

Перспективы долговременного хранения электронных документов

Кюнг Павел Алексеевич
к.и.н. заместитель директора
Российский государственный архив научно-технической
документации

Федеральный закон от 27 июля 2006 г. N 149-ФЗ "Об информации, информационных технологиях и о защите информации"

- электронный документ - 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

Федеральный закон от 22 октября 2004 г. N 125-ФЗ "Об архивном деле в Российской Федерации"

- В состав Архивного фонда Российской Федерации входят находящиеся на территории Российской Федерации архивные документы независимо от источника их происхождения, времени и **способа создания, вида носителя**, форм собственности и места хранения, в том числе ... **электронные** ... документы...
- **архивный документ** - материальный носитель с зафиксированной на нем информацией, который имеет реквизиты, позволяющие его идентифицировать, и подлежит хранению в силу значимости указанных носителя и информации для граждан, общества и государства

Росархив предлагает ввести в статью 3 Федерального закона «Об архивном деле в Российской Федерации» понятие

- **электронный архивный документ** – документированная информация, представленная в электронной форме на материальном носителе (материальных носителях) с возможностью его (их) изменения или замены, которая имеет метаданные (реквизиты), позволяющие ее идентифицировать, и подлежит хранению в силу значимости указанной информации для граждан, общества и государства».

электронный архивный документ

- Архивисты переходят в своей работе от приоритетного сохранения носителя информации, которая с ним непосредственно связана, к сохранению только информации.
- Исчезает понятие оригинальности документа т.к. для двух копий одного файла определить «оригинал» не представляется ВОЗМОЖНЫМ.

электронный архивный документ

Мы живем в мире электронных документов, которые не имеют электронной подписи, но тоже должны быть сохранены:

Аудиовизуальные документы

Электронные сообщения

Сайты

Базы данных

Документы личного происхождения

ГОСТ 2.102-2013 Единая система конструкторской документации. Виды и комплектность конструкторских документов.

- ГОСТ устанавливает возможность создания конструкторских документов в бумажном и электронном виде и их **равнозначность** и взаимозаменяемость и выполнение графических документов в виде двухмерных документов и в виде электронных моделей (3D).
- В случаях, когда одновременно применяют бумажные и электронные КД одного вида и наименования, стандарт допускает их взаимное преобразование друг в друга.
- взаимное соответствие между этими документами **обеспечивает разработчик.**

ГОСТ 2.051-2013 Единая система конструкторской документации. Электронные документы. Общие положения

Визуализация: отображение информации в пригодной и понятной для восприятия человеком форме.

Визуализация выполняется соответствующими программными и/или техническими средствами. Результат визуализации должен быть оформлен согласно требованиям стандартов ЕСКД. Результатом визуализации может быть бумажная копия электронного конструкторского документа.

Тезисы

Электронные документы в настоящее время стали основным средством документирования действительности и основным средством передачи информации.

Огромный пласт информации создается и передается осознанного контроля человека (метаданные, информация передающаяся с различных устройств), образуя «большие данные».

Гипотеза

- Существует реальная угроза сохранности значительных массивов исторических памятников, создаваемых и хранимых в электронном виде на долговременной перспективе.
- Угроза заключается в лишении возможности прочтения документов, в невозможности восстановить изделия целиком в связи с обновлением ПО или утратой библиотек данных.
- Также существенную опасность составляет лишение возможности подтверждения аутентичности документа.

Обозначение проблемы

- Одновременное присутствие на международном рынке ПО конкурирующих между собой решений, использующих разные подходы к работе с информацией.
- Крупные игроки не заинтересованы в кроссплатформенности своих форматов
- Многие системы предусматривают при работе с документом использование библиотек, например библиотек материалов, текстур и т.д. Сохранение моделей вне этих библиотек приводит к невозможности их корректного отображения.
- На настоящий момент вопросы долговременного хранения электронных документов производителями ПО не ставятся.

Опыт решения

- *Стандарт PDF/A (ISO 19005-1:2005) для архивного хранения.*
Закреплен в Правилах организации хранения, комплектования, учёта и использования документов Архивного фонда Российской Федерации и других архивных документов в органах государственной власти, органах местного самоуправления и организациях, Приказ Министерства культуры РФ от 31 марта 2015 г. № 526

ГОСТ Р 54989-2012. Обеспечение долговременной сохранности электронных документов

электронное архивирование, электронная архивация (electronic archiving): Сохранение электронной информации в обособленном физическом или логическом пространстве, где эта информация защищена от утраты, от внесения изменений и деградации.

