

DOI: 10.51790/2712-9942-2022-3-1-4

СОЗДАНИЕ АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ СКАНИРОВАНИЯ БУМАЖНЫХ ДОКУМЕНТОВ РОСГИДРОМЕТА

Е. В. Колесников, В. М. Шаймарданов

Всероссийский научно-исследовательский институт гидрометеорологической информации — мировой центр данных, г. Обнинск, Российская Федерация
evgeni-kolesnikov@mail.ru

Аннотация: настоящая статья посвящена вопросам обеспечения долговременной сохранности, оперативного обслуживания и удаленного многопользовательского доступа к архивным документам на бумажных носителях по гидрометеорологии и смежным с ней областям (метеорологии, гидрологии, аэрологии, океанологии, агрометеорологии, гелиогеофизики и др.). Данные, собранные в такого рода документах, представляют собой информацию о состоянии окружающей природной среды, её загрязнении, полученную в результате деятельности Федеральной службы России по гидрометеорологии и мониторингу окружающей среды (Росгидромет). Данная информация имеет не только историческую ценность, но и практический интерес для деятельности органов власти, физических и юридических лиц.

В работе представлена автоматизированная информационная технология (АИТ) сканирования архива документов на бумажной основе. АИТ имеет многоуровневую архитектуру и решает следующие задачи: автоматизация учета, структуризация информации, долговременное хранение и удаленный доступ к электронным копиям бумажных документов с помощью поискового модуля, интегрированного в автоматизированную информационную систему (АИС). АИТ использует современное программное и аппаратное обеспечение. В работе описан весь технологический процесс сканирования бумажных документов и проведена оценка полученных результатов в ходе эксплуатации АИТ на примере одной из организаций Росгидромета — Всероссийского научно-исследовательского института гидрометеорологической информации — мирового центра данных (ФГБУ «ВНИИГМИ-МЦД»).

Ключевые слова: бумажные документы, сканирование документов, автоматизированная информационная технология, автоматизированная информационная система, электронный архив, удаленный доступ, модель описания данных, метаданные, долговременное хранение, Госфонд Росгидромета.

Для цитирования: Колесников Е. В., Шаймарданов В. М. Создание автоматизированной информационной технологии сканирования бумажных документов Росгидромета. *Успехи кибернетики*. 2022;3(1):28–38. DOI: 10.51790/2712-9942-2022-3-1-4.

DEVELOPMENT OF AN AUTOMATED PAPER DOCUMENTS SCANNING TECHNOLOGY FOR THE RUSSIAN METEOROLOGICAL SERVICE

E. V. Kolesnikov, V. M. Shaimardanov

*Russian Research Institute for Hydrometeorological Information, World Data Center,
Obninsk, Russian Federation*
evgeni-kolesnikov@mail.ru

Abstract: this paper considers the issues of long-term preservation, maintenance, and remote multi-user access to paper records in hydrometeorology and related fields (meteorology, hydrology, aerology, oceanology, agrometeorology, heliogeophysics, etc.). Such records contain the environment and pollution status data acquired by the Russian Meteorological Service. This information is not only of historical value but also of practical interest for the authorities and other consumers.

The paper presents an automated information technology (AIT) for scanning historical paper records. AIT has a multi-level architecture and performs the following: automated document tracking, data structuring, long-term storage, and remote access to digital copies of paper documents with an integrated search module of an automated information system (AIS). The AIT uses advanced software and hardware. The paper describes the entire process of paper records scanning and evaluates the results of the AIT pilot operation at the Russian National Research Institute for Hydrometeorological Information, World Data Center.

Keywords: paper records, document scanning, automated information technology, automated information system, electronic archive, remote access, data models, metadata, long-term storage, Russian Meteorological Service National Archive.

Cite this article: Kolesnikov E. V., Shaimardanov V. M. Development of an Automated Paper Documents Scanning Technology for the Russian Meteorological Service. *Russian Journal of Cybernetics*. 2022;3(1):28–38. DOI: 10.51790/2712-9942-2022-3-1-4.

