

МЕТЕОРОЛОГИЯ

В.М. Шаймарданов

ОСНОВНЫЕ ПРИНЦИПЫ ОРГАНИЗАЦИИ ИНФОРМАЦИОННЫХ УСЛУГ МАТЕРИАЛАМИ РОССИЙСКОГО ГОСУДАРСТВЕННОГО ФОНДА ДАННЫХ О СОСТОЯНИИ ОКРУЖАЮЩЕЙ ПРИРОДНОЙ СРЕДЫ

V.M. Shaimardanov

MAIN PRINCIPLES OF THE ORGANIZATION OF INFORMATION SERVICES BY MATERIALS OF THE RUSSIAN STATE FUND OF THE DATA ABOUT OF ENVIRONMENT CONDITION

Накопление большого объема данных на различных носителях для упрощения обслуживания пользователей требует консолидации на унифицированной платформе с использованием общих программных средств.

Один из вариантов это перевод всей информации в электронный вид с последующим распознаванием. Что позволит обеспечить надежную и долгосрочную сохранность данных и оптимизировать процесс обслуживания.

Ключевые слова: данные, информационные услуги, обслуживание пользователей.

The accumulation of large volumes of data in different media to facilitate user service requires the consolidation using a unified platform and common software tools.

One of the options is to convert all information into electronic format with further recognition. It ensures a reliable and long-term data integrity and optimization of service process.

Key words: data, information services, service of users.

К настоящему времени в архивах Росгидромета накоплен большой объем данных на различных носителях. Например, только во ВНИИГМИ – МЦД хранится [2]:

- более 12 ТБ – на магнитных лентах ЕС ЭВМ (50 000 магнитных лент с данными о состоянии природной среды начиная с 1881 г.), на картриджах IBM 3480 и на CD-ROM (данные, полученные из НИИ и от лицензиатов);
- 1000 ТБ – данные, полученные в рамках международных проектов на различных технических носителях (магнитооптические диски, магнитные картриджи, CD и др.);
- 2500 ТБ – копии документов на бумажных и фотоносителях;
- большое количество специализированных и производных массивов.

Вся совокупность накопленных данных в фонде скомплектована по типу носителей в трех больших группах [1]:

- на технических носителях ЭВМ,
- данные, хранящиеся в листовых материалах на бумажной основе (таблицы, карты, тексты и др.),
- на фотоносителях (микрофильмы, микрофиши, фотоотпечатки).

Исходя из сложившейся ситуации выполнение работ по оказанию услуг пользователям необходимо организовывать с использованием накопленных данных на различных носителях и соответствующих программных средств. До сегодняшнего дня не автоматизированы процессы обслуживания потребителей этими данными, а также отсутствует оперативный мониторинг выполнения заявок. Не автоматизировано ведение процессов выполнения заявок на оказание информационных услуг. Поскольку значительная доля информации накоплена на фото- и бумажных носителях, это ограничивает оперативный доступ к данным и их поиск. Для достижения поставленной цели и совершенствования схемы обслуживания потребителей необходимо ее доработать.

На рис. 1 представлена создаваемая схема оказания информационных услуг.

Из схемы видно, что выполнение этих работ многоэтапная процедура.

В сложившейся системе оказания информационных услуг по предоставлению гидрометеорологической информации можно выделить ряд моментов, усовершенствование которых позволит повысить эффективность выполнения этих услуг.

Сформулируем основные принципы построения системы оказания информационных услуг по обеспечению потребителей гидрометеорологической информацией.

Система должна позволять:

- получать необходимые пользователю сведения оперативно: в требуемые сроки, с заданным качеством и соответствующей полнотой;
- обеспечивать оперативный мониторинг выполнения заявки, ведя оперативный учет и анализ оказанных услуг;
- предоставлять сведения (результаты услуги) предусмотренными средствами доведения информации;
- предоставлять сведения потребителям в требуемых форматах;
- предоставлять сведения на носителях по требованию потребителя;
- предоставлять возможность доступа телекоммуникационными средствами к информационным базам данных фонда;
- обеспечивать оперативное взаимодействие с пользователем на стадии заключения договора и его выполнения;
- обеспечивать информирование и рекламу по оказанию информационных услуг;
- в исключительных случаях предоставлять возможности работы пользователю непосредственно в фонде;
- оперативно реагировать на устранение недостатков, обнаруженных пользователем (реакция на рекламацию).

