различных ведомств, в Казахстане - нет. Хотя наличие такой сети могло бы позволить не только использовать совместные ресурсы, но и агрегировать имеющиеся вычислительные мощности. Например, совместно использовать для научных исследований супрекомпьютер при наличии такового на балансе одного из участников кластера. Или же создать сеть суперкомпьютеров – HPC Grid (Higher Perfomance Computer). Схожие идеи высказываются в работах [6, 7]. К сожалению, до появления суперкомпьютера в создаваемой сети еще достаточно далеко. Но, собственные ресурсы кластера после реализации данного проекта уже появились. Они доступны пользователям сети на портале https://acagor.kz/. На сегодняшний день авторы портала ставят перед собои цель создать в рамках портала многопроцессорный вычислительный кластер. С помощью которого будут решатся различные задачи производственные, природопользовательские, экономические и социальные сферы. Одним из таких задач является задача изучения и прогноза эпидемиологической ситуации от короновирусной инфекции COVID-19. Специалисты из разных областей науки разделились во мнении, сколько продлиться пандемия. Издание The Atlantic

проанализировала и составило 4 наиболее вероятных сценария развития событий [8]. Первый и второи сценарии предпологали что короновирус продлится от одного месяца до четырех месяцев, часть медицинского сообщества была уверена, что коронавирус окажется несерьезным патогеном. К сожелению за это время короновирус не пошел на спад, и пандемия продолжает наростать. Трети сценарии предполагал, что пандемия продлится до одного года, а четвертый больше года. В этих прогнозах нет четких сроков по окончанию пандемии, они сделаны аналитическим путем. Изучению и прогноза посвящены методологий совместного использования математических моделей и реальных данных, которая является эффективным инструментом для изучения сложных эпидемиологических процессов и решения на ее основе практических задач. Существенную и важную роль здесь играют сопряженные задачи. Прямые и обратные связи между моделями и ежедневными реальными данными о распространений коронавируса и системную организацию вычислительных технологий удобно строить с помощью вариационных принципов. Этот подход естественно приводит к комбинированному использованию методов прямого и обратного моделирования. Наиболее перспективным для решения задач оценки и прогнозирования процессов распространения вирусов является подход, базирующийся на классическом вариационном принципе Лагранжа с использованием сопряженных уравнений.

Для более точного прогноза эпидемиологической ситуации, необходимо разработать математическую модель распространения коронавируса COVID-2019 и разработать вычислительный алгоритм для многопроцессорного вычислительного кластера. Это позволит более точно описать эпидемиологическую ситуацию во всем мире. Схема взаимодействия портала Академгородка с вычислительной системой представлена на рисунке 1.

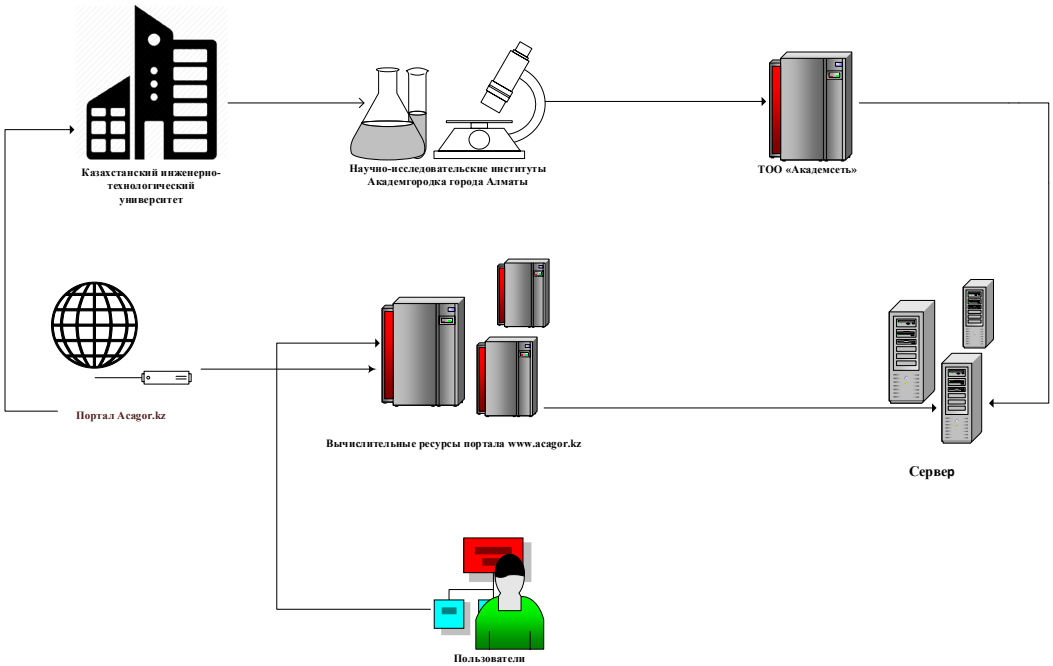

Рисунок 1-Структурная схема вычислительной системы на базе портала Академгородка

С ростом объемов накапливаемого цифрового контента вполне возможно создание собственного ЦОД например в КазИТУ. Также при достаточном финансировании возможно построение высокоскоростной ВОЛС сети. Как достижение первой космической скорости открыло перед человечеством совершенной новый класс задач и технологий, так и построение высокоскоростной сети передачи данных позволит использовать в НИИ большой спектр новых приложений. К тому же создаваемая сеть и централизованные ИТ ресурсы могут стать местом для практики студентов КазИТУ, а возможно, станет местом их будущего трудоустройства. Безусловно, на данный момент, в индустрии ИКТ лучшим по многим параметрам решением для создания распределенной корпоративной сети является оптоволоконная сеть. Срок амортизации такой сети составляет 25 лет. Но, уровень капитальных затрат при создании ВОЛС сети является достаточно высоким, особенно для экспериментальных проектов подобных описываемому, даже с учетом уникальной географической близости НИИ участников проекта. Поэтому рабочей группой проекта было принято решение воспользоваться относительно недорогими радио технологиями WLL на базе оборудования известной фирмы UBIQUITY. Преимущества радиосетей городского масштаба в сравнении с ВОЛС также общеизвестны.

Во-первых, относительно невысокая стоимость развертывания и стоимость подключения даже достаточно удаленных (до 30 км) точек. Это дает существенную гибкость при экспериментировании с топологией. Как следствие существенно снижается стоимость масштабирования апробированных решений. Нет оплаты за трафик. Есть только стоимость амортизации радио оборудования (антенн), стоящего на балансе КазИТУ и стоимость оплаты лицензии на радиочастоты. В данном проекте использовались радиочастоты 5.47 -5,49 ггц предоставленные со стороны ТОО Академсеть в безвозмездное пользование участникам проекта. При необходимости отказа от услуг ТОО «Академсеть» участники проекта могут подать собственную заявку на радиочастоты, либо проложить ВОЛС сеть с гораздо большей производительностью. Особенно в случае более масштабного использования дополнительных перспективных результатов проекта, которые, были получены в работе только после построения своей закрытой радиосети передачи данных. В частности , была апробирована технология VDI (Virtual desktop infrastructure – виртуальный десктоп), позволяющая перенести вычислительные мощности рабочих компьютеров проекта в используемый Дата-Центр. При таком подходе вся работа с данными производится централизованно в ЦОД, мощности которого позволяют задействовать средства защиты высокого уровня. А рабочие места в НИИ используются только как терминал виртуального десктопа и подключения сканера. Даже первичное размещение отсканированного контента было организовано не на десктопе в НИИ, а непосредственно на VDI физически размещенный в ЦОД. Такой подход избавляет от рисков дублирования контента в личных интересах оператора терминала в НИИ. В период реализации искомого проекта, в качестве

терминала использован недорогой ведомственный ПК, установленный непосредственно в НИИ, или КазИТУ. Наличие стабильного канала от КазИТУ до ЦОД позволило апробировать возможности технологии VDI для использования в учебном процессе одной из групп ИТ специальности университета. Для связи рабочих терминалов проекта с ЦОД использовалось оборудование компании UBIQUITY. Профессиональная точка доступа Ubiquiti Rocket M5 (Рисунок 2) является высокопроизводительным решением операторского уровня для создания базовых станция в диапазоне 5 ГГц.

Рисунок 2 - Профессиональная точка доступа Ubiquiti Rocket M5

Данный диапазон менее зашумлен и предпочтителен для использования в местах, где высока загрузка распространенного диапазона 2.4 ГГц. Мощность радиомодуля Rocket M5 составляет до 27 dBm в зависимости от выбранного режима работы. Поскольку точки доступа Rocket не имеют встроенных антенн, в качестве антенны используется узконаклонная антенна Ubiquiti RocketDish с параболическим отражателем (Рисунок 3).

Рисунок 3 - Узконаклонная антенна Ubiquiti RocketDish

Точка доступа и антенна полностью защищены от попадания влаги, не требуют дополнительных кожухов для зимней эксплуатации и могут быть смонтированы на стену, мачту или установлены на специальной треноге. Топология соединения сети НИИ и КазИТУ строилась на основе ПО Ubiquiti UniFi Network Planner по протоколу 802.11b, n. Монтажные работы по установке оборудования для беспроводного соединения НИИ и КазИТУ проводились сотрудниками ЦОД ТОО «Академсеть» (Рисунок 5). На основе предложенной технологии разработана и реализована распределенная библиотечная информационная система научно-образовательного кластера НИИ и КазИТУ, расположенных в Академгородке г. Алматы. С помощью технологии WiMAX оцифрованные наиболее важные научные труды НИИ и КазИТУ, расположенных в Академгородке г. Алматы сразу поступают на сервер баз данных. Для этого организовано прямое соединение НИИ и ЦОД ТОО «Академсеть». За счет него обеспечивается более высокий уровень информационной безопасности. Информационная система позволяет хранить колоссальные объемы информации при минимизации объема хранения, осуществлять их выбор, обеспечивая быстроту и точность поиска.

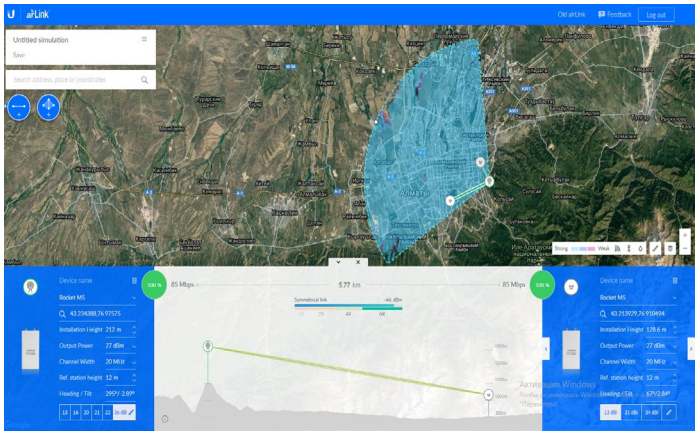

Рисунок 4 - Профессиональная точка доступа Ubiquiti Rocket M5
Ubiquiti UniFi Network Planner

Рисунок 5 – Монтажные работы по установке оборудования для беспроводного соединение WiMAX

1.3 Программная часть распределенной информационной системы

Программная часть распределенной информационной системы состоит из следующих подсистем, представленных на рисунке 6:

- Репозиторий цифровых объектов;
- Подсистема управления текущей научно-исследовательской информацией;
- Подсистема интеграции распределенных информационных ресурсов;
- Подсистема доступа к распределенным информационным ресурсам на основе Web-технологий (Web-портал).

