

УДК 55:681.3(571.6)

© В.В. Наумова, И.Н. Горячев, К.А. Платонов

В.В. Наумова, И.Н. Горячев, К.А. Платонов

WEB-ИНТЕГРАЦИЯ НЕОДНОРОДНЫХ НАУЧНЫХ ДАННЫХ И СЕРВИСОВ ПО ГЕОЛОГИИ ДАЛЬНЕГО ВОСТОКА РОССИИ НА ОСНОВЕ ПОРТАЛЬНОГО РЕШЕНИЯ

Территория Дальнего Востока России характеризуется высокой степенью геологической изученности. Здесь выявлено большое количество месторождений различных полезных ископаемых, в том числе уникальных по масштабам. Достижения Российской геологической науки, полученные в результате многолетних разноплановых исследований, являются существенным вкладом в мировой процесс изучения геологического строения, геологической эволюции и металлогении планеты Земля.

Многолетние исследования ученых из институтов Дальневосточного отделения РАН позволили собрать огромную информацию по геологии и геофизике Дальнего Востока. В институтах полученные данные систематизируются (Наумова, 2006). Создаются архивы и базы данных, ГИС, информационно-поисковые системы.

Наиболее крупными из них являются:

1. ГИС «Минеральные ресурсы, металлогенез и тектоника Северо-Восточной Азии». Система осуществляет хранение, обработку, доступ, отображение и распространение картографической и атрибутивной информации о геологических объектах Восточной и Южной Сибири, юга Дальнего Востока России, Монголии, Северо-Восточного Китая, Кореи и Японии (Наумова и др., 2006).
2. ГИС по геологии и полезным ископаемым Верхне-Колымского региона (Ворошин и др., 2006). Территория системы охватывает одну из крупных золоторудных провинций России.
3. ГИС «Геолого-геофизический атлас Области сочленения Центрально-Азиатского и Тихоокеанского складчатых поясов масштаба 1:5 000 000» (ИТИГ).
4. ГИС «Позднекайнозойские подводные вулканы Тихого океана» (Рашидов и др.).

Благодаря новым методам сбора данных неуклонно растет их объем, повышается оперативность их получения, завершается переход на качественно новые, цифровые технологии сбора, обработки, распространения и использования данных. Для получения исходных данных используются системы

дистанционного зондирования Земли из космоса, цифровые системы наземного и воздушного лазерного сканирования, другие цифровые и электронные геодезические приборы, цифровые аэросъемочные камеры, глобальные навигационные спутниковые системы (ГНСС) GPS/ГЛОНАСС. Новая цифровая и электронная среда существования геологических данных создает условия для использования современных информационных технологий.

Часть этой информации доступна пользователям Интернет. Информация очень фрагментирована и располагается:

1. На сайтах институтов Дальневосточного отделения РАН.
2. На сайтах научных журналов.
3. В электронных библиотеках.
4. В базах данных.
5. В ГИС и электронных картах.

В Интернет существуют стандартные методы поиска информации с использованием поисковых систем, таких как Google, Yandex, Bing, Yahoo, и др. Но, несмотря на разные алгоритмы поиска и предоставления информации в этих системах, основная идея во всех системах одна. Пользователь вводит ключевые фразы, и поисковая машина выдает список Интернет страниц, на которых наиболее часто встречаются искомые слова.

Поисковые машины используют для ранжирования сайтов рейтинговые оценки (Ханчук, Наумова, 2009). Наличие высоких рейтинговых оценок для сайта очень важно с точки зрения его более высокой доступности широкой аудитории пользователей Интернет, поскольку рейтинговые оценки используют практически все поисковые системы. Наличие рейтинговых систем для сайтов, разрабатываемых и применяемых крупными поисковиками: Google, Yandex и др. при выдаче результатов поиска пользователям Интернет позволяет пользователям всего мира быстрее находить наиболее качественную и отвечающую запросу информацию. Результаты запросов пользователей сортируются поисковыми машинами и предоставляются пользователям в порядке уменьшения их рейтингов. Таким образом, наличие

высоких рейтингов ставит сайт в лучшее положение по отношению к другим сайтам.