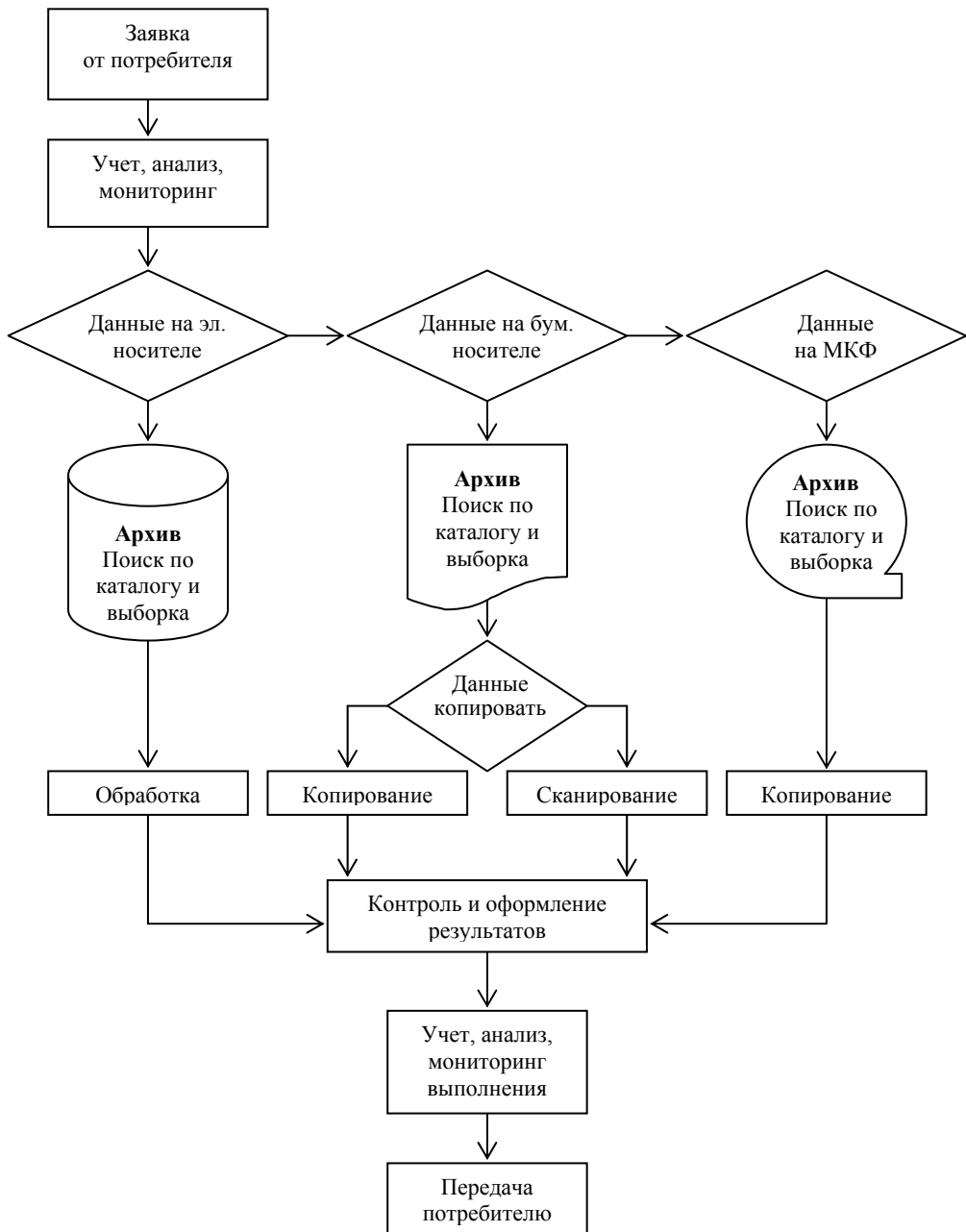


Рис. 1. Схема оказания информационных услуг с использованием различных типов носителей информации