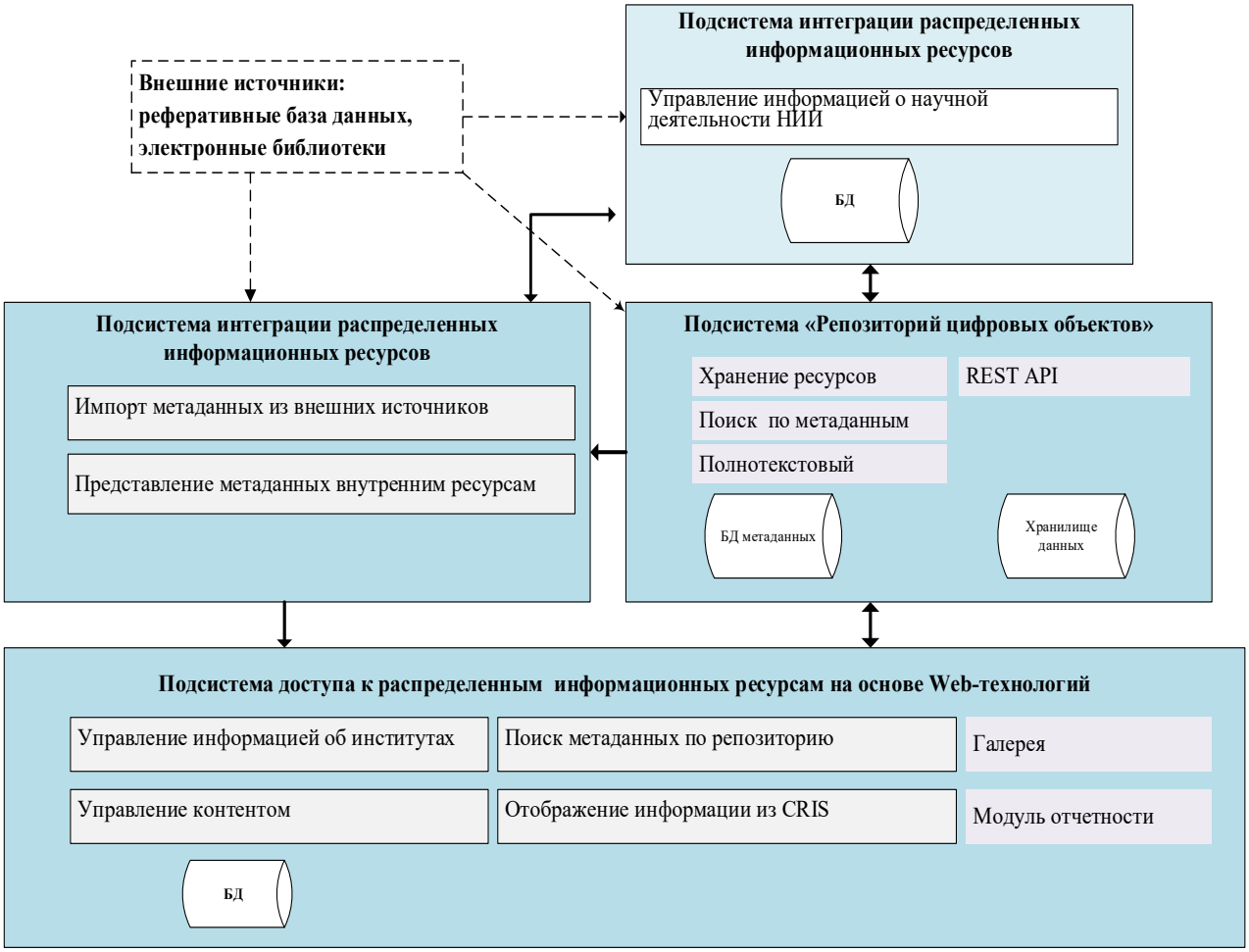

Рисунок 6 - Взаимосвязь подсистем распределенной информационной системы

Подсистема "Репозиторий цифровых объектов" предназначена для долгосрочного хранения результатов научной деятельности НИИ с возможностью поиска информационных ресурсов по метаданным, полнотекстового поиска, ведения статистики.

Подсистема управления текущей научно-исследовательской

информацией предназначена для хранения информации о НИИ, их сотрудниках, а также информации об их научной деятельности (участие в финансируемых проектах, конференциях, прохождение стажировок и др.).

Подсистема интеграции распределенных информационных ресурсов предназначена для импорта метаданных из внешних источников (реферативных баз данных), а также для предоставления метаданных внутренним ресурсам (порталу Академгородка) на основе стандартных протоколов.

Подсистема доступа к распределенным информационным ресурсам на основе Web-технологий предназначена для предоставления стандартизированного единого пользовательского интерфейса для всех функций и модулей, входящих в состав распределенной информационной системы. Портальная часть должна обеспечивать удаленный доступ к информационным ресурсам и услугам. Подсистема объединяет процессы предоставления доступа к информационным ресурсам всем заинтересованным субъектам взаимодействия.

Разработано техническое задание для разработки программной части распределенной библиотечной информационной системы (Приложение Е), в котором определены требования к отдельным компонентам системы, к ее архитектуре, описаны роли пользователей, требования к эргономике и технической эстетике, требования к поддержанию логической целостности данных и обеспечению обратимости изменений, требования к надежности и безопасности, требования к режимам функционирования, требования к защите информации от несанкционированного доступа, требования к эксплуатации, техническому обслуживанию и хранению компонентов системы и требования к патентной чистоте.

1.4 Модель функционирования распределенной информационной системы

На рисунке 7 описана модель функционирования распределенной информационной системы, согласно которой:

– программная часть (Web-сервер, хранилище данных и базы данных) расположены в ЦОД ТОО «Академсеть»;

– контент-менеджеры (сотрудники НИИ и КазИТУ) наполняют контент посредством Web-интерфейса на страницах репозитория и системы управления научными исследованиями;

– внешние пользователи со своих рабочих станций посещают Web-портал Академгородка и просматривают информацию об институтах, производят поиск по репозиторию и системе управления научными исследованиями на странице портала.

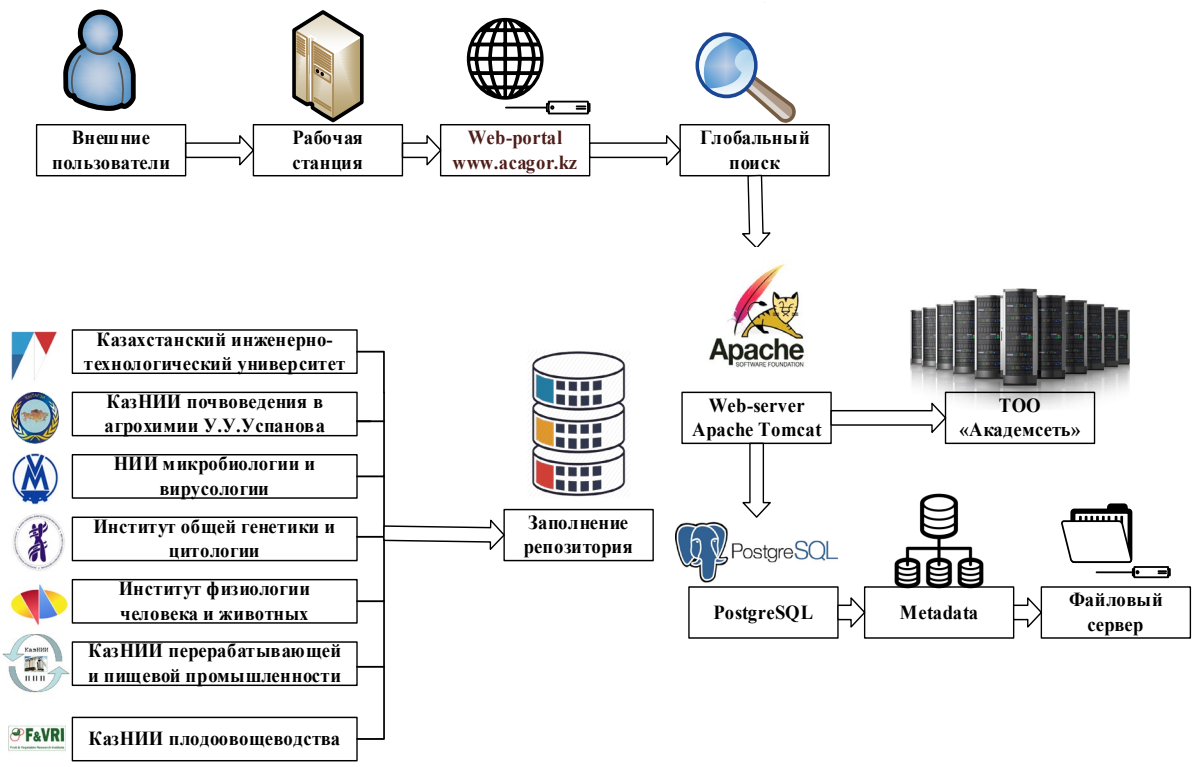

Рисунок 7 - Модель функционирования распределенной информационной системы

1.5 Заключение по разделу

Построенная модель распределенной информационной системы полностью соответствует потребностям участников научно-образовательного кластера как по информационному содержанию, так и по поддержке отраслевой и языковой специфике, поскольку решает основные задачи, предъявляемые к этим системам: обеспечение системы надежного долговременного хранения цифровых (электронных) документов с сохранением всех смысловых и функциональных характеристик исходных документов; обеспечение «прозрачного» поиска и доступа пользователей к документам, как для ознакомления, так и для анализа содержащихся в них фактов; организация сбора информации по удаленным цифровым репозиториям, поддерживающим протоколы OAI-PMH, SRW/SRU, Z39.50. Система в полной мере обеспечивает необходимыми вычислительными ресурсами проводимые научные исследования и образовательные процессы, упрощая перспективу дальнейшего ее развития, и позволяет построить передовую IT-инфраструктуру управления интеллектуальным капиталом, электронной библиотекой, в котором будут храниться все книги и научные труды Казахстанского инженерно-технологического университета и научно-исследовательских институтов Академгородка города Алматы.

2 ОТБОР И АДАПТАЦИЯ К ВЫБРАННОЙ АРХИТЕКТУРЕ НЕОБХОДИМОГО СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Согласно третьему пункту календарного плана реализован отбор необходимого свободно распространяемого программного обеспечения. При проектировании и разработке программной части распределенной информационной системы исполнители проекта придерживались следующих нормативных документов:

– Закон РК "О персональных данных и их защите" № 94-V (от 21.05.2013 г. с изменениями от 28.12.2017 г.);

– Закон РК "О связи" №567-II (от 05.07.2014 г. с изменениями от 09.01.2018 г.).

2.1 Выбор, установка и адаптация репозитория цифровых объектов

К программному обеспечению, положенному в основу подсистемы "Репозиторий цифровых объектов", были выдвинуты следующие основные требования:

1. Возможность работы с произвольными документами, географическими картами, аудио- и видеоматериалами;
2. Гибкая организация хранения ресурсов (разбиение на коллекции, подколлекции и пр.);
3. Гибкая система прав пользователей: создание групп пользователей; возможность задания доступа пользователей заданной группы к заданному множеству объектов по требуемому способу доступа (загрузка, просмотр, редактирование, удаление, изменение атрибутов); идентификация, аутентификация и авторизация пользователей;
4. Возможность интеграции распределенных информационных ресурсов на основе стандартных протоколов (Z39.50/SRU/SRW);
5. Наличие программного интерфейса (API) для осуществления интеграции с внутренними ресурсами;
6. Распознавание текста в оцифрованных материалах для организации полнотекстового поиска;
7. Сбор статистики и предоставление различной отчетности.

При выборе подсистемы репозитория цифровых объектов было изучены следующие институциональные репозитории и системы для создания электронных библиотек: Ambra, Digital Commons, DSpace, ePrints, Evergreen ILS, Greenstone, Fedora Commons, Invenio, RODA и VuFind, а также изучен опыт их использования организациями Канады, Новой Зеландии, Чехии и Шри-Ланки [5], [9-13].

Анализируя преимущества и недостатки перечисленных систем, были отобраны системы DSpace, ePrints и Greenstone, как наиболее удовлетворяющие установленным требованиям.

К сильным сторонам Greenstone можно отнести иерархическую

структуризацию каждого документа, автоматическое извлечение метаданных из документа при его загрузке. Однако, данная система поддерживает лишь ограниченное количество форматов: MS Word, Excel, RTF, HTML, Plain, PDF, ZIP, MP3. Хранение географических карт, а также других результатов научной деятельности, имеющих более сложную структуру, на взгляд исполнителей проекта, не предусмотрено. Система предоставляет широкие возможности поисковых запросов: помимо булевых операций и группировки слов с помощью скобок, поддерживается поиск слов в исходной форме.

К сильным сторонам DSpace можно отнести более совершенную систему прав пользователя по сравнению с рассмотренными системами: различные научно-исследовательские институты могут иметь свои области в пределах системы. В каждом институте могут быть назначены определенные сотрудники, ответственные за премодерацию, т.е. пользователи, имеющие возможность просматривать и редактировать материалы до того, как они будут включены в репозиторий. DSpace, как и другие рассмотренные системы, предоставляет интерфейсы интеграции с другими подсистемами на основе открытых международных стандартов. DSpace поддерживает более 70 форматов информационных ресурсов. Материалы, имеющиеся в DSpace индексируются в Google Scholar. Существует большое количество плагинов к системе DSpace, расширяющие его возможности, в том числе и система управления научными исследованиями DSpace-CRIS.