Введение

На сегодняшний день сеть организаций – структурных подразделений Росгидромета (НИУ, управления и центры Гидрометеорологической службы (УГМС, ЦГМС)) обладает большими объемами бумажных документов по гидрометеорологии и смежным с ней областям. На начало 2021 года в Госфонде Росгидромета накоплено 2 754 038 единиц хранения на бумажной основе, хранящихся в специализированных архивохранилищах, с поддержанием необходимых температурно-влажностных условий долговременного хранения [1, 2]. Ежегодно количество документов в Госфонде Росгидромета в среднем увеличивается на 25 000 единиц хранения. Таким образом, идет непрерывное пополнение архива по всем организациям.

Ведение Госфонда на бумажных носителях является надежным способом долговременного хранения при условии соблюдения всех методических рекомендаций архивного дела [3]. Однако существует проблема получения информации, связанная с ограниченным тиражом (1–3 экземпляра) бумажных документов, достаточно длительным по сегодняшним меркам процессом их поиска в архивохранилище с последующей выдачей конечному пользователю при личном посещении архива. Также существует проблема многократного использования, особенно актуально это для документов в единственном экземпляре, которая ведет к их физическому износу и потере ценной информации, как частичной, так и полной. Таким образом, обслуживание бумажными документами пользователей является ресурсозатратным процессом.

Данный вид документов является ценным и востребованным материалом для различной деятельности современного общества. Гидрометеорологическая информация используется в таких сферах деятельности, как строительство, сельское хозяйство, энергетика, авиация, нефтегазовая промышленность, наука и юриспруденция, что говорит об актуальности проблем долговременного хранения и удаленного доступа. В связи с этим необходимо разработать и организовать технологический процесс перевода информации с бумажных носителей на современные технические носители. Это позволит в полном объеме решить актуальные на сегодняшний день проблемы архивного дела, а именно: упрощение доступа к информации, обеспечение сохранности подлинников за счет их постепенного вывода из оборота, организация удаленного доступа к электронному фонду [4].

На сегодняшний день наиболее быстрым и эффективным способом перевода информации в электронный вид является процесс сканирования бумажных документов. В результате полученные электронные объекты можно будет классифицировать, описать и поместить в электронный архив для долговременного хранения. Благодаря клиент-серверной архитектуре АИС и метаданным объектов хранения будет организован удаленный многопользовательский доступ к необходимым пользователю электронным копиям бумажных документов.

В статье представлено полное описание всех этапов автоматизированной технологии оцифровки бумажных документов Госфонда Росгидромета и реализующей ее АИС ЭФОД.

Анализ исходного состояния архива бумажных документов

Прежде чем перейти к разработке технологического процесса, необходимо провести предварительный анализ исходного состояния архива бумажных документов Росгидромета, а именно: состава и качества бумажного архива, его объема (количества единиц хранения) – и провести классификацию по видам наблюдений.

Результаты анализа крайне важны, поскольку без стопроцентной оценки состава архива и вариативности видов документов нельзя переходить к подготовке технического задания и инженерно-техническим работам, т.к. ошибочное представление (оценка) повлечет за собой неправильные расчеты временных, вычислительных и человеческих ресурсов, а также ошибочный выбор необходимого для сканирования оборудования.

Например, при неправильном проведении классификации архивных документов в процессе разработки информационной модели описания будут учтены не все аспекты, что приведет в дальнейшем к серьезным правкам в базе данных и логике работы АИС. Вследствие чего уже проделанную работу по формированию электронного архива (могут быть уже сканированы, описаны и загружены несколько тысяч объектов хранения) придется выполнять заново. Поэтому была проведена тщательная, кропотливая, пусть и затратная по времени работа перед непосредственным процессом сканирования документов.

В этой работе были задействованы преимущественно сотрудники, ответственные за фонд данных: архивариусы, персонал, обслуживающий пользователей, специалисты смысловых гидрометеорологических подразделений и методического сопровождения сети Росгидромета.

В результате проведенного анализа сформирована информационная база, позволяющая перейти к разработке ключевой составляющей технологии оцифровки бумажных документов – модели описания бумажных документов. Модель описания представляет собой совокупность полей (метаданных), каждое из которых описывает конкретное свойство объекта хранения. Создание модели описания обеспечивает однозначную идентификацию каждого объекта хранения в АИС и определяет его как уникальную единицу хранения электронного и бумажного хранилищ.