Принципиальных изменений в системе обслуживания можно достичь путем организации распределенной системы доступа и использования накоплен-

ных баз данных. Например, доступ к данным на машиночитаемых носителях Централизованного гидрометеорологического фонда (ЦГМФ), в том числе и средствами телекоммуникации по согласованному регламенту непосредственно из научно-исследовательских учреждений (НИУ) или из управлений и центров по гидрометеорологии и мониторингу окружающей среды (УГМС, ЦГМС) по их зоне ответственности. На первом этапе пользователями по этой схеме могут быть организации Росгидромета, которые ведут обслуживание внешних потребителей, обозначим их удаленными узлами обслуживания.

В ходе реализации этого направления должны быть разработаны (развиты) необходимые технологии формирования баз данных, являющихся основой применения телекоммуникационных средств и Web-технологий.

В состав этой работы входит:

- разработка состава и структуры основной базы данных подсистемы;
- разработка и развитие технологий доступа к информационной продукции средствами удаленного доступа;
- разработка и развитие технологий подготовки ГИС приложений для улучшения обслуживания потребителей готовой продукцией.

В качестве концептуального решения может быть принята следующая схема:

- на уровне ЦГМФ обеспечивается подготовка и резервное копирование всех информационных баз данных для поддержки репликации и тиражирования;
- в удаленных узлах (УГМС, ЦГМС, НИУ) могут храниться наборы данных, необходимые для функционирования узла в соответствии с его специализацией, но обязанности хранить страховую копию на них не возлагаются.

Рекламации на полноту, качество и выявленных технических ошибок в информации сообщается в ЦГМФ. Все эти изменения после включения в состав фонда становятся доступными и для других пользователей.

Передача информационных ресурсов на каждый узел должна проводиться по согласованному регламенту в соответствии с согласованным перечнем источников. Эту загрузку должны обеспечивать соответствующие программные и технологические средства узла. Иными словами, каждый узел в УГМС, ЦГМС и НИУ должен быть обеспечен программными средствами усвоения информации, включая входной контроль качества информации и средства мониторинга их состава и содержания, в том числе каталогизации как основы работы удаленного узла.

Основной задачей развития этого направления является создание нового поколения информационных баз данных и разработка (развитие) технологий их ведения и эффективного доступа к данным, в том числе удаленного.

В целом подсистема формирования баз данных для обслуживания осуществляет следующие основные функции:

- формирование информационной базы данных посредством загрузки в нее всех необходимых сведений из базового архива;
- ведение информационной базы данных, включая поддержание ее в актуальном состоянии, защиту ресурсов, их регулярное пополнение.

Указанные функции должны поддерживаться соответствующими программно-технологическими комплексами, которые группируются в набор АРМ:

- АРМ Администратор удаленного доступа к архивам ЦГМФ;
- АРМ Администратор ведения информационных баз для обслуживания.
- АРМ Администратор удаленного доступа к архивам ЦГМФ предназначен для организации доступа к санкционированному сегменту архивов, а также формирования и ведения базы в отдельном узле системы на локальном уровне (УГМС, ЦГМС, НИУ).

АРМ Администратора Централизованного фонда должен быть ориентирован на работу в распределенной среде.

Функции АРМ Администратора Централизованного фонда:

- контроль структурных изменений;
- контроль схемы размещения информации в Централизованном фонде;
- мониторинг данных на уровне каталогов и адресов;
- получение справочной информации по заданному регламенту;
- классификация данных и формирование флагов уровня доступа к информации;
- задание уровней доступа классам пользователей;
- регламентное резервирование и восстановление данных;
- контроль производительности системы в целом;
- поддержка и обеспечение актуальности информационных баз в целом.