Система ePrints поддерживает больше форматов метаданных, однако не поддерживает расширенное дублинское ядро. В системе можно назначить следующие роли пользователей: пользователи, редакторы, администраторы, причем редакторы имеют доступ к материалам всех научно-исследовательских институтов, хранящихся в репозитории.

В результате проведенного анализа был выбран репозиторий цифровых объектов DSpace версии 6.2. Для корректной работы данного программного обеспечения, в системе дополнительно установлены следующие компоненты:

1. Комплект разработчика приложений на языке Java – Open JDK;
2. Apache Maven, Apache Ant для компиляции DSpace из исходных кодов;
3. Свободный веб-сервер Apache Tomcat для взаимодействия системы с пользователем посредством Web-технологий;
4. Открытая платформа полнотекстового поиска Apache SOLR.
5. Свободная СУБД PostgreSQL для хранения метаданных о цифровых объектах и другой информации.

Репозиторий доступен по адресу https://acagor.kz/resource/repository, главная страница которого показана на рисунке 8. При сборке программного обеспечения DSpace 6.2, внесены изменения в ее конфигурацию с целью адаптации к условиям, установленным в Республике Казахстан. Стандартная схема метаданных DSpace, основанная на схеме DCMI, расширена следующими полями: «Журнал в списке ВАК», «Номер названия, тома, выпуска журнала», «Полная библиографическая ссылка по ГОСТ», «Первая и последняя страница статьи» и др. Список типов ресурсов дополнен позициями «Статья в материалах конференций», «Статья в газете», «Отчет по НИР», «Патент»,

«Технический отчет», «Музейный объект», «Рабочие материалы». Кроме того, добавлена возможность работы с географической информацией согласно работе [14].

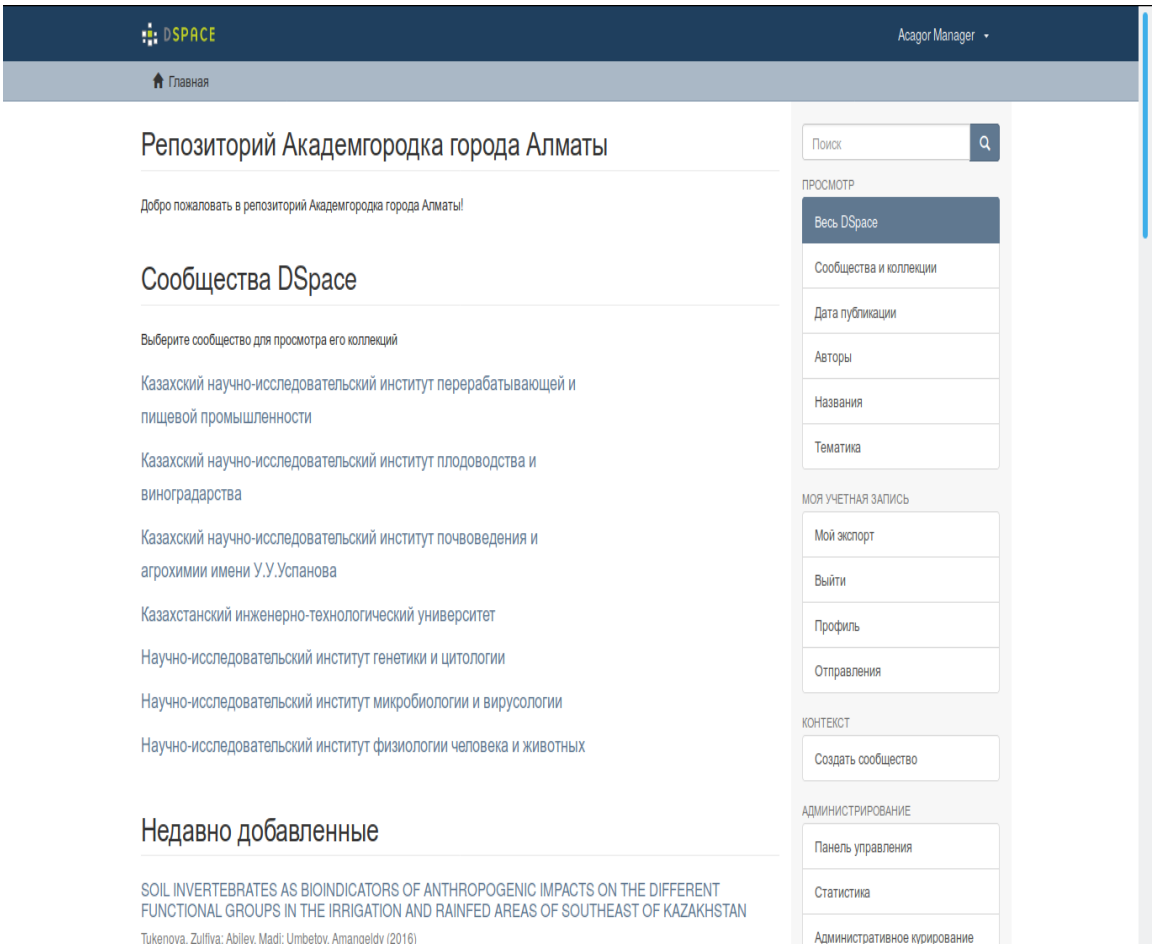

Рисунок 8 - Главная страница репозитория цифровых объектов https://acagor.kz/resource/repository

Для поддержки процесса наполнения полнотекстовых баз данных, созданные профили метаданных зарегистрированы в системе DSpace, и в соответствии с ними настроены рабочие процессы, а также пользовательский интерфейс системы.

Внутренняя организация хранения ресурсов в системе DSpace построена следующим образом: в репозитории созданы 7 сообществ, соответствующие научно-исследовательским институтам. Каждое сообщество, в свою очередь, состоит из нескольких коллекций, соответствующих типу ресурса (статьи, монографии, отчеты НИР и пр.)

На рисунке 9 приведен пример страницы репозитория с информацией о публикации.

Рисунок 9 - Пример страницы репозитория с информацией о публикации

Ввод данных в подсистему осуществляется:

– в интерактивном режиме через встроенные WEB-интерфейсы;

– заимствованием данных из других систем (DOI, PubMet, ArXiv, CiNii, CrossRef и др.);

– в пакетном режиме - импорт данных в форматах DIM, MEKOF, MARC21, DC и др.;

– синхронизацией данных по OAI-PMH с внешними источниками.

Доступ к данным репозитория возможен не только через WEB-интерфейсы DSpace, но и по протоколам OAI-PMH, SOLR, SRW/SRU, Z39.50. Последнее обеспечивается связью DSpace с системой ZooSPACE [15].

2.2 Выбор, установка и адаптация системы управления научными исследованиями

Изначально в качестве системы управления научными исследованиями [16, 17] было выбрано свободное расширение системы DSpace – DSpace-CRIS. Система позволяет хранить следующую информацию:

– информацию о научно-исследовательских организациях;

– информацию о сотруднике научно-исследовательских организаций, различные написания его имени, в том числе на разных языках;

– ссылки на профили в различных базах данных (Scopus, Researcher ID, ORCID);

– информацию о научной деятельности (участие в финансируемых проектах, конференциях, прохождение стажировок и др.).

Система ведет статистику публикаций по каждому ученому. Система интегрирована с репозиторием цифровых объектов, что позволяет просматривать публикации ученых. CRIS-система позволяет экспортировать

информацию о публикациях ученого в популярные форматы. Система допускает отслеживание изменений на странице с помощью технологий RSS.

Система доступна по адресу https://acagor.kz/resource/cris . Пример профайла ученого в системе DSpace-CRIS приведен на рисунке 10.

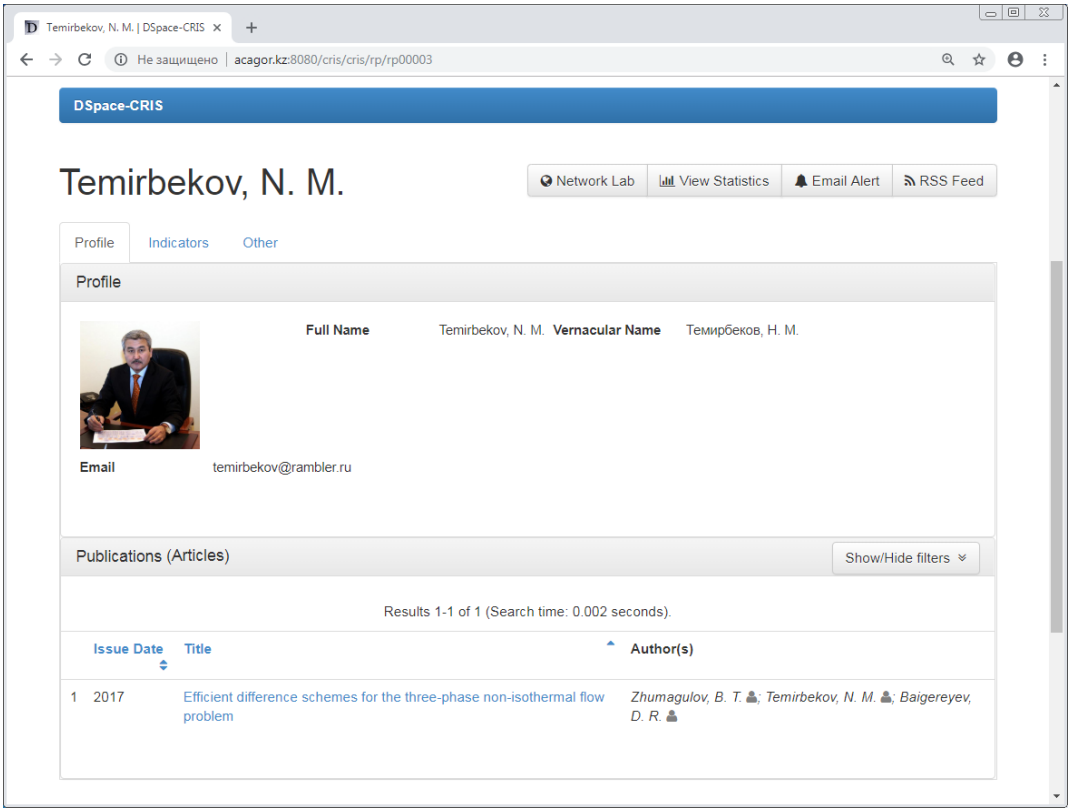

Рисунок 10 - Пример профайла ученого в системе CRIS

В последствии данное прогрммное обеспечение заменено программным модулем собственной разработки (см. раздел 4).

2.3 Выбор интегрирующего программного обеспечения

Изначально в качестве интегрирующего программного обеспечения выбрана распределенная информационная система ZooSPACE, разработанная научными сотрудниками ИВТ СО РАН [18-20].

Распределенная информационная система ZooSPACE интегрирует данные из различных информационных источников, обеспечивая доступ к разнородной распределенной информации в соответствии со стандартными протоколами (SRW/SRU, Z39.50). Система функционирует на основе оригинальных серверов ZooPARK-ZS, серверов LDAP и WEB-серверов Apache, обеспечивая сквозной поиск информации в разнородных базах данных, извлечение информации в стандартных схемах и форматах и ее отображение [15].

Адаптация данного программного обеспечения к распределенной информационной системе Академгородка находится на завершающей стадии,

после отладки которого будет осуществлена интеграция репозитория цифровых объектов как с внутренними ресурсами (ИРБИС-64 и др.), так и ресурсами сторонних организаций, доступных по протоколам SRW/SRU или Z39.50.

В последствии данное прогрммное обеспечение заменено программным модулем собственной разработки (см. раздел 4).