В ходе работы по классификации документов и оценке объемов были использованы нормативные и руководящие документы, проведены интервью с сотрудниками, работающими с фондом, и специалистами в области гидрометеорологии и смежных с ней дисциплин. В итоге получена следующая классификация видов массивов архива бумажных документов Росгидромета:

- метеорологические таблицы и ежемесячники;
- гидрологические ежегодники;
- отчеты научно-исследовательских судов (НИС);
- таблицы аэрологических наблюдений (ТАЭ);
- спутниковые снимки;
- документы по загрязнению окружающей среды;
- климатические справочники;
- исторические данные (летописи ГГО);
- зарубежная информация и т.д.

В следующем разделе подробно охарактеризован процесс создания модели описания, достаточной для точного определения каждого объекта в среде долговременного хранения.

Модель описания бумажных документов

На основе изучения архива бумажных документов (оценки объемов и классификации документов) разработана модель описания бумажных документов.

Основная идея заключается в том, чтобы выделить минимально достаточное количество свойств бумажных документов, необходимых для однозначного определения каждого объекта хранения как уникальной единицы и достаточного смыслового содержания, которые в дальнейшем может использовать пользователь, составив соответствующий поисковый запрос к базе данных АИС ЭФОД [4].

Выделенные критерии, составляющие модель описания бумажных документов фонда гидрометеорологических данных, раскрывают как свойства документа, так и контент — смысловое содержимое документа, что в совокупности позволяет идентифицировать каждый объект хранения.

Отметим, что перед тем, как пользователи смогут составлять поисковые запросы для получения интересующих их документов, необходимо проиндексировать каждый оцифрованный бумажный документ в единой базе мета-описаний объектов хранения АИС ЭФОД в соответствии с разработанной моделью описания.

Для реализации данного процесса разработан программный компонент, позволяющий через веб-интерфейс пополнить базу мета-информацией об объектах хранения. Пользователю предоставлены поля как с фиксированными, заранее заполненными списками, например, дисциплина, вид наблюдений, в которых он выбирает значение из списка, так и текстовые поля для заполнения в свободной форме, например, район наблюдений.

Использование полей с фиксированным набором данных (тезаурусов) позволяет пользователю облегчить задачу по составлению поискового запроса, поскольку ему не нужно будет задумываться, какие гидрометеорологические дисциплины и виды наблюдений используются для описания бумажных

документов, а также сократит время подготовки запроса.

В таблице представлена разработанная модель описания АИС ЭФОД. Столбец «Название поля» несет информацию о кратком названии атрибута модели описания. «Описание поля» вносит пояснение в значение названия поля. «Тип данных» – это тип данных, определяющий значение в базе данных (числовые, символьные, дата и время, логические и т.д.). «Обязательность к заполнению» указывает на обязательность заполнения конкретного атрибута модели описания в базе данных. «Несколько значений» обозначает, что данный атрибут может иметь несколько значений. Например, один документ может относиться к нескольким дисциплинам (метеорология и океанография).

Таблица

Перечень метаданных

№	Название поля	Описание	Тип данных
Блок «Идентификатор»			
1	Код документа (уникальный идентификатор объекта)	<i>Автоматически заполняется на основе заполненных полей:</i> <i>Категория документа</i> <i>Дисциплина</i> <i>Вид наблюдения</i> <i>Номер инвентарной книги</i> <i>Инвентарный номер</i>	текст
2	Номер инвентарной книги	Поле для формирования кода документа	числовой
3	Инвентарный номер	Поле для формирования кода документа	числовой
4	Категория документа	Поле для формирования кода документа <i>Заполняется из словаря</i>	текст
5	Код дисциплины	Поле для формирования кода документа <i>Содержит код гидрометеорологической дисциплины</i> <i>Заполняется из словаря</i>	текст
6	Код вида наблюдений	Поле для формирования кода документа <i>Содержит код вида наблюдений</i> <i>Заполняется из словаря</i>	текст
Блок свойств			
7	Организация	Название организации, выпустившей документ	текст
8	Название документа	Имя документа в АИС ЭФОД	текст
9	Авторы/редакторы	Авторы документа	текст
11	Язык издания	Язык(и), на котором издан документ <i>Заполняется из словаря</i>	текст
11	Страна	Страна <i>Заполняется из словаря</i>	текст
12	Место издания / издательство	Издатель документа	текст
13	Год издания	Год, в котором издан документ	числовой
14	Вид представления	Вид, в котором представлена структура документа (таблица, карта, атлас, ежегодник и т.д.) <i>Заполняется из словаря</i>	текст
15	Количество страниц/листов	Количество страниц документа	числовой
16	Том	Номер тома	текст
17	Выпуск	Номер выпуска	текст
18	Серия	Номер серии	текст
19	Часть	Номер части	текст
20	Доступ	Тип доступа (открытый, для служебного пользования) <i>Заполняется из словаря</i>	текст
21	Способ получения	Способ получения информации (платный, бесплатный) <i>Заполняется из словаря</i>	текст