Создателями Google была предложена мера популярности Web-страницы, не зависящая от запроса пользователя, которая называется PageRank. Можно придерживаться примерно такой градации: PageRank от 4 до 5 – наиболее типичный для большинства сайтов средней «раскрученности». 6 – очень хорошо «раскрученный» сайт. 7 – величина, практически недостижимая для множества сайтов, но иногда встречается. Значения 8, 9, 10 имеют исключительно популярные и значимые проекты. Например, в данный момент у сайта русской Википедии PR = 8, у английской Википедии, и у сайта Microsoft – 9. Значение 10 имеют всего несколько десятков сайтов. В их числе Google. Google использует показатель PageRank найденных по запросу страниц, чтобы определить порядок выдачи этих страниц посетителю в результатах поиска.

Нами были проведены различные вебометрические исследования в российском сегменте глобальной сети, а именно в ее Дальневосточном научном сегменте (Ханчук, Наумова, 2009). Геологические ресурсы Дальневосточного отделения РАН, исключая Информационный сервер Дальневосточного геологического института ДВО РАН имеют низкие значения рейтингов, а некоторые ресурсы имеют нулевые рейтинги.

Таким образом, поиск информации по геологии и геофизике Дальнего Востока России в Интернет по сайтам ДВО РАН затруднен, поскольку общедоступные поисковые серверы, и в частности Google, не выводят ссылки на их информацию на страницы результатов поисковых запросов.

Связность научного геологического Интернет-пространства России минимальна. Некоторый анализ связности структур научного Интернет-пространства России приведен в работе (Печников, Чуйко, 2008).

В Интернет-пространстве России практически полностью отсутствуют профессиональные научные тематические коммуникаторы, используя которые можно находить ту или иную геологическую информацию. Мы называем научным www-коммуникатором сайт, который может и не являться официальным сайтом научного учреждения и/или организации РАН и имеющий входящие и/или исходящие ссылки на множество официальных и других научных сайтов. Хотя во многих случаях создание серьезного коммуникатора может простыми средствами решить задачу организации поиска необходимой информации.

Таким образом, в настоящее время необходимо создать высокорейтинговый централизованный ресурс, который взял бы на себя роль www-коммуникатора и стал бы единой точкой входа к геолого-геофизическим ресурсам не только

Дальневосточного отделения РАН, но и к мировым ресурсам по геологии Дальнего Востока.

Портал по геологии Дальнего Востока создается на основе информационного сервера Дальневосточного геологического института Дальневосточного отделения РАН (ДВГИ ДВО РАН), долгоживущего (с 1998 г.) и достаточно раскрученного ресурса зоны RU. Сервер имеет высокий индекс цитирования Yandex = 2400, Page Rank Google = 6.

Портал создается на основе существующих и вновь строящихся тематических ресурсов Дальневосточного геологического института ДВО РАН в Интернет.

Авторами проведена работа по интеграции собственных и внешних тематических Интернет-ресурсов по геологии Дальнего Востока. Интеграция проведена двумя этапами. Сначала интегрированы между собой однородные ресурсы и созданы различные тематические сайты и порталы. Затем однородные по информации и программному обеспечению сайты и порталы интегрируются в единое порталное решение.

Интеграция библиотечных ресурсов

Электронная библиотека ДВГИ работает в Интернет с 2000 г. В 2009 г. была проведена коренная реконструкция библиотеки. В настоящее время ресурс реализован на модуле DOCMAN для Joomla.

Основные разделы библиотеки:

1. Монографии.
2. Статьи.
3. Путеводители.
4. Авторефераты диссертаций.
5. Учебные пособия.

