

Построение научно-образовательного кластера на базе интегрированной распределенной информационной системы Академгородка

Н. М. Темирбеков¹, О. Л. Жижимов², Д. Р. Байгереев³,
Б. Б. Омиржанова⁴, А. Н. Темирбеков⁵, А. М. Оразгалиева¹

¹ Казахстанский инженерно-технологический университет,
пр. Аль-Фараби 93 А, 050060 Алматы, Казахстан

² Институт вычислительных технологий СО РАН,
пр. Академика Лаврентьева, 6, 630090 Новосибирск, Россия

³ Восточно-Казахстанский государственный технический университет
им. Д. Серикбаева, ул. Серикбаева, 19, 070004 Усть-Каменогорск, Казахстан

⁴ КазНИИ перерабатывающей и пищевой промышленности,
пр. Гагарина, 238 Г, 050060 Алматы, Казахстан

⁵ Казахский национальный университет им. аль-Фараби,
пр. аль-Фараби, 71, 050040 Алматы, Казахстан
temirbekov@rambler.ru, zhizhim@mail.ru, dbaigereyev@gmail.com,
omirzhanova61@mail.ru, almas_tem@mail.ru, aorazgalieva@bk.ru

Аннотация В работе рассматриваются вопросы создания технологической модели интегрированной распределенной библиотечной информационной системы, объединяющей оцифровываемый книжный фонд и научные труды Казахстанского инженерно-технологического университета и ряда НИИ, расположенных в Академгородке г. Алматы (Казахский научно-исследовательский институт почвоведения и агрохимии имени У. У. Успанова, Казахский НИИ перерабатывающей и пищевой промышленности, Казахский НИИ плодоводства и виноградарства, НИИ микробиологии и вирусологии, НИИ физиологии человека и животных, НИИ генетики и цитологии), работающей на базе их единой облачной IT-платформы.

Ключевые слова: научно-образовательный кластер, распределенная информационная библиотечная система, Академгородок

В развитых странах информация стала важнейшим объектом производственной деятельности и потребления, поскольку она оказывает значительное или даже решающее воздействие на направления и результаты прогресса в научной, технической, культурной, экономической и других сферах жизни этих стран. Сказанное в полной мере относится также к деятельности отдельных организаций, фирм, учебных заведений, административных органов управления и физических лиц, являющихся производителями и потребителями разнородной информации [2,2,3]. Поэтому информация рассматривается как один из наиболее значимых ресурсов человеческого сообщества

и в государственном плане рассматривается как стратегический ресурс. Сохранение, рациональное использование и развитие этого ресурса является задачей огромного значения для любого сообщества, государства и организации.

В то же время гигантские объемы уже накопленной информации, непрерывно продолжающийся рост ее количества, разнородный и разобченный по многим признакам характер хранения и распространения, отсутствие унифицированного доступа к ней создают существенные и все возрастающие проблемы ее эффективного использования.

Осознание указанных проблем, а также качественные изменения в области развития современных информационных технологий и средств передачи данных привели к необходимости поиска новых подходов и решения проблем создания хранилищ информационных ресурсов, их организации, средств и способов доступа пользователей к ним. Такие подходы сегодня стали трактовать как создание «цифровых» или «электронных» библиотек. В дополнение к информационному обслуживанию на печатных носителях приходит обеспечение пользователей, основанное на электронном представлении самой разнообразной информации, тиражируемой в неограниченном количестве и оперативно доступной по глобальным компьютерным сетям, независимо от времени обращения к ней и местонахождения пользователей.

Впечатляющие успехи в развитии средств вычислительной техники и телекоммуникаций в последние десятилетия привели к существенным изменениям во всех областях человеческой деятельности, которые в совокупности принято относить к важным признакам постиндустриального или информационного общества. Эти изменения самым существенным образом влияют на развитие технических и программных средств, использующих в библиотеках и информационных органах, равно как и тех технологий, которые связаны с их деятельностью. Вполне очевиден тот факт, что проектировщики средств автоматизации библиотечной и информационной деятельности конкретных организаций опираются на реально существующие средства. Однако также ясно, что без учета перспективы развития этих средств и технологий, основанных на их использовании, проектировать новые системы нельзя. Последнее обстоятельство тем более важно в такой динамично развивающейся области, как информационные технологии XXI века.

