

*Секция «Проблемы моделирования, проектирования и разработки программных средств»,
научный руководитель – Рыбанов А.А., канд. техн. наук, доцент*

УДК 004.91

**ИССЛЕДОВАНИЕ МЕТОДОВ И РАЗРАБОТКА АВТОМАТИЗАЦИИ
СОЗДАНИЯ ИНТЕГРИРОВАННОГО АРХИВА ДОКУМЕНТОВ**

Агафонкин М.А., Рыбанов А.А., Свиридова О.В.

*Волжский политехнический институт, филиал ФГБОУ ВО «Волгоградский государственный
технический университет», Волжский, e-mail: mihash-vlz@mail.ru*

Так как многие предприятия, компании развиваются и расширяются, перед ними встаёт проблема в большом объёме документооборота и соответственно перевод с бумажных носителей в электронный вид. И в следствии с этим была написана эта статья. В ней рассмотрены основные проблемы и их решения, по внедрению технологий автоматизации создания интегрированного архива документов. Также описывается порядок развёртки оборудования на предприятии. В статье выявлены и описаны самые необходимые функции этих технологий. А именно: сканирование, распознавание и корректировка ошибок, создание и миграция электронных документов и образов, индексирование документов, оперативный поиск и отображение документов, управление функционированием системы. А так же рассмотрены аналоги программного обеспечения, которые соответствуют современным требованиям.

Ключевые слова: автоматизация создания электронного архива документов, корпоративный электронный архив, архив документов

**STUDY OF METHODS AND DEVELOPMENT OF AUTOMATION OF CREATING AN
INTEGRATED ARCHIVE OF DOCUMENTS**

Agafonkin M.A., Rybanov A.A., Sviridova O.V.

*Volzhsy Polytechnic Institute, branch Volgograd State Technical University, Volzhsky,
e-mail: mihash-vlz@mail.ru*

The article considers the main problems and their solutions for implementing automation technologies for creating an integrated archive of company or enterprise documents. The most necessary functions of these technologies have been identified and described. As well as analogues of software are considered.

Keywords: automation of creation of electronic archive of documents, corporate electronic archive, archive of documents

Электронные архивы активно используют, как в крупных, так и в мелких компаниях и предприятиях. И без корпоративного электронного архива уже нельзя представить работу в компании, так как на нём строится весь документооборот. Но с каждым годом количество документов увеличивается и требования к электронному архиву возрастают. Подобные факты подталкивают к использованию технологий, способных автоматизировать большинство процессов в архиве документов.

Постановка проблемы. В наше время компании всё чаще сталкиваются с большим объёмом документооборота [2]. Что приводит к увеличению времени обработки этих документов. И в связи с этим, компании внедряют технологии автоматизации в электронный архив. Что позволит сократить время обработки документов, пропадут проблемы с хранением, ускоренный поиск

по различным параметрам, автоматическое распределение документов в архиве и много других функций. Благодаря этим технологиям, работать становится проще, пользователь меньше заботится о хранении документов, и перенаправляет сэкономленное время непосредственно на обработку самого документа.

**Процесс автоматизации
электронного архива**

Автоматизации электронного архива это трудоёмкий и ресурсозатратный процесс. Общую концепцию можно обрисовать следующим образом. Вначале следует провести развёртывание высокопроизводительной сети, графические рабочие станции и мощные сервера ввода и обработки информации. Для перевода документов с бумажных носителей в электронный вид, применяют промышленные сканеры потокового вво-

да и подходящие к ним русифицированное программное обеспечение. Система снабжает высокоэффективное индексирование и полнотекстовой поиск неструктурированной информации большого объема. В высокопроизводительной и отказоустойчивой системе памяти, хранятся данные, которые используются для поиска документов. А данные представленные в графическом виде, хранятся на носителях, что характеризует продолжительным временем хранения и невысокой стоимости.

В итоге образуется ряд основных функций:

- сканирование;
- распознавание и корректировка ошибок;
- создание и миграция электронных документов и образов;
- индексирование документов;
- оперативный поиск и отображение документов;
- анализ документов;
- управление функционированием системы;

Понятие электронного архива документов состоит из определений современных автоматизированных систем управления документооборотом (СУД) и информационных хранилищ данных – DataWarehouse (ИХД). Отличающейся чертой внедрения системы электронного архивирования заключается в том, что в первую очередь необходимо заполнить базу данных системы весь объем документов [2]. Из-за того, что это очень длительный и трудоемкий процесс, то он потребует максимальной автоматизации, а именно: исключения оператора от любого участия в процессе ввода, распознавания, корректировки и индексирования документов. И с этим связана следующая отличительная особенность электронного архива документов – обеспечение полнотекстового поиска. Характерного для систем управления документооборотом, построение и поддержка системы атрибутивного поиска оказывается неприемлемым из-за временных и стоимостных затрат.

Касательно ИХД, то в электронном архиве также обрабатываются большие объемы данных, которые могут достигать больше одного терабайта. Так как структура базы данных системы направлена на объекты, а не на процессы, то архивные данные однозначные и их точность не зависит от времени, а также стабильны, в том плане, что они не удаляются и не модифицируются. Отличительная черта заключается в том, что в ИХД данные имеют интегральный вид, которые получены из множества разнотипных СУБД

и файловых систем [3]. Как правило, в КЭА утверждается единая технология ввода документов. Но на практике всё обстоит иначе: ряд архивов уже обладает конкретными наработками либо в организации документооборота, либо в автоматизации каталогов, описей и специфичных картотек. И следовательно на первом этапе имеющиеся базы данных интегрируются в электронный архив.

