

Конференция iPRES-2014: Интеграция систем электронного правительства с электронными архивами

 rusrim.blogspot.com/2015/04/ipres-2014.html

Статья Кульдара Ааса (Kuldar Aas, Национальные Архивы Эстонии), Джанет Делв (Janet Delve, университет Портсмута, Великобритания), Риккардо Виейра (Ricardo Vieira, университет Лиссабона, Португалия) и Росса Кинга (Ross King, Австрийский технологический институт) была опубликована в трудах конференции iPres-2014, доступных по адресу <http://ipres2014.org/sites/default/files/upload/iPres-Proceedings-final.pdf> (стр.267). Об этой конференции, прошедшей в октябре 2014 года в Мельбурне, Австралия, также см. <http://rusrim.blogspot.ru/2014/04/11-ipres-2014-2014.html> . Вниманию читателей предлагается перевод отдельных фрагментов этой статьи.

Джанет Делв (Janet Delve) и Риккардо Виейра (Ricardo Vieira)

Кульдар Аас (Kuldar Aas) и Росс Кинг (Ross King)

Если интероперабельность (возможность взаимодействия) между активными системами электронного правительства была в течение последнего десятилетия важной областью исследований, то необходимость сохранять в течение длительного времени большую часть этой информации после её первоначального создания игнорировалась, а вопросу повторного использования данных придавалось второстепенное значение. В данной статье рассматриваются потребности в обеспечении долговременной сохранности изначально-электронных данных «электронного правительства», и рассказывается о том, какие дальнейшие действия предлагаются в рамках финансируемого Евросоюзом проекта E-ARK для экономически-эффективного решения этой проблемы.

До настоящего времени основные усилия были сконцентрированы на активной фазе жизненного цикла информации (т.е. на этапах создания данных и управления ими в краткосрочной перспективе), и мало внимания уделялось более поздним стадиям (обеспечение долговременной сохранности и доступа). Для того, чтобы восполнить этот пробел, Европейская комиссия финансирует трехлетний проект под названием «Электронный ковчег» или E-ARK (о проекте E-ARK см. также подборку материалов: <http://rusrim.blogspot.ru/search/label/e-Ark> - Н.Х.).



В рамках проекта E-ARK в начале 2014 года было проведено комплексное исследование с целью узнать больше о технических деталях доступных решений. Результаты исследования показали, что даже тогда, когда такие решения имеются, они скорее представляю собой прагматичные подходы к решению общей проблемы, в том смысле, что они нацелены лишь на удовлетворение текущих насущных потребностей в части долговременной сохранности и не принимают во внимание вопросы повторного использования и обеспечения доступа,



По существу, типичные национальные требования к обеспечению долговременной сохранности, как правило, включают ряд требований к метаданным, которые должны быть выполнены государственными органами, передающими данные на архивное хранение, а также некоторые правила форматирования как метаданных, так и самих данных (т.е. устанавливаются архивные файловые форматы). Обычным способом доступа к сохраненным данным является использование архивных каталогов, при этом на пользователей в первую очередь возлагается бремя выявления соответствующих наборов данных, прежде чем они смогут начать поиск тех битов информации, которая им на самом деле нужна.



Потребность в стандартизации

Основной выявленной нами причиной применения таких фрагментированных подходов является отсутствие стандартизации.

Цель проекта E-ARK в том, чтобы развивать дальше высокоуровневые спецификации MoReq2010 и обновлять их, добавляя в них более детальные требования, сформулированные на основе уже имеющейся национальной передовой практики. Для поддержки большей интероперабельности требуются, в частности, следующие элементы:

- Схема метаданных, предписывающая обязательное использование её основных элементов (ядра) для автоматизации рабочих процессов экспорта, проверки и передачи;
- Возможности для расширения, позволяющие добавлять в ядро метаданные, специфические для определенной страны или отрасли (примером могут служить метаданные, специфические для обмена электронными медицинскими документами или электронными счетами-фактурами);
- Спецификации на повторное использование метаданных, определяющие, как метаданные, созданные в системах «электронного правительства», могут быть повторно использованы для архивных целей;
- Модель рабочих процессов на стадиях подготовки к передаче и передачи на архивное хранение, описывающая в основных чертах действия, играющие ключевую роль для проверки (валидации) метаданных, целостности данных и аутентичности;
- Механизмы передачи, поддерживающие эффективную и защищённую массовую

передачу документов и их метаданных из государственных органов в архивы.

