

Разработка концепции электронного архива

Наталья Храмцовская

к.и.н., ведущий эксперт по управлению документацией компании «ЭОС»,
член Гильдии Управляющих Документацией и ARMA International

С чем нужно определиться прежде, чем выбирать концепцию инфраструктуры электронного архива? Какие варианты ее построения бывают? Сколько это стоит? Ответы на вопросы даны на основе опыта построения государственных электронных архивов, многое из рассказанного может быть применимо в работе с корпоративными архивами.

См. статью
«Электронные
архивы и их
проблемы»
на стр. 66
журнала
№ 3' 2012

Электронный архив по большому счету создается так же, как и традиционный архив на бумажных носителях, только в данном случае приходится решать ряд дополнительных проблем, связанных с особенностями электронных документов.

Главным условием успеха проекта является поддержка высшего руководства. Электронный архив – удовольствие достаточно дорогое, поэтому с самого начала нужно понять, чего ожидает от него руководство, и так продумать ход проекта, чтобы первая существенная отдача была продемонстрирована уже через несколько месяцев и чтобы в дальнейшем на каждом этапе работ начальству можно было бы предъявить какой-либо яркий результат.

Пример 1

Центральный государственный электронный архив Украины (ЦГЭА), например, начал свою деятельность с создания архивных коллекций значимых для общества интернет-ресурсов.

В Центральном государственном архиве Украины на январь 2011 года хранилось несколько архивных коллекций интернет-ресурсов, в том числе коллекция материалов о выборах Президента Украины 2010 года. «Мы заключили соглашения с политическими партиями, с кандидатами, которые баллотировались на выборах, на право копирования их сайтов, – рассказывал корреспондентам Анатолий Гаврилович Лавренюк, директор Центрального государственного электронного архива Украины. – Наша коллекция содержит 27 сайтов, причем пять из них уже исчезли из Интернета»¹.

Как показывает опыт зарубежных проектов, от момента открытия проекта до ввода электронного государственного архива в эксплуатацию обычно проходит около 5 лет. Из этого времени где-то год уходит на создание инфраструктуры архива, установку оборудования и развертыва-

¹ Юрій Потапшій «Електронна Україна: зроблено лише перший крок» // сайт журналу «Віче», декабрь 2010, <http://www.viche.info/journal/2301/>.

ние программного обеспечения. Куда больше времени – примерно 2-3 года – уходит на разработку концепции архива, и это вполне понятно: неудачные решения на данной стадии могут повлечь многочисленные проблемы, с которыми придется бороться в течение длительного времени, и существенные дополнительные расходы.

Нужно тщательно оценить, какие ресурсы (финансовые, кадровые, материальные) потребуются для создания архива, какую нормативную и методическую базу придется разработать, спланировать этапы работы и определить ожидаемую отдачу.

Организационно-правовой статус архива

При разработке концепции электронного архива ключевым является вопрос о том, каким будет его правовой статус: *кому он будет подчиняться, какими законодательными и нормативными документами будет регламентироваться его деятельность, как будут строиться отношения архива со службами документационного обеспечения организаций-источников комплектования и бумажными архивами.* Во всех проектах государственных электронных архивов, которые нам удалось выявить и изучить, была выбрана стратегия создания самостоятельного архива, и только в нашей стране решено идти по пути создания Центра хранения электронных документов (ЦХЭД) на базе уже действующего государственного архива².

При разработке концепции электронного архива необходимо продумать и ответ на вопрос о том, *предполагается ли принимать на архивное хранение документы только постоянного и длительного срока хранения, или же в него будут поступать на хранение и документы временного срока хранения.* Есть, например, страны, которые собираются создать как государственные электронные архивы для документов постоянного хранения, так и отдельные центры хранения электронной документации, которые будут хранить законченные делопроизводством электронные документы органов государственной власти временного срока хранения, так называемые «промежуточные» архивы.

Очень важно помнить, что архив – это прежде всего его коллектив. Следует с самого начала создать группу профессионалов-единомышленников, вокруг которой и будет формироваться коллектив архива. На этапе создания архива эта группа может быть небольшой. Одной из ее задач будет разработка детальной концепции электронного архива.

Пример 2

Команда проекта создания электронного архива австралийского штата Новый Южный Уэльс состоит из 6 человек: руководитель проекта, два архивиста, двое разработчиков и специалист по вопросам безопасности и оборудованию.

