

Электронные издания: проблемы долговременного использования и пути их решения

Digital Publications: Problems of Long-Term Use and Ways of their Solution

Електронні видання: проблеми довготривалого використання та шляхи їх вирішення

А. А. Смородина

НТЦ «Информрегистр», Москва, Россия

Alina Smorodina

«Informregistr» Scientific and Technical Center, Moscow, Russia

А. А. Смородина

НТЦ «Информрегистр», Москва, Росія

В докладе рассматриваются вопросы сохранности и долговременного использования электронных изданий, приводятся пути решения проблем в данной области. Более подробно рассмотрен подход, использующий технологию виртуализации.

Preservation and long-term use of digital publications, as well as ways of resolving the associated problems are considered in this paper. In more detail is presented an approach using a virtualization technology.

У доповіді розглядаються питання збереження та довготривалого використання електронних видань, наведено шляхи вирішення проблем у цій галузі. Більш детально розглянуто підхід, що використовує технологію віртуалізації.

Объемы информации, представленной в цифровом виде, стремительно растут. Развитие компьютерной техники и информационных технологий способствует развитию информационного общества. На всех этапах развития общества одной из главных задач ставится сохранение электронной информации для нынешнего и будущего поколений.

Исследованием проблем сохранения электронных изданий занимаются на международном и национальном уровне, в 2003 году на конференции ЮНЕСКО была принята «Хартия о сохранении цифрового наследия»¹, в которой декларируется необходимость принятия срочных мер по сохранению цифрового наследия.

На сегодняшний день в России существуют автоматизированные библиотечные информационные системы (АБИС), предназначенные для поддержки библиотечных процессов, но они не обеспечивают хранение и предоставление доступа к цифровой информации.

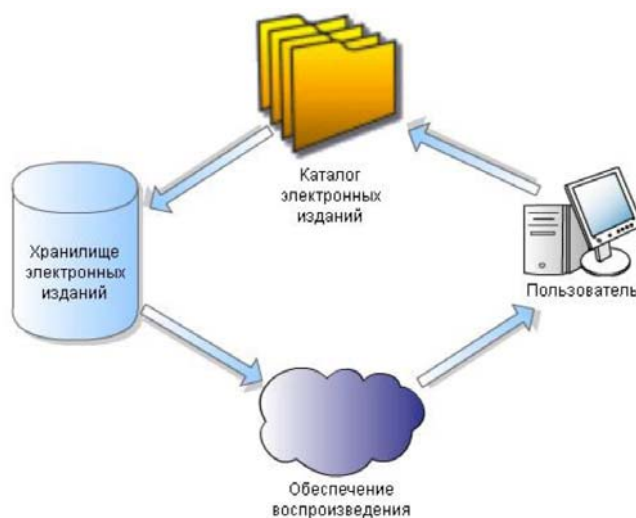
Одним из технологических решений обеспечения сохранности и воспроизводимости наряду с компьютерным залом в библиотеке или интернет-библиотекой с доступом к рабочей станции в читальном зале можно рассматривать технологии виртуальных машин и создание хранилища образов электронных изданий.

Для решения проблемы сохранения электронных изданий для будущих поколений необходимо создание и обеспечение работы хранилища электронных изданий с интерфейсом доступа к информационным ресурсам. В результате должны быть решены такие задачи, как:

1. сохранность информации,
2. неизменность хранимой информации,
3. актуальность хранимой информации,
4. воспроизводимость хранимой информации,

¹ Хартия о сохранении цифрового наследия // Программа ЮНЕСКО «Информация для всех» в России [Электронный ресурс]. – Электрон. дан. – сор. 2002–2009. – Режим доступа: <http://www.ifap.ru/ofdocs/unesco/digit.htm>

5. защиту хранимой информации от незаконного копирования,
6. доступ к хранимой информации через унифицированный интерфейс (каталог)²

