|  |  |  |
| --- | --- | --- |
| ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ  Федеральное государственное автономное образовательное  учреждение высшего профессионального образования  "Национальный исследовательский университет  "Высшая школа экономики"  Пермский филиал | | |
| Факультет бизнес-информатики | | |
| Кафедра информационных технологий в бизнесе | | |
| УДК 004.043+004.2 | | |
| **АНАЛИЗ НЕСТРУКТУРИРОВАННЫХ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ  КАК СРЕДСТВО ОПТИМИЗАЦИИ  КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ**  *Выпускная квалификационная работа бакалавра* | | |
|  |  | Работу выполнил студент  группы БИ-10-2  4 курса факультета бизнес-информатики  Э.А. Брезгина  Научный руководитель:  старший преподаватель кафедры информационных технологий в бизнесе  В.В. Ланин  “\_\_\_\_\_” 2014 г. |
| Пермь 2014 | | |

**ОГЛАВЛЕНИЕ**

[Список условных обозначений и сокращений 4](#_Toc390288599)

[Введение 5](#_Toc390288600)

[Глава 1. Управление электронными документами в корпоративной информационной системе 7](#_Toc390288601)

[1.1. Корпоративная информационная система 7](#_Toc390288602)

[1.1.1. Понятие КИС 8](#_Toc390288603)

[1.1.2. Архитектура КИС 10](#_Toc390288604)

[1.2. Электронный документооборот 15](#_Toc390288605)

[1.2.1. Понятие и сущность электронного документа 15](#_Toc390288606)

[1.2.2. Недостатки традиционного подхода к ЭД 19](#_Toc390288607)

[1.2.3. Структура электронного документа 21](#_Toc390288608)

[1.2.4. Жизненный цикл электронного документа 29](#_Toc390288609)

[1.2.5. Стандарты по управлению ЭД 30](#_Toc390288610)

[1.2.6. Понятие электронного документооборота 37](#_Toc390288611)

[1.3. Семантические технологии в КИС 39](#_Toc390288612)

[1.3.1. Обзор технологий классического подхода для проектирования архитектуры КИС 40](#_Toc390288613)

[1.3.2. Обзор технологий семантического подхода для проектирования архитектуры КИС 44](#_Toc390288614)

[1.3.3. Применение классических и семантических технологий в КИС 52](#_Toc390288615)

[Глава 2. Моделирование бизнес-процессов по управлению электронными документами 58](#_Toc390288616)

[2.1. Бизнес-процесс «Ввод документа в ИС» 61](#_Toc390288617)

[2.2. Бизнес-процесс «Регистрация документа в ИС» 63](#_Toc390288618)

[2.3. Процесс «Выполнение заявки на выдачу документа из архива» 64](#_Toc390288619)

[2.4. Процесс «Выполнение запроса на удаление документа из ИС» 65](#_Toc390288620)

[2.5. Процесс «Выполнение запроса на редактирование документа» 67](#_Toc390288621)

[2.6. Бизнес-процесс «Формирование, согласование, утверждение» 68](#_Toc390288622)

[Глава 3. Проектирование архитектуры КИС 71](#_Toc390288623)

[3.1. Разработка методики оценки эффективности архитектуры КИС 71](#_Toc390288624)

[3.3.1 Методы оценки архитектуры КИС 72](#_Toc390288625)

[3.3.2 Методика оценки архитектуры КИС 72](#_Toc390288626)

[3.2. Формирование требований к разрабатываемой архитектуре 74](#_Toc390288627)

[3.3. Проектирование архитектуры КИС 79](#_Toc390288628)

[3.3.1 Проектирование концептуального уровня архитектуры КИС 79](#_Toc390288629)

[3.3.2 Проектирование логического уровня архитектуры КИС 84](#_Toc390288630)

[3.3.3 Проектирование физического уровня архитектуры КИС 88](#_Toc390288631)

[Заключение 90](#_Toc390288632)

[Библиографический список 92](#_Toc390288633)

[Приложение A. Реквизиты оформления организационно-распорядительных документов по ГОСТ Р 6.30-2003 96](#_Toc390288634)

[Приложение B. Элементы схем БП в нотации ARIS eEPC 97](#_Toc390288635)

[Приложение C. Бизнес-процесс «Ввод документов в ИС (сканирование)» 98](#_Toc390288636)

[Приложение D. Бизнес-процесс «Регистрация документа в ИС» 99](#_Toc390288637)

[Приложение E. Бизнес-процесс «Выполнение заявки на предоставление документа из электронного архива» 101](#_Toc390288638)

[Приложение F. Бизнес-процесс «Выполнение запроса на удаление документа из ИС» 102](#_Toc390288639)

[Приложение G. Бизнес-процесс «Выполнение запроса на редактирование документа в архиве» 103](#_Toc390288640)

[Приложение H. Схема бизнес-процесса «Формирование, согласование, утверждение документа» 104](#_Toc390288641)

[Приложение I. Схема архитектуры КИС 105](#_Toc390288642)

# Список условных обозначений и сокращений

* ИТ – информационные технологии;
* ИС – информационная система;
* КИС – корпоративная информационная система;
* ЭД – электронный документ;
* ЭДО – электронный документооборот;
* СЭД – система электронного документооборота;
* БП – бизнес-процесс;
* СОА – сервис-ориентированная архитектура.

# Введение

Ежедневно каждая организация в процессе своей деятельности создает и обрабатывает десятки, сотни и тысячи документов. Естественно, что по мере роста предприятия неизбежен этап, когда процесс документооборота перестает быть эффективным. Для решения этой проблемы в управление предприятием вводятся определённые способы управления бизнес-процессами, которые связаны с потоком документов в организации. С учетом того, что современной экономике свойственен быстрый рост доли информационных технологий, а также интеллектуализации технологий бизнеса, все в большей степени повышается роль информации и знаний, находящихся в организации. В связи с этим компании стремятся эффективно управлять и распоряжаться имеющимися у них знаниями.

Предметные области управления знаниями и семантические технологии в настоящее время сильно развиваются и идея всеобщей структуризации информации становится все более возможной. А поскольку большинство неструктурированной информации содержится в электронных документах, электронный документооборот требует особого внимания. Именно поэтому, объектом исследования являются процессы электронного документооборота, протекающие в корпоративной информационной системе, а предметом исследования является применение методов семантического подхода к анализу электронных документов в архитектуре КИС.

Целью данной работы является проектирование архитектуры корпоративной информационной системы, учитывающей применение методов извлечения смыслового и структурного содержания электронного документа.

Осуществление перечисленных особенностей возможно за счет применения в архитектуре КИС методов извлечения из электронных документов метаданных и знаний, которые впоследствии могут быть применены на разных этапах прохождения жизненного цикла управления ЭД на предприятии. Таким образом, для достижения поставленной цели необходимо решить следующие задачи:

* анализ литературы, сбор, систематизация и обобщение материалов, обзор архитектуры КИС, электронного документооборота, семантических технологий, применяемых в КИС;
* моделирование бизнес-процессов, связанных с управлением электронными документами, их анализ;
* разработка методики оценки эффективности проектируемой архитектуры КИС.
* сбор и формирование требований к разрабатываемой архитектуре КИС;
* проектирование архитектуры КИС, ориентированной на работу с неструктурированными электронными документами.

Проведенный в ходе данной выпускной квалификационной работы анализ предметной области, формулировка требований и определение технологий семантического анализа, применимых к архитектуре КИС позволят спроектировать архитектуру КИС с ориентацией работы на семантические возможности системы в плане оптимизации процессов работы с ЭД.

# Управление электронными документами в корпоративной информационной системе

В настоящее время сфера применения электронного документооборота неуклонно расширяется и в любых организация большого и среднего масштаба каждый день выполняется множество операций с электронными документами. Эффективный и прозрачный документооборот во многом определяется его организацией и налаженностью в корпоративной информационной системе.

В данной главе представлен взгляд на организацию с точки зрения информационной системы масштаба предприятия, выявлена структура электронного документа, а также рассмотрен процесс документооборота в компании и место, которое он занимает в рамках всей корпоративной информационной системы предприятия.

## Корпоративная информационная система

В настоящее время, с точки зрения системной концепции все компании, независимо от области их деятельности, достаточно сильно похожи друг на друга. В структуру каждой из них входят многочисленные подразделения, непосредственно выполняющие тот или иной вид детальности. Подразделения организации пронизаны, как вертикальными, так и горизонтальными связями, выполняют свои бизнес-процессы в рамках единого бизнес-процесса компании и обмениваются различными потоками данных, как материальными, так и информационными. Таким образом, любая компания – это совокупность элементов, взаимодействующих между собой, и выполняющих тот или иной вид деятельности, в рамках стратегии компании. Учитывая, что последние десятилетия развитие компании протекает в ракурсе применения информационных технологий в автоматизации деятельности компаний, нельзя упускать из виду важность использования информационных технологий в оптимизации и совершенствовании деятельности компании.

Проектирование и построение комплексных систем автоматизации предприятия в XXI веке переросло в отдельно взятую область развития информационных технологий. Проектирование информационных систем на уровне всего предприятия достаточно трудоемкий и масштабный процесс. Но поскольку большинство организаций по структуре похожи между собой, общий взгляд на организацию позволил сформировать общие принципы построения корпоративных информационных систем.

### Понятие КИС

Информация в современном мире превратилась в один из наиболее важных ресурсов, а информационные системы (ИС) стали необходимым инструментом во всех сферах деятельности. Разнообразие задач, решаемых с помощью ИС, привело к появлению множества разнотипных систем, отличающихся принципами построения и заложенными в них правилами обработки информации [1]. На предприятиях используется достаточно большое количество классов систем: ERP (Enterprise Resource Planning – система планирования ресурсов предприятия), ECM (Enterprise Content Management – системы управления корпоративным информационным контентом), BPM (Business Process Management) – системы управления бизнес-процессами предприятия) и пр. Под этими системами понимаются такие системы, в которых функционально объединены, как системы для решения задач автоматизации, учета и управления производством, финансовыми, человеческими и информационными ресурсами.

Корпоративная информационная система – это инфраструктура предприятия, задействованная в процессе управления всеми информационно-документальными потоками и обеспечивающая основные бизнес-процессы предприятия [9]. Важно отметить, что использование КИС на предприятии предполагает использование информационной системы на основе концепции сквозной автоматизации бизнес-процессов, выполняемых взаимодействующими подразделениями организации.

Эта особенность позволяет сформулировать ряд требований, предъявляемых современным корпоративным информационным системам при их проектировании:

1. должно выполняться управление полным циклом выполнения заказов клиентов на всех этапах делового процесса;
2. должно быть доступно оперативное получение аналитической информации для обеспечения и повышения качества принятия управленческих решений;
3. необходимо внедрение системы управленческого учета затрат в разрезе различных видов деятельности, проектов, подразделений предприятия;
4. должна быть реализована технология бюджетирования и обеспечения динамической связи необходимых ресурсов с утвержденными проектами и направлениями деятельности;
5. необходимо наличие системы бизнес-планирования;
6. необходимо наличие систем электронного документооборота и управления рабочими потоками.

Для реализации КИС, удовлетворяющей вышеперечисленным требованиям многие методы и средства канонического проектирования информационных систем, предназначенные для локальной автоматизации процессов, не подходят, следовательно, применяются средства индустриального проектирования ИС на основе CASE (Computer-Aided Software/System Engineering - автоматизированная разработка программного обеспечения) и компонентных технологий. Также из приведенного перечня требований видно, что создание КИС – это очень сложный и трудоемкий процесс, требующий решения многочисленных задач и немалых затрат.

Однако в большинстве случаев выполнения проектов по созданию ИС, особенно больших и охватывающих организацию целиком, цель создания ИС достигается редко, а также сопровождается рядом больших временных и ресурсных затрат. Причиной возникновения таких проблем является не столько нечеткая постановка целей, сколько отсутствие комплексного системного подхода ко всем этапам жизненного цикла создания ИС.

Системный подход, в свою очередь, также оказывается недостаточным для выполнения сложных проектов масштаба предприятия. Поскольку требуется еще большее применение комплексных подходов, применение архитектурного подхода позволит повысить эффективность, масштабируемость и управляемость создаваемой ИС с точки зрения всех аспектов архитектуры ИС.

### Архитектура КИС

Корпоративная информационная система представляет собой более узкое понятие, чем просто информационная система, поэтому целесообразно первоначально разобраться с понятием архитектуры для ИС, а уже потом определить особенности архитектуры ИС масштаба предприятия. Следует отметить, что в настоящее время, как в российской, так и в зарубежной практике проектирования ИС, понятие «архитектуры системы» используется очень широко, и при этом имеет множество трактовок и объяснений, тем самым порождая неизбежную путаницу в терминологии. Все варианты определений вместе дополняют друг друга и помогают понять существенные аспекты архитектуры корпоративной информационной системы.

Одно из наиболее употребляемых формальных определений архитектуры приведено в стандарте ANSI / IEEE Std 1471 – 2000 Института инженеров-электриков и электронщиков, который предоставляет метамодель для определения архитектуры [38]. Согласно данному стандарту, архитектура – это фундаментальная организационная система, заключенная в своих компонентах, их взаимоотношениях и окружении, а также, включающая в себя принципы, определяющие проектирование, создание и развитие системы. В соответствии с данным стандартом система обладает архитектурой, которая может быть описана с различных точек зрения заинтересованных лиц, рассматривающих архитектуру системы. Однако стандарт не определяет саму структуру архитектуры ИС, иными словами говорится о том, что необходимо представлять архитектуру с разных ракурсов, но при этом не указывается, какие это должны быть представления.

Также заслуживает внимания стандарт, являющийся стандартом РФ, ГОСТ 34.320-96. Отечественный стандарт не использует термин «архитектура системы», но определяет вид структуры ИС и основные компоненты ИС. Данный стандарт интересен тем, что, в отличие от предыдущего, содержит конкретные знания об архитектуре ИС, которая состоит из трех уровней: внешняя схема, внутренняя схема и уровень концептуальной схемы, информационной базы и информационного процессора [19].

Один из самых известных специалистов, работающих в области архитектуры систем, Дж.Захман в своей методологии архитектуры предприятия представляет архитектуру ИС в виде структурированного набора моделей [32], которые отражают различные содержательные точки зрения на структуру системы того круга лиц, которые вовлечены в ее создание и развитие – собственника, менеджеров, проектировщика, конструкторов, субподрядчиков, пользователей. При этом различные точки зрения обращаются на различные структурные аспекты системы (структура данных, функции, сетевая инфраструктура, организация, время и пр.).

Анализ различных источников на предмет определения архитектуры ИС позволяет сделать вывод о том, что, несмотря на разнообразие понятий, подходов к определению, в большинстве случаев, термин «архитектура системы» является синонимичным к понятию «структура системы». Также, важно отметить, что большое количество методологий дает определения и обозначает рамки архитектуры, но на практике большинство организаций при проектировании архитектуры ИС не соблюдает все пункты методологий по ряду различных причин. Одной из ключевых проблем для организаций, в настоящее время, является построение связующего звена между существующими бизнес-процессами компании и средствами их автоматизации. Вопрос построения ИТ-архитектуры – один из наиболее приоритетных для развивающихся компаний, следящих за тенденциями развития комплексных информационных технологий масштаба предприятий.

Исходя из всего вышеперечисленного можно сказать, что архитектура КИС – это концепция, определяющая модель, структуру, выполняемые функции и взаимосвязь компонентов ИС масштаба предприятия. Для того чтобы определить, как модель, так и структуру КИС, необходимо выбрать способ организации архитектуры КИС.

Проследим текущее состояние архитектуры современной КИС, опираясь на эволюцию программных архитектур. Распределённые ИС, в которых компоненты распределены по нескольким компьютерам, по типу архитектуры разделяют на: файл-серверные и клиент-серверные ИС [13]. Далее под сервером следует понимать процесс, который обслуживает информационную потребность клиента, а задачей клиента является инициирование связи с сервером.

Файл-серверная архитектура КИСпредставляет наиболее простой случай распределенной обработки данных в ИС, согласно которой на сервере распределяются только файлы данных, а на клиентской части находятся приложения пользователей вместе с СУБД. Использование файл-серверов предполагает, что вся обработка данных выполняется на рабочей станции, а файл-сервер лишь выполняет функции накопителя данных и средств доступа. Взяв за основу файл-серверную архитектуру можно спроектировать, разработать и отладить простое, работающее с небольшими объемами информации и рассчитанное на применение в однопользовательском режиме файл-серверное приложение. Такая архитектура ИС подходит небольшой компании, например, для ведения кадрового учета достаточно иметь изолированную систему, работающую на отдельном компьютере. Минусом такого решения можно назвать риск нарушения надежности хранения и поддержания целостности данных.   
А в сложных случаях, требующих от такой ИС поддержание проекта, выполняемого группой, файл-серверной архитектуры становится недостаточно.

На смену файл-серверной архитектуре со временем пришла клиент-серверная архитектура. Такой вид архитектуры упрощает взаимодействие пользователей с информационной системой и между собой в процессе исполнения бизнес-процессов. Клиент-серверная архитектура в КИС может быть реализована по-разному, а выбор конкретной схемы реализации данной архитектуры основывается на различных наборах требований, связанных с территориальным распределением, эксплуатационной надежности, быстродействием, простотой обслуживания и др. Существует две схемы реализации клиент-серверной архитектуры КИС:

1. двухуровневая клиент-серверная архитектура;
2. трехуровневая клиент-серверная архитектура.

Двухуровневая клиент-серверная архитектура, в отличие от файл-серверной архитектуры основана на том, что данные и СУБД размещаются на серверах, в то время как остальные компоненты программного обеспечения (графический интерфейс пользователя и логика приложений) размещаются на настольных компьютерах («толстые» клиенты). Такая архитектуры ИС решает проблемы файл-серверной архитектуры. Здесь становится возможным полная поддержка многопользовательского режима работы с ИС, а также есть гарантия целостности данных. Масштабируемость, централизованность администрирования и способность архитектуры к развитию позволяют использовать такой тип архитектуры для быстрого создания приложения для небольшого количества пользователей. Двухуровневая клиент-серверная архитектура на сегодняшний день может считать устаревшей, и для построения распределенных КИС она не подходит, в силу существующих недостатков (ограниченная масштабируемость, ограниченное использование сетевых вычислительных ресурсов). Все это снижает быстродействие системы и негативно сказывается на стоимости аппаратных компонент, которые необходимы для ее функционирования.