№	Название поля	Описание	Тип данных
22	Владелец информации	Владелец документа (УГМС, НИИ)	текст
Блок «Контент»			
23	Дисциплина	Содержит название гидрометеорологической дисциплины <i>Заполняется из словаря</i>	текст
24	Вид наблюдений	Содержит название вида наблюдений <i>Заполняется из словаря</i>	текст
25	Параметры	Содержит перечень параметров наблюдений, входящих в состав документа	текст
26	Массив	Название массива, к которому принадлежит документ <i>Заполняется из словаря</i>	текст
27	Начало наблюдений	Дата начала наблюдений	дата
28	Конец наблюдений	Дата окончания наблюдений	дата
29	Район наблюдений	Район, в котором проводились наблюдения	текст
30	Дополнительная информация	Дополнительная информация включает в себя особенности документа, такие как дополнительные виды наблюдений, которые являются уточнением, пояснением <i>Заполняется из словаря</i>	текст

Структура модели описания бумажных документов состоит из 3 блоков и 30 полей.

Блок «Идентификатор» позволяет определить каждый оцифрованный документ как уникальную единицу хранения в АИС.

Поле «Код документа» позволяет идентифицировать, определять каждый сканированный документ, помещенный в среду долговременного хранения, и предоставлять пользователю по запросу. Данное поле имеет кодовую структуру, состоящую из совокупности соответствующих кодов полей: номера инвентарной книги, инвентарного номера документа, категории документа, кода дисциплины и кода вида наблюдений (рис. 1).

В основе кода документа лежит инвентарный номер. Было бы логично использовать его в качестве кода документа, так как он позволяет идентифицировать объект хранения в автоматизированной системе, однако в случае фонда гидрометеорологических бумажных документов инвентарный номер не является сквозным и обнуляет номерной ряд в каждой из инвентарных книг, имеющих в фонде. Поэтому к нему добавлен инвентарный номер книги, в состав которой он входит.

Кроме того, код документа был расширен добавлением в идентификатор категории документа, дисциплины и вида наблюдений. Это расширение позволяет организовать поиск по крупным категориям для удобства просмотра массивов однотипных документов в одном рабочем пространстве.

Например, по такой схеме можно получить массив востребованных документов, таких как гидрологические ежегодники или метеорологические ежемесячники. Также это снижает нагрузку на аппаратную часть технологии за счет увеличения размерности каталожной структуры и уменьшения количества файлов внутри каждого каталога.

В **блок свойств** помещены описания «паспортных» данных объекта хранения, которые включают: название документа, организации, подготовившей документ, год издания и т.д.

«Организация» — данное поле определяет организацию, департамент, структуру, которая выпустила данный документ, например, ФГБУ «Северное УГМС».

«Название документа» — определяет название книги.

«Язык издания» — сообщает, на каком языке издан документ.

«Страна» указывает территориальную принадлежность содержимого документа. Например, если это Россия/СССР, то данные, которые содержит документ, относятся к данной стране.

«Год издания» — год, в котором был издан документ (не путать с периодом наблюдений).

«Вид представления» документа определяет тип издания, например, таблица, карта, ежегодник, бюллетень, справочник и т.д.

«Количество страниц», «Том», «Выпуск», «Серия», «Часть» — поля, говорящие сами за себя.

«Способ получения» — поле определяет условия предоставления документа: на платной или бесплатной основе.

«Владельцем информации» является организация Росгидромета — НИУ или УГМС.