Дальнейшее развитие научных основ для построения интегрированных распределенных информационных систем соответствует целям, поставленным в рамках Государственной программы «Цифровой Казахстан» на 2017-2020 годы, и нацелена на выполнение следующих прикладных задач:

- развитие креативного общества (развитие человеческого капитала путем повышения цифровой грамотности населения, повышение квалификации специалистов в области инфокоммуникационных технологий, развитие креативного мышления и т.д.);
- цифровые преобразования в отраслях экономики (развитие цифровой индустрии путем, автоматизации транспортно-логистической системы стра-

ны; внедрения цифровых технологий в сфере сельского хозяйства; пищевой промышленности).

«Наука – цифровой экономике!» Под таким девизом во всем мире разрабатываются новые информационные технологии и системы, интеграция ресурсов и обеспечение информационной безопасности. Эти проблемы становятся особенно актуальными в свете решения задач построения цифровой экономики.

Основным направлением 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.

Одним из пионеров и, как следствие, одним из мировых лидеров в развитии научно-образовательной информационной системы и облака, является «Электронная библиотека Сибирского отделения РАН», который сегодня называется Порталом СО РАН (<http://www.sbras.ru/>). По данным Webometrics сайт занимает первое место в России среди научных организаций и 45-е место в мире. Автор этой электронной библиотеки является сотрудником ИВТ СО РАН [1,5,6,7,8].

Существует достаточно много технологических разработок ИС для электронных библиотек, например, Euro CRIS (<http://www.eurocris.org/>), Информика (<http://www.informika.ru/>), eLibrary (<http://elibrary.ru/>), MathNET (<http://www.mathnet.ru/>).

В Казахстане создана Республиканская межвузовская электронная библиотека (РМЭБ) на базе университетских образовательных ресурсов по единой технологии, предложенной Ассоциацией вузов РК и научно-образовательной компьютерной сетью Казахстана KAZRENA.

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.

Создаваемая авторами РИС будет полностью соответствовать потребностям участников научно-образовательного кластера как по информационному содержанию, так и по поддержке отраслевой и языковой специфике. Отличие от существующих проектов заключается в том, что для создаваемой ИС ставится требование обеспечения ссылочной целостности данных и соответствие международному стандарту ISO-1472 (Open Archive information System - OAIS) [9]. Это очень жесткое требование, которое тяжело соблюсти даже в хорошо формализованных СУБД. Жесткие гиперссылки будут заменены ассоциативными связями между документами.

С помощью предложенной технологической модели СУЭБ будет создана РБИС, позволяющая хранить все созданные ресурсы НИИ в актуальном виде и поддерживать к ним защищенный в силу закрытости сети доступ. РБИС будет предоставлять следующие типы информационных ресурсов:

- географические материалы (карты, спутниковые снимки, данные полевых наблюдений и т.п.), а также соответствующие базы метаданных;
- библиографические базы данных и электронные каталоги;
- полнотекстовые базы данных и цифровые репозитории;
- авторитетные базы данных (словари, справочники и т.д.);
- другие ресурсы (аудио и видеозаписи, электронные презентации и др.).

Все цифровые объекты представляют собой результат интеллектуальной деятельности НИИ почвоведения и агрохимии им. Успанова, НИИ перерабатывающей и пищевой промышленности, НИИ плодоводства и виноградарства, НИИ микробиологии и вирусологии, НИИ физиологии человека и животных, НИИ генетики и цитологии, образовательного учреждения КазИТУ, расположенные в Академгородке г.Алматы. Контентное наполнение распределенной библиотечной информационной системы – оцифрованная база знаний научно-образовательного агропромышленного кластера.

Требования к модели такой информационной системы обусловлены двумя принципами:

- информационные потребности научных работников и представителей бизнеса, экономики;
- обеспечение надежного и долговременного хранения информации.

Опыт разработки и эксплуатации информационных систем выдвигает следующие основные требования к подобным системам [6]:

1. системы должны разрабатываться на основе свободно распространяемого программного обеспечения;
2. системы должны иметь возможность импортировать информацию из других систем;
3. системы должны предоставлять информацию в соответствии со стандартными протоколами, схемами и форматами;
4. системы должны интегрировать информацию, специфичную для CRIS-систем (Current Research Information System) [7,10];
5. системы должны иметь полный комплект административных и пользовательских интерфейсов, желательно на основе WEB-технологий;
6. системы должны разграничивать доступ к хранимой информации;
7. системы должны иметь возможность расширения как по номенклатуре хранимых данных, так и по номенклатуре внешних источников информации.