Для решения этих проблем была предложена трехуровневая клиент-серверная архитектура. Основное ее отличие от двухуровневой состоит в том, что в ней присутствует физическое разделение программ (промежуточное звено), отвечающих за хранение данных (СУБД) от программ, обрабатывающих эти данные (сервер приложений). Такое разделение программных компонент позволило оптимизировать нагрузки, как на сетевое, так и на вычислительное оборудование комплекса. По сравнению с двухуровневой клиент-серверной или файл-серверной архитектурой можно выделить следующие достоинства трехуровневой клиент-серверной архитектуры, это: масштабируемость, конфигурируемость (изолированность уровней друг от друга), высокая безопасность, высокая надежность, низкие требования к скорости канала (сети) между терминалом и сервером приложений, низкие требования к производительности и техническим характеристикам терминалов, как следствие снижение их стоимости.

Трехуровневая архитектура сложнее двухуровневой, но благодаря тому, что функции распределены между серверами второго и третьего уровня, эта архитектура представляет, как высокую степень гибкости и масштабируемости, высокую безопасность (т.к. защиту можно определить для каждого сервиса или уровня), так и высокую производительность (т.к. задачи распределены между серверами).

Несмотря на то, что клиент-серверную архитектуру на сегодняшний день можно назвать классической, сейчас наблюдается устоявшийся интереса к сервис-ориентированной архитектуре (SOA – Service-Oriented architecture), связанный с постоянно возрастающим интересом к Web-сервисам, на технологиях которых основана эта концепция. Практические аспекты сервис-ориентированной технологии позволяют решить проблемы масштабируемости, интегрировать сети передачи данных, упростить процедуры проектирования и управления сетями, а также создать другие распределенные приложения, прозрачно взаимодействующие с ресурсами систем при помощи прикладных программных интерфейсов и открытых стандартов. Следует отметить, что сервис-ориентированная архитектура – это подход к разработке ПО, в основе которого лежат сервисы со стандартизированными интерфейсами. КИС, построенная на основе сервис-ориентированной архитектуре, состоит из набора сущностей, доступных через прикладные программные интерфейсы. Нередко сервис-ориентированную архитектуру называют третьей ступенью развития ИС, последующей за мэйнфреймами и клиент-серверной архитектурой.

Рассмотрим эволюцию архитектуры ИС на примере компании SAP и ее решений. Классическим примером первой ступени мэйнфрейма может служить система SAP R/2. В клиент-серверной архитектуре таким образцом стала широко распространенная SAP R/3, а в качестве сервис-ориентированного решения компания SAP позиционирует интеграционную платформу NetWeaver, в которой сервис-ориентированный подход позволит облегчить и ускорить внедрение КИС на предприятии.

Каждый тип архитектуры имеет ряд преимуществ и недостатков, а также справляется с решением тех или иных задач, продиктованных потребностями компаний на протяжении временного периода развития информационных технологий. Какой бы ни была архитектура ИС, в каждой системе происходит работа с различными ресурсами. Документ, в свою очередь, в не зависимости от его формата и вида, основной ресурс и источник информации для компаний.

Для того чтобы понять связь архитектуры КИС и документа, рассмотрим пример. Чаще всего документ создается в формате некоторого офисного приложения; затем после утверждения, документ публикуется на сайте в HTML-формате или преобразуется в PDF для дальнейшего распространения. С точки зрения потребителя информации это один и тот же документ, а с точки зрения архитектуры ИС – три разных. Содержимое документа, его семантика представляют собой важную составляющую документа, однако в современной архитектуре КИС не предусмотрено наличие семантических технологий, позволяющим усовершенствовать концепции существующих архитектур ИС.

## Электронный документооборот

Автоматизация документооборота компании, в настоящее время, стала одной из типичных задач компании, стремящейся к автоматизации своего производства, а также всех процессов, протекающих в ней. Данное явление обусловлено тем, что документы в компании – это самое распространенное средство поддержки выполнения бизнес-процессов, обеспечивающее фиксацию и перенос информации от одного исполнителя к другому. Таким образом, эффективность выполнения бизнес-процессов организации во многом определяется скоростью и качеством протекания процессов в рамках жизненного цикла документов.

Важность таких ключевых показателей при управлении бизнес-процессами компании, как количество и качество, совместно с развитием технологий автоматизации, неизбежно влечет за собой повсеместное использование электронных документов в организации.

### Понятие и сущность электронного документа

Документы содержат информацию, являющуюся ценным ресурсом и важным элементом деловой деятельности. Системный подход к управлению документами позволяет организациям и обществу защищать и сохранять документы в качестве доказательства действий.

С точки зрения ИС документ представляет собой мощный информационный поток, который нуждается в контроле и тщательной организации, если стоит задача совершенствования бизнес-процессов документооборота. Поскольку базовый термин, на котором фокусируется данная глава – это электронный документ,   
в первую очередь необходимо обозначить, что будет пониматься под понятием «Документ», и, основываясь на данном определении, выявить особенности понятия «Электронный документ».

За время своего существования термин «Документ» не получил единого, общепринятого определения. В различных государственных документноориентированных нормативных актах не содержится однозначного указания на то, что можно считать документом, а что – нет, однако, несмотря на различные толкования термина и его определение, можно опереться   
на “Национальный стандарт РФ. Система стандартов по информации, библиотечному и издательскому делу. Управление документами”. Таким образом, в соответствии с ГОСТ Р ИСО 15489-1-2007 документ – это зафиксированная на материальном носителе идентифицируемая информация любой формы, в виде текста, звукозаписи, изображения и (или) их сочетания, также имеющая реквизиты, созданная, полученная и сохранённая организацией или физическим лицом в качестве доказательства при подтверждении правовых обязательств или деловой деятельности [13].

В определении «Документа» характерны следующие моменты. Документом признается, прежде всего, материальный носитель, в то время как наличие зафиксированной на ней информации является вторичным признаком. Центр фокусировки перенесен на наличие реквизитов. Однако, учитывая наличие разрозненных определений данного понятия, трактовка понятия с той точки зрения, когда документ – это, прежде всего, информация, а материальный носитель – признак вторичный, имеет право на существование. Таким образом, следует отметить, что «Документ», как более обобщенное понятие «Электронного документа» позволяет перенести признаки «Документа» на понятие «Электронного документа» и также считать основой «Электронного документа» информацию.

В новейшей (с изменениями от 1 января 2011 г.) редакции Федерального закона «Об информации, информационных технологиях и о защите информации» (№ 149‑ФЗ) электронный документ понимается как «документированная информация, представленная в электронной форме, то есть   
в виде, пригодном для восприятия человеком с использованием электронных вычислительных машин, а также для передачи по информационно-телекоммуникационным сетям или обработки в информационных системах» [20].

В определении «Электронного документа», соответственно, акцент переносится на то, что это в первую очередь, это информация, которая представлена в электронной форме. Исходя из определения, формируется ряд требований, которым ЭД должен отвечать:

* содержать реквизиты, которые позволяют подтвердить его целостность и подлинность;
* создаваться, обрабатываться, храниться, приниматься и передаваться при помощи технических программных средств;
* быть воспроизведенным (отображенным) в понятной для восприятия форме.

Очевидно, что при использовании ЭД есть как положительные, так и отрицательные стороны. Преимущества применения ЭД:

* упрощается и ускоряется сбор, обработка и систематизация всей информации;
* существенно ускоряется компьютерная обработка докумен­тов;
* оперативная (практически мгновенная) передача документов на любые расстояния при рассылке их по электронной почте;
* при хранении электронных документов не требуются большие площади для архивов;
* создание электронных архивов позволяет использовать уже полученную информацию для подготовки новых документов;
* повышается эффективность работы;
* наличие базы шаблонов документов уменьшает время на создание новых;
* уменьшение расходования бумаги при работе с проектом электронного документа и внесении каких-либо изменений до вывода на бумагу;
* рассылка проекта документа одновременно нескольким визирующим;
* рассылка документа одновременно нескольким адресатам по электронной почте.

Недостатки ЭД:

* необходимость наличия определенных технических средств для создания и дальнейшей работы, а также хранения документов;
* защита информации, содержащейся в документах, от несанкционированных изменений.

Существует еще одна проблема, которая вытекает из особенности представления информации в электронном виде - копирование электронной информации. Одним из наиболее важных преимуществ, представленной в электронной форме информации, является легкость ее копирования. Поэтому при передаче по каналам связи электронного документа его необходимо защитить, во избежание инцидентов, от несанкционированного копирования. Однако на передающем конце остается контрольная копия. Она может рассматриваться как заверенная исполнителем копия документа. Но если заверить ее электронно-цифровой подписью, снабдив при этом всеми необходимыми атрибутами, тогда будет изготовлен один документ в двух экземплярах, которые имеют равную юридическую силу.

Следовательно, ЭД можно приравнять, проводя аналогию:

* к юридически значимому документу - электронный документ при передаче его со всеми соответствующими атрибутами по каналам связи;
* к рабочей копии документа - электронную копию документа, которая осталась у респондента;
* ко второму экземпляру или заверенной копии самого юридически значимого документа, которые равносильны первому экземпляру - электронную копию документа, оставшуюся у респондента и заверенную электронно-цифровой подписью;
* к заверенной копии документа - его распечатку, которая заверена в установленном законом порядке.

Кроме проблем, вытекающих из рамок правового статуса ЭД, есть еще один трудный вопрос, требующий разрешения, особенно в условиях текущего развития ИТ, в частности, развитие систем электронного документооборота. Проблема заключается в том, что, в настоящее время, КИС, работающие с ЭД, в большинстве своем, оперируют документом, как целостной единицей. Такая концепция принятия документа не позволяет извлечь из документа данные, которые содержаться непосредственно внутри неструктурированного ЭД. В свою очередь, потенциальные возможности технологий «интеллектуализации» содержат множество возможностей для обработки большого количества документов и данных, содержащихся в них.

Данные, находящиеся внутри документа являются неструктурированными, следовательно, для работы с ними, а также для последующего превращения человеком данных в информацию, необходим подход, позволяющий рассматривать документ, как набор структурированных данных.

Рассмотрев некоторые особенности ЭД, преимущества и недостатки его использования, а также некоторые юридические аспекты, следует обратить внимание также на структуру ЭД и базовые задачи управления документами, которые объяснят причину рассмотрения концепции документа как объекта, состоящего из множества составляющих, полезных для бизнес-процессов документооборота.

### Недостатки традиционного подхода к ЭД

Накопление документов, существующих в электронном виде, а также критически важных для бизнеса, растет в компаниях в быстрых темпах. Одной из основных проблем последнего десятилетия является не столько количество документов, сколько потребность во взаимосвязанной системе для работы с ними.

Раньше, в большинстве случаев, система управления документами поддерживалась за счет разрозненных средств, начиная от папок на файловой системе и электронной почты и заканчивая специализированными приложениями. При этом бизнес-логика, иными словами совокупность правил, принципов и зависимостей работы с документами в ИС, формально нигде не была закреплена, а с документами работали сотрудники отдела документационного отдела. Соответственно, возникла потребность в системе, которая решает проблему организации информации в рамках единого процесса.

Первые три десятилетия развития информационных технологий основное внимание уделялось структурированным данным, процессам их сбора, хранения и обработки. Для работы с неструктурированной информацией использовались простые инструменты, на примере текстовых редакторов. Однако в конце 90-ых годов сложилась ситуация критической точки, когда более 80-ти % всей корпоративной информации стало храниться в виде неструктурированных документов, при этом системы, которые должны были обеспечивать надежное хранение и доступ отсутствовали.

Таким образом, следует вывод о том, что накопленные массивы электронной информации, хранящейся в электронных документах, являются ценнейшим интеллектуальным активом организации, ее корпоративной памятью, следовательно, нуждаются в механизмах анализа содержимого документов. При этом интеллектуальные активы должны быть легкодоступны и надежно сохранены. Обеспечить эти потребности позволяет концепция управления корпоративной информацией призванная объединить разрозненные технологии, поддерживающие работу с неструктурированной информацией.

КИС, обладающая такой концепцией обработки корпоративной информации, должна быть настраиваема и адаптируема к окружающим внешним факторам. Невыполнение этого требования может привести к тому, что ИС, по истечении некоторого времени, перестанет выполнять свои функции из-за изменений в организационной структуре компании, ее внешнем окружении, спектре выполняемых работ. Все эти изменения неизбежно ведут к соответствующим изменениям системы понятий, которыми оперирует компания, тематике и структуре документов, классификации документов и способах организации архива.

Проблемы, возникающие при создании такой КИС, не охватываются традиционными подходами к автоматизации документооборота в компании и не решаются статистическими и лингвистическими методами полнотекстового поиска в базе данных документов.

В таком случае требуется применение интеллектуальных решений, семантических технологий с ориентацией задач управления документами (поиск, анализ, классификация, каталогизация, хранение) на понимание информационной системой содержания документа, его семантики. Одной из главных проблем анализа неструктурированных электронных документов является большое количество разнородного содержимого в документе. Если каждый элемент (слово или строку) подвергать анализу, то время извлечения данных и знаний резко возрастет и, в таком случае, не будет удовлетворять требованиям пользователей.

В то же время очевидно, что не все элементы документа несут полезную информацию, а в силу гибкости естественных языков формально различные слова (синонимы и пр.) на самом деле означают одинаковое понятие. Все эти процессы проходят в рамках выполнения процесса семантического индексирования – механизма, с помощью которого поисковый инструмент определяет не только соответствие содержимого документа ключевым запрашиваемым словам, но и определяет неявные семантические связи ключевого запроса. Таким образом, устранение описанных проблем выполняется на этапе предварительной обработки документа.

В связи с этим, первым этапом семантического индексирования, предшествующим непосредственному извлечению данных из документа и последующей их обработки, является этап структурирования данных документа. При системном подходе к вопросу извлечения из ЭД знаний, необходимо иметь соответствующий подход к определению структуры документа и его элементов.

### Структура электронного документа

Согласно определению «документа», формулировку которого в 1997 году дает спецификация организации DMA (Document Management Alliance) [32]: «Документ – это набор информации, который имеет отношение к выделенной теме или к родственным темам». Это определение является обобщающим, не привязывает документ к наличию носителя или реквизитов. Объект неструктурированных данных, другими словами, называемый документом, может принимать множество различных форм: это может быть закодированный текст, оцифрованное изображение бумажного документа, графика, выходные файлы САПР (система автоматизированного проектирования), цифровое видео. Таким образом, любой типы неструктурированных данных, которые могут быть автоматизированы, храниться в цифровой форме и воспроизводиться в виде, воспринимаемом человеком – это документ.

Документам, как живым и интеллектуальным объектам, принадлежит три основные характеристики, за которые ответственны три различных технологических средства. Во-первых, это объекты хранения всевозможных данных от текста до графики и звука. Во-вторых, это связи, которые «оживляют» документы. В-третьих, это атрибуты и индексы, появляющиеся в результате обработки документов в ИС. Все три характеристики можно назвать синтаксической нотацией документа, с которой также связано понятие стиля документа – его форма, проявляющаяся при воспроизведении документа при его обработке текстовыми редакторами.

По всем признакам документ является структурированным объектом, однако структурированным частично. Определение документа, как неструктурированного позволяет перейти от концепции построения архитектур ИС, ориентированных на построение полнотекстовых баз данных хранения документов к архитектуре с гибким и универсальным хранилищем данных с интеллектуальными технологиями анализа содержимого документов. Для этого необходимо учесть следующие характеристики структуры документа, накладываемые понятием ЭД:

1. структура документа, то есть его формат и взаимосвязи между составляющими документ элементами, должна оставаться неизменной;
2. в документе должен быть отражен контекст его создания, получения и использования документа (в том числе процесс деловой деятельности, частью которого является данная операция, дата и время данной операции и ее участники);
3. должны быть представлены связи между отдельными записями, составляющими в совокупности документ.

Вышеописанные характеристики структуры ЭД отображены на схематическом изображении модели ЭД (см. рис. 1.1). Рассмотрим схему более подробно.

|  |
| --- |
|  |
| Схема структуры ЭД |

Как видно на схеме структуры ЭД, структура документа имеет несколько основных составляющих: содержимое и метаданные. Содержимое документа состоит из элементов, которые, с вою очередь, могут быть связаны друг с другом и состоять из различных реквизитов. Набор реквизитов – обязательных элементов оформления официальных документов, зависит от вида документа [11] и в первую очередь позволяет определить структуру ЭД для человека, тем самым являясь важной частью документа.

Состав реквизитов, использующиеся, к примеру, при подготовке и оформлении организационно-распорядительных документов (см. приложение A), определяется ГОСТом Р 6.30-2003 «Унифицированные системы документации. Унифицированная система организационно-распорядительной документации. Требования к оформлению документов» [20]. Кроме того, данный стандарт устанавливает требования к оформлению реквизитов и закрепляет схемы их расположения на макете документа. Несмотря на то, что в стандарте описаны 30 реквизитов, необходимо помнить, что различные документы состоят из разного набора реквизитов. Более того, некоторые реквизиты являются взаимоисключающими, например, «Адресат» и «Гриф утверждения». Данные реквизиты располагаются в одной и той же части макета документа и, с точки зрения документационного управления, занимают одну и ту же площадь в отведенной им части, тем самым, являясь взаимоисключающими.

Содержательный элемент документа может иметь два типа: структурированные и полуструктурированные (неструктурированные). Особенностью структурированных документов является наличие фиксированной структуры для сбора данных, четко обозначенных полей, некоторых служебных элементов. Примером структурированного документа может служить бланк регистрации, изображенный на рис. 1.2, а. Использование такого типа документов значительно повышает время и качество ввода документа в ИС, распознавания образа, обработки документа.

Полуструктурированные (неструктурированные) документы отличаются от структурированных тем, что их поля не имеют фиксированного размера и расположения. В таком случае при распознавании ИС ориентируется на статический текст, осуществляя его поиск по всему документу, что требует значительных затрат времени и снижает качество распознавания изображения. Пример полуструктурированного документа в виде извещения на оплату представлен на рис. 1.2, b.

|  |  |
| --- | --- |
|  |  |
|  |  |
| Типы форм документов: a) пример структурированной формы,  b) Пример полуструктурированной формы | |

Независимо от типа формы документа, «машиночитаемая форма» содержит следующие основные элементы: информационные поля, идентификаторы, сервисные поля и декоративные элементы. Каждый элемент формы имеет свое значение.

Информационные поля служат для ввода текста при заполнении формы, в них размещаются группы знакомест, метки, группы меток. Сервисные поля позволяют выявлять принадлежности формы к определенному образцу и служат для ориентации документа во время ввода в ИС, за счет фокусировки на статическом тексте и реперных точках (черных квадратов). Идентификаторами служат штрих-коды, однозначно характеризующие принадлежность формы к определенному образцу (в том случае, если реперные точки у форм различных образцов совпадают). Декоративные элементы, такие как подписи и печати не подлежат распознаванию.