Основные темы:

- геммология;
- геодинамика и тектоника;
- геоинформатика;
- геохимия;
- гидрогеология;
- гляциология, инженерная геология;
- информатика;
- литология и седиментология;
- математические методы;
- металлогения, рудные месторождения;
- минералогия;
- общая и региональная геология;
- петрология;
- стратиграфия и палеонтология;
- физико-химические методы.

Интеграция экспертных геологических знаний

100 ученых из институтов наук о Земле Дальневосточного отделения РАН: Дальневосточного

геологического, Тектоники и геофизики, Геологии и природопользования, Северо-Восточного комплексного, Вулканологии и сейсмологии, Морской геологии и геофизики, Тихоокеанского института географии являются авторами монографии «Геодинамика, магматизм и металлогения Востока России», на основе которой авторами построена Интернет-энциклопедия «Геология Дальнего Востока России».

Монография представляет собой наиболее полную современную сводку по тектонике, геодинамике, сейсмичности, магматизму и полезным ископаемым Дальневосточной окраины России. Основой для характеристики магматизма, металлогенического районирования и типизации месторождений полезных ископаемых послужила концепция аккреционной тектоники. Детально охарактеризованы основные структуры и магматические пояса, показана современная геодинамика и сейсмичность территории, расшифровано ее глубинное строение. Наиболее полно описаны явления сдвиговой тектоники и процессы формирования трансформных окраин континентов. Всесторонне освещены разнообразные рудные и россыпные месторождения благородных, цветных, редких и черных металлов, а также редких земель. В книге рассмотрена тектоническая и металлогеническая эволюция активных, пассивных и трансформных континентальных окраин и островных дуг с позднего архея до современности.

Портальное решение основано на технологии Wiki. В результате работ построен портал, структуру и содержимое которого пользователи могут сообщать изменять с помощью инструментов, предоставляемых порталом. Разрешение на регистрацию пользователей и соответственно на возможность внесения изменений, дополнений дается администратором.

Сущность концепции Wiki (BoLeuf, Ward Cunningham, 2001) заключается в следующем:

- Wiki предлагает всем пользователям редактировать любую страницу или создавать новые страницы на Wiki-сайте, используя обычный Web-браузер без каких-либо его расширений;
- Wiki поддерживает связи между разными страницами за счет почти интуитивно понятного создания ссылок на другие страницы и отображения того, существуют данные страницы или нет;
- Wiki не является тщательно изготовленным сайтом для случайных посетителей. Напротив, Wiki стремится привлечь посетителей к непрерывному процессу создания и сотрудничества, который постоянно меняет вид сайта.
- Wiki характеризуется такими признаками:
- возможность многократно править текст посредством самой Wiki-среды (сайта), без

применения особых приспособлений на стороне редактора;

- особый язык разметки – так называемая Wiki-разметка, которая позволяет легко и быстро размечать в тексте структурные элементы и гиперссылки; форматировать и оформлять отдельные элементы;
- учет изменений (версий) страниц: возможность сравнения редакций и восстановления ранних;
- проявление изменений сразу после их внесения;
- разделение содержимого на именованные страницы;
- гипертекстовость – связь страниц и подразделов сайта через контекстные гиперссылки;
- множество авторов – некоторые Wiki могут править все посетители сайта.

Базовые каталоги энциклопедии:

- теоретические основы тектонического, геодинамического и металлогенического анализа;
- общая характеристика рельефа Дальнего Востока России;
- глубинное строение и сейсмичность Востока России;
- кратоны и орогенные пояса Востока России;
- магматические пояса и зоны типовых геодинамических обстановок;
- осадочные бассейны Востока России;
- типы месторождений полезных ископаемых;
- металлогенические пояса и рудные районы Востока России;
- сдвиговые дислокации и их роль в процессах магматизма и рудообразования в условиях трансформной окраины Азии;
- геодинамические реконструкции и металлогения Востока России;
- тектоническая схема Дальнего Востока России;
- литература;
- фотогалерея.