Как обстоят дела с электронными документами в государственных архивах сейчас? Кто окажет методическую помощь в организации хранения, учета и использования электронных документов? Каковы планы Росархива по созданию Центра хранения электронных документов? Сколько Правительство готово выделить денег и на что? Ответы на эти вопросы вы найдете в статье «Центр хранения электронных документов: планы и перспективы» на стр. 67 журнала № 2' 2012

² Храмцовская Н.А. Центр хранения электронных документов: планы и перспективы // Делопроизводство и документооборот на предприятии. – 2012. – № 2. – С. 67–76.

Если в архиве планируется хранить сложные объекты, такие, как базы данных и веб-сайты, то для обеспечения их сохранности потребуется приобретение специального *программного обеспечения*. В идеале архив должен иметь возможность опробовать различное программное обеспечение и выбрать то, которое ему подходит. Следует отметить, что современные электронные архивы основное программное обеспечение, как правило, закупают у известных поставщиков, но в то же время активно используют для решения отдельных задач и свободное, и условно-свободное программное обеспечение (например, для идентификации форматов, для преобразования одних форматов в другие).

Очень важно при разработке архива также продумать вопросы, связанные с *правами на хранимую в архиве информацию* и правами на хранение и использование коммерческого *программного обеспечения*, необходимого для доступа к документам в специфических форматах.

На кого будет работать электронный архив?

Необходимо по возможности быстрее *определиться с источниками комплектования* архива и с тем, *какого рода документы и материалы будут приниматься* на хранение. Будут ли, помимо документов федеральных органов власти, приниматься на хранение документы региональных и муниципальных органов власти, коммерческих организаций, частные коллекции? Будет ли архив обслуживать все ветви власти или только исполнительную власть? Будут ли в архив поступать только юридически значимые электронные подлинники или это будет более широкий круг материалов, включающий заверенные и заверенные электронные копии бумажных документов и иную полезную электронную информацию?

Пример 3

Государственный электронный архив США ERA к настоящему времени уже принял на хранение электронные документы президентских администраций и ряда федеральных органов исполнительной власти. Ведется подготовка к приему документов Конгресса США. Для каждого из этих комплексов документов приходится учитывать их специфику и особенности как в процессе передачи документов на архивное хранение, так и при организации их хранения и предоставления к ним доступа.

Большое значение имеет вопрос о том, *будет ли перед архивом поставлена задача поддержки оперативной работы фондообразователей, стыковки с имеющимся у них системами* электронного документооборота, деловыми системами, электронной почтой и интернет-порталами, а также задача предоставления публичного доступа к части документов, в том числе в режиме он-лайн.

Учитывая повышенный интерес к использованию *облачных технологий* при управлении документами, стоит подумать и о том, нельзя ли

на базе электронного архива создать некое облачное решение для менее развитых в техническом отношении и ограниченных в средствах органов государственной власти, не способных самостоятельно обеспечить сохранность своих электронных документов.

Выбор варианта инфраструктуры архива

Известные проекты создания электронных архивов достаточно сильно отличаются по используемым технологиям, функциональным возможностям и стоимости. По итогам их анализа можно условно выделить четыре наиболее характерных подхода к созданию необходимой инфраструктуры. Это:

1. Создание архивной информационной системы, обеспечивающей прием на хранение и постоянную сохранность документов всех ветвей власти федерального уровня (включая информацию, отнесенную к государственной тайне), а также решение ряда других дополнительных задач, таких, как оперативное представление государственным органам необходимой им информации, обеспечение публичного доступа к государственной информации, в том числе и в онлайн-режиме, поддержка электронного согласования сроков хранения документов, образующихся в деятельности государственных органов и др.³

Пример 4. Государственный электронный архив США (программа ERA)

Примером такого проекта может служить государственный электронный архив США (программа ERA) – самый амбициозный и самый дорогостоящий в мире, концепция которого включает в себя чуть ли не все мыслимые функциональные возможности.

Первоначально планировалось создать систему, способную решать все возможные задачи, начиная от электронного согласования сроков хранения документов и автоматизации процесса сдачи электронных документов и заканчивая постоянным хранением (с проведением по мере необходимости миграции) документов всех ветвей власти и предоставлением оперативного доступа к этим документам федеральным органам власти, организациям и гражданам. Однако по ходу проекта сроки его реализации затягивались, стоимость росла, к тому же за это время уже успели смениться два поколения компьютерных технологий, появились новые концепции, на которые проект изначально рассчитан не был. В итоге проект ERA столкнулся со множеством проблем.