Кроме содержимого, обладающего набором элементов и реквизитов, в ЭД необходимо наличие метаданных об управлении документами, а также объясняет принципы, лежащие в основе управления ими. Сформулировать требования и основы для создания и управления метаданными позволяет национальный стандарт Российской Федерации «Система стандартов по информации, библиотечному и издательскому делу. Процессы управления документами. Метаданные для документов» ГОСТ Р ИСО 23081-1-2008 [13].

В стандарте [13] говорится, что управление метаданными – это неотъемлемая часть управления документами, обеспечивающая выполнение множества различных функций и целей. В рамках управления документами метаданные определяются как данные, описывающие контекст, содержание и структуру документов, а также процессы управления ими на протяжении всего жизненного цикла (ЖЦ) документа. Сами по себе метаданные являются структурированной или полуструктурированной информацией, позволяющей создавать, регистрировать, классифицировать, обеспечивать доступ, хранение и размещение документов, в рамках и за пределами одной или нескольких областей деятельности.

В схеме структуры ЭД (см. рис. 1.1) обозначены два типа метаданных: метаданные свойств содержимого и метаданные контекстной связи. Метаданные свойств содержимого допускается использовать для идентификации, подтверждения подлинности и статуса документа. Чем больше и качественнее определен набор метаданных свойств содержимого, тем больше возможности использования метаданных в процессе управления документами в ИС. Метаданные контекстной связи, в свою очередь, содержат информацию о контекстуализации документов, процессов и систем, в рамках которых создаются, управляются и используются ЭД. Такие метаданные важны при анализе и оптимизации бизнес-процессов. В целом, все метаданные представляют интерес в рамках управления ЭД на протяжении ЖЦ ЭД.

Управление документами всегда включает в себя управление метаданными. Однако электронная среда требует иного выражения традиционных требований и иных механизмов для идентификации, ввода, определения свойств и использования метаданных. В электронной среде документы нормативного характера – это документы, сопровождаемые метаданными, определяющими их нормативный статус. Этот нормативный статус должен быть четко документирован, а не подразумеваться, как в некоторых процессах управления бумажными документами. С ЭД очень важно предоставить гарантию того, чтобы в ИС, которые создают, осуществляют ввод и управление документами, осуществлялось создание и ввод метаданных об управлении документами. И наоборот, электронная среда предоставляет новые возможности для определения и создания метаданных, а также для обеспечения полного, своевременного ввода документов.

Важно отметить, что структура ЭД может рассматривать с нескольких точек зрения. Структура документа, воспринимаемая человеком, и структура, необходимая ИС для обработки документа (машиночитаемая), различаются. Более того, в документе содержится, как структурированная информация, так и неструктурированная. Для работы с неструктурированными данными, их необходимо первоначально обработать. Также отмечено, что каким бы ни был тип документа, структурированным или полуструктурированным, при управлении ЭД в ИС должны учитываться метаданные.

На основе построенной схемы структуры ЭД, а также описанной информационной модели ЭД в работе [19], была построена информационная модель структуры ЭД (см. рис. 1.3), которая отображает основные информационные блоки ЭД, с которыми предстоит работа в ИС. Информационная модель – это модель объекта, в нашем случае ЭД, представленная в виде информации, описывающей существенные для данного рассмотрения параметры и переменные величины, связи между ними, входы и выходы объекта и позволяющая путём подачи на модель информации об изменениях входных величин моделировать возможные состояния объекта.

|  |
| --- |
|  |
| Информационная модель структуры ЭД |

В целом, модель состоит из трех блоков, соответствующих составным частям схемы структуры ЭД. Соответствие элементов схемы и информационной модели представлено в табл. 2.1. Далее рассмотрим предназначение каждой модели, входящей в состав всей информационной модели структуры ЭД.

##### Соответствие элементов схемы и информационной модели структуры ЭД

|  |  |
| --- | --- |
| Элемент схемы  структуры ЭД | Элемент информационной модели  структуры ЭД |
| Содержание | Модель семантического описания |
| Метаданные свойств содержимого | Модель описания свойств содержимого |
| Метаданные контекстной связи | Модель описания контекстной связи |

Модель семантического описания представляет собой набор конечного числа семантических сущностей, иными словами описание следующих элементов: объектов, сущностей, понятий содержания документа. Также в этой модели учитываются отношения между этими элементами, состояния, причины и следствия.

В свою очередь, модель описания свойств содержимого выражается в описании статических характеристик, как общих для всех типов документов (создатель, формат, дата создания), так и дополнительных, характерных для определенного контекста документов. Так, например характеристика «Идентификатор» или «Дата создания документа» являются обязательными для всех документов, в то время как такая характеристика, как «Приложение, в котором создан документ» будет встречаться только у документов, созданных с помощью приложений.

Модель описания контекстной связи необходима для описания движения документа и принадлежности его к бизнес-процессам. Для этого в данной модели предусмотрены такие характеристики, как «переход», «варианты» и «состояния». Последовательность переходов, которой обладает документ, выражается в показателях: «время», «характер», «причина» и др., а характеристика «варианты» отображает потенциальные способы переходов, заложенные в маршруты движения ЭД. В свою очередь характеристика «состояния» отображает перечень состояний документа на пути его движения. Более того, вся модель описания контекстной связи функционирует в условиях временной шкалы.

Таким образом, в данном разделе была обозначена схема структуры ЭД и модель структуры ЭД, что позволит в последующих главах проектировать архитектуру КИС с базовым объектом – документ, в противовес стандартным системам электронного документооборота, в которых основным объектом является запись в базе данных. Кроме того, исходя из понятия и структуры ЭД, целесообразно также обозначить жизненный цикл ЭД, который отражает все фазы ЭД, которые присутствуют во время работы с ЭД в ИС.

### Жизненный цикл электронного документа

В любой организации и вне ее происходит непрерывный процесс движения ЭД, а также, что немаловажно информации и знаний. Жизненный цикл ЭД отличается от традиционного «бумажного» документа большей гибкостью и прозрачностью.

Проводя анализ электронного документооборота (ЭДО), в качестве объекта исследования можно рассматривать отдельный документ (бумажный / электронный), документопоток, документационный технологический процесс. Однако основополагающим объектом исследования и совершенствования можно назвать документопоток.

Согласно ГОСТ Р 51141-98 [5] документопоток – это поток документов, циркулирующих между пунктами обработки и создания информации (руководителями организации и структурных подразделений, специалистами) и пунктами технической обработки документов: секретариатом, канцелярией и др.

В каждой организации можно выделить следующие потоки документов:

1. поток входящей документации, состоящий из поступающих в организацию документов;
2. потоки внутренней документации, состоящие из документов, созданных в организации и не предназначенных к выходу за ее пределы;
3. поток исходящей документации, состоящий из документов, предназначенных для отправки в другие организации [5].

**Основные этапы ЖЦ ЭД**

Исходя из определения ЭД, можно сделать вывод о том, что управление корпоративными ЭД – это технология, используемая для ввода, управления, хранения, сохранности, доставки и вывода, как документов, так и содержимого этих документов, относящихся к важнейшим процессам организации. За основу жизненного цикла ЭД возьмем принятую в концепции систем типа ECM структурную схему управления корпоративным контентом, предложенную ассоциацией по вопросам управления информацией и изображениями (AIIM ‑ Association for Information and Image Management) [33].

ЖЦ ЭД в ИС будет состоять из следующих этапов:

1. ввод документа в ИС;
2. хранение документа в ИС;
3. обеспечение сохранности документа в ИС;
4. управление документом;
5. доставка документа;
6. вывод документа из ИС.

Кроме этого, следует обозначить, что ЖЦ, состоящий из шести вышеописанных этапов, также должен включать в себя целый перечень состояний документа, что является очень важным компонентом в понимании управления ЭД. Список состояний документа, принятый в системе описания организации – ARIS [13] представляется достаточным и наиболее полным и включает в себя: планирование; создание (разработка); согласование; визирование; утверждение; регистрация; публикация; регламентация доступа; извещение об изменениях; оперативное хранение; выполнение основных функций; архивное хранение информации; экспертиза ценности; уничтожение.

На основе ЖЦ ЭД становится возможным определение таких понятий как управление ЭД и электронный документооборот (ЭДО).

### Стандарты по управлению ЭД

Согласно основным этапам ЖЦ ЭД обратимся к нормативным документам и стандартам, относящимся к управлению электронными документами в организации, содержащими четко сформулированные требования к работе с документами и информацией. В настоящее время общепризнанными международным сообществом документами, устанавливающими требования к управлению ЭД, являются следующие стандарты: ГОСТ Р ИСО 15489-1 – 2007, ГОСТ Р ИСО 23081-1 – 2008, ISO 15836:2003, ISO/TR 15801:2004, ISO/TR 18492:2005, ISO/TR 12037:1998, ISO 2142:2001, ISO 19005-1:2005, XKMS, MoReq2.

Рассмотрим каждый стандарт более подробно.

1. ГОСТ Р ИСО 15489-1 – 2007 «Система стандартов по информации, библиотечному и издательскому делу. Управление документами. Общие требования» [16] идентичен международному стандарту ISO 15489‑1:2001 (Information and documentation. Records management. General). Данный стандарт устанавливает требования, предъявляемые к документам, создаваемым в процессе осуществления деловой деятельности: аутентичность (подлинность), достоверность, целостность (неизменность), пригодность для использования.

Стандарт ГОСТ Р ИСО 15489-1 – 2007 является высокоуровневым нормативным документов, в котором отражены основные принципы современного управления документами. С практической точки зрения стандарт может быть полезен для того, чтобы обозначить для начальства и сотрудников подразделений важность управления документами в деловой деятельности. К тому же, с точки зрения ЖЦ ЭД, стандарт устанавливает требования ко всем этапам ЖЦ.

Кроме того, следует обратить внимание на такие характеристики, которые устанавливают требования к проектированию систем управления документами, в соответствии с которыми система должна обладать функциональными характеристиками, позволяющими выполнять и поддерживать различные процессы документооборота.

1. ГОСТ Р ИСО 23081-1 – 2008 «Система стандартов по информации, библиотечному и издательскому делу. Управление документами. Процессы управления документами. Метаданные для документов» [13] идентичен международному стандарту ISO 23081 (Metadata for records).

В стандарте подчеркивается значимость метаданных об управлении документами в различных процессах деловой деятельности и различных ролей и типов метаданных, а также объясняются принципы и структуру, лежащие в основе управления ими.

1. ISO 15836:2003 (The Dublin Core metadata element set) [34] – «Набор элементов метаданных “Дублинского ядра”» представляет собой универсальный набор метаданных, содержащихся в Дублинском ядре. Указания по применения Дублинского ядра содержатся во многих государственных стандартах метаданных, основанных на данном стандарте.
2. ISO/TR 15801:2004 (Information Stored electronically – Recommendations for trustworthiness and reliability) [23] – «Информация, хранимая в электронном виде – Рекомендации по обеспечению достоверности и надежности» является международным стандартом. Стандарт регламентирует внедрение и эксплуатацию систем управления информацией, которые хранят информацию в электронном виде и для которых важны достоверность, надежность, аутентичность и целостность информации. Стандарт охватывает не полный ЖЦ ЭД, затрагивая только этапы ввода, использования и управления документами.
3. ISO/TR 18492:2005 (Long-term preservation of electronic document-based information) [25] – «Обеспечение долговременной сохранности электронной документированной информации» является техническим отчетом и первым международным руководством по решению проблемы обеспечения долговременной сохранности электронных документов.

В отчете предлагается концепция разработки стратегий обеспечения долговременной сохранности, а также описываются зарекомендовавшие себя приемы практической работы, которые можно применять в организациях для обеспечения долговременной доступности и аутентичности электронных документов.

1. ISO/TR 12037:1998 (Recommendations for the expungement of information recorded on write-once optical media) [28] – «Рекомендации по удалению информации, записанной на оптических носителях однократной записи». В данном стандарте рассматривается достаточно узкая проблема частичного уничтожения информации на носителе однократной записи.
2. ISO 12142:2001 (Media error monitoring and reporting techniques for verification of stored data on optical digital data disks) [26] – «Мониторинг ошибок при передаче и методики отчетности для верификации сохраненных данных на оптических цифровых дисках» концентрируется на узкой проблеме передачи данных для сохранности.
3. ISO 19005-1:2005 (Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1)) [23] – «Файловый формат для долговременного хранения электронных документов – Часть 1: Использование формата PDF 1.4 (PDF/A-1)». Данный международный стандарт разработан для описания принципов и методов использования формата PDF для долговременного хранения документа, как в организациях, так и в архивах.

Значение стандарта заключается в том, что коммерческие и государственные организации, библиотеки, архивы, другие учреждения во всем мире используют формат PDF для представления важной информации в неизменном виде. Большая часть такой информации должна храниться длительное время, некоторая доля информации должна храниться постоянно. Цели, которые реализует стандарт:

* определение формата файла, основанного на PDF, который обеспечит механизм для представления электронных документов таким способом, который позволит сохранить их визуальное изображение с течением времени;
* определение формы для представления логической структуры истории электронных документов в метаданные в рамках соответствующих файлов;
* обеспечение структуры для документирования контекста и другой семантической информации об электронных документах в пределах формата.

1. XKMS (XML Key Management Specification) [37] – «Спецификация управления ключами XML» является набором протоколов, разработанный W3C, который описывает распространение и регистрацию открытых ключей, применимый для использования совместно со стандартом XML, определенным W3C и IETF.

Данный стандарт подходит, например, для того, чтобы СЭД имела разработанный на основе стандартов интерфейс, позволяющий подключать новые технологии электронных подписей по мере их появления.

1. MoReq2 (Model Requirements for the Management of Electronic Records) [27] – общеевропейская функциональная спецификация по управлению электронными документами, описывающая функциональные возможности систем управления документами и являющаяся инструментом тестирования систем на соответствие требованиям.

Требования, обозначенные в данном стандарте охватывают все этапы ЖЦ ЭД, и изложены таким образом, что они одинаково и однозначно понятны сотрудникам-документоведам и ИТ-специалистам, в связи, с чем стандарт MoReq2 завоевал популярность в качестве «протокола обмена» между пользователями и разработчиками.

Особенности спецификации MoReq2:

* унификация со стандартом ISO 15489 терминология в области ЭД;
* учет ряда национальных стандартов таких стран, как: Великобритания, Германия, Норвегия, США и др.;
* наличие систем тестов для оценки и сертификации СЭД;
* разделение спецификации на две части: базовый и опциональные модули;
* содержит «типовые» требования и носит общий характер (вопросы специфичные для секторов экономики не рассматриваются).

В данном разделе приведена таблица, которая показывает, как ключевые стандарты связаны с процессами управления ЭД. Согласно обозначенным в п.1.2.4. этапам ЖЦ ЭД, сравним рассмотренные выше стандарты с точки зрения наличия требований к этапам. Первый этап ввода документа в ИС, в рамках сравнительного анализа, разделим на два подпункта: «создание» и «ввод»; этап управления дополним подпунктом «использование». Таким образом, сводная таблица стандартов управления ЭД представлена в табл. 2.2.

##### Сравнительный анализ стандартов по управлению ЭД

| № п/п | Стандарт | Создание | Ввод | Обеспечение сохранности | Управление | Использование | Хранение | Передача | Уничтожение |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | ГОСТ Р ИСО 15489-1 – 2007 | **+** | **+** | **+** | **+** | **+** | **+** | **+** | **+** |
| 2 | ГОСТ Р ИСО 23081-1 – 2008 |  | **+** | **+** | **+** | **+** |  | **+** |  |
| 3 | ISO 15836:2003 |  | **+** | **+** | **+** | **+** |  | **+** |  |
| 4 | ISO/TR 15801:2004 |  | **+** |  | **+** | **+** |  |  |  |
| 5 | ISO/TR 18492:2005 |  |  | **+** |  |  |  |  |  |
| 6 | ISO/TR 12037:1998 |  |  |  |  |  |  |  | **+** |
| 7 | ISO 12142:2001 |  |  |  |  |  | **+** |  |  |
| 8 | ISO 19005-1:2005 | **+** | **+** | **+** |  |  |  |  |  |
| 9 | XKMS |  | **+** | **+** |  |  |  |  |  |
| 10 | MoReq2 | **+** | **+** | **+** | **+** | **+** | **+** | **+** | **+** |

Как видно из сводной таблицы стандарт ГОСТ Р ИСО 15489-1 – 2007 и спецификация MoReq2 обладают набором требований для каждого этапа ЖЦ ЭД и являются обобщающими, в то время как остальные стандарты являются более характерными для разных этапов ЖЦ. Так, к примеру, стандарты метаданных имеют отношение к процессам ввода, использования, обеспечения сохранности, передачи и управления. В их число входят ГОСТ Р ИСО 23081-1 – 2008, охватывающий все виды метаданных для управления документами, и ISO 15836:2003 «Дублинское ядро», специфицирующее стандартный набор метаданных для целей поиска.

Для процесса создания документов наибольший интерес представляют стандарты форматов документов. Примером такого стандарта форматов был рассмотрен стандарт ISO 19005-1:2005, в котором разработано и описано подмножество формата PDF (Portable Document Format – «Переносимый формат для информационных материалов») – PDF-A для использования при архивном хранении.

К процессу ввода имеют отношение различные стандарты метаданных. К этому процессу также относятся некоторые из стандартов форматов, - с точки зрения возможности автоматического извлечения значений метаданных; и стандарты, затрагивающие правовые вопросы, например ISO/TR 15801:2004.

Ключевым стандартом в области обеспечения сохранности ЭД и информации является ISO/TR 18492:2005, содержащий концепцию разработки и управления процессами обеспечения сохранности. Кроме того, большая часть деятельности по обеспечению сохранности существенно опирается на использование стандартов метаданных, а также стандарта ISO 19005-1:2005, определяющего формат, подходящий для длительного хранения документов. Более того, вопрос обеспечения сохранности затрагивается и в стандарте XKMS.

Стандарты метаданных могут поддерживать процессы управления доступом и срокам хранения, а также процессы передачи документов. Также на этапах управления и использования применимы стандарты, затрагивающие правовые вопросы, а именно ISO/TR 15801:2004.

Для этапов процесса ЖЦ ЭД хранение и уничтожение рассмотрены индивидуальные для каждого этапа стандарты. В качестве стандарта в области хранения рассмотрен ISO 12142:2001, фокусирующийся на отдельных аспектах процесса хранения, связанных с хранением информации на оптических дисках. А для процесса уничтожения рассмотрен стандарт ISO/TR 12037:1998, фокусирующийся на отдельных аспектах процесса уничтожения, а именно, удалении.