2.4 Заключение по разделу

Таким образом, согласно третьему пункту календарного плана, выполнены следующие работы:

- произведен анализ институциональных репозиториев и систем для создания электронных библиотек, а также изучен опыт их использования другими организациями;
- произведен анализ соответствия указанных систем установленным требованиям;
- произведена установка репозитория цифровых объектов DSpace;
- произведена адаптация DSpace к условиям, установленным в Республике Казахстан: расширена стандартная схема данных DSpace;
- учитывая пожелания руководств научно-исследовательских институтов добавлена возможность работы с географической информацией;
- определена внутренная структура организации хранения ресурсов в DSpace;
- произведена установка системы управления научными исследованиями DSpace-CRIS;
- на завершающей стадии разработки находится адаптация интеграции репозитория DSpace с системой DSpace-CRIS согласно выбранной архитектуре распределенной информационной системы;
- выбрано интегрирующее программное обеспечение – распределенная информационная система ZooSPACE;
- проводится отладка и адаптация данной системы к распределенной информационной системе Академгородка.
- используется телекоммуникационная технология WiMAX, разработанная с целью предоставления универсальной беспроводной связи между НИИ на больших расстояниях для широкого спектра устройств. На основе предложенной технологии разработана и реализована распределенная библиотечная информационная система научно-образовательного кластера НИИ и КазИТУ, расположенных в Академгородке г. Алматы. С помощью технологии WiMAX оцифрованные наиболее важные научные труды НИИ и КазИТУ, расположенных в Академгородке г. Алматы сразу поступают на сервер баз данных.

3 СОЗДАНИЕ УНИКАЛЬНЫХ ПРОГРАММНЫХ КОМПОНЕНТ, НЕОБХОДИМЫХ ДЛЯ ФУНКЦИОНИРОВАНИЯ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ В ВЫБРАННОЙ АРХИТЕКТУРЕ

3.1 Разработка портала Академгородка

Данная подсистема распределенной информационной системы предназначена для предоставления стандартизированного единого пользовательского интерфейса для всех функций и модулей, входящих в состав распределенной информационной системы. Портальная часть должна обеспечивать удаленный доступ к информационным ресурсам и услугам. Подсистема объединяет процессы предоставления доступа к информационным ресурсам всем заинтересованным субъектам взаимодействия. Представление услуг предполагает:

– навигацию по сервисам;

– доступ к информации о проекте и нормативным правовым и методическим материалам;

– идентификацию, аутентификацию и авторизацию пользователей;

– предоставление основной информации о научно-исследовательских институтах, их сотрудниках;

– предоставление информации о последних достижениях научно-исследовательских институтов, о предстоящих событиях, конференциях;

– наличие фото- и видеогалереи;

– личный кабинет, анализ времени работы на сайте, хранение истории запросов;

– управление запросами по поиску авторов, названию публикаций;

– управление запросами по полнотекстовому поиску по ресурсам;

– экспорт данных в различные форматы;

– предоставление различных отчетностей.

К порталу выдвинуты следующие основные требования:

– Серверная часть портала должна использовать шаблон проектирования Model-View-Controller (MVC) – схему разделения данных приложения, пользовательского интерфейса и управляющей логики на три отдельных компонента: модель, представление и контроллер таким образом, что модификация каждого компонента может осуществляться независимо.

– Клиентская часть портала должна использовать адаптивную верстку для удобства просмотра сайта как на настольных компьютерах, так и на мобильных устройствах.

Портал Академгородка, главная страница которого приведена на Рисунке 11, доступен по адресу https://acagor.kz.

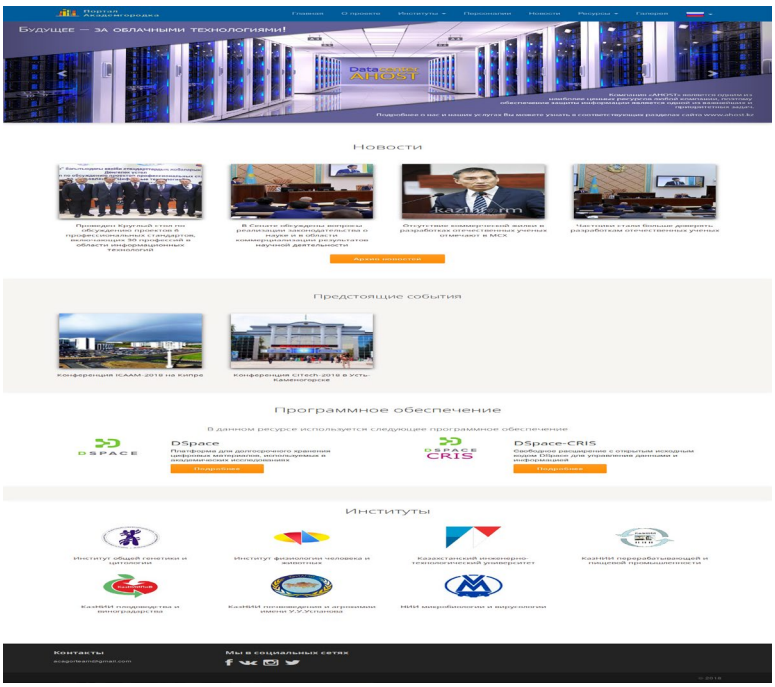

Рисунок 11 - Главная страница портала Академгородка https://acagor.kz

3.1.1 Архитектура портала Академгородка

Сайт разработан с использованием веб-фреймворка Django 2.0.5 (на языке Python 3.6). Выбор данного веб-фреймворка определен возможностями быстрого прототипирования, автоматического создания миграций для деплоймента. Схема взаимодействия портала Академгородка с остальными составляющими распределенной информационной системы представлена на Рисунке 12.

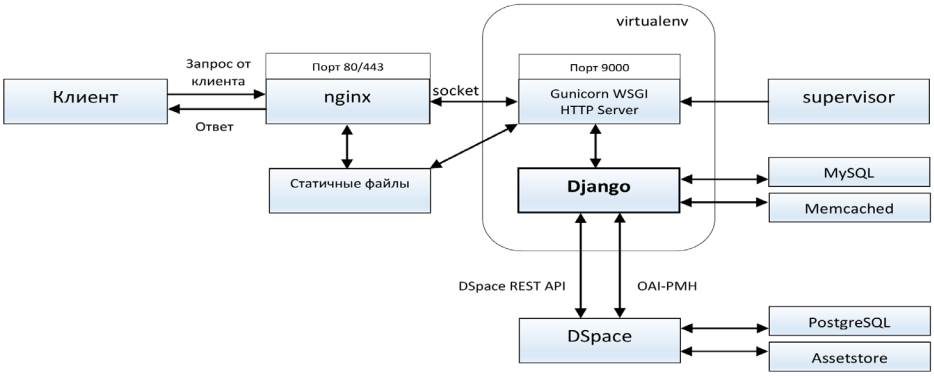

Рисунок 12 - Схема взаимодействия портала Академгородка с составляющими распределенной информационной системы

Сайт функционирует на WSGI HTTP-сервере Gunicorn 19.8.1 с HTTP-сервером nginx 1.14, установленным в качестве обратного прокси-сервера.

Сайт использует программное обеспечение Memcached 1.5.6 для реализации временного хранения (кэширования) данных в оперативной памяти

с целью разгрузки базы данных. Для хранения данных портала используется система управления базами данных MySQL 5.7.23.

Django установлен в виртуальном окружении Python с целью избежания возможных конфликтов версий пакетов, возможности обновления необходимого вспомогательного программного обеспечения независимо от пакетов, установленных в операционной системе. Кроме бинарных пакетов python, pip и wheel, в виртуальном окружении установлен веб-сервер Gunicorn. В Django установлены следующие дополнительные пакеты: пакет mysqlclient для работы с базой данных, пакет memcached для работы с сервисом кэширования, пакет PIL для работы с изображениями.

Для запуска и мониторинга Gunicorn используется система контроля процессами supervisor, которая запускает сокет UNIX, прослушивает его и автоматически запускает Gunicorn в ответ на запрос.

Установлен SSL-сертификат от центра сертификации Let's Encrypt.

3.2 Заключение по разделу

Согласно четвертому пункту календарного плана выполнены следующие работы:

- разработана архитектура портала Академгородка, схема его взаимодействия с остальными составляющими распределенной информационной системы;
- произведено описание программного обеспечения, использованного для функционирования портала;
- разработан механизм автоматического деплоймента портала Академгородка;
- осуществлена интеграция портала Академгородка с репозиторием цифровых объектов с использованием DSpace REST API. В результате на странице профиля ученого отображается список его материалов; на странице научно-исследовательского института отображается список всех ресурсов данного института; имеется возможность производить поиск информации репозитория по метаданным.
- произведена отладка и адаптация программного обеспечения ZooSPACE и работа по написанию собственных модулей для интеграции с остальными внутренними библиотечными системами и внешними реферативными базами данных;
- разработан модуль оплаты онлайн-услуг пользования ресурсами системы;
- разработана система автоматического резервного копирования.

4 СОЗДАНИЕ НЕОБХОДИМЫХ УНИКАЛЬНЫХ ПРОГРАММНЫХ КОМПОНЕНТ, НЕОБХОДИМЫХ ДЛЯ ФУНКЦИОНИРОВАНИЯ РИС В ВЫБРАННОЙ АРХИТЕКТУРЕ

4.1 Интеграция DSpace и портала Академгородка

В целях предоставления пользователю единого пользовательского интерфейса ко всем функциям интегрированной распределенной информационной системы, осуществлена интеграция репозитория цифровых объектов DSpace и Портала Академгородка [21-26].

Рассмотрено два подхода интеграции. Программно реализован первый подход, который основан на использовании API-интерфейса REST DSpace, обеспечивающий программный интерфейс для сообществ, коллекций, метаданных элементов и файлов. Для этого развернуто веб-приложение REST. При каждом запросе пользователя портала Академгородка осуществляется запрос в DSpace REST с помощью утилиты curl. В DSpace создан непривилегированный пользователь, от имени которого производится поиск по репозиторию. Ответы от DSpace REST возвращаются в формате JSON, парсинг которых производится в разработанном программном модуле Integrator проекта Django.

Первое обращение в DSpace REST осуществляется следующим запросом:

curl -v -X POST --data ”email=acagormanager@yandex.kz &password=****” https://acagor.kz:8080/rest/login.

Ответ возвращается в следующем виде:

HTTP/1.1 200 OK

Set-Cookie: JSESSIONID=***; Path=/rest/; Secure; HttpOnly

Следующая последовательность запросов возвращает список коллекций (научно-исследовательских институтов):

curl -v ”http://acagor.kz:8080/rest/communities” –cookie ”***”

curl -v ”http://acagor.kz:8080/rest/communities/ i/collections” –cookie ”***”, где i – идентификатор (handle) сообщества, полученного в предыдущем запросе.

Следующий запрос возвращает список трудов Казахского научно-исследовательского института перерабатывающей и пищевой промышленности:

curl -v ”http://acagor.kz:8080/rest/collections/7/items” –cookie ”***”.

Анализ ответов показал, что в зависимости от текущей загруженности репозитория, каждый запрос в DSpace REST осуществляется от 10 до 30 секунд. Это создает временный отказ в обслуживании (DoS) репозитория цифровых объектов, а при выполнении запросов из нескольких рабочих станций может вызвать распределенный отказ в обслуживании (DDoS).

Программно реализован второй подход, который заключается в выгрузке метаданных и кэшировании результатов во временной базе данных в течение некоторого времени. Создана временная база данных со схемой, представленной на Рисунке 17.

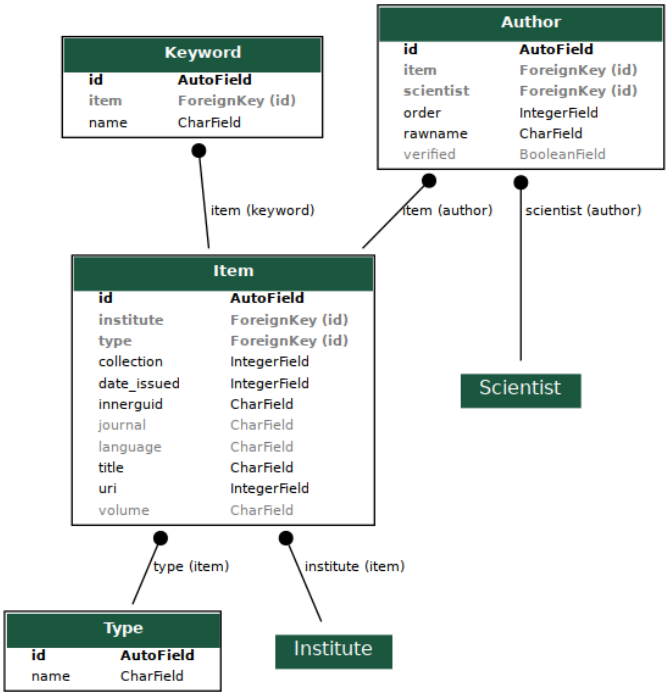

Рисунок 17 - Схема базы данных

Данный подход позволил существенно сократить время отклика. При использовании закэшированных данных время отклика составило в пределах 0,01-0,08 секунд (Рисунок 18).