Блок «Контент» включает в себя смысловую часть документа, а именно:

- гидрометеорологические дисциплины;
- виды наблюдений с указанием параметров и периода наблюдений;
- районы (территорию), где проведены наблюдения;
- массив гидрометеорологических документов, к которому принадлежит текущий объект хранения.



Рис. 1. Процесс формирования кода документа

Отметим, что поля «Дисциплина» и «Вид наблюдений» – связанные поля отношением «один ко многим», т.е. у одной дисциплины может быть несколько видов наблюдений. Например, у дисциплины «Гидрология» имеются следующие виды наблюдений:

- наблюдения на реках и каналах;
- наблюдения на озерах и водохранилищах;
- наблюдения на болотах.

В поисковом веб-интерфейсе все связанные списки меняют свои значения в зависимости от выбора значения в родительском списке. Таким образом, использование заполненных ранее словарей существенно облегчает процесс составления поискового запроса.

Поле «Дополнительная информация» введено в связи с некоторыми особенностями документов, а именно, когда один документ совмещает в себе несколько видов наблюдений, что не позволяет хранить в данном поле агрегированные данные.

В таблицу модели описания не вошли так называемые служебные поля:

- «Имя оператора, формировавшего описание документа»;
- «Дата создания мета-описания»;
- «Статус объекта»;
- «Дата загрузки объекта хранения в автоматизированную систему»;
- «Количество просмотров и загрузок документов».

Эти поля включены в структуру базы данных, несут информацию для обслуживания системы и используются для подготовки отчетности и контроля за жизнедеятельностью АИС ЭФОД.

Разработанная информационная модель описания бумажных гидрометеорологических документов позволяет формализовать процесс наполнения базы данных и реализовать разработку автоматизированной информационной технологии.

Описание этапов технологии

Работа с документами выполняется с помощью АИС ЭФОД. Она предоставляет инструменты для каждого из этапов оцифровки бумажных документов Единого государственного фонда данных о состоянии окружающей среды, ее загрязнении (ЕГФД).

Первым этапом работ по оцифровке бумажного фонда данных является подготовка документов. На этом этапе формируются описи передаваемых из архивохранилища на участок сканирования доку-

ментов. Участок состоит из сканеров различных форматов (A0-A5) для сканирования разных типов бумажных документов. На этом же этапе специалист формирует уникальный идентификатор, который входит в опись. На основе уникальных идентификаторов формируются штрих-коды, наклеиваемые на документы, которые объединяются в партии по форматам для оптимизации процесса сканирования.

После завершения процесса сканирования контролер образов проводит проверку отсканированных образов. В случае когда документ соответствует всем критериям качества, электронный образ документа становится доступен для индексирования — формирования мета-описания в базе данных, иначе документ возвращается на повторное сканирование и процесс проверки повторяется до тех пор, пока не будет достигнут приемлемый результат.

Индексатором является специалист, владеющий предметной областью, способный грамотно описать документ. С целью минимизации ошибок каждый документ индексируют два специалиста независимо друг от друга, при этом не имея доступа к информации, кто из группы индексаторов описывает текущий документ. Как показала практика, такой подход позволил на порядок сократить количество ошибок.

Далее описанный документ проходит этап верификации. На этом этапе верификатор проверяет мета-описание, при этом он уже знает, кто из группы индексаторов описывал документ, соответственно, может указывать на ошибки в работе индексаторов и улучшать результат их работы.

Одновременно с переходом на этап индексирования физический бумажный документ возвращается в архивохранилище с очень важным нововведением, а именно: каждому документу присваиваются координаты стеллажа, ряда и полки в специализированной базе данных.

Ранее местоположение документов в архивохранилище можно было найти в книге учета и с помощью алфавитного указателя, что требовало некоторого количества времени. Теперь же этот процесс в разы проще и быстрее, поскольку имеется база данных координат на технических носителях.

На заключительном этапе верифицированная партия документов проходит экспертизу, после чего принимается в фонд. Параллельно подготавливаются соответствующие документы.

Описание автоматизированной информационной системы электронного фонда оцифрованных бумажных документов Росгидромета

АИС ЭФОД предназначена для автоматизации всех этапов процесса оцифровки архивных бумажных документов различных форматов о состоянии окружающей природной среды и её загрязнении и формирования электронного фонда документов.