Излишние амбиции и ошибки при планировании проекта привели к тому, что неоднократно срывались сроки выполнения его ключевых этапов. В результате для приема на хранение президентских документов администрации Дж. Буш-младшего пришлось в срочном порядке задействовать резервный вариант и до-

³ См.: Храмовская Н.А. Современные идеи и опыт в области государственного управления: Межведомственное электронное взаимодействие. – М. – 2010. – С. 127–134.

полнительно закупать «коробочную» систему. Более того, в связи с тем, что развертывание и этой системы запаздывало, пришлось больше года поддерживать в рабочем состоянии унаследованные системы, из которых документы передавались на хранение. При этом на репутации проекта негативно отразилось то, что реально сработавшие «аварийные» меры оказались на порядок дешевле стоимости этапа основного проекта, и это вызвало дополнительное внимание и вопросы со стороны контролирующих органов.

2. Создание чисто архивной информационной системы, опирающейся на единые обязательные правила подготовки документов к передаче на хранение источниками комплектования и на использование специального контейнерного формата, применяемого для сдачи, хранения и распространения информации.

Пример 5. Электронный архив штата Виктория в Австралии

Более скромный по масштабам и решаемым задачам проект регионального электронного архива штата Виктория в Австралии. Этот архив дольше всех работает в штатном режиме, его концепция наиболее глубоко концептуально проработана и по ней имеется довольно много публикаций. Особенно интересен опыт использования стандартного контейнерного формата для сдачи, хранения и распространения документов, а также варианты применения криптографии и электронных цифровых подписей⁴.

3. Применение широко распространенного программного обеспечения и оборудования дает возможность значительно удешевить проект за счет относительно низкой стоимости таких программных продуктов и аппаратных средств и возможности привлекать для их обслуживания менее высокооплачиваемый ИТ-персонал. Используется дешевая память на жестких дисках, резервирование на магнитных лентах. Частично применяется и «музейный» подход, когда закупается оборудование, позволяющее читать информацию с носителей устаревших видов.

Пример 6. Электронный архив штата Вашингтон в США

Электронный архив штата Вашингтон, США – пример дешевого, быстро развертываемого решения на основе «коробочного» программного обеспечения общего назначения. Это решение в настоящее время тиражируется в ряде других штатов США. Первичное развертывание заняло менее года, во многом благодаря использованию хорошо известных «коробочных» продуктов компании Майкрософт. К настоящему времени архив содержит порядка 100 миллионов электронных документов, значительная часть которых доступна через Интернет.

⁴ См.: Храмовская Н.А. Современные идеи и опыт в области государственного управления: Межведомственное электронное взаимодействие. – М. – 2010. – С. 75–86.

4. Использование специализированного «коробочного» программного обеспечения «Safety Deposit Box» фирмы Tessella. Это экономичное, быстро развертываемое решение, основной особенностью которого является наличие развитых средств планирования и автоматизации работ по обеспечению долговременной сохранности электронных документов (таких, например, как реестр файловых форматов, автоматическое выявление файлов в устаревших форматах и автоматизированный выбор для них путей миграции). О качестве данного продукта свидетельствует и тот факт, что Национальные Архивы Великобритании и фирма Tessella получили в 2010 году за его разработку престижную государственную премию.

Пример 7. Использование специализированного программного обеспечения компании Tessella

Данным недорогим вариантом создания электронного архива уже воспользовались в ряде стран: в Великобритании, Австрии, Швейцарии, Финляндии, Эстонии.

Опыт всех этих проектов заслуживает изучения, тем более что практически в каждом из них есть свои «изюминки» – оригинальные идеи и решения.

При разработке концепции электронного архива важно помнить и о том, что система должна быть способна развиваться и что в какой-то момент может потребоваться провести замену оборудования и/или программного обеспечения. В этой связи следует по возможности избегать чрезмерной зависимости от конкретных поставщиков.

Сколько это стоит?