Интеграция пространственных данных и сервисов

Современные компьютерные технологии предоставляют широкие возможности для организации хранения и экспрессного поиска геологической информации. В настоящее время геологические ведомства всех развитых стран мира активно используют ГИС в своей деятельности. И это не удивительно. Ведь они оперируют данными, имеющими отчетливую, зачастую детальную, пространственную привязку. А для хранения этих данных, быстрого и удобного доступа к ним на основе местоположения

и создания на их основе высококачественных карт разного назначения технология ГИС подходит наилучшим образом. Внедрение ГИС в геологическую отрасль в мире началось несколько десятков лет назад, шло параллельно и в значительной мере способствовало развитию самих ГИС-технологий.

Одним из мощнейших путей обмена геологической информацией в настоящее время является Интернет, позволяющий осуществлять доступ к информационным ресурсам через большое число поисковых систем. Пространственные геологические данные – один из видов информационных ресурсов, имеющих свои особенности, которые определяют специфику их размещения в Интернет, поиска, отображения, обмена и использования. К этим особенностям относятся: графическое представление пространственных карт в виде цифровых карт, их координатная привязка к земной поверхности и множество характеристик, связанных с графическими объектами.

С точки зрения пользователей, ГИС-портал является единой точкой доступа прежде всего к метаданным. Он обеспечивает поиск необходимой пространственной информации по ее описанию, а также непосредственное получение геоданных и работу с цифровыми картами. С другой стороны, ГИС-портал – это технология и программное обеспечение одношлюзового Web-доступа для поиска, передачи и использования геоданных и сервисов в любом пункте глобальной сети Интернет, а также размещения информации об имеющихся у пользователей данных. Портальное решение позволяет решить три основных задачи:

1. Объединение пространственных ресурсов от разных производителей на всех уровнях интеграции: от глобального до территориального или локального.
2. Обеспечение поиска и доступа к необходимой информации простыми средствами, не требующими специализированного программного обеспечения и подготовки.
3. Упорядочение пространственной информации в общедоступные каталоги, пригодные для автоматического формирования и исследования.

Структура организации портала должна включать три основных компонента:

- каталог метаданных, где пользователи производят поиск данных и размещают сведения об имеющихся у них данных;
- ГИС-узлы, где опубликованы пространственные данные;
- ГИС-пользователи, которые осуществляют поиск данных, а затем соединяются с ГИС-узлами, где эти данные расположены, для использования доступных данных и сервисов.

Первым геологическим ГИС-порталом в Российской Академии наук стал портал «ГеоМЕТА» (Вершинин и др., 2008). Основным принципом построения системы является выделение единой (базовой) схемы метаданных, реализованной в виде OWL – онтологии на основе стандартов ISO 19115-2003 «Географическая информация. Метаданные» и ISO 19119-2005 «Географическая информация сервисы».

Необходимо отметить еще один ГИС-портал в области наук о Земле РАН. Это распределенная информационно-аналитическая система для поиска, обработки и анализа пространственных данных Сибирского отделения РАН (Шокин и др., 2007).

В настоящее время в глобальной сети Интернет находится большое количество пространственной информации по различным аспектам геологии Дальнего Востока на всех уровнях масштабности. Это результаты глобальных проектов Геологических служб США, Японии, КНР, Канады и др. А также пространственная информация геологических институтов РАН по этой территории. Поэтому задача интеграции пространственных данных по геологии Дальнего Востока на основе единого ГИС-портала является актуальной.

В 2010 г. авторами реализован ГИС-портал «Геология Дальнего Востока России». Портал разработан на основе свободно-распространяемого программного обеспечения с открытым программным кодом – GeoNetwork Opensource и GeoServer. Обобщенная функциональная схема ГИС-портала приведена на рис. 1.

Географический регион – Дальневосточный федеральный округ:

- Республика Саха (Якутия);
- Камчатский край;
- Приморский край;
- Хабаровский край;
- Амурская область;
- Магаданская область;
- Сахалинская область;
- Еврейская автономная область;
- Чукотский автономный округ.