Структурно система представляется в виде компонент (рис. 2), каждой из которых соответствует определенное множество логически связанных между собой функций по обработке и представлению данных, относящихся к той или иной сущности предметной области.

Система состоит из следующих структурных компонент:

- средств хранения файлов различного типа;
- контрольно-принимающей системы;
- среды обеспечения сетевого взаимодействия компонентов.

Средства хранения файлов в зависимости от их назначения включают в себя различные программные и технические средства:

- базу данных с метаданными, предназначенную для хранения атрибутивных карт со ссылками на каталог файлов с электронными копиями различных видов бумажных документов;
- ленточную библиотеку, предназначенную для хранения многостраничных tiff-файлов (образы документов) и csv-файлов, содержащих атрибутивные карты [5];
- систему хранения данных, обеспечивающую хранение многостраничных PDF-файлов образов документов и csv-файлов, содержащих атрибутивные карты;
- контрольно-принимающую систему, которая представляет собой непосредственно программные компоненты АИС ЭФОД, обеспечивающие реализацию функций обработки и подготовки данных для последующего их сохранения на средствах, описанных выше;
- среду обеспечения сетевого взаимодействия, необходимую для коммутации структурных компонентов системы между собой с целью передачи необходимых данных.

Система реализована в виде трёхзвенной архитектуры:

- сервер базы данных;
- сервер приложений;

- клиентское приложение.

Клиентское приложение развёрнуто на специализированных рабочих местах: специалиста предварительной обработки документов, технолога, контролера документов, контролера образов, специалиста группы сканирования, индексатора, специалиста по обработке бумажных документов (ОБД), администратора системы, аналитика, верификатора. Внешние пользователи подключаются к АИС ЭФОД посредством сети Интернет через брандмауэр, настроенный по сетевым правилам. АИС ЭФОД предоставляет несколько АРМов, каждый из которых предназначен для выполнения определённых функций.

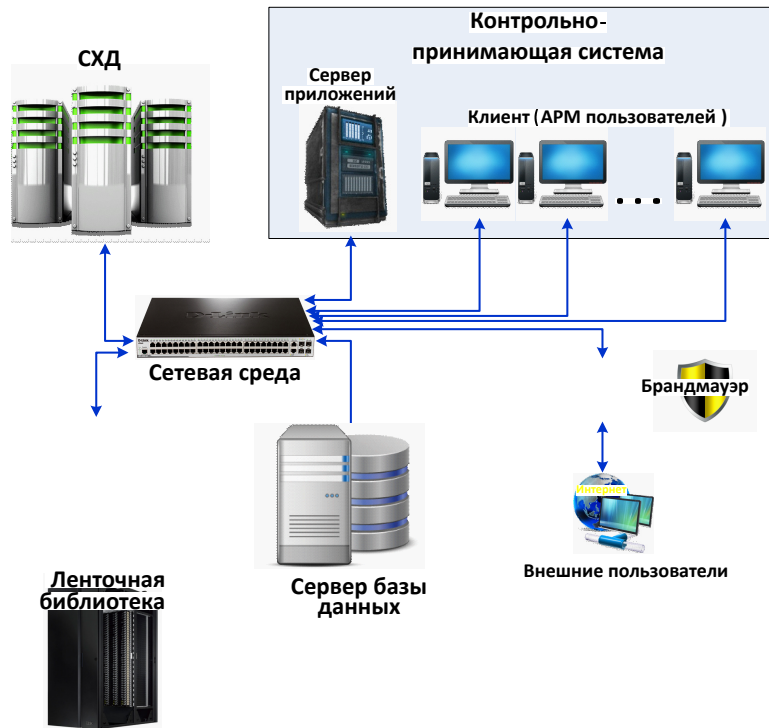


Рис. 2. Структурная схема АИС ЭФОД

АРМ специалиста предварительной обработки документов предоставляет возможность фиксации выдачи из архива материалов и возврата, добавления в систему новых (на основе уникального идентификатора) документов, печати штрих-кодов на основании вышеуказанного идентификатора, назначения документу профиля.