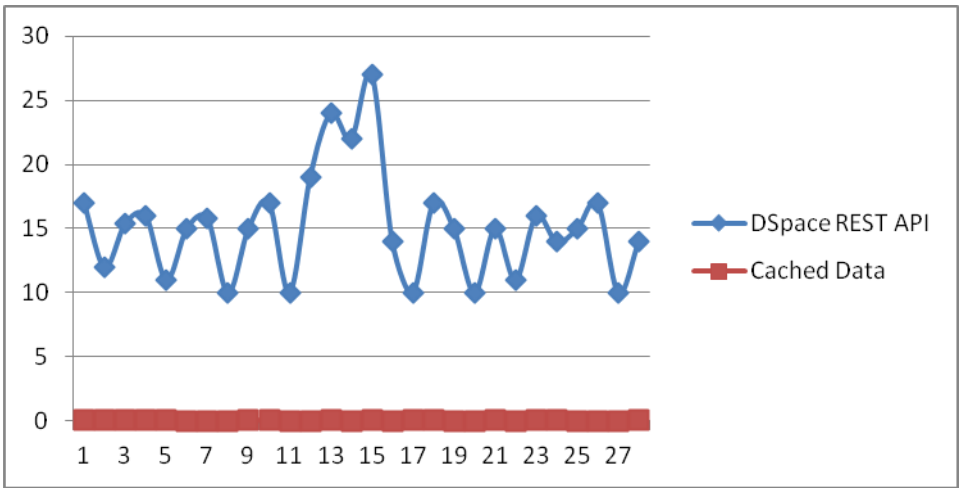

Рисунок 18 - Среднее время отклика при использовании DSpace REST API и закэшированных данных

Кроме того, проанализирован модуль для CMS Drupal, реализованный Джорджтаунским университетом.

Результат интеграции репозитория цифровых объектов DSpace и портала Академгородка прослеживается на следующих страницах веб-портала.

а) Поиск ресурсов по названию ресурса или его авторов (Ресурсы >

Поиск). В левой части страницы отображается статистика найденных ресурсов в разрезе научно-исследовательских институтов, даты публикации, типа ресурса и ключевых слов. Реализована фильтрация найденных ресурсов по указанным фильтрам. На Рисунке 19 приведен результат поиска по названию ресурсов, содержащих слово «почва». Числа в скобках указывают количество найденных ресурсов, соответствующих данному критерию.

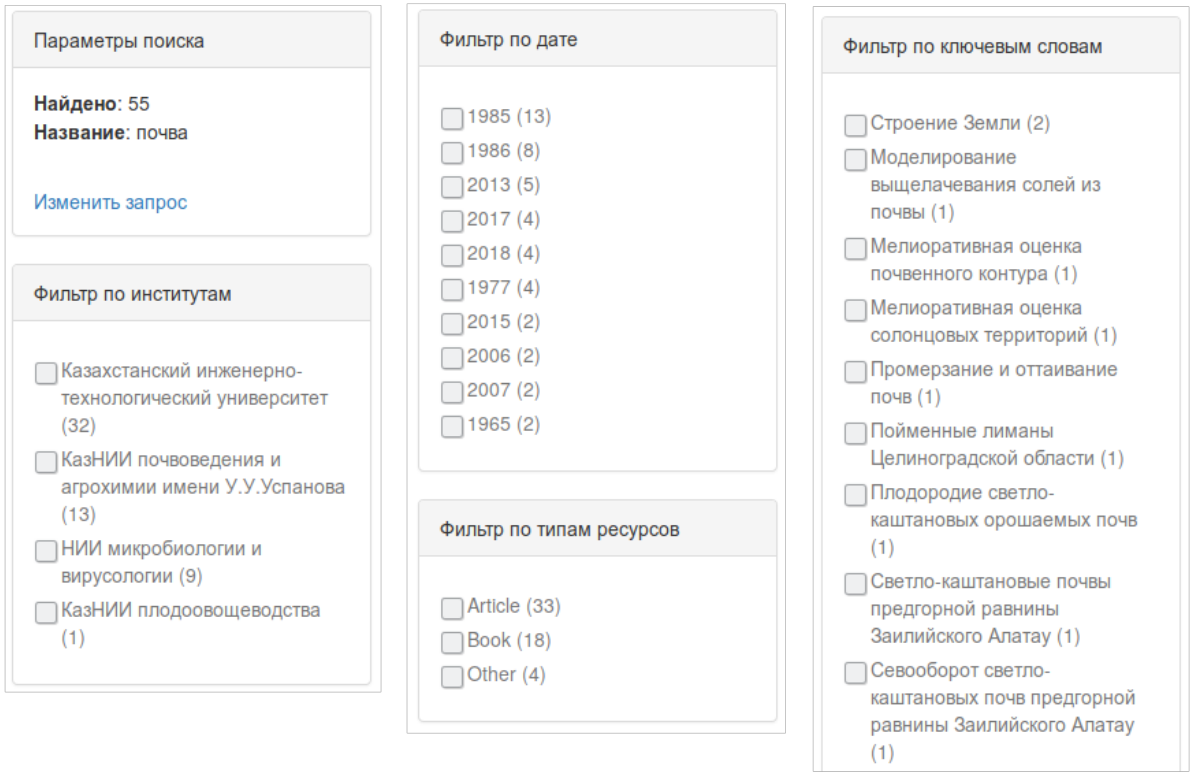

Рисунок 19 - Блок фильтров в разрезе научно-исследовательских институтов, даты публикации, типа ресурса и ключевых слов

При нажатии на ссылку с именем ученого в результатах поиска открывается его профиль (Рисунок 20), который содержит краткую информацию об его научном вкладе, ссылки на профили в различных базах данных (при наличии), а также перечень публикаций, загруженных в репозиторий цифровых объектов (Рисунок 21). При наличии ссылок на реферативные базы данных (ORCID, Researcher ID), указывается данная информация. Список всех ученых, чьи труды загружены в репозиторий цифровых объектов отображается в разделе «Персоналии» (Рисунок 22).

Темирбеков Нурлан Муханович

Ученая степень: доктор физико-математических наук

Звание: профессор

Основной вклад

Темирбеков Н.М. является известным специалистом в области вычислительной математики и вычислительной гидродинамики. Им проведены фундаментальные исследования математических вопросов аппроксимаций, устойчивости и сходимости разностных схем для уравнений Навье-Стокса с сильно меняющимися коэффициентами. Численно решены уравнения Навье-Стокса преобразованные в произвольные ортогональные криволинейные координаты. Предложены эффективные методы численного решения задач гидродинамики в областях со сложной геометрией;

Участвовал в работе создания автоматизированной информационной системы организаций учебного процесса в вузе при кредитной системе обучения. Получено свидетельство о государственной регистрации объекта интеллектуальной собственности под названием «Автоматизированная информационная система организации учебного процесса» (программа для ЭВМ);

Создана новая геоинформационная система (ГИС) для оценки влияния антропогенных источников загрязнения на качество атмосферного воздуха промышленного города. Информационная система внедрена в Восточно-Казахстанском филиале «Казгидромет»;

По результатам научных исследований опубликованы две монографии «Приближенные методы решения уравнений вязкой жидкости в областях со сложной геометрией», «Математические модели и информационные технологии пограничного слоя атмосферы» и более 90 научных статей в научных журналах;

Для науки и высшей школы Казахстана под его руководством подготовлено 10 кандидатов наук и докторов PhD. Руководитель нескольких научных проектов в области фундаментальной и прикладной математики;

Рисунок 20 - Профиль ученого

2013г. Награжден благодарственным письмом министра образования и науки МОН РК Саринжипова А.Б. за личный вклад в развитие системы образования РК, дипломом акима ВКО Сапарбаева Б.М. «Лучший ректор – руководитель, поддерживающий все инициативы студенческой молодежи»;

2014г. Награжден золотой медалью А.Байтурсынова, нагрудным знаком 20 лет маслихатам Республики Казахстан и почетной грамотой Председателя Сената Парламента РК Токаева К. за вклад в развитие отечественной законодательной системы;

Публикации в цифровом репозитории

- An Iterative Method for Solving Nonlinear Navier-Stokes Equations in Complex Domains Taking into Account Boundary Conditions with Uniform Accuracy
- Efficient difference schemes for the three-phase non-isothermal flow problem
- Fractional flow formulation for three-phase non-isothermal flow in porous media
- INFORMATION TECHNOLOGY FOR NUMERICAL SIMULATION OF VISCOUS INCOMPRESSIBLE FLOW IN BICONNECTED DOMAINS
- Numerical solution of an optimal control problem governed by three-phase non-isothermal flow equations
- Numerical solution of the problem of optimizing the process of oil displacement by steam
- Probabilistic Statistical Modeling of Air Pollution from Vehicles
- Study of Stability of the Difference Scheme for the Model Problem of the Gaslift Process
- Построение научно-образовательного кластера на облачных вычислительных технологиях
- Построение научно-образовательного кластера на облачных вычислительных технологиях 2018
- Приближенные методы решения уравнений вязкой жидкости в областях со сложной геометрией

Рисунок 21 - Перечень публикаций, загруженных в репозиторий цифровых объектов, на странице профиля ученого

Поиск успешно производится также с учетом различных окончаний искомого слова. Из руководства к DSpace выяснено, что на текущий момент возможность указания разного написания имени ученого не реализована. Другими словами, Иванов и Ivanov считаются разными учеными. При поиске ресурсов по автору учитывается различное написание его имени, в том числе на разных языках. Пример результатов поиска по автору приведен на Рисунке 23. Из приведенного примера видно, что указанные ресурсы попали в выборку, хотя имя автора имеет разное написание.

б) Отображение списка научно-исследовательских институтов, описание научных направлений и достижений, списка их сотрудников, а также метаданных публикаций, имеющихся в репозитории цифровых объектов (Рисунок 24).

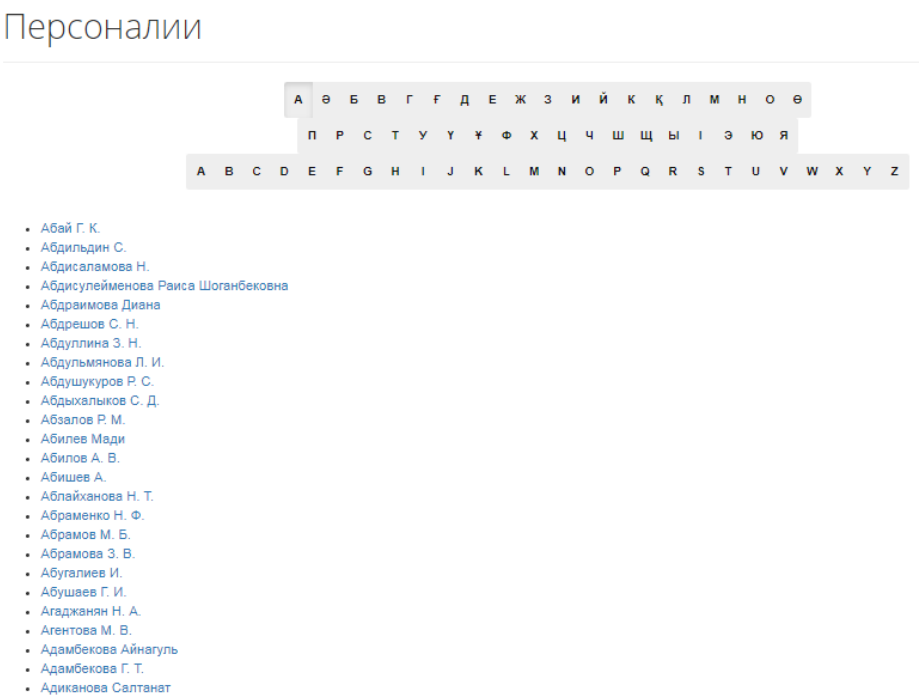

Рисунок 22 - Раздел «Персоналии»

Рисунок 23 - Результат поиска по автору

Рисунок 24 - Список публикаций сотрудников научно-исследовательского института на веб-портале

Разработан механизм сопоставления автора с уже имеющимися в базе данных авторами при загрузке нового ресурса с целью представления более релевантной информации пользователям. При появлении новых статей алгоритм перебирает все введенные имена и его вариации (например, с помощью транслитерации), и в случае нахождения, присваивает его к данной статье. Кроме того, в административной части веб-портала разработана страница для контент-менеджера (Рисунок 25), позволяющая управлять сопоставлением авторов: переназначать профиль, если алгоритм неправильно определил ученого, добавлять новые альтернативные имена и др.