В результате обзора и анализа, рассмотренных в данном разделе стандартов, можно сделать вывод о том, что выбранные стандарты позволяют очертить общую картину взаимосвязи процессов, затрагивающих ЭД, и стандартов, регламентирующих и предлагающих требования для работы с ЭД. Кроме того, формирование требований к разрабатываемой архитектуре будет основываться на анализе требований, предъявляемых стандартами на определенном этапе ЖЦ ЭД. Таким образом, можно сказать, что за основу требований можно взять обобщающие стандарты, а для определения более характерных этапам требований ссылаться на стандарты, отвечающие на определенные для этапа вопросы.

### Понятие электронного документооборота

В настоящее время в Российской Федерации, в частности, действует ряд общегосударственных нормативно-правовых актов, регламентирующих наиболее общие правила подготовки, оформления и организации работы с документами в организации.

В проекте Федерального закона «Об электронном документе» №107599­3 [12], который будет регулировать отношения, связанные с формированием и использованием электронных документов, а также, определяющий основные требования, предъявляемые к ЭД в электронном документообороте (ЭДО), дается следующее понятие ЭДО: электронный документооборот – процесс формирования, обработки, хранения и обмена электронными документами. Несмотря на прозрачность и понятность определения термина «Электронный документооборот», в рамках данной работы особого внимания требует обозначение места ЭДО в процессах КИС.

В определении ЭДО делается упор не только на движение документов, но и на то, что это процесс должен быть организован в электронной среде. Важно также и то, что документооборот зависит от системы управления, является вторичным по отношению к ней, но в то же время именно документооборот, отражая систему управления, позволяет ее наглядно увидеть. Поэтому вторичность порядка движения документов по отношению к целям и задачам организации, документооборот рекомендуется нормировать и регулировать.

Независимо от того, какая организация использует ЭДО, в рамках каждой компании образуется набор ЭД, характерный для вида деятельности конкретной компании. В связи с этим существует достаточное количество видов электронного документооборота, такие как:

* производственный документооборот;
* управленческий документооборот;
* архивное дело (совокупность процедур архивного документооборота);
* кадровый документооборот (процедуры кадрового учета);
* бухгалтерский документооборот и др.

Очевидно, что систем электронного документооборота может быть столько же, сколько существует видов деятельности, как следствие, ИС, автоматизирующие частные виды документооборота, развиваются по направлению массовости.

Более детально, какие базовые виды детальности, понимаются под ЭДО, можно понять, если разделить понятия делопроизводство и документооборот. Электронным делопроизводством называется совокупность операций над документами, а это:

* регистрация входящего документа (автоматическое или ручное задание реквизитов);
* создание исходящего документа;
* сохранение в архиве (публикация);
* внесение изменений (дополнений) в документы и их реквизиты;
* поиск документов;
* задание поручений и контроль исполнения.

В свою очередь, делопроизводство, снабженное дополнительными операциями:

* функциями маршрутизации;
* регистрацией содержания документов;
* выполнением бизнес-процессов, связанных с документом,

называется электронным документооборотом[1]. Это определение охватывает большое число реально функционирующих СЭД - систем автоматизации работы с документами на протяжении всего их жизненного цикла, однако, его осмысление в большей степени зависит от определения жизненного цикла ЭД.

СЭД являются старейшим ИТ-компонентом КИС. Это подтверждает тот факт, что, в частности, широкое внедрение ИТ двадцать лет назад в России начиналось именно с решения автоматизации делопроизводственных задач [9]. Сейчас, в большинстве случаев, СЭД является интегрированной частью общей ИС предприятия. Однако само понятие СЭД со временем претерпевает эволюционные изменения, поскольку расширяется спектр решаемых такими системами задач, выходя за традиционные для России СЭД рамки. Переход от систем для документооборота, решающих ограниченный набор задач (организация распорядительного документооборота и автоматизация канцелярии) к единым технологиям работы документами, пронизывающим все аспекты деятельности организации. Такие технологии называются ECM-технологиями. ECM-система, как и СЭД является частным случаем КИС.

Более того, несмотря на наличие СЭД и ECM-систем, в данной работе принимаются во внимание ИС масштаба всей организации, поскольку СЭД – системы ограничивающиеся ЭДО, в то время как в данном исследовании представляет интерес движение ЭД в рамках всей КИС компании. Именно поэтому в следующем разделе будет рассматриваться применение различных технологий в КИС для работы с ЭД и их анализа.

## Семантические технологии в КИС

Основное направление данной работы – это переход от стандартных методов обработки электронных документов (создание документа, поиск по ключевым словам, редактирование) к методам работы с семантикой (обнаружение смысловых связей, формирование аннотаций) ЭД. Методы работы с семантикой, а также их программная реализация, называются семантическими технологиями. Использование семантических технологий в архитектуре КИС рассматривается в качестве средства повышения качества обработки электронных документов, увеличение функциональности проектируемой архитектуры КИС.

Обзор литературных источников по проблеме проектирования и создания КИС [8,9,18], а также по интеллектуальным методам извлечения знаний из неструктурированных документов [11,22,31] обнаруживает два значительно отличающихся друг от друга подхода к проектированию архитектуры КИС и ее функционалу.

Первый подход – классический, строится на основе комбинирования существующих технологий для поддержки различных процессов работы с документами. К таким технологиям относятся корпоративные порталы, форумы, системы управления документами. Второй подход – семантический, основан на использовании методов и технологий работы со смыслом, семантикой данных, информации и документов (системы, основанные на технологии построения онтологий предметной области, системы поддержки принятия решений, экспертные системы, в которых реализован семантический поиск, автоматическое аннотирование).

Оба подхода не противоречат друг другу, и в рамках данной работы будут использоваться совместно. В настоящем разделе рассмотрим, чем каждый из подходов может быть полезен в процессе совершенствования архитектуры КИС, в рамках ЭДО.

### Обзор технологий классического подхода для проектирования архитектуры КИС

В качестве технологий, принадлежащих первому подходу (классическому), целесообразно рассмотреть такие технологии, которые являются составляющими комплексной системы автоматизации документооборота. Рассмотрим СЭД как композицию из трех информационных технологий, реализующих работу с ЭД, таких как:

* корпоративный информационный портал;
* система управления документами;
* WorkFlow‑система (система автоматизации бизнес-процессов).

**Корпоративный информационный портал** (Enterprise Information Portal, EIP) – это информационная автоматизированная система, предоставляющая сотрудникам компании в едином порядке необходимые ресурсы. Первое определение термина «Корпоративные информационные порталы» позволяет отметить назначение, для которого порталы созданы: «Корпоративные информационные порталы – это приложения, которые позволяют раскрывать информацию, хранящуюся внутри и вне организации и предоставить каждому пользователю единую точку доступа к предназначенной для него информации, необходимой для принятия обоснованных управленческих решений» [30].

Порталы первого периода использовались в качестве инструмента для публикации информации от руководства для сотрудников компании. На корпоративном портале можно было ознакомиться с последними организационными изменениями, новостями. Кроме того, портал выполнял функции просто файлового хранилища.

С развитием технологий функции порталов расширялись. Порталы стали использоваться для поиска необходимой информации, таким образом, портальные решения стали развиваться в направлении структурирования и классификации данных, а также повышения скорости и эффективности поиска необходимой информации. Основываясь на вышеописанных задачах, корпоративные порталы стали выполнять функции баз знаний в различных предметных областях связанных с деятельностью компании.

Решая ряд самостоятельных задач, корпоративные информационные порталы являются также ключевым элементом рынка средств управления корпоративным контентом, обеспечивая функции коллективной работы и поддержки документно-ориентированных бизнес-процессов. Упрощенно процесс работы корпоративного портала можно разбить на несколько этапов:

1. В портал из различных источников поступает информация.
2. Портал производит первичное распознание информации и предоставляет ей доступ.
3. При использовании систем управления знанием формируются метаданные.
4. Метаданные проходят через «фильтр», установленный пользователем, при этом ненужные данные отбрасываются.
5. После этого, отфильтрованные данные передаются пользователю.

Кроме того, корпоративный портал обычно разделен на две части:

* Внутренний портал (интранет), в котором расположена защищенная информация, предназначенная для использования ограниченным числом пользователей (здесь используется внутренняя интрасеть).
* Внешний портал, который доступен для общественного использования и в котором расположена публичная информация (здесь используется глобальная сеть Интернет).

Таким образом, можно сделать вывод о том, что современные корпоративные порталы – это мультифункциональные системы, решающие множество технологических и деловых задач. Однако, несмотря на то, что порталы предоставляют возможность оперативного поиска информации и дают эффективные инструменты работы с электронными документами в рамках взаимодействия сотрудников компании, данная технология не обеспечена интеллектуальными средствами для работы с электронными документами.

**Система управления документами** (Document Management System, DMS) –это информационная система, использующаяся для хранения и отслеживания электронных документов или образов бумажных документов.

Первоначально системы управления документами предназначались для обеспечения несложных функций хранения файлов документов, таких как: хранение и доступ к файлам, просмотр документов, разграничение прав доступа к фалам, ведение протокола доступа к фалам, отслеживание истории обработки документов, управление версиями.

С развитием ИТ, к вышеописанным функциям добавились следующие:

* ведение картотеки документов и инструменты быстрой разработки электронных форм;
* поддержка справочников с информацией для заполнения карточек;
* описание процесса обработки документов (фиксированных маршрутов);
* навигация и организация представления учётной информации о документах;
* свободная маршрутизация документов и поддержка персональных очередей пользователей;
* описание жизненного цикла обработки документа;
* управление маршрутизацией и средствами мониторинга процессов.

Наличие подобных средств управления процессами контроля, движения и обработки документов делает систему применимой для решения гораздо более широкого круга задач, нежели только ведение архива документов. Именно функциональность систем управления документами позволяет добавить в СЭД столь необходимые функции обработки слабоструктурированных данных.

Несмотря на то, что системы управления документами в общем случае предоставляют возможности хранения, версионирования, пометку метаданными и безопасность по отношению к документам, а также индексирование и развитие возможности поиска документа, они так же, как и корпоративные порталы оперируют электронными документами, как единым объектом, не беря во внимание содержание документа, логические связи элементов документа, смысл.

**WorkFlow‑система** (система автоматизации бизнес-процессов) – это информационная система, предоставляющая средства для автоматизации деловых процедур, включая разработку маршрутов, контроль исполнения и пр.

Концепция WorkFlow-системы была предложена компанией Staffware около 20 лет назад. Данная концепция рассматривает весь комплекс задач автоматизации бизнеса как совокупность бизнес-процессов. Инструменты, предоставляемые WorkFlow-системой, обеспечивают формирование описаний процессов, данных, а также содержат средства описания электронных форм для обработки этих данных. Технология WorkFlow подразумевает чёткое исполнение процесса согласно его описанию. Сервисы системы поддерживают реализацию бизнес-процессов, формирование и подготовку очередей заданий к обработке, автоматически обеспечивают нужную активность стадий процесса, контроль своевременности исполнения этапов процесса и различные способы реакции на возникающие проблемы.

Современная промышленная WorkFlow-система обладает гибким инструментарием моделирования процессов актуальных для компании, отслеживания состояния процессов и получения информации о несоответствии их течения, накопления статистики об отклонениях. Таким образом, WorkFlow‑система позволяет непрерывно улучшать и реструктурировать [бизнес-процессы](http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%91%D0%B8%D0%B7%D0%BD%D0%B5%D1%81-%D0%BF%D1%80%D0%BE%D1%86%D0%B5%D1%81%D1%81) компании.

Наличие средств семантического анализа данных позволило бы такой системе существенным образом повлиять на скорость реализации бизнес-процессов, а также на производительность и качество труда сотрудников. Однако в данной концепции не предусмотрено наличие таких технологий. Для того чтобы применить в архитектуре КИС концепцию семантических технологий анализа ЭД, тем самым совершенствовать процесс ЭДО, а также повсеместного использования документов в КИС, необходимо обозначить технологии, с помощью которых можно совершенствовать архитектуру.

### Обзор технологий семантического подхода для проектирования архитектуры КИС

Семантические технологии представляют собой направление развития информационных технологий, основывающееся на переходе от потокового представления данных (изображения, гипертекст и пр. поточные способы представления данных) к семантическому, позволяющему определять содержание и смысл данных. Таким образом, основная идея семантических технологий сводится к тому, чтобы все данные в ИС были бы пригодны для анализа.

Для того чтобы данные проанализировать и применить результаты в совершенствовании деятельности компании, необходимо определить метод представления знаний. В семантических технологиях встречается множество методов представления знаний, таких как: деревья решений, семантические сети, нейронные сети, фреймы и сети фреймов, онтологии, правила продукции и др.   
В рамках данной работы, наибольший интерес представляет онтологический метод, поскольку использование именно этого подхода позволит реализовать в архитектуре КИС требуемые функциональные возможности по управлению ЭД. Следует отметить, что для машинной обработки данных, в семантическом подходе предусмотрено использование нескольких основополагающих принципов, входящих в стек понятий семантических технологий [35].

Во-первых, необходимо использование такой системы идентификации объектов, которая обеспечит однозначную и единообразную идентификацию какого-либо ресурса (документа, изображения и пр.). С такой задачей справляется глобальная схема имен (Uniform Resource Identifier, URI).

Во-вторых, данные, представленные в семантической форме, записываются в форме «субъект – отношение - объект». Такая форма представления данных называется триплетом и позволяет понять пользователю факт, который сообщает система. Например, может быть передано такое сообщение: «сотрудник имеет № банковской карты, равный 111111». Для того чтобы любую информацию можно было записать в такой форме триплетов, можно воспользоваться моделью описания данных и метаданных (Resource Description Framework, RDF), предложенную консорциумом всемирной паутины (World Wide Web Consortium, W3C).

В-третьих, необходимо установить все возможные виды объектов, их свойства и связи и утверждения в формате RDF в дальнейшем можно интерпретировать и представить в виде онтологий, описанных с помощью языка описания онтологий (Web Ontology Language, OWL).

Таким образом, основываясь на принципах семантических технологий, в данном разделе будут рассмотрены основные технологии семантического подхода:

* системы, основанные на технологии построения онтологий предметной области;
* экспертные системы, в которых реализован семантический поиск, автоматическое аннотирование;
* системы управления знаниями.

А в качестве таких систем, в которых реализованы технологии семантического подхода, направленные на работу с семантикой данных ЭД будут рассмотрены:

* портал знаний, реализованный на платформе Microsoft SharePoint;
* портал знаний, реализованный на платформе Oracle Portal;
* система извлечения знаний из документов «Аналитический курьер» компании «Ай-Теко».

**Системы, основанные на технологии построения онтологий предметной области,** проектируются при помощи онтологического моделирования. Онтологическое моделирование – это процесс построения, развития, обработки и использования онтологии предметной области. Согласно общепринятому определению под онтологией (в широком смысле) понимается база знаний специального типа, которая может «читаться» и пониматься, отчуждаться от разработчика и/или физически разделяться ее пользователями [18].

Анализ литературных источников в области онтологического моделирования показывает, что научно-техническими предпосылками данного направления являлись исследования и разработки в таких областях, как искусственный интеллект, Data Mining (интеллектуальный анализ данных), Semantic Web (семантическая паутина). Следовательно, использование онтологического моделирования, в качестве семантического подхода для проектирования архитектуры КИС будет являться способом всеобъемлющей и детальной формализации некоторой области знаний с помощью концептуальной схемы. Обычно такая схема состоит из структуры данных, содержащей все релевантные классы объектов, их связи и правила (ограничения), принятые в этой области.

Помимо стандартных ролей, онтологии могут взять на себя ещё одну роль – наполнение информацией пользовательских интерфейсов, в дополнение к стандартной роли по интеграции информации. В этой связи такие структуры называются адаптивными онтологиями. Некоторые из дополнительных применений адаптивных онтологий: названия атрибутов и подсказки, навигация и просмотр структур и деревьев, структуры меню, автоматическое завершение при вводе данных, контекстные выпадающие списки выбора, проверка правописания, и т.д. Иными словами, то, что делает онтологию адаптивной, это добавление к стандартной машиноориентированной цели онтологии дополнительных средств для понимания человеком - названий, синонимов, определений и прочего. Системы, основанные на онтологиях – это модульные приложения общего плана, предназначенные для работы в соответствии со спецификациями, содержащимися в адаптивной онтологии.

Системы, основанные на онтологиях, предназначены для выполнения конкретных общих задач: импорт и экспорт в различные форматы, создание и управление наборами данных, отчетность, просмотр, поиск, визуализации данных, распределение прав доступа пользователей, и тому подобное. Эти приложения определяют свои специфические функции в соответствии с содержащимися в онтологии спецификациями. Основным преимуществом является то, что одна система может реализовать некоторые общие функции, основанные на любой правильно построенной адаптивной онтологии.

**Экспертные системы, с семантическим поиском** также являются примером систем, реализующих семантический подход в работе с неструктурированными документами. Упомянутые выше адаптивные онтологии, которые предназначены именно для человеческого понимания, покрывают бизнес-терминологию конечных пользователей, включая синонимы и омонимы. Это позволяет внедрять продвинутые технологии поиска данных, такие как фасетный поиск.

Фасетный поиск – это поиск с использованием фасетной классификации (совокупность нескольких независимых классификаций, осуществляемых одновременно по различным основаниям), многочисленных фильтров, основанных на классификациях информации по различным признакам. Реализуется на основе онтологий (фасеты генерируются на основе иерархии концептов в онтологии), учитываются отношения между сущностями в онтологии [21].

**Системы управления знаниями (СУЗ)** существенно отличается от ИС организаций, предназначение которых эффективное хранение данных, обработка и предоставление информации. СУЗ представляет собой систему, реализующую функции увеличения объема и повышения уровня использования знаний организации, используя эффективный поиск и фильтрацию данных, знаний.

Основной объект, которым оперируют СУЗ – знания. Система управления знаниями обладает такими технологиями, с помощью которых сотрудники организации могут взаимодействовать на уровне моделирования бизнес-процессов с помощью теории, концепций, достоверных знаний сотрудников. Все это достигается за счет использования в СУЗ языка описания профессионального общения, который опирается на метаописания и онтологии предметной области. Назначение СУЗ, предложенное в [19], иллюстрирует схема, представленная на рис. 1.4.

|  |
| --- |
|  |
| Назначение систем управления знаниями |

Если предназначением учетной ИС является эффективное хранение, обработка и предоставление пользователям по регламенту или запросу искомой информации, то предназначением СУЗ является увеличение объема и повышение уровня использования знаний организации за счет их эффективного поиска и фильтрации, включения в коллективную память компании профессиональных знаний и опыта, в том числе хранящегося в головах специалистов. Как видно на схеме, СУЗ взаимодействует с бизнес-процессами компании, а знания накапливает из различных источников, таких как: базы знаний, базы данных, метаописания, онтологии. Особенный интерес в данных системах представляет использование онтологических моделей для представления знаний.