АРМ технолога позволяет формировать список документов, отобранных для коррекции метаданных, выполнить его редактирование из списка, просмотреть метаданные, введенные каждым из индексаторов, указать, какой вариант является правильным, либо задать свой вариант, сохранить документ. Кроме того, АРМ технолога обеспечивает возможность редактирования справочных данных, управления профилями постобработки, управления документами, просмотра и печати статистики.

АРМ контролера документов обеспечивает:

- отображение списка выданных из архива документов, предоставление возможности ручного добавления документа в список, вывод на печать описи переданных и принятых из архива бумажных документов;
- отображение списка неразобранных документов, предоставление возможности назначения документу профиля, отображение списка профилированных документов с возможностью внесения корректив;
- автоматическое распределение документов по специалистам группы сканирования на основании форматов сканеров и загрузки (и плана на смену), предоставление возможности печати описи переданных на сканирование документов;
- отображение списка отсканированных документов;
- отображение списка переданных и обработанных на этапе ОБД документов, предоставление возможности печати описи;
- отображение списка возвращенных в архив документов, предоставление возможности печати

описи возвращаемых документов.

АРМ контролера образов позволяет сформировать список документов, отобранных для их контроля, открыть для просмотра, выполнить просмотр облегченного образа. В случае необходимости АРМ предоставляет возможность корректировки страницы документа, получения для редактирования исходного (облегченного) образа данной страницы, внесения изменений каким-либо внешним графическим редактором, замены исходного образа страницы в системе с последующим обновлением многостраничного pdf-документа.

АРМ специалиста группы сканирования предоставляет пользователю список документов, переданных ему для работы, а также возможность отметки о ее начале/завершении.

АРМ индексатора предоставляет пользователю для просмотра список документов, переданных ему на индексирование. Данный АРМ обеспечивает параллельную индексацию одного и того же документа разными индексаторами, а также предоставляет возможность внести значения атрибутов в ее карточку для сохранности.

АРМ администратора предоставляет возможность работы с системой, включая:

- управление пользователями системы и их ролями;
- редактирование справочных данных;
- управление профилями постобработки;
- управление профилями сканеров;
- управление документами, включая редактирование атрибутов, изменение статуса, постановку/снятие отметок о передаче на этапы сканирования, сшивки/расшивки, отправку на проверку, отметку о готовности для записи в хранилище;
- просмотр и печать статистики.

АРМ верификатора позволяет пользователю сформировать список документов, отобранных для коррекции метаданных, открыть на редактирование документ из списка, просмотреть метаданные, введенные каждым из индексаторов, указать, какой вариант является правильным, либо задать свой вариант, сохранить документ.

АРМ аналитика предоставляет возможность вывода на печать или экран всех статистических отчетов, предусмотренных в системе.

Порядок работы с автоматизированной информационной системой

Работа начинается с авторизации в системе путём ввода логина и пароля в форму (рис. 3). После чего пользователю будет предоставлен соответствующий интерфейс для работы с системой. В зависимости от роли пользователя интерфейс подразделяется на: специалиста предварительной обработки документов, технолога, контролера документов, контролера образов, специалиста группы сканирования, индексатора, специалиста по обработке бумажных документов, администратора системы, аналитика, верификатора, поиск.

ФГБУ «ВНИИГМИ-МЦД»

Электронный фонд
отсканированных документов

Вход в систему

Имя пользователя:

Пароль:

Войти

Версия: 1.18.353

© 2018 Росгидромет

Рис. 3. Форма авторизации в АИС ЭФОД

Для поиска документов необходимо заполнить поля поисковой формы (рис. 4). После чего АИС обработает значения, введенные пользователем, и сформирует запрос к базе данных.

Дисциплина: Вид наблюдений:

Период наблюдений: с: по: Любой год из указанного периода:

Параметры среды (ключевые слова): Все значения:

Дополнительная информация:

[Расширенный поиск](#)

Название книги:

Категория документа:

Страна:

Организация:

Место издания, издательство:

Год издания:

Вид представления:

Район наблюдений:

Доступ:

Способ получения:

Язык издания:

Искать документы:

Рис. 4. Краткая и расширенная формы поиска документов

В соответствии с запросом пользователю формируется и выводится результат в виде списка документов (рис. 5).