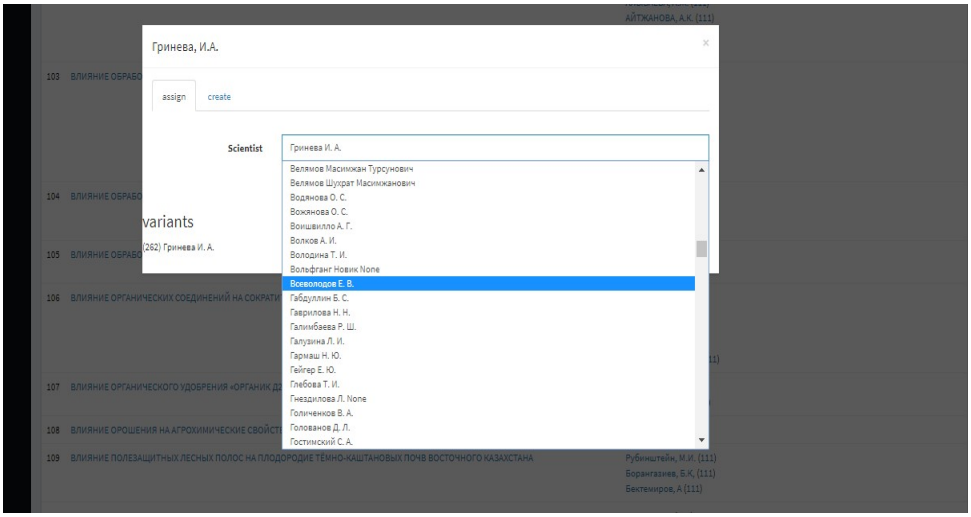

Рисунок 25 - Управление процессом сопоставления авторов

4.2 Административная часть портала

Административная часть портала представляет собой приложение Django, которое позволяет осуществлять действия, требующие повышенные привелегия. Реализованы следующие разделы:

- домашняя страница, содержащая информацию о текущем состоянии системы, залогиненных пользователях, управление пользователями и др.;
- научно-исследовательские институты;
- модуль интеграции;
- модуль конференции;
- управление баннерами на главной странице;
- галерея;
- новости;
- персоналий;
- публикации;
- утилиты.

Общий вид приложения приведен на Рисунке 26. В левой части страницы отображены разделы административной части. Технически, разделы представляют собой модели Django, помеченные специальным флагом. Доступ к каждому разделу назначается индивидуально каждому пользователю приложения.

На Рисунке 26 отображена страница для редактирования списка научно-исследовательских институтов. На рисунке 27 приведена форма создания и редактирования записи для научно-исследовательского института.

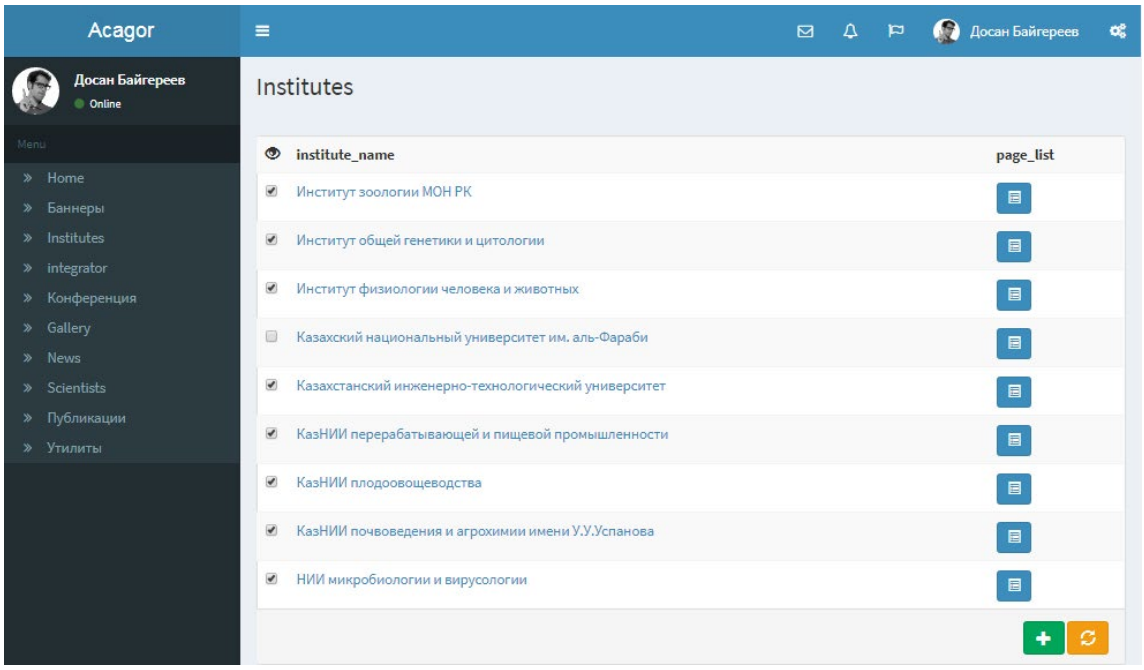

Рисунок 26 - Административная часть портала: список научно-исследовательских институтов

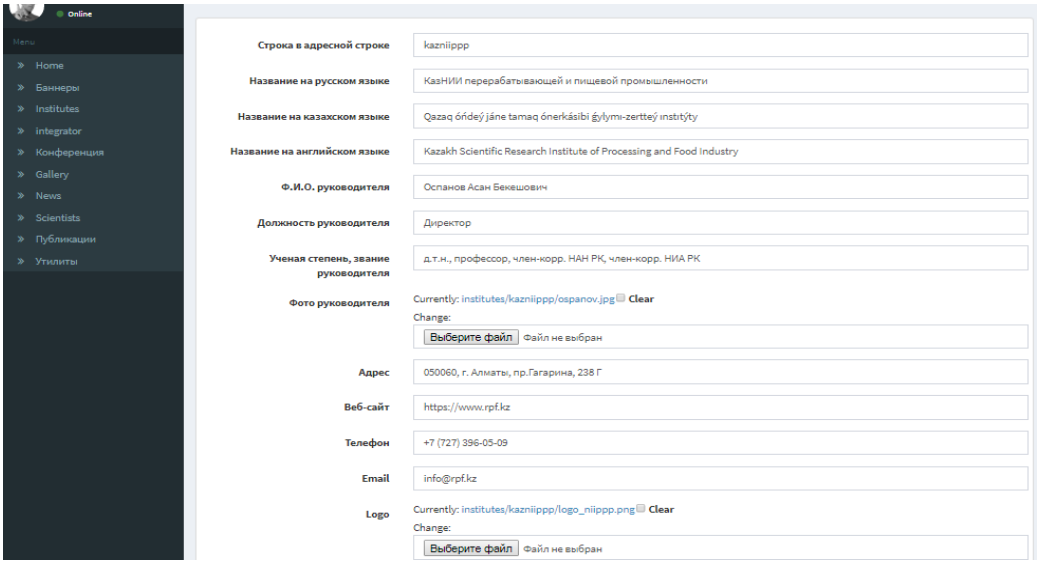

Рисунок 27 - Административная часть портала: редактирование информации о научно-исследовательском институте

4.3 Модуль конференций

Разработан модуль конференций, позволяющий проводить организационную работу по регистрации участников конференции, сбору материалов, объявлению решений организационного комитета, загрузке квитанций по оплате и др. Главная страница модуля для пользователей (участников конференции) представлена на Рисунке 28.

Каждый участник конференции проходит обязательную регистрацию на портале Академгородка. Данные пользователи являются локальными пользователями приложения Django и не вносятся в службу каталогов.

Процесс подачи заявки состоит из трех этапов. На первом этапе вводится название, язык доклада и форма участия. На втором этапе вводится информация об авторах доклада, место работы, контактная информация. На третьем этапе производится загрузка файла с докладом. Уведомление о решении организационного комитета автоматически отправляется на электронную почту участника.

В административной части модуля конференции (Рисунок Е.1) имеется возможность установить один из статусов заявки (Принята, Отклонена, Возвращена на доработку), просмотреть историю изменения статусов, а также текст отправленных электронных сообщений (Рисунок Е.2), просмотреть статистику в разрезе секций, скачать загруженные материалы конференции в zip-архиве (Рисунок Е.3).

Кроме того, реализована возможность подачи заявки на бронирование гостиницы, указания времени прибытия участника.

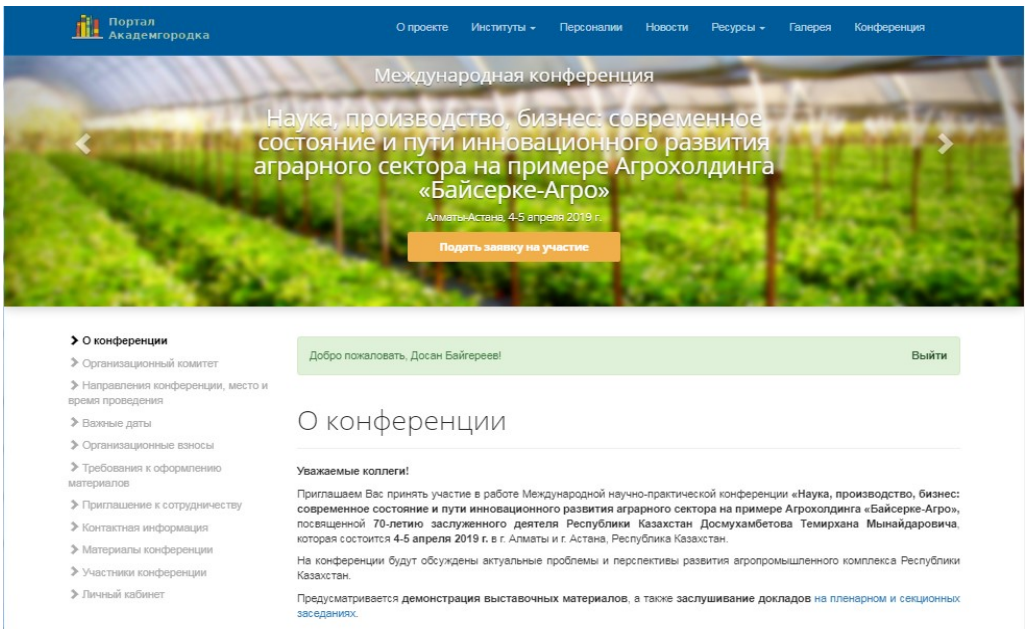

Рисунок 28 - Главная страница модуля конференций

Модуль успешно прошел апробацию на следующих конференциях:

1. На международной конференции «Наука, производство, бизнес: современное состояние и пути инновационного развития аграрного сектора на примере Агрохолдинга «Байсерке-Агро», проведенной 4-5 апреля 2019 года в г. Алматы https://acagor.kz/conference/baiserke-agro. Всего на конференцию поступило 347 работ, из них 299 приняты, 36 отклонены, в 22 случаях подача заявки не завершена.

2. На международной научно-практической конференции «Computational and Information Technologies in Science, Engineering and Education» (CITech–2020), посвященная 1150-летию аль-Фараби 9-10 октября в г. Алматы https://acagor.kz/conference/citech-2020. Всего на конференцию поступило 82 работ, из них 55 приняты, 22 отклонены, отправлено на доработку 5, 43 работы приняты для опубликования в совместном выпуске журналов "Вестник НИА РК" (входит в перечень КОКСОН РК) и "Вычислительные технологии" (входит в перечень ВАК РФ).

Схема данных модуля приведена на Рисунке Е.4. Модуль может использоваться для любого количества конференций, причем для каждой конференции может быть задана различная конфигурация (правила приема докладов, уровни доступа и др.).