**Портал знаний, реализованный на платформе Microsoft SharePoint,** представляет собой попытку практического применения теории управления знанием. Обеспечивая прямое взаимодействие пользователей с необходимым интеллектуальным капиталом, портал знаний ускоряет обработку информации и снижает объем повторяемой работы в организации.

Каждый портал знаний может иметь иерархическую или сетевую структуру, т.е. состоять из более специализированных порталов, связанных некоторыми отношениями, например, отношением “общее-частное”. Такая архитектура делает портал знаний гибким и легко расширяемым. В основе порталов знаний лежат онтологии, содержащие описание устройства и типологии соответствующих сетевых ресурсов. Основными компонентами портала знаний являются: онтология, коллекционер онтологической информации о ресурсах, конструктор запросов и формирователь ответов.

Основная цель порталов знаний - объединение ряда возможностей по управлению знаниями в компании. Портал должен охватывать все уникальные аспекты и компоненты предприятия, позволяя отказаться от отдельных, специализированных и несовместимых порталов. Архитектура порталов знаний должна совмещать возможности различных приложений и источников информации и служить единственной точкой доступа, обеспечивая общие метаданные и уровни безопасности. Для реализации этого следует обратиться к объектно-ориентированным методам. Способность вместить и содержимое, и правила, создать специализацию посредством подклассов делает объекты безальтернативным средством представления источников, отношений и взаимодействий, содержащихся внутри портала знаний.

Существует несколько платформ, на которых ведется реализация порталов знаний. Платформа Microsoft SharePoint представляет собой рабочую платформу для предприятий и работы в сети Интернет, интегрированные функции которой предоставляют широкие возможности коммуникации между людьми. Благодаря унифицированной инфраструктуре SharePoint Server 2013 позволяет снизить расходы и дает возможность быстро реагировать на потребности предприятия независимо от способа развертывания: как на месте, так и в качестве размещенной службы. Портал знаний, реализованный на платформе SharePoint Server 2013 предназначен для:

* общения сотрудников и обмена информацией между сообществами, что дает возможность делиться знаниями и идеями;
* управления корпоративным содержимым и балансирование между обеспечением взаимодействия с пользователем и соответствия политике и процессу;
* реагирования на изменение потребностей компании путем принятия решений на основе данных;
* помощи пользователям в поиске необходимого содержимого, сведений и других пользователей с помощью поиска в корпоративной среде.

**Портал знаний, реализованный на платформе Oracle Portal** выполняет все те же задачи портала знаний, которые были перечислены выше. Однако платформа реализации портала Oracle Portal обладает отличными от SharePoint Server 2013 особенностями. Основные составляющие порталов знаний, реализованных на платформе Oracle Portal:

* библиотеки готовых компонентов;
* средства управления содержанием с поддержкой коллективной работы;
* средства для сквозного полнотекстового поиска информации на портале, в интернет/интранет, в базах данных и в хранилищах неструктурированной информации;
* средства интеграции внешних систем, приложений, различных видов источников данных;
* средства для анализа бизнес-информации (BI - Business Intelligence);
* средства однократной регистрации и управления информационной безопасностью;
* средства поддержки мобильного беспроводного доступа;
* интегрированное средство разработки веб-приложений.

Объединяя существующие и развивающиеся технологии совершенствования порталов знаний, вне зависимости от платформы реализации поратала, можно сказать, что корпоративные порталы знаний стремятся к интеграции таких информационных ресурсов и технологий, как: электронная почта и Интернет-ресурсы; системы управления базами данных (СУБД) и сами базы данных (БД); средства создания хранилищ данных; системы поддержки групповой работы; локальные системы автоматизации и документооборота; системы документооборота; экспертные системы и базы знаний; знания специалистов и др.

**Система извлечения знаний из документов «Аналитический курьер» компании «Ай-теко»** представляет собой инструмент для исследования событий, проблем или объектов путем мониторинга ресурсов сети интернет и других источников. Данная система реализует следующие возможности:

* анализ мнений субъектов относительно объектов;
* анализ тональности в отношении объектов;
* анализ связей между объектами и темами в объеме всего хранилища документов;
* выявление источников информационных кампаний и атак, поиск перепечаток;
* мониторинг наиболее обсуждаемых тем и сюжетов в потоке поступающих документов;
* социометрия авторов и сообщений, пространственная локализация авторов;
* анализ динамики развития событий, частотный анализ тем;
* прогноз появления событий нужных типов и определение их значимости;
* тематическая классификация входных сообщений и др [14].

Система используется в области прогнозной аналитики, а инструменты семантического анализа текста в составе системы применяются для прогнозирования появления определенных событий, для выявления их информационных профилей и др. Кроме того, в состав системы входит отказоустойчивое масштабируемое хранилище Big Data, позволяющее надежно и эффективно обрабатывать миллиарды документов. Оно может использоваться отдельно от системы, входит в состав многих приложений, разработанных для различных пользователей.

Все рассмотренные в данном разделе системы, так или иначе, имеющие в своем составе семантические технологии, являются достаточно узконаправленными на применение семантических технологий и не решают весь перечень функциональных задач проектируемой КИС. Однако технологии онтологического подхода, позволяющие построить такие системы, которые основаны на онтологии предметной области или в которых реализован семантический поиск и аннотирование являются наиболее подходящими для использования в проектируемой архитектуре КИС.

Из анализа некоторых возможностей систем, реализующих семантический подход, представляется возможным сформировать сводную таблицу (см. табл. 2.3), отображающую применение семантических технологий в частности, использующих онтологический подход представления знаний на разных этапах ЖЦ ЭД.

##### Задачи, решаемые семантическими технологиями на разных этапах ЖЦ ЭД

| № | Этап ЖЦ ЭД | Задача |
| --- | --- | --- |
| 1 | Ввод | * классификация документов (тематическое индексирование) кластеризация документов * описание документа * описание элементов предметной области |
| 2 | Хранение | * хранение элементов документа (содержимое, метаданные) |
| 3 | Обеспечение сохранности | * оценка полноты, целостности и непротиворечивости документов |
| 4 | Управление | * интеллектуальная интеграция информации * информационный поиск * поиск зависимостей между документами * обеспечение доступа к содержательным знаниям и данным из документов в ИС * визуализация знаний и данных * генерация ответа на запрос к ИС |
| 5 | Доставка | * управление маршрутами движения документов |
| 6 | Вывод | * экспорт в различные форматы |

Применение онтологий в архитектуре КИС дает возможность выполнения большого спектра задач представления, обработки знаний, содержащихся в ЭД, а также исполнение запросов к данным, извлеченным из ЭД. Кроме онтологического подхода в архитектуре КИС, можно использовать технологии Text-Mining, непосредственно занимающиеся извлечением данных из неструктурированных документов.

### Применение классических и семантических технологий в КИС

На основе определенного в п. 1.2.4. ЖЦ ЭД в КИС (см. рис. 1.5), обозначим классические технологии, которые применяются на каждом из этапов и определим те этапы ЖЦ ЭД, в которых применение семантических технологий позволит совершенствовать движение ЭД в КИС.

|  |
| --- |
|  |
| Этапы жизненного цикла ЭД в КИС |

Рассмотрим каждый этап ЖЦ подробно.

**Этап 1. Ввод документа в ИС.** На данном этапе ЖЦ информация создается или импортируется в ИС извне. С этого первого шага начинается движение информации на предприятии, и как следствие возникает задача управления ею. Условно можно выделить два класса информации: создаваемая человеком и создаваемая приложениями. К первому классу относятся бумажные и электронные документы, различные заполненные формы и анкеты, мультимедийная информация (аудио, видео, изображения и т. п.), а также информация на архивных носителях, например, на микрофильмах. Ко второму классу (информация, создаваемая приложениями) относятся данные, выгружаемые из других ИС (ERP, CRM, специализированные системы), электронные формы, электронные данные, полученные по каналам электронного обмена.

На этапе ввода используются следующие технологии:

* **Сканирование документов.** Эта технология позволяет работать в системе с образами бумажных документов, что не только многократно повышает удобство работы, но и делает возможным параллельную и удаленную работу с одними и теми же документами.
* **Распознавание образов документов (технологии OCR/HCR/ICR/OMR/IDR).** Это целая совокупность технологий, включающая не только оптическое распознавание символов, но и такие технологии, как интеллектуальное распознавание символов и целых документов, распознавание рукопечатных символов (используется при распознавании бумажных форм), распознавание меток.
* **Обработка форм.** Формы могут быть бумажные, электронные или веб-формы. Заполненные формы пересылаются на участок обработки, где их содержимое заносится в систему.
* **Агрегация.** Информация может поступать из разных источников, но в ИС она должна быть представлена как логически целая.
* **Интеграция.** Информация в ИС может поступать из других ИС. Интеграционному решению необходимо решать задачи — захвата и преобразования данных и метаданных, сохранение юридической значимости информации и др.
* **Классификация.** Полученная информация индексируется, классифицируется и категоризируется, получает регистрационные номера, чтобы в дальнейшем в ИС была возможность эту информацию найти, а также на нее ссылаться.

**Этап 2. Хранение документа в ИС.** Информация, поступившая в систему, хранится в ней в соответствии с установленными для этого класса информации политиками. Вопросы хранения могут быть рассмотрены на нескольких уровнях абстракции, от физических методов до политик и общей стратегии управления информацией на предприятии.

Технологии, применяющиеся на данном этапе:

* **SAN, NAS, CAS, магнито-оптические диски, CD/DVD, дисковые и ленточные библиотеки, SSD, RAID.** Это целый парк технологий и периферийных устройств, предназначенных для хранения информации. Выбор той или иной технологии обуславливается объемом информации, сроком хранения, параметрами доступа к ней пользователей, требуемым уровнем надежности и безопасности;
* **файловые системы, CMS, базы данных, хранилища данных, репозитории.**На этом уровне решаются вопросы организации хранения данных на уровне файлов, записей в базе данных, контента в хранилище и CMS.

Перечисленные технологии призваны решать следующие задачи, возникающие в процессе хранения:

1. поиск;
2. контроль версий;
3. регистрация прибытия/убытия документов и информации;
4. блокировка;
5. аудит действий пользователей;
6. извлечение документов и информации.

На логическом уровне можно выделить задачу управления политиками хранения.

**Этап 3. Обеспечение сохранности**. Этот этап ЖЦ можно также определить как отдельное направление деятельности, включающее в себя ряд задач:

* обеспечение юридической значимости информации;
* перевод системы в специальные режимы (аудит, проверка и т. п.);
* восстановление информации, пострадавшей в результате катастроф природного и техногенного характера;
* своевременный перевод в новые форматы и на новые носители (это особенно актуально при долговременном хранении информации).

Для решения этих и других задач могут быть использованы различные инструменты и технологии:

* электронная подпись;
* специальные архивные форматы, например PDF-A [23];
* вывод информации на архивные носители (бумага, микрофильмы);
* резервное копирование и восстановление информации;
* контролируемая конвертация и перенос на другие носители.

**Этап 4. Управление документом.** Информация, которая обращается в организации и хранится в **ИС**, только абстрактно представляет собой что-то логически целое. На практике выделяют несколько классов и соответственно несколько отдельных видов деятельности или дисциплин:

* управление документами;
* управление информацией, значимой с точки зрения законодательства или бизнеса;
* управление цифровыми активами;
* управление электронной почтой;
* управление веб-контентом;
* организация взаимодействия;
* управление бизнес-процессами.

Стоит отметить, что эти задачи могут пересекаться и включать в себя другие, и выделение перечисленных дисциплин носит условный характер.

**Этап 5. Доставка документа.** Хранение информации — не самоцель, информация постоянно используется в повседневной деятельности. Поэтому в сферу деятельности ИС входит организация ее доставки потребителям.

На этапе доставки данных можно выделить следующие задачи:

1. **Управление правами доступа.** В системе, как правило, хранится информация, предназначенная для разных групп пользователей. Поэтому невозможно представить себе **систему документооборота** без механизма управления правами доступа.
2. **Управление отчетами, печатью и публикациями.** Результатом ряда бизнес-процессов является бумажный или электронный документ, доступ к другой информации осуществляется через интернет- и интранет-порталы. Либо затребованные документы отправляются электронной почтой или по другим каналам связи.
3. **Организация доступа с мобильных устройств.** Удаленная и мобильная работа с корпоративной информацией для ряда сотрудников становится не вспомогательным, а основным сценарием.
4. **Межведомственный электронный документооборот (МЭДО).** Отказ от бумаги и переход на электронный документооборот между организациями помимо того, что даёт экономию на бумаге, печати и расходов на доставку, позволяет организовать совместную работу над документами, увеличить скорость и эффективность взаимодействия.

Для решения этих и других задач используются технологии:

* управление правами доступа к цифровой информации, «водяные знаки»;
* электронная подпись, инфраструктура публичных ключей подписи;
* преобразование, форматирование, сжатие и объединение информации;
* XML, JSON, SOAP, REST.

**Этап 6. Вывод документа из ИС.** Информация, хранящаяся в ИС, может быть отправлена в архив или удалена по истечении определенного периода времени согласна установленным нормам. Технологии, использующиеся на этапе вывода информации из ИС, связаны с организацией полного удаления данных из ИС и представляют собой встроенные в архитектуру модули вывода из ИС.

Каждый из вышеописанных этапов ЖЦ, в совокупности с технологиями, может быть дополнен возможностями семантического подхода в ключе управления ЭД. В качестве основной семантической технологии, в связи с перспективностью ее использования в архитектуре КИС, в разделе 1.3.2 определен онтологический подход. Данный подход можно использовать для реализации различных задач, перечисленных в таблице 2.3, например, таких как:

* хранение элементов документа (содержимое, метаданные)
* интеллектуальная интеграция информации;
* информационный поиск;
* поиск зависимостей между документами;
* обеспечение доступа к содержательным знаниям и данным из документов в ИС;
* визуализация знаний и данных;
* генерация ответа на запрос к ИС.

Перспективным преимуществом при внедрении в архитектуру КИС семантических технологий является получение удобной интегрированной КИС, которая включает в себя функции существующих в настоящее время технологий классического подхода, а также проектируется с учетом применения методов семантического подхода, таких как: онтология предметной области, семантический поиск, аннотирование, классификация и др.

# Моделирование бизнес-процессов по управлению электронными документами

В качестве нотации моделирования бизнес-процессов, в силу удовлетворяющих в данной работе особенностей моделирования бизнес-процесса, выбрана нотация ARIS, являющаяся методологией предназначенной для моделирования бизнес-процессов организации, разработанная компанией IDS Sheer. Аббревиатура ARIS расшифровывается как «Архитектура интегрированных информационных систем» (Architecture of Integrated Information Systems). В данной методологии понятие «архитектура» служит для описания типа и функциональных свойств информационных технологий и взаимоотношений между отдельными составляющими ИС. Кроме того, учитывая развивающуюся и укрепляющуюся связь информационных технологий с бизнес-процессами компании, модели, построенные на базе концепции ARIS, позволяют описать бизнес-процессы достаточно полно и подробно.

Для описания бизнес-процессов в методологии ARIS предусмотрено большое количество типов моделей, каждая из которых принадлежит тому иному аспекту. Для описания бизнес-процессов работы с ЭД подходит модель ARIS eEPC (extended Event Driven Process Chain), предназначенная для создания моделей бизнес-процессов на нижнем уровне. Бизнес-процесс, созданный по данной модели, позволяет наглядно отразить поток работ, который протекает внутри подразделения, выявить связи между функциями и организационными единицами, и что немаловажно, в данной работе, отразить связи с ИС. Для каждой функции внутри процесса определяются конечные и начальные события, ответственные исполнитель, материальные и документальные потоки, сопровождающие, какую либо работу.

Также, важно отметить, что модель бизнес-процесса ARIS eEPC построена на определенных семантических правилах описания:

1. каждая функция должна быть инициирована событием и должна завершаться событием;
2. в каждую функцию не может входить более одной стрелки, «запускающей» выполнение функции, и выходить не более одной стрелки, описывающей завершение выполнения функции.

Поскольку принято решение о моделировании бизнес-процессов по методологии ARIS, целесообразно выбрать такой инструмент для моделирования, который поддерживает данную нотацию. Воспользуемся программным продуктом ARIS Toolset специально созданным для построения бизнес-процессов в нотации ARIS, представляющим собой бесплатную версию программного продукта ARIS.   
В инструменте моделирования бизнес-процессов ARIS Toolset предусмотрен большой набор элементов, для отображения потоков работ. Элементы схемы бизнес-процесса, использованные в рамках данной работы, отображены в приложении C.

Перед выполнением этапа моделирования бизнес-процессов необходимо отметить, что процессы по управлению и организации работы с ЭД, протекающие в компании, не ограничиваются процессами ЭДО, обозначенными в разделе 1.2.6, однако составляют некоторую часть среди всех процессов, связанных с управлением ЭД.

1. Процессы управления ЭД, протекающие в КИС:
2. систематизация, регламентация работы с ЭД;
3. создание, обработка ЭД по шаблонам;
4. автоматизация учетной деятельности с ЭД, в т.ч.:
   * классификация ЭД по различным критериям;
   * регистрация документов по заданным схемам, шаблонам;
   * учет сроков хранения документов.
5. автоматизация поиска ЭД;
6. электронная рассылка документов;
7. автоматизация процедур коллективной работы с ЭД:
   * совместная разработка проекта документа;
   * согласование документа;
   * экспертиза документа;
   * исполнение документа.
8. обеспечение защиты от несанкционированного доступа, искажения, удаления информации из ИС.

Рассмотрим процессы в рамках стандартных архитектурных и программных решений систем, занимающихся управлением ЭД (пример укрупненной модели некоторых процессов представлен на рисунке 2.1).

|  |
| --- |
|  |
| Укрупненная модель процессов в СЭД |

В силу того, что в модели представлены только те процессы, которые позволяют очертить картину процессов, иллюстрирующих ЖЦ ЭД в КИС, связи между процессами условны. Это означает, что в промежутке между процессами, связанными на схеме пунктирной линией могут располагаться другие процессы, которые не несут важности в рамках отображения ЖЦ ЭД. Таким образом, условно обозначены три входа у процесса «Формирование, согласование, утверждение документ». Этот процесс отражает типичный и характерный для большинства средних/крупных предприятий процесс ЭДО.

Далее каждый процесс будет рассмотрен и описан подробно в модели «как-есть», а также в виду применения семантических технологий в архитектуре будет осуществлен анализ изменения бизнес-процессов, будут определены такие моменты в бизнес-процессах, которые могут быть совершенствованы, изменены, ускорены за счет применения семантических технологий на тех или иных этапах бизнес-процессов.