Когда речь заходит о хранении электронных документов, то достаточно часто в качестве основного аргумента в его пользу выдвигается утверждение о его дешевизне по сравнению с хранением бумажных документов. Однако при подсчете экономии разработчики различных программ как-то не любят вспоминать о том, что, в отличие от бумажных документов, электронные документы нельзя просто положить на полку и забыть там на столетие-другое – они требуют непрерывного ухода и внимания. И хотя удельные расходы в расчете на один электронный документ невелики и постоянно снижаются, **общая сумма затрат на создание и содержание электронного архива на достаточно длительном интервале времени, скорее всего, окажется существенно больше, чем при бумажном хранении.**

Решения о финансировании проектов принимаются исходя из финансовых возможностей той или иной страны или региона. В Таблице приведены данные о расходах на создание и текущую деятельность федеральных и региональных электронных архивов семи стран. Стоимость проектов колеблется в пределах от нескольких миллионов долларов до почти миллиарда долларов.

Таблица. Стоимость проектов электронных архивов

Страна (регион), название проекта	Информация о стоимости проекта
США, «Архив электронных документов» (Electronic Records Archive, ERA)	430 млн (2005–2011) + 240 млн (2011–2030) Подсистема хранения президентских документов – 40 млн долларов
США, штат Вашингтон, «Электронный архив штата Вашингтон» (Washington State Digital Archives)	Всего за 8 лет (оценка 2004 г.) – 18,8 млн долларов, в т.ч.: – Здание – 14,5 млн – Оборудование 1,5 млн первоначально + 7,1 млн в течение 8 лет. – ПО 1,1 млн первоначально + 6 млн в течение 8 лет. – Ежегодно 2,5 млн долларов
Великобритания, Национальный электронный архив баз данных (NDAD)	Ежегодно 1 млн ф.ст.
Австрия, «Электронный архив Австрии» Государственного архива Австрии (Digitales Archiv Österreich)	4,6 млн евро (2009-2010)
Швейцария, Электронный архив Федерального архива Швейцарии (Schweizerisches Bundesarchiv)	1,3 млн швейцарских франков (2008-2009)
Чехия, Государственный электронный архив Чехии (Národní digitální archiv)	466 млн крон (25 млн долларов) (2010–2013)
Эстония, Электронный архив Национальных Архивов Эстонии (Rahvusarhiiv)	8 млн евро (2010-2011)
Австралия, штат Новый Южный Уэльс, электронный архив штата	3,4 млн австралийских долларов (проект на 3 года)
Новая Зеландия, «Государственный электронный архив» (New Zealand Government Digital Archive)	12,6 млн новозеландских долларов (это примерно 7,6 млн долл. США) (2011–2013)

Компоненты электронного архива

Разработку концепции электронного архива желательно начать со знакомства с международным стандартом OAIS, который давно уже стал «Библией» электронной архивации. Он описывает типовую архитектуру и функции электронного архива. Практически все известные проекты электронных архивов реализованы на основе именно этой модели.

В предыдущем номере журнала мы уже рассказывали в общих чертах об этом стандарте, который в 2003 году был утвержден в качестве международного стандарта ISO 14721:2003 «Системы передачи данных и информации о космическом пространстве. Открытая архивная информа-

ционная система. Эталонная модель» (Space data and information transfer systems – Open archival information system – Reference model (OAIS))⁵.

В терминологии OAIS широко используются **понятия 3 видов информационных пакетов:**