Цели стратегии обеспечения долговременной сохранности

- Читаемость электронных документов
- Интерпретируемость электронных документов
- Идентифицируемость электронных документов
- Доступность электронных документов
- Понятность электронных документов
- Аутентичность электронных документов

ГОСТ Р ИСО 13008-2015 Информация и документация. Процессы конверсии и миграции электронных документов

валидация (validation): Подтверждение посредством представления объективных доказательств, выполнения требований, установленных в отношении конкретного вида использования или применения

Целью процесса валидации является проверка успешности выполнения конверсии/миграции электронных документов с тем, чтобы третьи стороны в достаточной степени могли быть уверены в сохранении документами аутентичности, надежности, целостности и пригодности к использованию.

Международный опыт решения

- *Long Term Archiving and Retrieval (LOTAR)* Международный проект с участием консорциума аэрокосмических и оборонных компаний из США и Европы, запущенный с целью разработки стандарта для архивирования 3D-моделей САПР.
- *JT*. Единый формат описания 3D данных ISO 14306:2012. Был первоначально разработан компаниями Engineering Animation и Hewlett Packard, в дальнейшем развивался Siemens PLM Software.
- VDA Recommendation 4958. Рекомендации по архивации технических электронных документов, разработанные Ассоциацией немецкой автомобильной промышленности, расширяющие стандарт ISO 14721:2012.

Международный опыт решения

- Проводить отбор «моделей», ориентированных на долговременного хранения, в зависимости от их потенциальной востребованности в будущем.
- Определять набор элементов (библиотек) необходимых для корректного отображения каждого изделия и сохранять также и их. Проверять их совместимость при каждой миграции в новый формат и при использовании различного программного обеспечения, в том числе и при использовании нейтральных форматов данных.
- Сохранять документы в исходном формате так долго, как это возможно при используемом программном обеспечении. При преобразовании файлов использовать сразу несколько нейтральных форматов, соответствующих STEP (ISO 10303).
- Для «моделей» состоящих из множества файлов, архивировать каждый файл в отдельности, сохраняя также все связи между файлами и создавая инструкцию по восстановлению изделия в целостности.
- Работать в тесном контакте с создателями документов, чтобы гарантировать, что вся информация, необходимая для их понимания архивируется вместе с ней.

Важное

- SHA-1 (Standard Hash Algorithm 1) - стандартный алгоритм вычисления хэш-функции N 1;
- WORM (Write Once Read Many [times]) - носитель информации однократной записи и многократного считывания;
- XML (Extensible Markup Language) - расширяемый язык разметки.

Что не решено

- На нормативном/методическом уровне не зафиксированы форматы для аудиовизуальных документов, для научно-технических документов, баз данных.
- Отсутствует нормативно закреплённая практика гарантирования подлинности документов без электронной подписи

Вопросы

- Должен ли архив верифицировать документы и проверять их подлинность?
- Нужна ли электронная подпись архивам – проверяем ли мы подписи в документах?
- Должен ли он гарантировать их читаемость?
- При невероятно быстром изменении технологий не рискуем ли мы потерять документы, дожидаясь пока изобретут вечное их хранение?
- Имеем ли мы право перекладывать решение проблем на будущие поколения и принимать в архивы то, что прочитать и сохранить сегодня не в состоянии?
- Может быть нужно разрабатывать способы консервации электронных документов, без текущих затрат на хранение?

Осознание

- Электронные документы имеют столь изменчивую природу, что мы не можем ждать от них принятия определенной формы.
- Мы непрерывно будем сталкиваться с новыми формами существования информации в электронном виде, которые нужно будет сохранять для потомков.
- И это потребует от теоретиков и практиков непрерывной работы по совершенствованию.

К первоочередным проблемам нормативно-методического характера следует отнести

- Разработку требований к системам электронного документооборота, обеспечивающих выделение электронных документов, подлежащих передаче на хранение в архив, из массива электронных документов, находящихся на оперативном хранении, к порядку их передачи на архивное хранение;
- Разработку рекомендаций по организации конверсии и миграции электронных документов, находящихся на хранении в архивах органов власти, государственных и муниципальных архивах, в случае устаревания программно-аппаратной среды, технологической смены форматов архивного хранения, выхода из строя носителей архивных документов.

Спасибо за внимание!