Категории каталога метаданных ГИС портала ДВГИ ДВО РАН:

- Вулканология.
- География, рельеф.
- Геодинамика и тектоника.
- Геофизика.
- Геохимия.
- Гидрогеология.
- Глубинное строение, сейсмичность.
- Литология и седиментология.
- Металлогения, рудные месторождения.
- Общая и региональная геология.

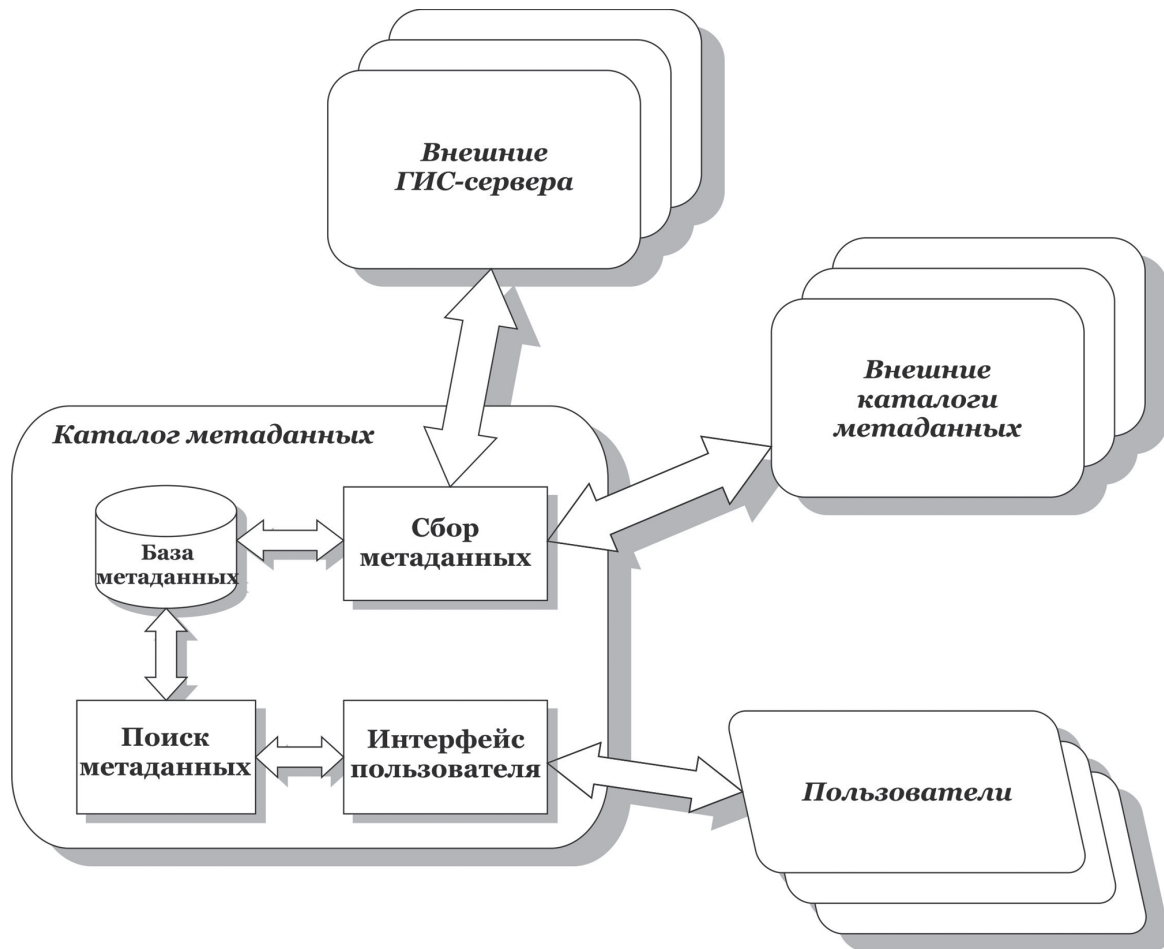


Рис. 1. Обобщенная функциональная схема ГИС-портала «Геология Дальнего Востока России»

- Петрология.
- Стратиграфия и палеонтология.