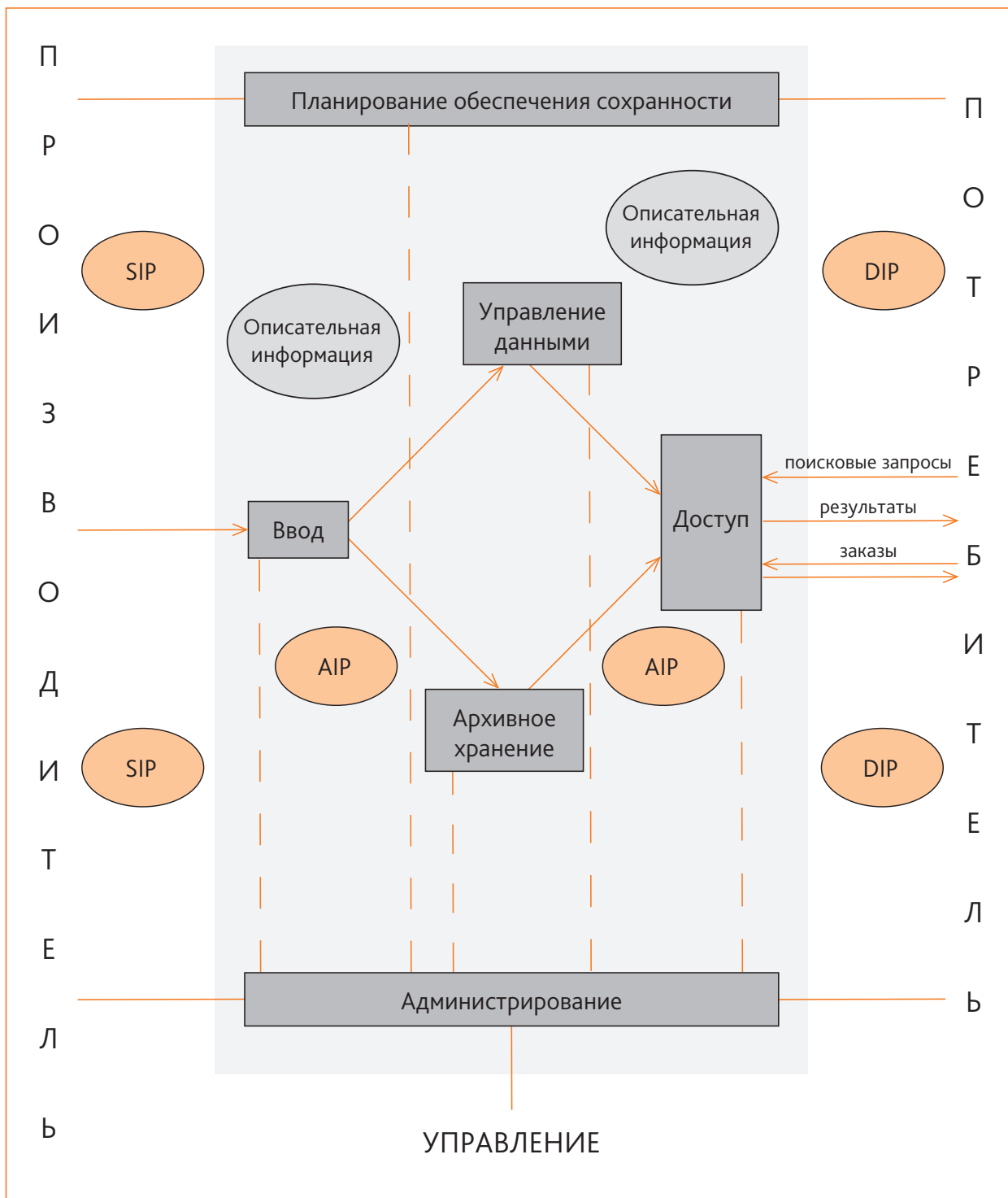
- *Сдаточный информационный пакет* (Submission Information Package, *SIP-пакет*) – формат, в котором документы и их метаданные поступают в архив от источника комплектования и используется затем для формирования одного или нескольких архивных информационных пакетов. Иногда SIP-пакеты представляют собой специально разработанные контейнерные форматы, в которые «упаковываются» документы и их метаданные, целостность и аутентичность которых обеспечивается с помощью электронных цифровых подписей или аналогичных средств. В других случаях SIP-пакетом может быть «дерево» папок заданной структуры, содержащее как файлы документов, так и файлы с метаданными;
- *Архивный информационный пакет* (Archival Information Package, *AIP-пакет*) – формат, используемый для долговременного хранения информации в архивной системе. В состав AIP-пакета, помимо информационного контента, включается информация, необходимая для организации долговременного хранения документов;
- *Дистрибутивный информационный пакет* (Dissemination Information Package, *DIP-пакет*) – формат, в котором данные предоставляются пользователю в ответ на его запрос. DIP-пакет может формироваться на основе одного или нескольких AIP-пакетов.

Внешними по отношению к архивной системе являются:

- *Производители (источники) информации* – источники комплектования архива и их информационные системы, которые создают информацию и передают ее на архивное хранение;
- *Пользователи* – лица или системы, взаимодействующие с сервисами электронного архива с целью поиска и получения сохраненной информации. Специальной группой пользователей является, согласно стандарту, так называемое «целевое сообщество» – множество тех пользователей, которые должны быть способны понять сохраненную информацию;
- *Высшее руководство организации-владельца архива*, определяющее его стратегию, цели и задачи.