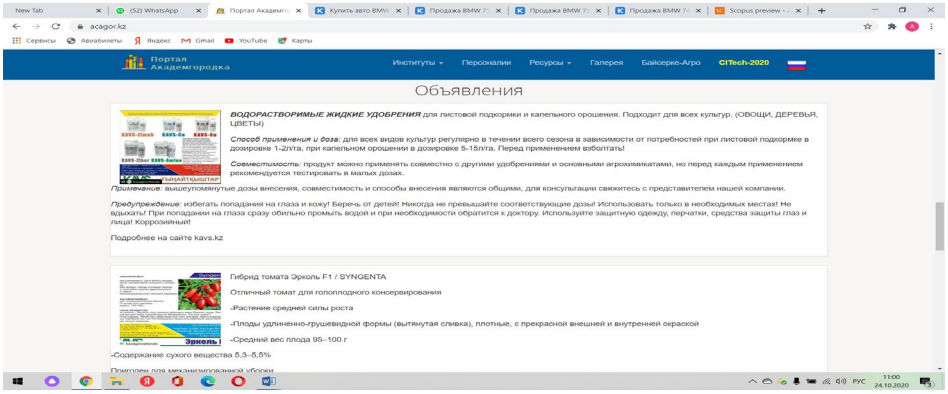

Рисунок 29 - Модуль размещение объявления

Система позволяет сохранить результаты интеллектуальной деятельности научно-исследовательских институтов в актуальном виде и предоставлять к ним доступ на основе Web-технологий. Созданный портал соответствует стратегии «Индустрия 4.0» целью которого является трансферт новых технологий в агропромышленный комплекс республики Казахстан и поиск практических инновационных решений для устранения факторов, ограничивающих развитие сельского хозяйства в Казахстане. На портале разработан модуль размещение информации новостного, делового и рекламного характера и возможности обмена информацией (Рисунок 29).

5 СОЗДАНИЕ ПОЛНОФУНКЦИОНАЛЬНОГО ПРОТОТИПА РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ НА БАЗЕ ОГРАНИЧЕННОГО ЧИСЛА УЧАСТНИКОВ

На доменном имени https://acagor.kz/ действует полнофункциональный прототип распределенной информационной системы, обеспечивающий идентификацию и управление ресурсов, идентификацию, аутентификацию и авторизацию пользователей, сбор статистики, предоставление различной информации. В 2018 году программное обеспечение, положенное в основу подсистем информационной системы, было установлено на сервере, расположенном в дата-центре ahost компании «Академсеть». Дата-центр оборудован по стандарту TIER-2, имеет шесть гермозон площадью 200 м^2, мощность энергоузла дата-центра составляет 630 кВА, 450 кВт. Дата-центр оснащен:

– 4 источниками бесперебойного и резервного энергоснабжения, каждый по 160 кВА/128 кВт, схема резервирования 2N + 1;

– системой гарантированного энергопитания: дизель-генератор 825 кВА/660 кВт, подогрев топливного бака, подогрев топливной системы, подогрев антифриза;

– системой охлаждения;

– системой контроля доступа;

– системой противопожарной и информационной безопасности.

В рамках выполнения задач научно-исследовательской работы, в 2019 году закуплен сервер со следующими техническими характеристиками, обеспечивающий хранение постоянно пополняющейся базы данных и бесперебойную работу фронт-интерфейса, публичного доступа к базе данных: Сервер 2x - процессорный стоечный Supermicro CPU 2xE5-2650V4 2.2 GHz / Fan 2x1U / DDR4 4x8GB 2666MHz / MB 8xDDR4 - KVM - 10xSATA - RAID SW 0,1,5,10 / Контроллер RAID 8-int / Корпус стоечный 1U 4xHS 2PSU / SSD 2x480GB / Опции для HDD/SSD 2xMCP / Net 2xSFP 10Gb. Данный сервер также размещен в дата-центре ahost, где на безвозмездной основе в защищенных условиях обеспечен доступом к сети Интернет и технологическим микроклиматом. Характеристики сервера обеспечивают его актуальность на достаточно длительную перспективу даже с учетом возможной смены поколений программного обеспечения. Кроме того, структура портала позволяет масштабировать и совершенствовать аппаратные средства что даст возможность при необходимости существенно расширить вычислительные мощности портала.Фронт-интерфейс ограниченного публичного доступа к базе данных построен на базе портала (сайта) с доменным именем https://acagor.kz/.

Учитывая значимость оцифрованной информации и её возможный уровень секретности, доступ к накопленным данным строго регламентируется владельцами информации; в данное время, это администрация НИИ, участников проекта. В то же время, внутренняя, служебная часть портала обеспечивает связь рабочих мест в НИИ с сервером посредством VPN канала.

6 ОЦИФРОВКА НАИБОЛЕЕ ВАЖНЫХ НАУЧНЫХ ТРУДОВ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ ИНСТИТУТОВ

Согласно шестому пункту календарного плана производится оцифровка и загрузка в репозиторий цифровых объектов наиболее важных научных трудов сотрудников НИИ общей генетики и цитологии (далее – ИОГиЦ), НИИ физиологии человека и животных (далее – ИФЧЖ), НИИ перерабатывающей и пищевой промышленности (далее – КазНИИ ППП), НИИ плодоовощеводства (далее – КазНИИПО), НИИ почвоведения и агрохимии имени У. Успанова (далее – КазНИИ ПиА), Казахстанского инженерно-технологического университета (далее – КазИТУ) и НИИ микробиологии и вирусологии (далее – НИИ МиВ). Кроме того производится загрузка статей в журналах, имеющих ненулевой импакт-фактор, материалы международных конференций, монографии и отчеты научно-исследовательских работ.

На момент подачи отчета (по состоянию на 24 октября 2020 года) в репозитории хранится 1011 объекта.

В таблице 1 приведена статистика по наполненности репозитория цифровых объектов по типам ресурсов.

Производится оцифровка более ранних, не имеющих цифровых копий результатов научной деятельности указанных институтов, оцифровка карт, а также подготовка аудио- и видеоматериалов.

Таблица 1 - Статистика по типам ресурсов

Тип ресурса	ИОГиЦ	ИФЧЖ	КазНИИ ППП	КазНИИПО	КазНИИ ПиА	КазИТУ	НИИ МиВ	Итого
Статьи	57	72	66	42	74	171	78	560
Монографии	45	38	59	29	53	84	136	444
Отчеты НИР	-	-	2	-	-	5	-	7
Итого	102	110	127	71	127	260	214	1011

В таблице 2 приведена статистика по годам публикаций. Согласно статистике, наиболее ранняя работа датируется 1898 годом, наиболее поздняя – 2020 годом.

Таблица 2 - Статистика по годам публикаций

Период публикаций	Количество публикаций
1898 – 1899	1
1900 – 1999	159
2000 – 2020	851

В таблице 3 приведена статистика по языкам публикации.

Таблица 3 - Статистика по языкам публикации

Язык публикации	ИОГиЦ	ИфЧЖ	КазНИИ ППП	КазНИИ ПиВ	КазНИИ ПиА	КазИТУ	НИИ МиВ	Итого
английский	10	7	24	2	10	26	1	80
казахский	2	2	3	4	2	5	2	20
русский	90	101	100	65	115	229	211	911
Итого	102	110	127	71	127	260	214	1011

ЗАКЛЮЧЕНИЕ

В результате выполнения научно-исследовательской работы получены следующие результаты:

1. Проведен анализ потребностей и возможностей научно-исследовательских институтов и КазИТУ расположенных в Академгородке г. Алматы.
2. Определена оптимальная архитектура распределенной информационной системы, приведены описание ее подсистем и схема их взаимодействия.
3. Разработана и реализована распределенная библиотечная информационная система научно-образовательного кластера НИИ и КазИТУ, расположенных в Академгородке г. Алматы. С помощью технологии WiMAX оцифрованные наиболее важные научные труды НИИ и КазИТУ, расположенных в Академгородке г. Алматы сразу поступают на сервер баз данных. Для этого организовано прямое соединение НИИ и ЦОД ТОО «Академсеть».
4. Рассмотрены вопросы использование ресурсов портала как многопроцессорный вычислительный кластер. С помощью которого будут решатся различные производственные, природопользовательские, экономические и социальные задачи.
5. Разработана модель функционирования распределенной информационной системы.
6. Произведен отбор и адаптация необходимого открытого и свободного программного обеспечения – компонентов подсистем распределенной информационной системы.
7. Проведен анализ институциональных репозиториев и систем для создания электронных библиотек, изучен опыт их использования другими организациями. В результате выбрана и установлена распределенная информационная система DSpace, произведена модификация и адаптация DSpace к условиям, установленным в Республике Казахстан.
8. Произведена установка системы управления научными исследованиями DSpace-CRIS.
9. Выбрано интегрирующее программное обеспечение – распределенная информационная система ZooSPACE, производится адаптация данной системы к распределенной информационной системе Академгородка.
10. Разработана архитектура портала Академгородка, схема его взаимодействия с остальными составляющими распределенной информационной системы, произведено описание программного обеспечения, использованного для его функционирования; произведено описание разделов портала.
11. Осуществлена интеграция портала Академгородка с репозиторием цифровых объектов с использованием программного интерфейса DSpace REST API.
12. Разработаны собственные модули для интеграции с другими

внутренними подсистемами.

13. Проведено наполнение РИС актуальной информацией и информацией, созданной в процессе оцифровки документальных источников.

14. По результатам научных исследований в 2018, 2019, 2020 годах опубликованы статьи в научных журналах «Вычислительные технологии» (Россия), Вестник Казну им.аль-Фараби (Казахстан), Przeglad Elektrotechniczny (Польша), International Journal of Electronics and Telecommunications (Польша).

15. Получено авторское свидетельство на созданную распределенную информационную систему научно-исследовательских институтов Академгородка города Алматы выданное РГП «Национальный институт интеллектуальной собственности» МЮ РК.

Результаты, полученные в рамках данной научно-исследовательской работы, могут быть использованы при построении распределенных интегрированных информационных систем, используемых в научных и образовательных организациях. Результаты могут быть коммерциализированы посредством предоставления сервисов: благодаря модульности системы, ее функциональность может быть расширена для представления платных услуг, таких как заказ на оцифровку редкого контента, размещение объявлений.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Ильин В. П. Как реорганизовать вычислительные науки и технологии// Вестник Российской Академии наук. - 2019. – Т. 89, № 3. – С. 232–242.

2 РКС и "РУСАГРО"объединяют усилия для цифровизации сельского хозяйства России [Электрон. ресурс]. - 2017. - URL: https://www.roscosmos.ru/24458/ (дата обращения 15.08.2020).

3 Агросигнал - система ГЛОНАСС/GPS контроля спецтехники и управления агро-бизнесом [Электрон. ресурс]. 2019. - URL: https://7gis.ru/podderzhka/platformyi/agrosignal.html (дата обращения 15.08.2020).

4 Saini O. P. The emergence of institutional repositories: a conceptual understanding of key issues through review of literature // Library Philosophy and Practice. – 2018. – Vol. 1774. – 19 p.

5 Franchke H., Gamalielsson J., Lundell B. Institutional repositories as infrastructures for long-term preservations // Information Research. – 2016. – Vol. 22 (2), No. 757. – 27 p.

6 Баранов А. В., Киселёв Е. А. Облачные сервисы для научных высокопроизводительных вычислений на базе платформы Proxmox// Вычислительные технологии. – 2019. – Т. 24. – №6. – С. 5-12.

7 Такабаев Т.М., Калимолдаев М.Н., Гречко С.М., Амирханова Г.А. Облачное развитие IT ресурсов МОН и их коммерческая эксплуатация // Международная научно практическая конференция «Қоғамды ақпараттандыру» III Халықаралық ғылыми-практикалық конференция – 2012. – С. 178-180.

8 Живая Кубань интернет-дневник Краснодарского края [Электрон. ресурс]. - 2020. - URL: livekuban.ru/news/obshchestvo/skolko-budet-dlitsya-epidemiya-koronavirusa/(дата обращения 15.08.2020).

9 Hippenhammer C. Comparing institutional repository software: pampering metadata uploaders // The Christian Librarian. – 2016. – Vol. 59, No. 1 – 6 p.

10 Baughman K., McDowell. Institutional repositories in the Czech republic // Gleeson Library Librarians Research. – 2016. – Vol. 10. – 29 p.

11 Ravikumar M. N., Ramanan T. Comparison of greenstone digital library and DSpace: Experiences from digital library initiatives at eastern university, Sri Lanka // Journal of University Librarians Association of Sri Lanka. – 2014. – Vol. 18, No. 2. – P. 76–90.