Единица информации – ГИС или электронная карта. Сбор информации осуществляется из собственных ресурсов ДВГИ ДВО РАН, а также из внешних ресурсов: мировых центров данных, мировых геологических служб, университетов, институтов РАН и из других источников. Стандарт метаданных – ISO 19115.

Стандартные возможности GeoNetwork предоставляют пользователям возможность осуществления простого и расширенного поиска метаданных.

Простой поиск:

- полнотекстовый поиск по всем атрибутам метаданных;
- по географическому региону из списка;
- по местоположению, указанному на карте.

Многокритериальный поиск может осуществляться по:

- атрибутам (названию, аннотации, ключевым словам);
- пространственному охвату;
- временному охвату;
- каталогам (локальному и/или удаленным);
- категориям.

На ГИС-портале стандартными средствами GeoNetwork реализовано отображение метаданных, средствами InterMap Viewer – отображение карт.

Медиа-сервер

Необходимость создания Медиа-сервера ДВГИ ДВО РАН связана с организацией хранения видеoinформации, получаемой при проведении научных видеоконференций, а также для организации потоковой трансляции в Интернет различных научных конференций с использованием технологий видеоконференцсвязи (Наумова и др., 2009).

Интеграция неоднородных данных по геологии Дальнего Востока на основе Web-интеграции

Необходимость создания целостного интегрального информационного поля пользователя, которое состоит из совокупности инструментальных средств, аналитических методов, описаний и геоданных, необходимых для проведения прикладных и фундаментальных исследований в науках о Земле отмечалась в статье Н.П. Лаверова с соавторами (Лаверов и др., 2008): «Подход к созданию целостного интегрального информационного поля пользователя в области наук о Земле реализует архитектуру

системы взаимосвязанных порталов, каждый из которых отвечает либо за определенную широкую предметную область, либо за отдельные аспекты технологического функционирования всей системы. Система порталов связана гиперссылками и представляет собой единую распределенную структуру, взаимодействие в которой основано на эффективном обмене метаинформацией и ресурсами».

Авторами данной статьи сделан дальнейший шаг в области интеграции неоднородных геологических порталов между собой. Предлагаемое решение основано на технологиях Web-интеграции.

Развитие Интернет сделало Web-браузеры доминирующим программным обеспечением для доступа к содержанию, приложениям и системам. Уже сложилась тенденция предоставлять доступ ко всем типам информации и сервисов посредством Web. Однако в Интернет функционирует огромное число разнородных приложений, созданных в различное время, различными организациями, на базе различных технологий. Задача Web-интеграции объединить разнородные Web-приложения и системы в единую среду на базе Web.

Типы интеграции:

Интеграция на уровне представления. Уровень представления – Web-базируемый пользовательский интерфейс, платформозависимый графический пользовательский интерфейс (GUI) или консоль терминала. Данный уровень позволяет пользователю взаимодействовать с приложением. Интеграция на уровне представления дает доступ к пользовательскому интерфейсу удаленных приложений.

Интеграция на уровне функциональности. Данная интеграция подразумевает обеспечение прямого доступа к бизнес-логике приложений. Это достигается непосредственным взаимодействием приложений с API (программному интерфейсу приложений) или же взаимодействием посредством Web-сервисов.

Интеграция на уровне данных. В данном случае предполагается доступ к одной или нескольким базам данных, используемых удаленным приложением.

Комплексная интеграция. Коммерческие решения по Web-интеграции, как правило, включают все три типа интеграции.

Основная цель создания портала «Геология Дальнего Востока России» состоит в организации единой точки входа к неоднородным тематическим ресурсам Дальневосточного геологического института ДВО РАН, что реализуется в разработке поискового механизма по всем сайтам, входящим в портал (рис. 2). Поиск данных осуществляется в трех режимах: поиск по ключевым словам, по

географическим областям и по каталогу. Поиск реализуется обращениями к поисковым средам сайтов. Визуализация поисковых запросов осуществляется как в среде сайтов, так и на портале. Таким образом, нами реализуются два типа интеграции: на уровне данных и на уровне представления.