⁵ Храмовская Н.А. Электронные архивы и их правовые проблемы // Делопроизводство и документооборот на предприятии. – 2012. – № 3. См. также <http://public.ccsds.org/publications/archive/650x0b1.PDF>.

Рисунок 1. Функциональная модель (блоки) открытой архивной информационной системы (OAIS)



SIP – пакет сдаваемой информации;
 AIP – пакет архивной информации;
 DIP – пакет распространяемой информации.

Во внутренней архитектуре электронного архива выделяются модули:

- *Модуль ввода (приема/передачи на архивное хранение).* Данный модуль обеспечивает сервисы и функции по приему сдаточных информационных SIP-пакетов от источников комплектования (или, иногда, от внутренних подразделений организации-владельца архива) и осуществляет подготовку информации к хранению и управлению в пределах архива. В число функций модуля входит:
 - получение SIP-пакетов, контроль их качества,
 - создание архивных информационных AIP-пакетов (там, где это нужно – во многих архивных системах AIP-пакеты идентичны SIP-пакетам),
 - извлечение из AIP-пакетов описательной информации для включения ее в научно-справочный аппарат архива и передача обработанных данных на хранение;
- *Модуль архивного хранения.* Данный модуль обеспечивает сервисы и функции для хранения, поиска и извлечения AIP-пакетов. Функции данного модуля включают:
 - прием AIP-пакетов от модуля ввода и помещение их в систему постоянного хранения,
 - управление иерархической системой хранения,
 - обновление носителей, на которые записаны архивные материалы, проведение регулярных и специальных проверок на наличие ошибок,
 - обеспечение возможности восстановления в случае техногенных и природных катастроф и
 - передача AIP-пакетов по запросам в модуль доступа;
- *Модуль управления данными.* Данный модуль обеспечивает сервисы и функции для управления и доступа как к описательной информации, идентифицирующей и документирующей фонды архива, так и к используемым для управления архивом административным данным. Функции модуля включают:
 - администрирование базы данных архива,
 - обновление базы данных (загрузка новой описательной информации или новых административных данных),
 - выполнение запросов к данным и выдача отчетов по полученным результатам;
- *Модуль администрирования.* Данный модуль обеспечивает сервисы и функции оперативного управления работой всего архива. Функции модуля включают:
 - проведение переговоров и заключение договоров с источниками комплектования архива,
 - аудиты поставляемых партий информации на предмет соответствия стандартам архива,
 - управление конфигураций оборудования и программного обеспечения,

- этот модуль также обеспечивает системные функции по мониторингу и улучшению работы архива, по инвентаризации, составлению отчетности, миграции и обновлению контента архива,
- модуль также отвечает за разработку и поддержание стандартов и политик архива, поддержку пользователей архива;
- *Модуль планирования деятельности по обеспечению сохранности.* Данный модуль обеспечивает сервисы и функции для мониторинга архивной системы и выдачи рекомендаций с целью обеспечения постоянной доступности сохраненной в архиве информации – даже в тех случаях, когда первоначальная архивная среда устареет. В число функций модуля входят:
 - инспектирование контента архива и периодическая выдача рекомендаций по организации миграции фондов архива, разработка рекомендаций для стандартов и политик архива, а также
 - мониторинг изменений в технологиях и в потребностях пользователей.В рамках этого модуля проектируются шаблоны информационных пакетов и оказание помощи по разработке на их основе SIP- и AIP-пакетов. Здесь же разрабатываются детальные планы миграции, прототипы программного обеспечения и планы тестирования;
- *Модуль доступа.* Данный модуль обеспечивает сервисы и функции, помогающие пользователям архива установить факт существования архивной информации, найти ее описание, определить ее местоположение и доступность, и дающие пользователям возможность запрашивать и получать нужную им информацию. В число функций модуля входят:
 - взаимодействие с пользователями по приему запросов,
 - контроль над доступом к информации,
 - координирование процесса исполнения запросов,
 - подготовка ответов и доставка информации пользователям.

В следующем материале цикла об электронных архивах мы расскажем о выборе форматов электронных документов, подходящих для длительного хранения информации разных видов. Вы узнаете о мировом опыте, современных тенденциях и существующих стандартах.