Количество документов: 28

Государственный водный кадастр. Ежегодные данные о режиме и ресурсах поверхностных вод суши. (ГВК. ЕДС)

Количество документов в массиве: 27

ЕДС Азербайджанское УГМС 1985г. т.VII ч.1,2
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно: Наблюдения на озерах и водохранилищах.
Период наблюдений: 1985.01.01 - 1985.12.31

ЕДС Армянское УГМС 1985г. т. XIII ч. 1,2
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно: Наблюдения на озерах и водохранилищах.
Период наблюдений: 1985.01.01 - 1985.12.31

ЕДС Верхне- Волжское УГМС 1984г. т.I вып.23 ч.1
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно:
Период наблюдений: 1984.01.01 - 1984.12.31

ЕДС Верхне- Волжское УГМС 1985г. т.I вып.23 ч.1
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно:
Период наблюдений: 1985.01.01 - 1985.12.31

ЕДС Грузинское УГМС 1985г. т.VI ч.1,2
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно: Наблюдения на озерах и водохранилищах.
Период наблюдений: 1985.01.01 - 1985.12.31

ЕДС Забайкальское УГМС 1984г. т.I вып.20 ч.1,2
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно: Наблюдения на озерах и водохранилищах.
Период наблюдений: 1984.01.01 - 1984.12.31

ЕДС Западно-Сибирское УГМС 1985г. т.I вып.10 ч.1
ГИДРОЛОГИЯ Наблюдения на реках и каналах

ЕДС Верхне- Волжское УГМС 1985г. т.I вып.23 ч.1
СССР/Россия / Верхне-Волжское территориальное управление по гидрометеорологии и контролю природной среды
Издано: Обнинск: ВНИИГМИ- МЦД / Верхне- Волжское УГМС /1987
ГИДРОЛОГИЯ Наблюдения на реках и каналах
Дополнительно:
Период наблюдений: 1985.01.01 - 1985.12.31
Район наблюдений: Бассейн Волги (верхнее течение).
Параметры среды (ключевые слова): Мутность воды; Ледовые явления; Толщина льда; Температура воды; Оценка надежности подсчета стока воды; Высота снега на льду; Наносы; Уровень воды; Расход воды;
Вид представления: Ежегодник
Объем: 470
Место хранения документов:
Доступ: Открытый фонд
Способ получения документа: Платный
Документ из ВНИИГМИ-МЦД
Количество просмотров:
Количество скачиваний:

[Загрузить](#) [Смотреть](#)

Рис. 5. Результат работы поискового модуля

Документ можно скачать или открыть для просмотра в браузере. Для этого достаточно кликнуть по необходимому документу в списке.

Заключение

В данной статье было представлено описание текущего состояния фонда бумажных документов Росгидромета, обозначены проблемы и способы их решения, а также охарактеризована автоматизированная информационная технология электронного фонда оцифрованных архивных бумажных

документов Росгидромета. Новизна и актуальность данной разработки заключается в том, что она позволяет решить для Росгидромета существующие на сегодняшний день задачи: долговременное хранение архивных документов и доступ, в том числе и удаленный сетевой, к одному и тому же документу большому числу пользователей с применением современных программных средств и технологий.

ЛИТЕРАТУРА

1. Ведение ЕГФД, архивация данных, обслуживание данными и информацией ЕГФД. Сведения о составе Госфонда Росгидромета. Режим доступа: <http://meteo.ru/egfd/142-about-egfd>.
2. Шаймарданов В. М., Шаймарданов М. З. Развитие автоматизированной архивной системы Росгидромета. *Ученые записки РГГМУ*. 2014;36:60–66.
3. РД 52.19.568-2010 Организация комплектования, учета, хранения и использования документов архивного фонда данных о состоянии окружающей среды, ее загрязнении.
4. Шаймарданов В. М., Долгих С. Г. Создание технологии управления данными в автоматизированной архивной системе. *Труды Главной геофизической обсерватории им. А.И. Воейкова*. 2014;572:162–174.
5. Шаймарданов В.М. Развитие архивной системы данных о состоянии окружающей природной среды на базе современных технических средств. *Труды Главной геофизической обсерватории им. А.И. Воейкова*. 2011;563:179–190.