На портале проектом предусмотрен каталог геологических ресурсов Интернет, основные разделы которого:

1. Управление природными ресурсами.
2. Наука.
3. Образование.
4. Горнодобывающая промышленность.
5. Музеи.
6. Аналитические центры.
7. Электронные библиотеки.
8. Научные журналы.
9. Тематические проекты.
10. Базы данных.
11. ГИС и электронные карты.
12. Новости.
13. Конференции.

Преимущества Web-интеграции:

- Web-интеграция позволяет развертывать информационные системы на базе сторонних приложений без необходимости разбираться в их родительских системах, программных средах и архитектурах баз данных.
- SOA и Web-сервисы используют программный язык и платформонезависимые интерфейсы между приложениями корпоративной инфраструктуры ИТ. Это дает очевидные преимущества в поддержке, управляемости, развертывании информационных сетей.
- Web-интеграция позволяет конструировать комплексную функциональность, комбинируя разнородные компоненты посредством протоколов Web-сервисов.
- Web-интеграция позволяет использовать Web-сервисы разработчиков.
- Web-интеграция позволяет развивать программные интерфейсы приложений через протоколы Web-сервисов без программирования.

Для Web-интеграции используется коммерческое ПО или популярные технологии, такие как PHP/Python/Perl, XForms, SOAP и т.д.

Как правило, говоря о Web-интеграции, подразумевают построение порталов (front office), насыщенных различной функциональностью, которые, в свою очередь, взаимодействуют с другими информационными системами (back office), агрегируя и визуализируя полученную информацию.

В большинстве случаев порталы оказываются в итоге витринами для отображения и вся их заслуга