Очевидно, что процесс экспорта документов из систем-источников с целью их передачи в долговременные хранилища концептуально не сильно отличается от обмена информацией между любыми другими системами, поэтому данная работа выполняется в тесном сотрудничестве с финансируемым Евросоюзом проектом электронной e-SENS (Electronic Simple European Networked Services – «Простые электронные европейские сетевые сервисы»), целью которого является разработка принципов и спецификаций для трансграничных услуг, использующих данные различных европейских государств.

Открытые форматы для долговременной сохранности

Еще одним важным вопросом, который необходимо исследовать, является стандартизация архивных информационных пакетов (Archival Information Package, AIP) и дальнейшее развитие их структур и принципов.

Был проведен подробный анализ преобладающих в данный момент принципов создания AIP-пакетов. Он показал, что уже имеется большое количество стандартов и спецификаций, которые можно использовать в качестве отправной точки (например, PREMIS, EAD, описание модуля метаданных для управления документами в MoReq2010, METS и BagIt).

В ходе проекта E-ARK будет продолжен анализ уже имеющиеся стандартов и будет сформирован ограниченный набор (ядро) ключевых элементов, обязательных для всех AIP-пакетов.

Доступ к заархивированным данным

В настоящее время большинство архивов предоставляет доступ к своим электронным фондам посредством специальных архивных каталогов. Кроме того, архивы склонны организовать свой контент в соответствии с архивными иерархическими классификационными схемами и правилами описания. Это означает, что «богатыми» описательными метаданными, как правило, снабжены лишь агрегации (коллекции) данных, - но не их отдельные элементы (т.е. документы), которые как раз в основном и представляют интерес для общественности.

В рамках проекта E-ARK идёт работа над серией решений, позволяющих преодолеть перечисленные проблемы. Первый из подходов предполагает использование семантического описания и технологий «хранилищ данных» (warehousing). В результате точкой доступа к сохраняемым данным будут просто семантически обогащенные OLAP-кубы (Online Analytical Processing, OLAP – *согласно Википедии, это «технология обработки данных, заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу», см. <https://ru.wikipedia.org/wiki/OLAP> - Н.Х.*), вместо снимков реляционных баз данных с их очень сложными структурами данных.

Ещё одним способом доступа является доступ к архивным документам через «внешние» системы (такие, как порталы государственных услуг или веб-сайты государственных органов).

Наконец, практическое внедрение «Прикладного профиля CMIS для архивов» (CMIS Application Profile for Archives – *о стандарте CMIS см. подборку <http://rusrim.blogspot.ru/search/label/CMIS> – Н.Х.*) позволит учреждениям передавать свои данные в электронные хранилища, одновременно сохраняя возможность и дальше предлагать удобные для пользователей услуги по доступу к данным.

Гармонизация знаний

Одной из целей проекта E-ARK является проектирование и реализация онлайн-центра открытого доступа к знаниям, дающего возможность загрузки, управления и консолидации существующих лучших практик, стандартов и других материалов, касающихся не только основной области управления информацией, но и взаимосвязанных областей.

В качестве ещё одной инновации, проект E-ARK будет способствовать повторному использованию архивных данных, содействуя выработке единого общеевропейского подхода к выполнению простых и сложных запросов от исследователей, организаций государственного и частного секторов, а также от граждан.

Кульдар Аас (Kuldar Aas), Джанет Делв (Janet Delve), Риккардо Виейра (Ricardo Vieira), Росс Кинг (Ross King)

Источник: труды 11-й международной конференции по электронной сохранности iPRES-2014
<http://ipres2014.org/sites/default/files/upload/iPres-Proceedings-final.pdf>