12 Castagné M. Institutional repository software comparison: DSpace, ePrints, Digital Commons, Islandora and Hydra (Report). – University of British Columbia, 2013. – 15 p.

13 Cullen R., Chawner B. Institutional repositories in New Zealand: comparing institutional strategies for digital preservation and discovery // Proceedings of the IATUL Conference. – 2008. – Vol. 18. – 11 p.

14 Скачков Д. М., Жижимов О. Л. Технология географического поиска информации в ”негеографических” информационных системах // Научные и организационно-технологические основы интеграции цифровых информационных ресурсов. – 2013. – С. 74-101.

15 Шокин Ю. И., Жижимов О. Л., Федотов А. М. Информационные

системы ИВТ СО РАН // Труды XVI Всероссийской конференции DICR-2017. – 2017. – С. 11-18.

16 Chudlarsky T., Dvorak J. A National CRIS Infrastructure as the Cornerstone of Transparency in the Research Domain. In: Jeffery, Keith G; Dvorak, Jan (eds.): E-Infrastructures for Research and Innovation: Linking Information Systems to Improve Scientific Knowledge Production: Proceedings of the 11th International Conference on Current Research Information Systems (June 6-9, 2012, Prague, Czech Republic). – 2012. - P. 9-17.

17 Guskov A.E., Zhizhimov O.L., Kikhtenko V., Skachkov D.M., Kosyakov D. RuCRIS: A Pilot CERIF based System to Aggregate Heterogeneous Data of Russian Research Projects // Procedia Computer Science. - 2014. - Vol.33. - P.163-167.

18 Жижимов О.Л., Федотов А.М., Федотова О.А. Построение типовой модели информационной системы для работы с документами по научному наследию // Вест. НГУ. Информационные технологии. - 2012. - Т.10, №3. - С.5-14.

19 Шокин Ю.И., Федотов А.М., Жижимов О.Л., Федотова О.А. Система управления электронными библиотеками в ИРИС СО РАН // Инфраструктура научных информационных ресурсов и систем: Сборник научных статей Четвертого Всероссийского симпозиума / Под редакцией доктора технических наук Е.Б. Кудашева, доктора физико-математических наук В.А. Серебрякова. - 2014. - Т.1. - Москва: Вычислительный центр РАН. - С.11-39.

20 Жижимов О.Л., Федотов А.М., Шокин Ю.И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2013. - Т.11. - № 1. - С.24-41.

21 Шокин Ю. И., Темирбеков Н. М., Жижимов О. Л., Темирбеков А. Н., Байгереев Д. Р. Модель интегрированной распределенной библиотечной информационной системы Академгородка города Алматы // Вычислительные технологии. - 2018. - Т. 23. - №5. - С. 110-119.

22 Temirbekov N., Baigereyev D., Temirbekov A., Smolarz A. Architecture of the distributed information system of the Almaty Academgorodok // Przeglad Elektrotechniczny. - 2019. - No. 5. - P. 75-78.

23 Темирбеков Н. М., Жижимов О. Л., Байгереев Д. Р., Омиржанова Б. Б., Темирбеков А. Н., Оразгалиева А. М. Построение научно-образовательного кластера на базе интегрированной распределенной информационной системы Академгородка // Совместный выпуск журналов "Вестник ВКГТУ им. Д. Серикбаева"и "Вычислительные технологии". - 2018. - Т. 1., №3., Ч. 1. - С. 120-127.

24 Темирбеков Н. М., Байгереев Д. Р., Темирбеков А. Н. Облачные технологии для организации научно-образовательного агропромышленного кластера Академгородка города Алматы // Вестник КазНУ им. аль-Фараби, серия «математика, механика, информатика», №3(103), 2019 г., -С. 75-89.

25 Temirbekov, N., Baigereyev, D., Temirbekov, A., Omirzhanova, B. Development of an information system for storing digitized works of the Almaty

Academgorodok research institutes(Conference Paper). AIP Conference Proceedings. Volume 2183, Article number 080005.-2019.

26 Temirbekov N.,Baigereyev D., Temirbekov A., Omirzhanova B. Development of a Distributed Information System of the Almaty Academgorodok. Journal of Engineering Science and Technology Review Special Issue on Telecommunications, Informatics, Energy and Management 2019.-p.127-130.-2020.

ПРИЛОЖЕНИЕ А
Схема данных и административная часть модуля конференции

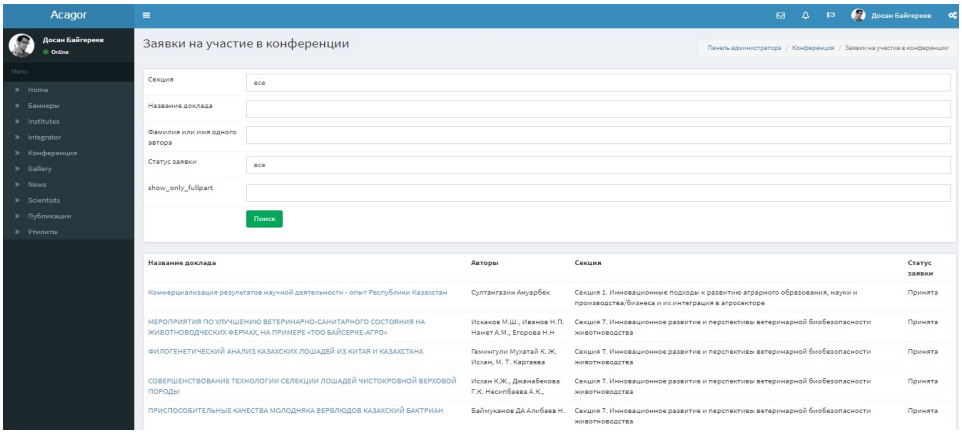

Рисунок Е.1 - Административная часть модуля конференции: список заявок

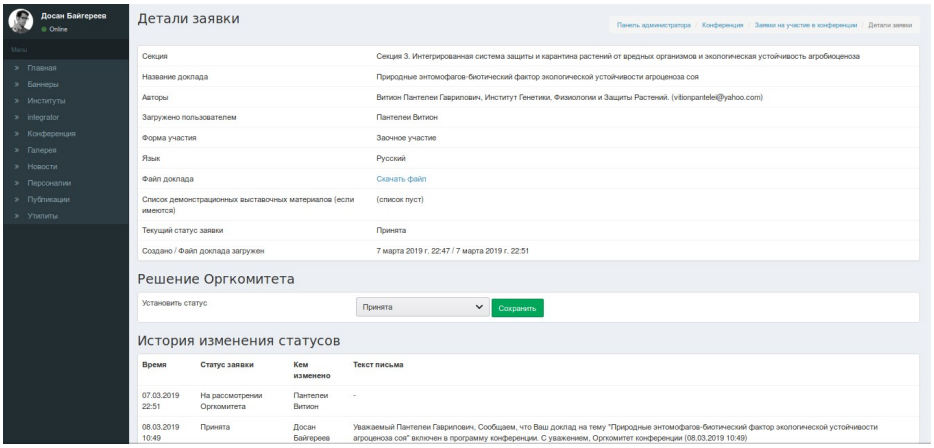

Рисунок Е.2 - Административная часть модуля конференции: детали заявки

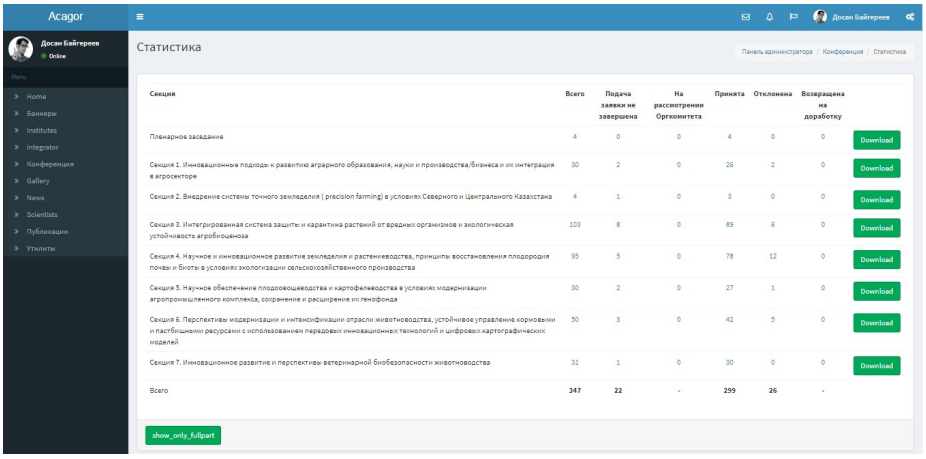

Рисунок Е.3 - Административная часть модуля конференции: статистика заявок

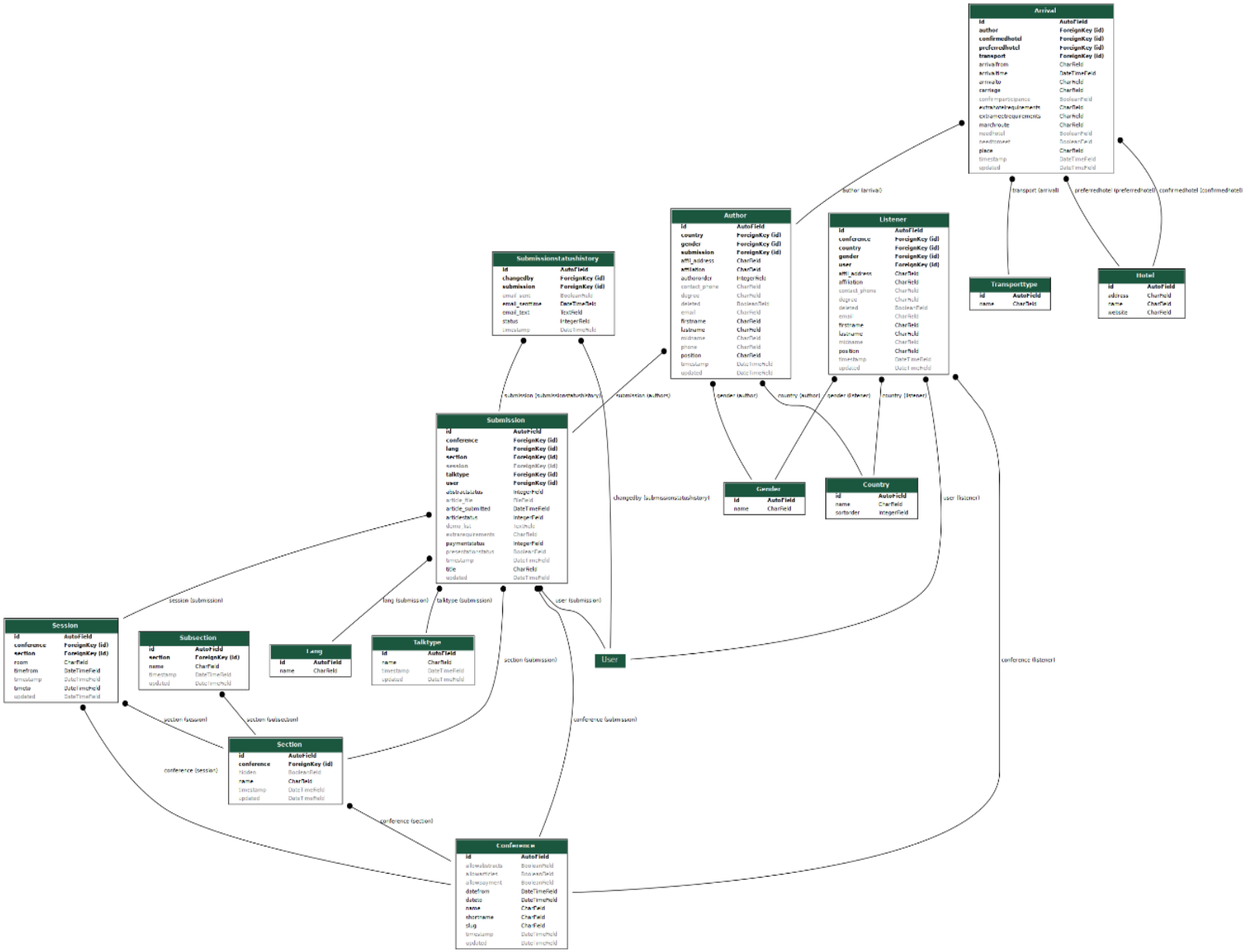

Рисунок Е.4 - Схема данных модуля конференции