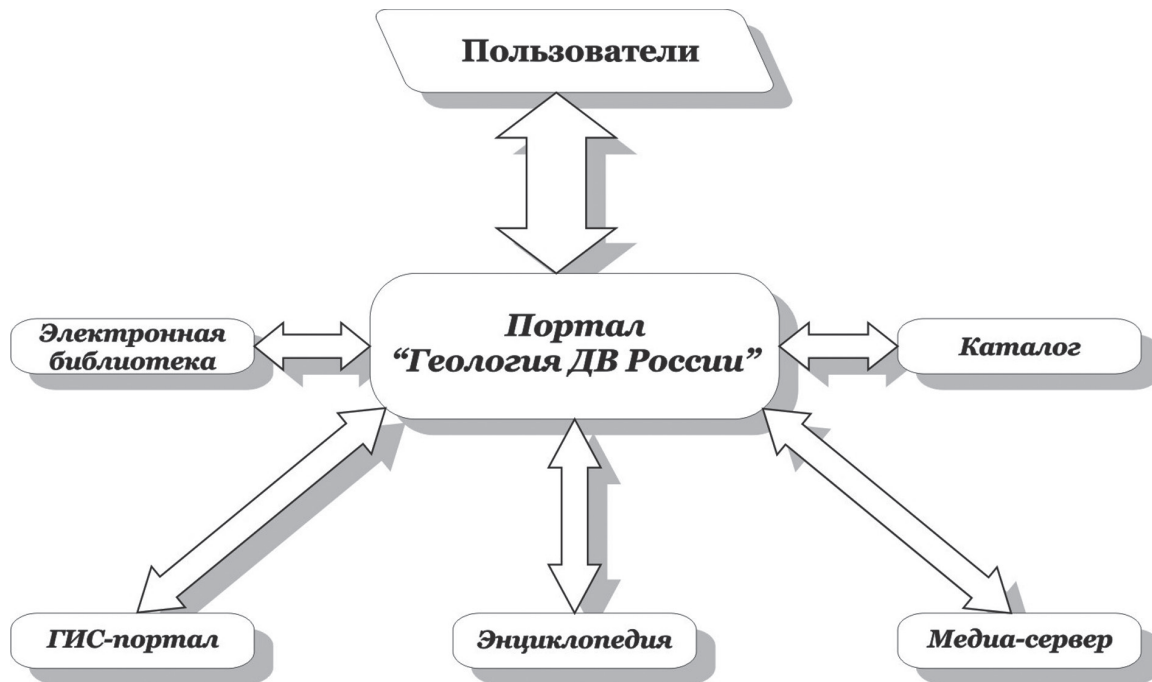


Рис. 2. Портал – единая точка входа к неоднородным геологическим ресурсам по геологии Дальнего Востока России

в том, что информация собрана в одном месте, структурирована, обеспечена средствами поиска и другими полезными сервисами (зачастую, в принципе, этого бывает достаточно).

Ключевые слова: интеграция неоднородных данных и сервисов, инфраструктура пространственных данных, порталные решения, геология Дальнего Востока России.

ЛИТЕРАТУРА

1. Вершинин А.В., Серебряков В.А., Ряховский В.М., Дьяконов И.А., Динь ле Дат, Шкотин А.В., Шульга Н.Ю. Создание среды интеграции пространственных данных и приложений // Открытое образование. – 2008. – № 4. – С. 9-16.
2. Ворошин С.В., Зинкевич А.С., Тюкова Е.Э. Региональные геоинформационные системы для геологических исследований: опыт создания и анализа // Тихоокеанская геология. – 2006. – № 5. – С. 22-38.
3. Геодинамика, магматизм и металлогения Востока России / отв. ред. А.И. Ханчук; чл. ред. колл.: С.М. Родионов, Н.А. Горячев, В.К. Попов, В.В. Голозубов, В.В. Наумова. – Владивосток : Дальнаука, 2006. – 981 с.
4. Лаверов Н.П., Арский Ю.М., Савин Г.И., Жижченко А.Б. Интегральное информационное поле в науках о Земле // Вестник РАН, 2008. – Т. 78. – № 10. – С. 875-879.
5. Наумова В.В. Состояние и перспективы развития геоинформатики в геологических науках на Дальнем

Востоке России // Тихоокеанская геология – 2006. – Т. 25. – № 5. – С. 3-8.

6. Наумова В.В., Миллер Р.М., Ноклеберг В.Дж., Патук М.И., Капитанчук М.Ю., Парфенов Л.М., Ханчук А.И., Родионов С.М. ГИС «Минеральные ресурсы, минералогенезис и тектоника Северо-Восточной Азии» // Тихоокеанская геология – 2006. – Т. 25. – № 5. – С. 8-22.

7. Наумова В.В., Сорокин А.А., Горячев И.Н. Видеоконференцсвязь – мультимедийный сервис Корпоративной сети Дальневосточного отделения РАН // Информационные технологии. – 2009. – № 4. – С. 66-70.

8. Рашидов В.А., Романова И.М., Бондаренко В.И., Палуева А.А. Web- и ГИС-технологии в геомагнитных исследованиях поднекайнозойских подводных вулканов Тихого океана // Материалы Международной Конференции «Итоги электронного геофизического года», 3-6 июня 2009 г., Переяславль-Залесский, Российская Федерация. – 69 с.

9. Ханчук А.И., Наумова В.В. Информационное пространство Дальневосточного отделения РАН // Вестник ДВО РАН. – 2009. – № 4. – С. 122-130.

10. Распределенная информационно-аналитическая система для поиска, обработки и анализа пространственных данных / Ю.И. Шокин, О.Л. Жижимов, И.А. Пестунов и др. // Вычислительные технологии. – 2007. – Т. 12, спецвыпуск 3. – С. 108-115.

11. BoLeuf, WardCunningham. The Wiki Way: Collaboration and Sharing on the Internet. AddisonWesley, 2001. – ISBN 020171499X.