Для реализации этих требований проще всего использовать проверенные временем технологические подходы, основанные на использовании свободно распространяемого программного обеспечения и на стандартных сетевых

протоколах. При этом речь может идти не об одной информационной системе, а о целом комплексе относительно самостоятельных информационных систем, взаимодействующих между собой в соответствии с фиксированными правилами (протоколами).

На рисунке 1 представлено архитектурное решение ИС, на рисунке 2 – ее технологическая модель.



Рисунок 1. Структура библиотечной информационной системы.

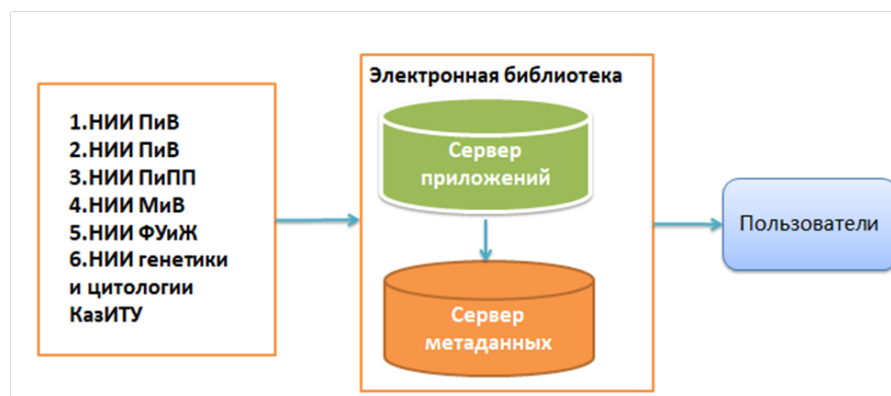


Рисунок 2. Архитектура электронной библиотеки.

РИС будет содержать следующие подсистемы:

- Подсистема – репозиторий цифровых объектов, предоставляющий пользовательские и административные WEB- интерфейсы для доступа к цифровым объектам и коллекциям, а также интерфейсы интеграции с другими подсистемами на основе открытых международных стандартов (Рисунок 3).
- Подсистема управления текущей научно-исследовательской информацией (CRIS), включающей информацию о публикациях сотрудников, их участии в конференциях и в исполнении исследовательских проектов. Подсистема будет включать пользовательские и административные WEB-интерфейсы, а также интерфейсы интеграции с другими подсистемами на основе открытых международных стандартов (Рисунок 4).
- Подсистема интеграции распределенных информационных ресурсов на основе технологий Z39.50/ SRU /SRW.
- Подсистема доступа к распределенным информационным ресурсам на основе WEB-технологий (WEB-портал) (Рисунок 5).

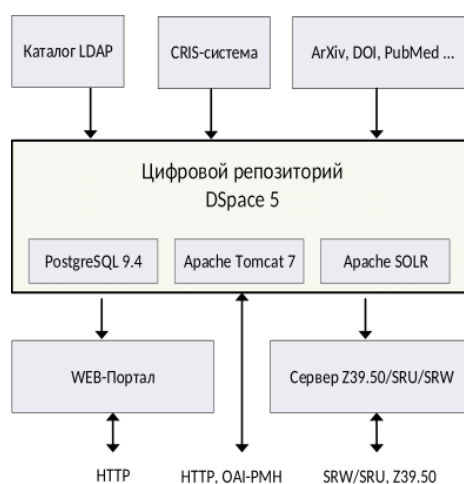


Рисунок 3. Цифровой репозиторий.

Указанные подсистемы в совокупности должны обеспечить:

- идентификацию информационных ресурсов;
- идентификацию, аутентификацию и авторизацию пользователей;
- управление метаданными;
- управление информационными ресурсами;
- сбора статистики;
- мониторинг доступности сервисов и ресурсов.

Реализация подсистем ИС должна основываться на открытых спецификациях, связанных с международными стандартами и рекомендациями.