До настоящего времени обслуживание информацией, находящейся на бумажных и фотоносителях, предполагает выполнение традиционных рутинных действий по типу поиска документов в библиотечных системах. Полученный результат в виде требуемых сведений может быть представлен в виде копии документа на бумажном носителе, либо в электронном виде отсканированного текста.

В процессе создания электронного архива основная операция – сканирование документов с бумажных и фотоносителей. Это должно составить основу программно-аппаратной среды для перевода бумажных и фотодокументов Госфонда в электронный вид, а именно, процессы сканирования, распознавания и занесения в электронное хранилище. Создание электронного архива документов, хранящихся на бумажных и фотоносителях, обеспечит повышение надежности, снижение времени доступа к данным, реализацию многопользовательской работы с документами, а также возможность автоматизированного учета, администрирования, пополнения и использования данных.

При необходимости, весь отсканированный документ или прицельно выбранные области документа (содержащие номера, даты, фамилии или, например, идентифицирующие штрих-коды) должны переводиться в символьный вид с помощью средств оптического распознавания (OCR). При этом структурированные данные, представленные в виде табличных материалов, требуют переноса на машиночитаемые носители как в форматах графических файлов (имиджей страниц), так и оцифровки (ручной или с помощью алгоритмов оптического распознавания символов). Оцифровка позволит вовлечь в обработку на ЭВМ

значительную часть фонда данных, которая на сегодняшний день недоступна для обработки. Следует учесть, что возможности применения универсальных алгоритмов и программных средств оптического распознавания символов для задач массовой оцифровки данных, особенно табличных и имеющих нестандартные символы, ограничены. Поэтому должен предусматриваться заказ специализированного программного обеспечения, применимого для таких специфических задач. Неизбежен также перенос значительной части документов в электронный архив путем ручной оцифровки.

Росгидромет, как видно из сказанного выше, обладает огромными, пополняемыми на регулярной основе, информационными ресурсами и, следовательно, владеет мощным средством потенциального информационного воздействия на экономические и социальные процессы. В связи с этим вопросы повышения эффективности использования гидрометеорологической информации являются весьма актуальными.

Для развития надежных систем обслуживания потребителей информацией, в том числе и изменениях климатической системы необходимо:

- расширение возможностей потребителей по доступу к данным для решения научных и практических задач (за счет разработки и внедрения в работу фонда современных информационных технологий, в том числе интернет-технологий и защиты данных от несанкционированного доступа);
- обеспечение деятельности по созданию справочной базы фондов непосредственно в процессе накопления, обработки и архивации данных при их включении в фонд;
- автоматизированное ведение семейств справочных массивов по основным информационным ресурсам;
- обеспечение удаленного интерактивного доступа потребителей к сведениям о составе, структуре фонда и удаленный санкционированный доступ к данным, находящимся на отсканированных документах;
- обеспечение оперативной подготовки и издания бюллетеней, отчетов, справок, позволяющих осуществлять оперативное и ретроспективное обслуживание потребителей сведениями о наличии и характеристиках информационных ресурсов фондов данных.

Для ускорения процесса обслуживания потребителей, как указано выше, создается компонент информационной продукции, который обеспечит обработку, обобщение и анализ информации, получение информационной продукции по регламентированному набору показателей: состав и содержание продукции, пространственно-временное разрешение, временной регламент подготовки продукции, формы представления продукции и способы доставки потребителям.

Литература

1. Шаймарданов М.З., Веселов В.М., Стерин А.М., Шаймарданов В.М. Концепция модернизации архивной системы Мирового метеорологического центра // Тр. ВНИИГМИ-МЦД. 2007. Вып. 172, с. 3-37.
2. Шаймарданов М.З., Шаймарданов В.М. Единый государственный фонд данных // Тр. 6-го Всеросс. метеорологич. съезда, 2011, СПб., с. 74-85.