Обозначенные в укрупненной модели бизнес-процессы компании по управлению ЭД необходимо рассмотреть подробно с точки зрения модели «как‑есть», обозначив такие пункты: описание процесса, начало выполнения, результат процесса, требования к срокам выполнения процесса, владелец процесса, исполнители процесса, информационные технологии, поддерживающие процесс. Некоторые пункты описания носят приближенный характер.

## Бизнес-процесс «Ввод документа в ИС»

**Описание процесса.** Бизнес-процесс «Ввод документа в ИС (сканирование)» является самым первым этапом движения документа в ИС. После сканирования бумажного документа и проверки на содержание ошибок бизнес-процесс сканирования переходит в бизнес-процесс регистрации документа в ИС. Далее, документ является уже электронным и в последующих этапах ЖЦ документа ИС оперирует ЭД.

Необходимо отметить, что процесс сканирования бумажных документов, бизнес-процесс которого рассмотрен в качестве примера, не единственный способ ввода документа в ИС. Неструктурированные данные попадают в КИС из разных источников, таких как: бумажные документы; электронные формы; электронная почта, мгновенные сообщения, текстовые документы.

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Принят пакет бумажных документов для сканирования.

**Результат процесса.** Результатом данного процесса является наступление следующих событий (см. табл. 2.1):

##### Результат выполнения процесса «Ввод документа в ИС»

|  |  |  |
| --- | --- | --- |
| № | Событие | Предшествующее действие |
|  | Основные события | |
| 1 | Документ подлежит регистрации в ИС | Перевод документа на этап регистрации в ИС |
|  | Неосновные события | |
| 2 | Документы возвращены клиенту | Документы не пригодны для сканирования;  Документы отложены для возврата клиенту |

**Требования к срокам выполнения процесса.** Для данного бизнес-процесса отсутствуют точные сроки выполнения, поскольку длительность выполнения процесса зависит от объема принятого пакета бумажных документов на сканирование. Условно, обозначим срок выполнения процедуры равный 5 часам.

**Владелец процесса.** Владельцем процесса является начальник отдела документационного управления.

**Исполнители процесса.** Основными исполнителями процесса ввода документа в ИС (сканирования) являются сотрудники документационного управления - 2 человека (см. табл. 2.2).

##### Исполнители процесса «Ввод документа в ИС»

| № | Организационная единица | Подразделение | Предмет деятельности |
| --- | --- | --- | --- |
| 1 | Сотрудник отдела документационного управления -1 | Отдел документационного управления | * прием бумажной документации; * регистрация входящей документации; * сортировка документации для сканирования * взаимодействие с поставщиками процесса |
| 2 | Сотрудник отдела документационного управления -2 | Отдел документационного управления | * сканирование бумажной документации; * контроль сканирования; * перевод документации на этап регистрирования в ИС |

**Информационные технологии, поддерживающее выполнение процесса** отображены в табл. 2.3.

##### Информационные технологии процесса «Ввод документа в ИС»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | Потоковое сканирование бумажных документов | МФУ для потокового сканирования | Офисное оборудование |
| 2 | Сканирование документов по одному | МФУ для единичного сканирования | Офисное оборудование |
| 3 | Проверка реквизитов документов | Система ввода | Модуль ИС |
| 4 | Перевод документа на этап регистрации в ИС | Система ввода | Модуль ИС |

Схема бизнес-процесса представлена в приложении D. На схематическом изображении бизнес-процесса видно, что он разделен пунктирной линией на ту часть, где события и процессы происходят вне ИС, и в ИС. Особенное внимание следует обратить на события, протекающие внутри ИС.

## Бизнес-процесс «Регистрация документа в ИС»

**Описание процесса.** Бизнес-процесс по регистрации документа в ИС представляет собой второй этап ввода документа в ИС. Независимо от используемого файлового формата документа, метода кодировки и других технологических характеристик, регистрация документа в ИС должна производиться без внесения каких-либо изменений в их содержимое. Более того, данный бизнес-процесс является одним из наиболее важных процессов, в рамках совершенствования бизнес-процесса в данной работе, поскольку именно на этом этапе представляется возможным применения широкого перечня технологий, направленных на анализ документов.

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Документ подлежит регистрации в ИС.

**Результат процесса.** Результатом данного процесса является наступление следующего события: Документ зарегистрирован в ИС

**Требования к срокам выполнения процесса.** Длительность выполнения бизнес-процесса, при благоприятных условиях выполнения всех функций, должна составлять 30 минут. В случае, когда требует корректировка регистрируемого документа длительность может составлять 60 минут.

**Владелец процесса.** Владельцем процесса является начальник отдела документационного управления.

**Исполнители процесса.** Основными исполнителями процесса регистрации документа в ИС является сотрудник отдела документационного управления   
(см. табл. 2.4).

##### Исполнители процесса «Регистрация документа в ИС»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Организационная единица | Подразделение | Предмет деятельности |
| 1 | Сотрудник отдела документационного управления | Отдел документационного управления | * обработка документа в ИС на всех этапах процесса; * регистрация документа в ИС |

**Информационные технологии, поддерживающее выполнение процесса** отображены в табл. 2.5.

##### Информационные технологии процесса «Регистрация документа в ИС»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | Извлечение из документов атрибутов, ключевых слов | Система автоматического индексирования | Модуль ИС |
| 2 | * Обработка документа в ИС; * Верификация документа; * Извлечение из документа атрибутов, ключевых слов; * Присвоение атрибутов, ключевых слов документу; * Формирование формата выходного документа; * Формирование карточки документов в ИС | Система ввода | Модуль КИС |

Схема бизнес-процесса представлена в приложении E на рис. E.1.

## Процесс «Выполнение заявки на выдачу документа из архива»

**Описание процесса:** Выполнение заявки на выдачу документа из архива представляет собой процесс, протекающий в ИС. Сотрудник отдела документационного обеспечения должен отреагировать на заявку о выдаче документа и предоставить документ запрашивающему сотруднику. Процесс отражает работу с ЭД на этапе ЖЦ «хранение».

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Поступила заявка на предоставление доступа к ЭД из архива.

**Результат процесса.** Результатом данного процесса является наступление следующих событий (см. табл. 2.6):

##### Результаты процесса «Выполнение заявки на удаление»

|  |  |  |
| --- | --- | --- |
| № | Событие | Предшествующее действие |
|  | Основные события | |
| 1 | Запрашиваемый пользователем документ предоставлен | Предоставление запрашиваемого пользователем документа |
| 2 | Запрашиваемый пользователем документ не предоставлен | Определение причин несоответствия |
|  | Неосновные действия | |
| 3 | Заявка не принята на рассмотрение | Направление пользователю сообщения об отсутствии документа в архиве;  Направление пользователю сообщения о несоответствии уровня доступа. |

**Требования к срокам выполнения процесса.** Длительность выполнения бизнес-процесса составляет не более 5 минут.

**Владелец процесса.** Владельцем процесса является начальник отдела документационного управления.

**Исполнители процесса.** Основными исполнителями процесса регистрации документа в ИС является сотрудник отдела документационного управления.

**Информационные технологии, поддерживающее выполнение процесса** отображены в табл. 2.7.

##### Информационные технологии процесса «Выполнение заявки на удаление»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | Извлечение из документов атрибутов, ключевых слов | Система автоматического индексирования | Модуль ИС |
| 2 | * Обработка документа в ИС; * Верификация документа; * Извлечение из документа атрибутов, ключевых слов; * Присвоение атрибутов, ключевых слов документу; * Формирование формата выходного документа; * Формирование карточки документов в ИС | Система ввода | Модуль ИС |

Схема бизнес-процесса представлена в приложении F на рис. F.1.

## Процесс «Выполнение запроса на удаление документа из ИС»

**Описание процесса**: Выполнение проса на удаление документа из ИС представляет собой процесс, протекающий в ИС. Сотрудник отдела документационного обеспечения должен отреагировать на запрос об удалении документа и выполнить запрос или отклонить. Процесс отражает работу с ЭД на этапе ЖЦ «вывод из ИС» и «обеспечение сохранности».

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Поступил запрос на удаление документа из ИС.

**Результат процесса.** Результатом данного процесса является наступление следующего события (см. табл. 2.8):

##### Результат процесса «Выполнение запроса на удаление»

|  |  |  |
| --- | --- | --- |
| № | Событие | Предшествующее действие |
|  | Основные события | |
| 1 | Запрос на удаление документа выполнен | Сохранение изменений в ИС |
| 2 | Запрос на удаление документа отклонен | Определение причин запрета на удаление |
|  | Неосновные действия | |
| 3 | Запрос не может быть выполнен | Направление сообщения пользователю о несоответствии уровня доступа;  Направление сообщения пользователю об отсутствии запрашиваемого документа |

**Требования к срокам выполнения процесса.** Длительность выполнения бизнес-процесса составляет не более 5 минут.

**Владелец процесса.** Владельцем процесса является начальник отдела документационного управления.

**Исполнители процесса.** Основными исполнителями процесса регистрации документа в ИС является сотрудник отдела документационного управления.

**Информационные технологии, поддерживающее выполнение процесса**

отображены в табл. 2.9.

##### Информационные технологии процесса «Выполнение запроса на удаление»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | * идентификация пользователя; * поиск запрашиваемого документа в ИС; * направление сообщений пользователю; * определение уровня доступа пользователя к запрашиваемому документу; * регистрация запроса для исполнения; * определение уровня контроля над документом; * удаление документа из ИС | Система электронного документооборота | Модуль ИС |

Схема бизнес-процесса представлена в приложении G.

## Процесс «Выполнение запроса на редактирование документа»

**Описание процесса**: выполнение процесса на редактирование документа представляет собой процесс, протекающий в ИС. Сотрудник отдела документационного обеспечения должен отреагировать на запрос о редактирвоании документа и выполнить запрос или отклонить. Процесс отражает работу с ЭД на этапе ЖЦ «управление» и «обеспечение сохранности».

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Поступил запрос на редактирование документа, находящегося в архиве

**Результат процесса.** Результатом данного процесса является наступление следующих событий (см. рис. 2.10):

##### Результаты процесса «Выполнение заявки на редактирование»

|  |  |  |
| --- | --- | --- |
| № | Событие | Предшествующее действие |
|  | Основные события | |
| 1 | Запрос на редактирование документа выполнен | Сохранение изменений в ИС |
| 2 | Запрос на редактирование документа отклонен | Определение причин запрета на удаление |
|  | Неосновные действия | |
| 3 | Запрос не может быть выполнен | Направление сообщения пользователю о несоответствии уровня доступа;  Направление сообщения пользователю об отсутствии запрашиваемого документа |

**Требования к срокам выполнения процесса.** Длительность выполнения бизнес-процесса составляет не более 5 минут.

**Владелец процесса.** Владельцем процесса является начальник отдела документационного управления.

**Исполнители процесса.** Основными исполнителями процесса регистрации документа в ИС является сотрудник отдела документационного управления.

**Информационные технологии, поддерживающее выполнение процесса** отображены в табл. 2.11.

##### Информационные технологии процесса «Выполнение запроса на редактирование»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | * идентификация пользователя; * поиск запрашиваемого документа в электронном архиве; * направление сообщений пользователю; * определение уровня доступа пользователя к запрашиваемому документу; * регистрация запроса для исполнения; * определение уровня контроля над документом; * переформирование документа. | Система электронного архива | Модуль ИС |
| 2 | * идентификация пользователя | КИС | КИС |

Схема бизнес-процесса представлена в приложении H.

## Бизнес-процесс «Формирование, согласование, утверждение»

**Описание процесса:** Бизнес-процесс «формирование, согласование, утверждение документа» отражает ЖЦ проекта документа и его последующую обработку сотрудниками разных подразделений согласно бизнес-требованиям к данному процессу.

**Начало выполнения процесса.** Началом выполнения процесса является наступление следующего события: Сформирован документ, требующий утверждения.

**Результат процесса.** Результатом данного процесса является наступление следующего события: документ утвержден.

**Требования к срокам выполнения процесса.** Длительность выполнения бизнес-процесса не регламентирована, зависит от внешних факторов (занятость согласующих, важность документа, количество сотрудников в организации и пр.). Условно, обозначим срок выполнения равный 5 рабочим дням.

**Владелец процесса.** Владельцем процесса является начальник отдела делопроизводства.

**Исполнители процесса.** Основными исполнителями процесса выполнения запроса на редактирование в электронном архиве является сотрудник отдела документационного управления (см. табл. 2.12).

##### Исполнители процесса «Согласование и утверждение документа»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Организационная единица | Подразделение | Предмет деятельности |
| 1 | Сотрудник отдела делопроизводства | Отдел делопроизводства | * регистрация документа в СЭД; * перевод документа на этап согласования; * отправка сообщений; * отправка документа |
| 2 | Автор документа | Любое | * формирование документа, требующего утверждения |
| 3 | Согласующий | Любое (профильное по компетенции положений в документе) | * формирование замечаний к документу |
| 4 | Начальник подразделения | Любое | * согласование документа |

**Информационные технологии, поддерживающее выполнение процесса** отображены в табл. 2.13.

##### Информационные технологии процесса «Согласование и утверждение документа»

|  |  |  |  |
| --- | --- | --- | --- |
| № | Функция | Наименование информационных технологий | Тип программного обеспечения |
| 1 | * регистрация документа в СЭД; * отправка документа; * формирование листа согласования. | СЭД | Модуль ИС |

Схема бизнес-процесса представлена в приложении I.

Смоделированные бизнес-процессы, связанные с управлением ЭД, носят обобщенный характер, однако каждый из рассмотренных выше процессов может быть совершенствован с точки зрения применения в них возможностей семантического подхода. Проанализируем каждый процесс подробнее.

Бизнес-процесс «Ввод документа в ИС» является самым первым этапом в ЖЦ ЭД в КИС и в большей степени зависит от сотрудников осуществляющих данный бизнес-процесс. Однако, после того как документ отсканирован он сразу же может быть передан подсистеме обработки форм. В этом бизнес-процессе никаких существенных изменений с применением семантических технологий не произведено, за исключением передачи выполнения функции проверки реквизитов и передачи документа на регистрацию подсистеме обработки форм.

Бизнес-процесс «Регистрация документа в ИС» после применения семантических технологий на некоторых этапах, сокращается по времени, удаляются некоторые этапы, в большей степени связанные с аннотирование, выделением ключевых слов, индексированием. Регистрация документа и извлечение данных из них является самым основополагающим этапом для применения семантических технологий. Схема бизнес-процесса «Регистрация документа в ИС» в модели «как-должно быть» представлена в приложении D на рис. D.2.

Используя новые информационные технологии, в частности семантический подход) для поддержки бизнес-процессов управления ЭД позволит сократить непроизводительные затраты рабочего времени. К примеру, за счет автоматического аннотирования, сотруднику не понадобится вручную вносить информацию в ИС, семантический поиск в архиве ускорит время выборки необходимых документов, а также позволит искать документы, схожие по тематике. Настройка уровней доступа для документов схожей тематики позволит сократить временные затраты на управление правами доступа, а время согласования документа может сократиться за счет учета в системе индивидуальных особенностей и знаний сотрудников компании.

# Проектирование архитектуры КИС

В рамках данной работы, рассматривается архитектура КИС предприятия. Поскольку целью работы является оптимизация системы за счет применения технологий анализа электронных документов в архитектуре КИС, необходимо разработать методику оценки эффективности разрабатываемой архитектуры; сформулировать требования, которым архитектура должна удовлетворять, и на основании требований спроектировать архитектуру с измененной концепцией подхода к ЭД.

## Разработка методики оценки эффективности архитектуры КИС

Одна из основных проблем, с которой сталкиваются проектировщики и разработчики КИС крупной организации, связана, во-первых, со сложностью объекта автоматизации, а во-вторых со сложностью внесения изменений в структуру данных и функции КИС по ходу выполнения жизненного цикла. Эти изменения обусловлены, как внешними факторами (изменение нормативного законодательства, изменение понятийного аппарата, принятого в компании, расширение сферы ответственности), так и внутренними процессами (изменение технологии, изменение организационной структуры).

Для решения проблемы сложности необходима разработка жизнеспособной архитектуры КИС, т.е. такой среды, которая обладает характеристиками адаптируемости и адаптивности. Адаптируемость предполагает возможность изменения КИС в соответствии с вышеперечисленными требованиями (внешними, внутренними). При этом архитектура КИС должны обеспечивать масштабируемость по разным направлениям, таким как: данные, пользователи, серверы, приложения. Под адаптивностью, в свою очередь, понимается способность КИС настраиваться на изменения в структуре данных, функциях, т.е. автоматически перестраиваться для обеспечения ЖЦ КИС.

В связи с тем, что необходима разработка методики оценки эффективности, требуется выполнение задачи обоснования критериев и показателей, определяющих эффективность разработанной архитектуры КИС, включая оценку соответствия архитектуры КИС жизнеспособности. Полученные результаты можно использовать для управления процессом внедрения архитектуры КИС, мониторинга и для дальнейшего повышения эффективности архитектуры.

### Методы оценки архитектуры КИС

Оценка любой ИС может быть проведена как количественными, так и качественными методами. А с точки зрения объекта оценки, методики оценки ИС могут быть разделены на методики, которые оценивают ИС в целом и методики, которые оценивают ИТ-процессы в компании. Более того, методики оценки могут быть направлены на оценку отдельных характеристик системы или строить обобщенные и комплексные оценки ИС.

В любом случае для оценки системы необходим набор критериев, определяющий существенные свойства системы. Обобщенные критерии позволяют рассматривать ИС по различным направлениям, фокусируясь на сильных и слабых сторонах системы. В свою очередь, комплексный критерий позволяет сравнить ИС с некоторой идеальной системой. Определение критериев не имеет общих методик и зависит от объекта исследования и цели, требующей достижения в рамках оценки информационной системы [2].

Оценки ИС с точки зрения оценки качества рассматриваются в многочисленных работах и стандартах, например [35]. Основной характеристикой качества на основе стандарта ISO 9126 определена функциональная пригодность ИС для пользователей. Также рассматриваются такие характеристики как защищенность, способность к взаимодействию и корректности. К факторам качества, согласно [13], относят: эффективность, надежность, сопровождаемость, удобство, применение, универсальность. Все эти характеристики применимы, как в целом ко всей ИС, так и к архитектуре ИС.

### Методика оценки архитектуры КИС

Аудит ИС проводится на основании различных стандартов и методик [12], которые разработаны в основном за рубежом. Среди зарубежных методик оценки ИС следует остановиться на стандарте COBIT. Данный стандарт нацелен на аудит ИТ с точки зрения бизнеса, поэтому финансовые и стратегические критерии занимают в нем значительно место. Кроме этого, в стандарте COBIT делается акцент на оценку процессов аудита ИТ в совокупности всех процессов.