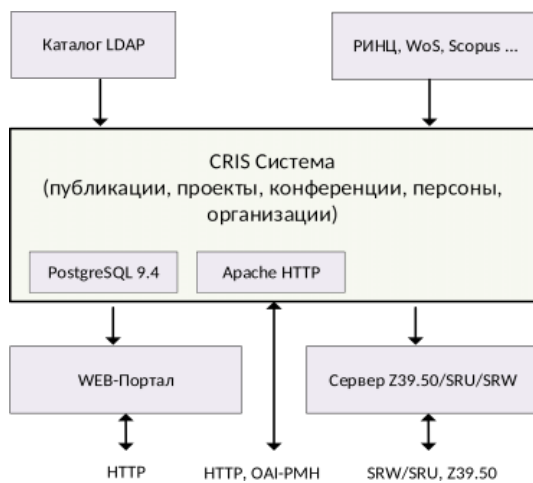


Рисунок 4. Подсистема CRIS.

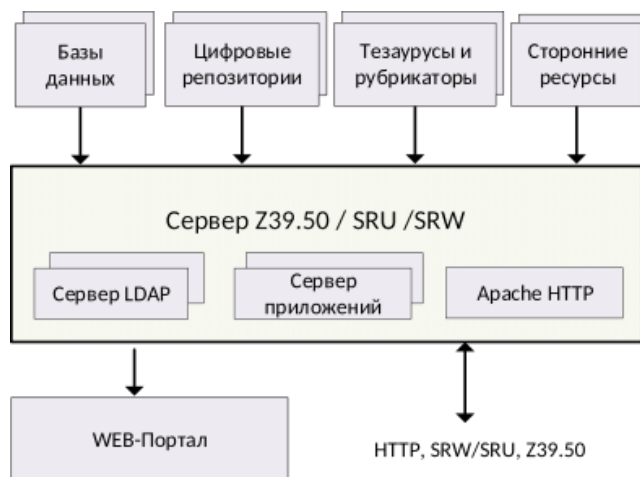


Рисунок 5. Подсистема Z39.50/SRW/SRU.

Создаваемая 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. Создаваемая система в полной мере обеспечивает необходимыми вычислительными ресурсами проводимые научные исследования и образовательные процессы, упрощая перспективу дальнейшего ее развития. Информационная система позволит построить передовую IT-инфраструктуру, управление интеллектуальным капиталом, электронной библиотекой, в которой будут храниться все книги и научные труды КазИТУ и НИИ.

Список литературы

1. Темирбеков, Н.М., Балова, Т.Г., Мокеров, В.О.: Проектирование сервисов портала вуза для интеграции информационных ресурсов образовательных программ. Информатика в образовании, 87–91 (2012)
2. Жумагулов Б.Т., Темирбеков Н.М., Денисова Н.Ф., Нугуманова А.Г., Архипова О.Г., Турганбаев Е.М.: Разработка информационной образовательной статистики Республики Казахстан. Вестник НИА РК 44, 75–80 (2012)
3. Temirbekov, N.M., Tokanova, S., Malgazhdarov, Ye.: Information technology for numerical simulation of viscous incompressible flow in biconnected domains. Journal of Theoretical and Applied Information Technology 88, 441–448 (2016)
4. Шокин, Ю.И., Федотов, А.М., Жижимов, О.Л.: Технологии создания распределенных информационных систем для поддержки научных исследований. Вычислительные технологии 20, 251–274 (2015)
5. 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 10, 5–14 (2012)
6. 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 1, 11–39 (2014)
7. Guskov, A.E., Zhizhimov, O.L., Kikhtenko, V., Skachkov, D.M., Kosyakov, D.: A Pilot CERIF based System to Aggregate Heterogeneous Data of Russian Research Projects. Procedia Computer Science 33, 163–167 (2014)
8. Жижимов, О.Л., Федотов, А.М., Шокин, Ю.И.: Технологическая платформа массовой интеграции гетерогенных данных. Вестник Новосибирского государственного университета. Серия: Информационные технологии 11, 24–41 (2013)
9. ISO – 14721 Reference Mode for an Open Archival Information System (OAIS): Recommended Practicle: CCSDS 650.0-M-2 (Magenta Book) (2012) <http://public.ccsds.org/publications/archive/650x0m2.pdf>
10. 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, 9–17 (1981)