Иной характерной и отличительной особенностью КЭА является то, что он имеет как средства оперативного поиска информации (On-line Time Processing – OLTP), так и средства оперативного анализа информации (On-line Analyze Processing – OLAP), которые являются главными средствами ИХД.

При всем разнообразии программных продуктов, на сегодняшний день не существует как таковых продуктов СУД и ИХД, которые позволяют обеспечить всеми необходимыми функциями электронного документооборота и архивирования для работы с огромными объемами документов. С другой стороны, наибольшее количество компонентов КЭА, как технических, так и программных, являются уникальными и настраиваемыми образцами.

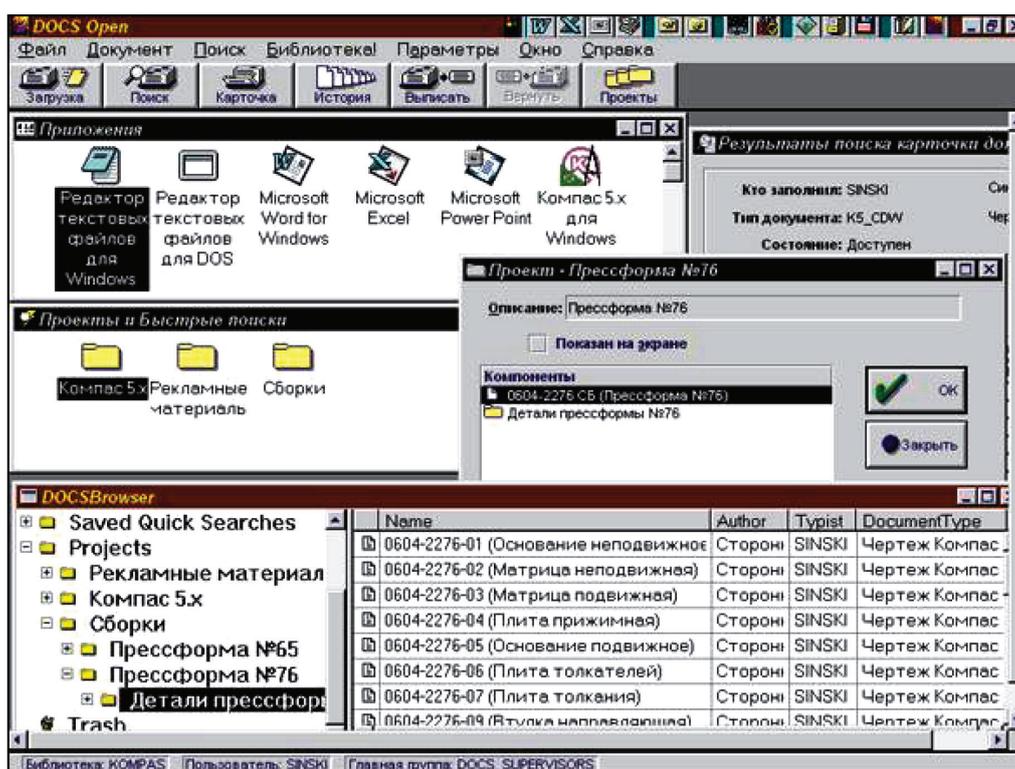
Поэтому, при проектировании электронного архива компании появляются, как минимум, две взаимосвязанные проблемы, а именно: оценка и выбор компонентов; интегрирование технологий, программных продуктов и технических средств.

Выбор отдельных компонентов, например, высокопроизводительной сети, конкретных моделей серверов или RAID-массива сходится с современными технологическими решениями построения традиционных автоматизированных систем сбора, хранения и обработки информации.

Обзор существующих решений

На сегодняшний день существует не так много программных продуктов на рынке для решения автоматизации электронного архива документов. Самые популярные и одновременно традиционные из них это:

«DOCS OPEN» (рис. 1) – издателем данного продукта является компания Hummingbird, которая более 15 лет разрабатывает решения масштабов предприятия. Это программное обеспечение является мощным решением по управлению документами, которое позволяет организовать корпоративный электронный архив документов, управлять им, назначать права доступа к информации и многое другое. Распространяется на операционные системы Windows.



DOSC Open

«Excalibur EFS» – компания Excalibur Technologies. Разработала и представила на рынке технологию адаптивного распознавания образов APRP. Данная технология основана на нейронных сетях, что обеспечивает обход проблемы ошибок связанных с распознаванием текстов, но и даёт возможность автоматического индексирования и поиска разных типов неструктурированной информации в электронной форме.

Excalibur EFS является другим примером по управлению электронным архивом. Серверное ПО данного средства работает в операционной системе Unix, а пользовательские рабочие места могут работать на операционной системе MS Windows. За основу базы данных учетных карточек можно использовать такие средства, как: СУБД Oracle, Informix, Sybase и Ingres. И поэтому как такового интерфейса программы

не существует. Разрабатывается для каждого клиента индивидуально и не распространяется в интернете.

Вывод

В результате исследовательской работы было принято решение создания упрощённого аналога автоматизированного электронного архива для общего доступа. Допускается использование программного обеспечения в маленьких компаниях.

Список литературы

1. Сыроедова Т.И. Выбор СЭД для организации / Т.И. Сыроедова // Современные технологии делопроизводства и документооборота. 2015. – № 6 (54). – С. 33.
2. Чесноков С.В. Внедрение электронного документооборота в Российских организациях / С.В. Чесноков // Электронный документооборот. – 2015. – С. 45.
3. Кузнецов И.Н. Документационное обеспечение управления и делопроизводство / И.Н. Чесноков // Вопросы современного документооборота и делопроизводства. – 2012. – С. 126.