Тем не менее, не существует такой методики оценки, которая учитывала бы специфику разрабатываемой архитектуры КИС с акцентом на управление ЭД и движение ЭД в рамках КИС, а также оценивающей архитектуру с точки зрения жизнеспособности. Поэтому в данной работе будет применена методика оценки архитектуры КИС, базирующаяся на стандартах качества ИС [15] COBIT, а также обеспечивающая оценку архитектуры со стороны жизнеспособности.

Согласно COBIT для оценки эффективности ИТ-процессов можно воспользоваться показателями результативности и показателями эффективности. Данные показатели подойдут и для оценки архитектуры КИС. Показатель результативности представляет собой ключевой показатель достижения цели и может быть измерен только после совершения факта. В свою очередь, показатель эффективности в первую очередь отвечает на вопрос об общей вероятности достижения цели и может быть измерен до получения результата. Таким образом, можно сказать, что проектируемая архитектура, перед этапом определения требований к ее реализации, должна быть оценена с помощью показателей эффективности. Однако уже существующие решения должны быть оценены с помощью показателей результативности.

Таким образом, перед формированием требований целесообразно провести оценку показателя эффективности для проектируемой архитектуры КИС. Формируя показатели обратимся не только к стандарту COBIT, а также к ГОСТ Р ИСО/МЭК 12207–2010, стандарту определяющему ЖЦ ИС, в частности к разделу описывающему процесс проектирования архитектуры.

Таким образом, основываясь на два вышеописанных стандарта, при оценке эффективности проектируемой архитектуры необходимо выполнить следующие шаги.

1. Определить основную бизнес-цель, выполнение которой поддерживается теми или иными технологиями в архитектуре КИС.
2. Выстроить иерархию целей, поддерживающих бизнес-цель (предусмотреть ИТ-цели, цели процесса, цели действий).
3. Выделить выходы процессов всех уровней.
4. Скоординировать все выходы и процессы относительно друг друга.
5. Для каждого процесса определить показатели оценки результативности и эффективности, на основании которых будет выполняться оценка.
6. Определить важность каждого критерия, используя метод анализа иерархий, внутри группы целей.
7. Посчитать суммарные значения, сделать выводы.

Методика оценки эффективности сводится к двум крупным этапам: определение показателей и определения суммарных значений показателей.

## Формирование требований к разрабатываемой архитектуре

Основная цель разрабатываемой архитектуры КИС состоит в решении задач поддержки выполнения основных бизнес-процессов компании, однако акцент в архитектуре делается на работу в КИС с таким объектом, как электронный документ. Для того чтобы разработанная архитектура осуществляла выполнение требующихся от нее задач, необходимо обозначить процессы, которые КИС будет автоматизировать, а затем сформулировать требования, которые предъявляются к разрабатываемой архитектуре, таким образом, в КИС должны быть автоматизированы основные процессы документооборота, перечисленные в   
главе 2 и дополненные процессами, решающими задачи, обозначенные в табл. 2.3.

Основные требования к КИС формулируются исходя из основных автоматизируемых процессов и задач определения жизнеспособности архитектуры КИС. Будучи адаптируемой, архитектура КИС должна удовлетворять следующим ряду требований: требования к программной архитектуре, требования к функциональной архитектуре и требования к технической архитектуре.

**Требования к программной архитектуре**: КИС должна обеспечивать одновременную многопользовательскую работу с рабочих станций, объединенных в локальную вычислительную сеть. КИС не должна иметь технических ограничений на число одновременно работающих пользователей, которое должно зависеть только от технических характеристик используемых аппаратных средств и сетевого оборудования, а также количества приобретенных лицензий.

Для организации взаимодействия между территориально удаленными подразделениями архитектура КИС должна позволять организовать автоматический обмен документами и осуществлять синхронизацию справочников в формате XML. Автоматический обмен документами должен исключать необходимость повторного ручного ввода получателями регистрационной информации о документах.

**Требования к функциональной архитектуре**: архитектура КИС на объектах автоматизации должна состоять из ряда подсистем, представляющих собой совокупность модулей, подсхем данных, компонент доступа к этим данным и компонент логики управленческих и деловых процессов. Основные функции системы должны делиться на 5 групп, согласно ЖЦ ЭД (этапы ввода/вывода – одна группа).

Предъявляемые функциональные требования:

* Должна быть сформирована модель данных, предусматривающая, что для каждого вида документа, подлежащего помещению в КИС должен быть настроен набор атрибутов (метаданные свойств содержимого), описывающих этот документ и его ЖЦ (метаданные контекстной связи), в рамках которого описывается процедура его прохождения через все стадии его жизни.

Требование обеспечивается за счет ввода в информационную модель ЖЦ ЭД блока метаданных.

* Модель данных в архитектуре должна обеспечивать возможность совместного использования вместе с другими компонентами КИС одной и той же информации.

Требование обеспечивается за счет ввода в архитектуру КИС функциональных модулей ввода, вывода, управления доступом.

* Для упрощения администрирования и сопровождения модель данных функциональной архитектуры КИС должна предусматривать объектно-ориентированный подход с возможностью настройки родительских и дочерних типов документов, наследующих основные свойства родительских.
* Все помещенные в КИС материалы должны автоматически индексировать по атрибутам и содержанию, что впоследствии упростит поиск документов и информации в архиве.
* КИС должна обеспечивать разграничение и администрирование доступа к базе данных в соответствии с компетенцией пользователей и определять авторизацию.
* Благодаря применению авторизации на уровне объектов каждый объект содержания, версия и представление содержания, а также каждый контейнер (объект, содержащий другие объекты), начиная с папок и заканчивая репозиториями, должен управляться средствами КИС в течение всего жизненного цикла документа.

Система безопасности архитектуры КИС должна обеспечивать решение следующих задач:

* целостность (предотвращение возможности несанкционированных изменений электронных документов);
* конфиденциальность (разграничение прав доступа к электронным документам);
* аутентичность (подтверждение авторства электронных документов);
* юридическая значимость.

Юридическая значимость документов обеспечивается с помощью использования сертифицированных средств криптозащиты информации (СКЗИ) – электронно-цифровой подписи (ЭЦП).

Должны обеспечиваться:

* идентификация и проверка подлинности субъекта доступа при входе в систему по паролю условно-постоянного действия;
* доступ к информации в соответствии с правами пользователя, назначаемыми администратором при регистрации пользователя в системе;
* регистрация входа (выхода) субъектов доступа в систему (из системы);
* запись в базе данных информации о пользователе, дате и времени создания, изменения и корректировки документа (протоколирование).

**Требования к технической архитектуре**: основой хранения документов в КИС должно быть хранилище корпоративного содержания. Хранилище должно состоять из трех основных компонентов, которые для приложений являются единым целым: файловое хранилище, где находятся элементы содержания, реляционная база данных с таблицами атрибутов и полнотекстовые индексы.

Файловое хранилище должно быть построено на базе корпоративной инфраструктуры хранения данных. На уровне файлового хранилища должно обеспечиваться хранение и миграция файлов содержания на основании определенных политик хранения информации. Например, при изменении статуса документа на «архивный» и настроенной политики долговременного хранения архивной документации, файл содержания должен автоматически перемещаться в архивное файловое хранилище, где он будет храниться в соответствии с регламентами.

База данных должна обеспечивать хранение метаданных, связанных с объектами содержания. Также в базе данных должны храниться списки пользователей, ролей, групп, политики безопасности, применяемые к объектам содержания, настройки системы и пр.

Сервис управления процессами КИС должен отвечать за исполнение документов в рамках определенных, настраиваемых моделей процессов регистрации, согласования, исполнения и другой работы с документами. Данный сервис должен предполагать возможность простого моделирования/редактирования означенных процессов и контроля за эффективностью их исполнения.

Полнотекстовые индексы КИС должны обеспечивать функциональность интеллектуального поиска, как по атрибутам объектов содержания, так и по тексту внутри файлов содержания с учетом семантики русского языка. Необходимо обеспечить возможность полнотекстовой индексации по основным офисным форматам документов (doc, pdf, xls и др.).

Интеграция – система должна обладать развитым интеграционным инструментарием, в частности, должна быть реализована поддержка модульной сервис-ориентированной архитектуры (СОА). Так как использование данной методологии к проектированию архитектуры КИС позволит более гибко организовать развитие систем, адаптацию к изменяющимся требованиям бизнеса, а также максимизировать экономический эффект от использования системы за счет управления жизненными циклами разрабатываемых в рамках ИС служб.

Кроме того, архитектура КИС предусматривает наличие семантических технологий для возможности интеллектуальной работы с ЭД. Серверы должны отвечать за совместное использование онтологической базы знаний и решения базовых задач по работе с онтологической моделью (редактирование и пополнение онтологии), метаданными (формирование семантических метаданных — аннотирование, хранение базы знаний) и информационными ресурсами (поиск, категоризация, навигация по метаданным).

## Проектирование архитектуры КИС

На основе сформированных требований можно перейти к этапу проектирования архитектуры КИС представляющему собой многоэтапный процесс, требующий системного подхода к проектированию архитектуры. Для проектирования полноценной архитектуры КИС необходимо рассмотреть ее на разных уровнях абстракции:

* концептуальный уровень;
* логический уровень;
* физический уровень.

Каждый из уровней архитектуры описан и спроектирован согласно сформированным требованиям к архитектуре, а также ориентированности на управление ЭД.

### Проектирование концептуального уровня архитектуры КИС

На этапе обзора существующих типов архитектур ИС было отмечено, что сервис-ориентированная архитектура (СОА) является наиболее новой концепцией и инструментом проектирования архитектуры КИС. Более того, СОА представляет собой не коробочный продукт или решение, а идеологию информатизации бизнеса, основанную на процессном подходе и методологии управления бизнес-процессами. Таким образом, в виду своей бизнес-ориентированности, а также модульному принципу проектирования, принято решение об использовании концепции СОА для проектирования архитектуры КИС.

Как отмечалось в разделе 1.1.2, особенность концепции СОА заключается в разработке такого программного обеспечения, которое основано на использовании распределенных, слабо связанных заменяемых компонентах, которые оснащены стандартизированными интерфейсами для взаимодействия по стандартизированным протоколам. В рамках СОА следует выделять три основных компонента: поставщик сервисов, потребитель сервисов и реестр сервисов. Под сервисом будем понимать законченный функциональный компонент системы, который может многократно использоваться в программной поддержке различных бизнес-процессов. Иными словами, в архитектуре КИС, спроектированной в такой концепции унификация и автоматизация бизнес-процессов будет осуществлена за счет использования типовых компонентов – сервисов.

Поставщик сервисов обеспечивает реализацию и описание сервиса. Реестр сервисов содержит информацию о существующих сервисах. А потребителем сервисов в свою очередь, являться такой компонент системы, который нуждается в функциональных возможностях поддержки своих бизнес-целей сервисом того или иного поставщика. Кроме того, потребитель сервиса при обращении к реестру для нахождения описания сервиса может напрямую использовать универсальный идентификатор ресурса (URI). Схема взаимодействия компонентов ИС изображена на рис. 3.1. Цифрами обозначена последовательность этапов взаимодействия.

|  |
| --- |
|  |
| * + 1. Схема с этапом сообщения адреса сервиса потребителю |
|  |
| * + 1. Схема с этапом поиска описания в реестре по URI |
| Схемы взаимодействия компонентов ИС в концепции СОА |

Обращая внимание на то, что в проектируемой архитектуре должна учитываться возможность использования семантических технологий, вышеописанная модель может быть усовершенствована с учетом особенностей семантических технологий. В семантической схеме взаимодействия компонентов, определим следующие моменты: a) поставщик сервиса регистрирует в реестре описание в обычном и онтологическом формате; b) потребитель сервиса передает в реестр запрос и онтологическое описание запроса. С учетом этих изменение, схема семантического взаимодействия компонентов ИС примет вид, изображенный на рис. 3.2.

|  |
| --- |
|  |
| Схема семантического взаимодействия компонентов ИС в концепции СОА |

При семантическом взаимодействии компонентов ИС с реестром сервисов взаимодействует семантический сервис, который осуществляет проверку на соответствие семантических совпадений из онтологии запроса от потребителя и онтологий сервисов, состоящих в реестре сервисов. Благодаря такой организации взаимодействия компонентов системы, при котором как в описании сервиса есть смысловое вложение, так и в запросе потребителя, модель СОА обеспечит более релевантное взаимодействие поставщика и потребителя сервиса. А семантический сервис, находящийся на уровне реестра сервисов, организовывает автоматический поиск сервисов и создает единую базу знаний.

Переход к семантической модели СОА вызван наличием недостатков технологий, участвующих во взаимодействии поставщика и потребителя сервисов. В работе [3] обозначены следующие недостатки технологий: спецификация WSDL (Web Services Description Language) – язык описания Web-сервисов и доступа к ним. Основан на языке XML, использующаяся для описания сервисов, жестко регламентирует формат сообщений, используемые протоколы и адрес, по которому находятся сервисы, и не позволяет отразить семантику; реестр UDDI (Universal Description Discovery & Integration) – инструмент для расположения описания Web-сервисов для последующего их поиска слабо интегрируется в существующие ИС и не содержит семантики, необходимой потребителям сервисов.

В связи с вышеперечисленными недостатками технологий, применяющихся в архитектуре СОА, в семантической модели СОА для придания описания сервиса семантической окраски будет использоваться расширение WSDL в виде описания на языке OWL-S. На рис. 3.3 отображены стандартная и семантическая модели СОА КИС с использующимися технологиями.

|  |  |
| --- | --- |
|  |  |
| a) | b) |
| Модели СОА КИС: a) Стандартная модель СОА КИС, b) Семантическая модель СОА КИС | |

Рассмотрим модели СОА КИС более подробно. Уровень потребителей сервисом в семантической модели отличается от стандартной наличием OWL описания запроса. Для обмена информацией между потребителями сервисов и сервисами по транспортному протоколу HTTP используется протокол SOAP, являющийся рекомендованным и самым широко используемым коммуникационным протоколом для Web-сервисов. На уровне сервера ИС включает в себя такие компоненты как:

* корпоративная сервисная шина (Enterprise Service Bus, ESB), отвечающая за взаимодействие элементов ИС (клиентов и сервисов);
* реестр сервисов, хранящий данные о предоставляемых поставщиками сервисах, а также содержащий информацию о бизнес-процессах, находящихся в ИС;
* модуль оркестровки сервисов, предназначенный для хранения серии действий по управлению потока работ (обычно имеющих одну линию выполнения);
* модуль хореографии сервисов необходимый для отражения видимого обмена сообщениями, а также правил взаимодействия между сервисами.

Работа реестра сервисов возможна за счет использования реестра UDDI, хранящего в себе множество точек соединения различных Web-служб. Поставщики сервисов самостоятельно регистрируют сервисы в реестре UDDI, а клиенты самостоятельно запрашивают необходимые сервисы. Также, следует отметить, что в семантической модели присутствует семантический сервис, обрабатывающий семантическое описание сервисов, и при обращении потребителя сверяющий семантическое описание запрашиваемого сервиса с существующими сервисами поставщиков.

Уровень поставщиков сервисов в семантической модели отличается от стандартной тем, что вместе с WSDL-файлом описания сервиса поставщик передает OWL-S описание (онтологию сервиса), в то время как в стандартной модели WSDL-файл – это единственный вариант описания сервиса.

Таким образом, проектируемая архитектура приобрела изменения уже на этапе определения уровней трех участников (поставщик, потребитель, реестр) модели СОА. Далее следует рассмотреть архитектуру на более низком уровне, где подробно будут разложены функциональные модули архитектуры, необходимые для реализации управления ЭД с учетом применения семантической концепции.

### Проектирование логического уровня архитектуры КИС

Проектирование логического уровня архитектуры предполагает определение принципов построения архитектуры, функциональных возможностей по обработке и обмену информацией, форматами представления данных и по организации информационной поддержки бизнес-объектов. Также на данном уровне архитектуры необходимо обозначить типы информационных процессов, распределенных по элементам КИС в соответствии с функциональными потребностями компании.

Поскольку, в рамках данной работы, архитектура КИС ориентируется на управление ЭД, все технологии совершенствования направлены на изменение концепции работы ИС с документами. Естественно, что документы, данные и знания в компаниях хранятся в электронной среде в виде корпоративной памяти [19]. Целесообразно рассмотреть структуру корпоративной памяти в архитектуре КИС на трех уровнях: онтологический уровень (т.к. было принято решение о выборе онтологического подхода семантических технологий), содержательный уровень, уровень физического хранения (см. табл. 3.1). Уровень физической хранения является более детальным и относится к этапу проектирования физического уровня архитектуры, поэтому в данном разделе будут рассмотрены первые два уровня корпоративной памяти.

##### Структура корпоративной памяти

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | Уровень | Данные | Знания | Документы |
| 1 | Онтологический уровень | Метаданные | Онтологии | Структура архивов |
| 2 | Содержательный уровень | Справочники, каталоги | Правила выбора | Отчеты, методики, технологии |
| 3 | Уровень программной реализации | БД, файлы, Web-страницы | Базы знаний | Электронные документы, чертежи и др. |

Как видно из приведенной таблицы, на онтологическом уровне данные представляются в виде метаданных, в нашем случае, с помощью метаданных содержания и метаданных семантической связи. Содержание документов представляется в виде онтологий. В качестве онтологий, содержащих знания, в данной работе используются: онтологии описания структуры документов; онтологии описания документов; онтологии описания предметной области. При масштабировании архитектуры, вызванном различными потребностями, количество онтологий может быть увеличено. Что касается документов, то на данном уровне документы организуются в некоторый набор архивов, по тем критериям, которые продиктованы непосредственно потребностями бизнеса.

Описав форматы представления и хранения данных, целесообразно перейти к уровню бизнес-логики архитектуры (см. рис. 3.4).

|  |
| --- |
|  |
| Уровень бизнес-логики архитектуры КИС |

Данный блок содержит бизнес-логику предметной области (управление ЭД), а также иерархию бизнес-процессов компании. К бизнес-объектам можно отнести: документы, роли, события, задачи и др. Также в бизнес-логике содержится понятие коллекций, они могут быть общие, ролевые, персональные; могут быть организованы в виде журналов, папок и другим способом. Немаловажным элементом данного уровня является обозначение функциональных задач, требующих решения. Все элементы бизнес-логики предметной области, так или иначе соединяются в иерархии бизнес-процессов компании.

Кроме уровня бизнес-логики, на этапе формирования логического уровня архитектуры, необходимо выделить уровень инфраструктуры (см. рис. 3.5).

|  |
| --- |
|  |
| Уровень инфраструктуры архитектуры КИС |

На данном уровне определяется сервисная поддержка организации Web-сервисов, использующихся в КИС; сами сервисы; подсистемы, каждая из которых включает в свой состав свой набор сервисов; модули, в которые организованы подсистемы по схожести подсистем к решению задач по управлению ЭД на разных этапах ЖЦ ЭД.

Особенное внимание следует обратить на модуль семантического ядра, реализующий весь функционал по семантической обработке ЭД. Остальные модули включают в свой состав такие подсистемы, которые пользуются подсистемами модуля семантического ядра, как переходным объектом к хранению и представлению данных с помощью онтологического подхода. В составе подсистем перечислены такие модули с подсистемами, которые носят характер управления ЭД, естественно, что в зависимости от потребностей компании, количество модулей может быть увеличено для решения тех или иных задач, носящих специфический для компании характер.

К числу функциональных модулей архитектуры относятся такие модули, которые удовлетворяют функциональным и техническим требованиям, предъявляемым к архитектуре и входящие в рамки ЖЦ ЭД (исключая хранение и вывод из ИС). Функциональные модули КИС с составляющими подсистемами перечислены ниже.

1. Ввод документов в ИС (модуль ввода):
   1. подсистема распознавания образов;
   2. подсистема обработки форм;
   3. подсистема аннотирования;
   4. подсистема индексирования;
   5. подсистема категоризации.
2. Обеспечение сохранности документов (модуль безопасности):
   1. подсистема идентификации пользователей;
   2. подсистема контроля версиями;
   3. подсистема контроля доступа к контенту;
   4. подсистема контроля взаимоисключений.
3. Управление (модуль управления):
   1. подсистема управления документами;
   2. подсистема управления записями;
   3. подсистема управления web-контентом;
   4. подсистема управления электронной почтой;
   5. подсистема управления цифровыми активами.
4. Доставка (модуль доставки):
   1. подсистема интеграция контента;
   2. подсистема выборки;
   3. подсистема синдикации;
   4. подсистема локализации;
   5. подсистема публикации.

Рассмотрим состав и структуру модуля семантического ядра. Предлагаемая семантическая подсистема поддерживает следующие варианты использования онтологии: аннотирование объектов; семантический поиск; формирование списка объектов, связанных с исходным объектом; категоризация; индексирование.

Онтологическая модель предметной области задает семантику понятий, которые используются для описания информационных объектов архитектуры КИС. Такие описания называются семантическими метаданными, и они позволяют:

1. устранить лексическую многозначность терминов, используемых для описания информационных объектов;
2. определять соответствие между различными информационными объектами, используя онтологию.

Более того, семантические метаданные могут описывать объект с разных точек зрения: структуры, контекста и контента. А в рамках проектируемой архитектуры КИС описание контента наиболее важно, поскольку в архитектуре предусмотрена реализация функции полнотекстовой обработки информации.

### Проектирование физического уровня архитектуры КИС

Физический уровень архитектуры КИС характеризует физическую организацию техниче­ских средств, реализующих ту или иную совокуп­ность информационных процессов и оснащенных каналами передачи и хранения информации. На данном уровне проектирования архитектуры КИС необходимо обратиться к третьему уровню структуры корпоративной памяти, представленной в табл. 3.1. Предлагаемые уровень обработки данных и организации хранения представлен на рис. 3.6.

|  |
| --- |
|  |
| Физический уровень архитектуры КИС |

Функциональность семантической подсистемы по работе с онтологиями расположена в сервере онтологий, а функциональность по работе с семантическими метаданными - в сервере семантических метаданных. Сервер онтологий – это отдельно функционирующее приложение, хранящее множество онтологий и предоставляющее к ним доступ. Сервер онтологий, в свою очередь, используется различными приложениями и модулями, которым нужны различные функции в процессе ЖЦ ЭД. Связь подсистемы работы с онтологией с другими подсистемами и модулями возможна через подсистему, хранящую правила преобразования данных ЭД в элементы онтологий.

Для реализации в семантической подсистеме выбранных вариантов использования онтологии, сервер онтологий предоставляет следующие функции: хранение онтологий; извлечение онтологий; логический вывод; поиск в онтологии запрашиваемых понятий и отношений.

Сервер семантических метаданных – это отдельно функционирующее приложение, хранящее семантические метаданные, предоставляющее к ним доступ и обрабатывающее их. Для реализации выбранного варианта использования онтологии сервер семантических метаданных, тесно взаимодействует с сервером онтологий и предоставляет следующие функции:

* составление семантических метаданных;
* хранение семантических метаданных;
* извлечение семантических метаданных;
* сравнение семантических метаданных.

Спроектированные уровни архитектуры КИС соединенные вместе, позволяют увидеть все взаимосвязи между уровнями. Архитектура КИС, спроектированная для управления ЭД с помощью функций семантического ядра, а также в своей концепции имеющая изменение стандартной архитектуры СОА концептуально меняют подход к обработке ЭД в КИС, работая с ЭД одновременно, как с логической единицей системы, так и с объектом, наполненным различными блоками хранения информации, а также метаданными. Принципиальная архитектура КИС представлена в приложении I.

# Заключение

Основным направлением данной работы являлось определение теоретической базы для проектирования архитектуры КИС, формулировка требований к создаваемой архитектуре КИС и проектирование архитектуры КИС, ориентированной на работу с электронными документами с применением методов анализа неструктурированных электронных документов. В работе были исследованы такие предметные области, как электронный документооборот, типология архитектур корпоративных информационных систем, а также технологии и подходы архитектурных решений к организации управления электронными документами в ИС.

В результате решенных задач, поставленных в рамках выпускной квалификационной работы, можно сделать следующие выводы:

* концепция организации управления электронными документами в ИС должна учитывать содержание документа, метаданные свойств содержимого, метаданные семантической связи;
* существующие решения по применению семантических технологий в архитектурах ИС решают узконаправленные задачи, большей частью связанные со структуризацией электронных документов;
* из существующих методов представления знаний, онтологический подход предоставляет большой набор возможностей, применяемых в архитектуре КИС, а также является одним из самых перспективных в условиях направленности развития ИТ;
* типовые бизнес-процессы, связанные с управлением электронными документами на данной стадии не поддерживаются мощными решениями семантического подхода, в свою очередь, применение этого подхода позволяет ускорить некоторые процессы, связанные с предварительной обработкой электронных документов перед этапом анализа данных;
* методики по оценки эффективности архитектур КИС, реализующих управление ЭД, отсутствуют. Существуют общие методики, в совокупности, позволяющие создать основу для оценки эффективности проектируемой, в рамках данной работы, архитектуры;
* требования, предъявляемые к архитектуре КИС должны полностью диктовать правила для управления ЭД в ИС, с учетом применения в архитектуре онтологического подхода;
* проектируемая архитектура, основанная на модели сервис‑ориентированной архитектуры должна быть дополнена семантическим сервисом для семантической поддержки web-сервисов.

Кроме того, можно сказать, что архитектура КИС, спроектированная для управления электронными документами с помощью функций семантического ядра, а также в своей концепции имеющая изменение стандартной архитектуры СОА концептуально меняют подход к обработке электронных документов в КИС, работая с электронным документов одновременно, как с логической единицей системы, так и с объектом, наполненным различными блоками хранения информации, а также метаданными.

В будущем данная тема является перспективной, развитие и расширение поставленной цели может вестись в направлении совершенствования, как архитектуры КИС, так и применения методов семантического подхода (кроме онтологического). Кроме этого, разработанная методика оценки эффективности архитектуры также имеет потенциал дальнейшего развития.

# Библиографический список

1. *Арлазаров В.Л.* Системы обработки документа. Основные компоненты. // Управление информационными потоками: сб. тр. ИСА РАН / под ред. В.Л. Арлазаров, Н.Е. Емельянов. – М.: УРСС. 2002. с. 3-20.
2. *Баринов В.А*. Теория систем и системный анализ в управлении организациями / В.А. Баринов, Л.С. Болотова, В.Н. Волкова / под ред. В.Н. Волковой А.А. Емельянова – М.: Финансы и статистика. 2006. – 848 с.
3. *Гонтарь Н*.*А*. Семантическое представление сервис-ориентированной архитектуры информационных систем // Материалы VII международной научно-технической конференции «Информатика и компьютерные технологии». – Донецк. 2011. – с. 343-345.
4. *Грекул В.И.* Проектирование информационных систем / В.И. Грекул, Г.Н. Денищенко, Н.Л. Коровкина. – М.: Интернет-Уни-т информ. технол.,  
   2005. – 304 с.: ил.
5. Делопроизводство и архивное дело. Термины и определения [Текст]: ГОСТ Р 51141-98. – М.:1998. 15 с.
6. *Ланин В*.*В.* Онтологии как основа функционирования систем обработки электронных документов // Материалы Всероссийской конференции «Знания-Онтологии-Теории». – Новосибирск. 2009. – Т.2. с. 173-177.
7. *Ланин В*.*В.* Онтология электронных документов как основа извлечения знаний о предметной области при ее моделировании / В.В. Ланин, Л.Н. Лядова // Материалы Всероссийской конференции «Знания-Онтологии-Теории». – Новосибирск. 2013.
8. *Макаров С.Б.* Что такое ECM? // Директор Информационной Службы. – 2003. – № 4.
9. Межгосударственный стандарт. Информационные технологии. Система стандартов по базам данных. Концепция и терминология для концептуальной схемы и информационной базы [Текст]: ГОСТ 34.320 –96. – Минск.: 1996. 46 с.
10. *Пагонин В*.*А.* Корпоративные информационные системы / В.А. Погонин, А.Г. Схиртладзе, С.И. Татаренко, С.Б. Путин. – Тамбов:   
    Изд-во ФГБОУ ВПО «ТГТУ». 2012. – 144 с.: ил.
11. *Пантелеев М.Г*. Интеллектуальные образовательные среды на основе технологий Семантического Web / М.Г Пантелеев, Д.В. Пузанков, П.В. Сазыкин, Д.А. Сергеев // Труды Международн. конф. «Искусственные интеллектуальные системы». – М.: Изд-во Физ.-мат. Литературы. 2002. – с. 236 - 243.
12. Проект Федерального закона № 107599-3 «Об электронном документе» [Текст].
13. *Репин В.В*. Бизнес-процессы: построение, анализ, регламентация. – М.: РИА «Стандарты и качество». 2007. – 240 с.: ил.
14. Сайт компании «Ай-Теко», Программный продукт «Аналитический курьер» [Электронный ресурс] [Режим доступа: http: http://www.i-teco.ru/solutions/business\_intelligence\_products/analytical\_courier/] [Проверено: 14.04.2014].
15. Система стандартов по информации, библиотечному и издательскому делу. Процессы управления документами. Метаданные для документов [Текст]: ГОСТ Р ИСО 23081-1-2008. – М.: Стандартинформ. 2008. 23 с.
16. Система стандартов по информационному, библиотечному и издательскому делу. Управление документами. Общие требования [Текст]: ГОСТ Р ИСО 15489-1-2007. – М.: Стандартинформ. 2007. 34 с.
17. *Таненбаум Э.* Распределённые системы. Принципы и парадигмы / Э. Таненбаум, М. ван Стеен. Пер. с англ. СПб.: Питер. 2003. 877 с.: ил.
18. *Тарасов В.Б*. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС. 2002. 180 с.
19. *Тузовский А.Ф*. Системы управления знаниями (методы и технологии) // А.Ф. Тузовский, С.В. Чириков, В.З. Ямпольский / под общ. ред. В.З. Ямпольского. – Томск: Изд-во НТЛ. 2005. – 260 с.
20. Унифицированные системы документации. Унифицированная система организационно-распорядительной документации. Требования к оформлению документов [Текст]: ГОСТ Р 6.30-2003. – М.: 2003. 19 с.
21. Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» [Текст]: М.: 2006.
22. *Elder J*. Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. – London, UK: Springer. 2012. 812 p.
23. Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1) [Text]: ISO 19005-1:2005. –ISO: 2005. 29 p.
24. Electronic Imaging – Information Stored electronically – Recommendations for trustworthiness and reliability [Text]: ISO/TR 15801:2004. –ISO:2004. 15 p.
25. Long-term preservation of electronic document-based information [Text]: ISO/TR 18492:2005. – ISO: 2005. 18 p.
26. Media error monitoring and reporting techniques for verification of stored data on optical digital data disks [Text]: ISO 12142:2001. – ISO: 2001. 25 p.
27. Model Requirements for the Management of Electronic Records. MoReq2 Specification. Update and extension. – European Commission. 2008. 212 p.
28. Recommendations for the expungement of information recorded on write-once optical media [Text]: ISO/TR 12037:1998. – ISO: 1998. 12 p.
29. Recommended Practice for Architectural Description of Software-Intensive Systems. Control Objectives for Information and related Technology (COBIT 4.0).: ANSI / IEEE Std 1471-2000. IT Governance Institute (ITGI). 2005. 124 p.
30. *Riis J. O*. Design of Enterprise Information Systems: Roots, Nature and New Approaches // 5th IFIP WG 8.9 Working Conference. – Aalborg, Denmark. 2011.
31. *Shilakes C.С.* Enterprise Information Portals. // IS Journal. November, 1998.
32. *Shwartz D. G.* Encyclopedia of Knowledge Management. – London, UK: Idea Group Reference. 2006. 946 p.
33. The AIIM Document Management Alliance (DMA) specifications [Электронный ресурс] [Режим доступа: http://dmatech.info/] [Проверено: 14.04.2014].
34. The Dublin Core metadata element set [Text]: ISO 15836:2003. – ISO: 2003. 14 p.
35. The Global Community of Information Professionals [Электронный ресурс] [Режим доступа: http:// www.aiim.org/] [Проверено: 14.04.2014].
36. The World Wide Web Consortium (W3C) an international community [Электронный ресурс] [Режим доступа: http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#(24)] [Проверено:14.04.2014]
37. XML Key Management Specification (XKMS) [Электронный ресурс] [Режим доступа: http://www.w3.org/TR/xkms2/] [Проверено: 14.04.2014].
38. *Zachman A.* A framework for Information Systems Architecture // IBM Systems Journal. 1987. Vol. 26. № 3.

###### Реквизиты оформления организационно-распорядительных документов по ГОСТ Р 6.30-2003

##### Реквизиты оформления организационно-распорядительных документов

|  |  |
| --- | --- |
| № реквизита | Наименование реквизита |
| 01 | Государственный герб Российской Федерации |
| 02 | герб субъекта Российской Федерации |
| 03 | эмблема организации или товарный знак (знак обслуживания) |
| 04 | код организации |
| 05 | основной государственный регистрационный номер (ОГРН) юридического лица |
| 06 | идентификационный номер налогоплательщика / код причины постановки   на учет (ИНН/КПП) |
| 07 | код формы документа |
| 08 | наименование организации |
| 09 | справочные данные об организации |
| 10 | наименование вида документа |
| 11 | дата документа |
| 12 | регистрационный номер документа |
| 13 | ссылка на регистрационный номер и дату документа |
| 14 | место составления или издания документа |
| 15 | адресат |
| 16 | гриф утверждения документа |
| 17 | резолюция |
| 18 | заголовок к тексту |
| 19 | отметка о контроле |
| 20 | текст документа |
| 21 | отметка о наличии приложения |
| 22 | подпись |
| 23 | гриф согласования документа |
| 24 | визы согласования документа |
| 25 | оттиск печати |
| 26 | отметка о заверении копии |
| 27 | отметка об исполнителе |
| 28 | отметка об исполнении документа и направлении его в дело |
| 29 | отметка о поступлении документа в организацию |
| 30 | идентификатор электронной копии документа |

###### Элементы схем БП в нотации ARIS eEPC

##### Обозначения элементов схемы БП в нотации ARIS eEPC

| № п/п | Элемент схемы | Наименование | Обозначение |
| --- | --- | --- | --- |
| 1 | Пример функции «Регистрация документа в СЭД» | Функция | Предметно-ориентированное задание или действие, выполняемое над объектом, в результате которых достигается одна или несколько целей, в рамках БП, стоящих перед компанией. |
| 2 | Пример события «Замечания по документу отсутствуют» | Событие | События переключают функции и могут быть результатом выполнения функции. |
| 3 | Пример должности «Сотрудник отдела делопроизводства» | Должность (Роль) | Тип организационной единицы, обозначающий исполнителя (конкретную должность) заданий, которые реализуются для достижения целей деятельности компании (в рамках описываемого БП). |
| 4 | Пример документа «Проект документа» | Документ | Компонент схемы, описывающий носитель информации (документ), который отображает входные и выходные данные в форме документов, журналов, сообщений. |
| 5 | Пример ИС «СЭД» | Информационная система | Компонент схемы, обозначающий ИС, программное обеспечение, которое используется для достижения целей деятельности компании (в рамках описываемого БП). |
| 6 |  | Оператор XOR (исключающее ИЛИ) | Оператор переключает события и функции по правилу «может быть выполнено только одно событие/функция из всех» |
| 7 |  | Оператор AND (И) | Оператор переключает события и функции по правилу «должны быть выполнены все события/функции» |
| 8 |  | Оператор OR (ИЛИ) | Оператор переключает события и функции по правилу «должно быть выполнено по крайней мере одно событие/функция» |

##### Обозначения типов связи между элементами схемы БП

|  |  |  |  |
| --- | --- | --- | --- |
| № п/п | Что соединяется | Тип связи | С чем соединяется |
| 1 | Функция | «activates» - является результатом | Событие |
| 2 | Событие | «activates» - переключает | Функция |
| 3 | Функция | «has output of» - порождает выход | Документ |
| 4 | Документ | «is input of» - порождает вход | Функция |
| 5 | Информационная система | «supports» - сопровождает | Функция |
| 6 | Должность | «carries out» - выполняет | Функция |

###### Бизнес-процесс «Ввод документов в ИС (сканирование)»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Ввод документов в ИС (сканирование)»,  модель «как-есть» |

###### Бизнес-процесс «Регистрация документа в ИС»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Регистрация документа в ИС»,  модель «как-есть» |

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Регистрация документа в ИС»,  модель «как-должно быть» |

###### Бизнес-процесс «Выполнение заявки на предоставление документа из электронного архива»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Выполнение заявки на предоставление документа из электронного архива», модель «как-есть» |

###### Бизнес-процесс «Выполнение запроса на удаление документа из ИС»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Выполнение запроса на удаление документа из ИС», модель «как-есть» |

###### Бизнес-процесс «Выполнение запроса на редактирование документа в архиве»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Выполнение запроса на редактирование документа», модель «как-есть» |

###### Схема бизнес-процесса «Формирование, согласование, утверждение документа»

|  |
| --- |
|  |
| 1. Схема бизнес-процесса «Формирование, согласование, утверждение документа», модель «как-есть» |

###### Схема архитектуры КИС

|  |
| --- |
|  |
| Архитектура КИС |