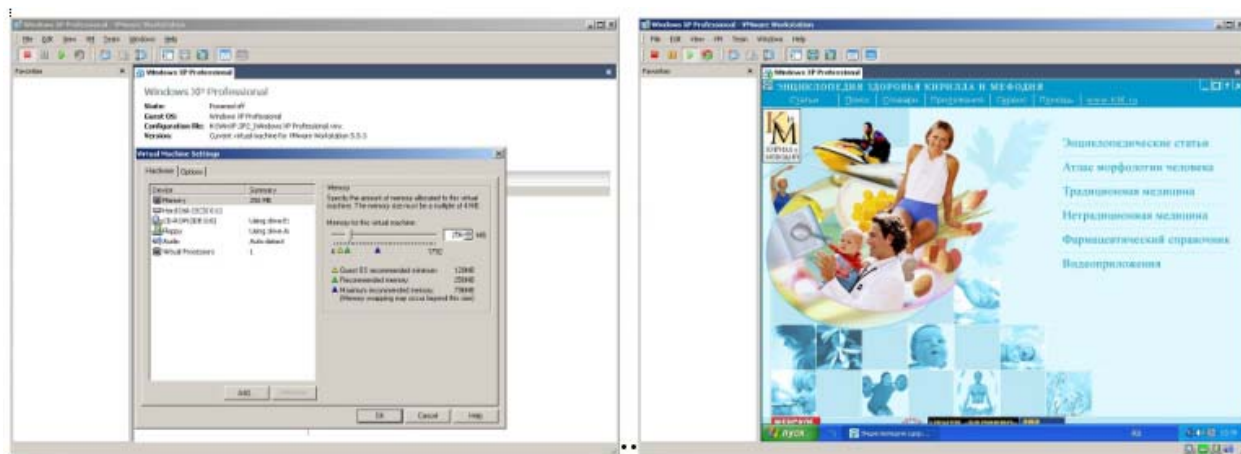


Также важным является вопрос о критериях отбора электронных изданий для сохранения и ценностях (контентные, алгоритмические, дизайнерские и др.), которые они представляют.

Проблемы сохранности и воспроизводимости электронных изданий актуальны для фонда Депозитария электронных изданий НТЦ «Информрегистр».

С обязательных экземпляров электронных изданий, регистрируемых в НТЦ «Информрегистр», делаются электронные образы, которые сохраняются в базе данных вместе с описанием их содержимого (в т. ч. служебной информацией). Актуальность архива электронных изданий поддерживается благодаря постоянному дополнению образами носителей изданий.

Таким образом, оригинал носителя сохраняется в архивных целях, а пользователи обращаются к образам носителя электронного издания, что позволяет обеспечить сохранность и воспроизводимость оригинала.



² Исследование вопросов сохранности электронных документов для ведения хранилища национального библиотечно-информационного фонда на базе ФИЦ: отчет по научно-исследовательской работе: № 12-1-2007 / № ГР 01. 2. 00614717 Федеральное агентство по информационным технологиям, Федеральное государственное унитарное предприятие Научно-технический центр «Информрегистр». – М., 2007. – 212 с.

Для каждого электронного издания заполнен специальный набор метаданных, с помощью которых описывают системные требования, необходимые для воспроизведения издания. В соответствии с этими характеристиками выбирается подходящая система, в которой данное электронное издание может быть воспроизведено в том виде, в котором оно было создано. После определения типа загрузки нужной системы происходит реализация загрузки. Таким образом, выполняется виртуальный запуск практически любой из требующихся программно-аппаратных платформ, используя при этом не огромный парк машин, а одну вычислительную систему (с возможностью в дальнейшем разделить ее на две и более).

Перечень системных характеристик, необходимый для определения платформы воспроизведения электронного издания, выглядит следующим образом:

Поле «Тип компьютера» – указание аппаратной платформы, для которой предназначено ЭИ. Например, кроме PC, это может быть PDA, Wearable PC и другие.

Поле «Марка процессора» – поколение процессора (часть аппаратного обеспечения компьютера или программируемого логического контроллера, отвечающая за выполнение основной доли работ по обработке информации — вычислительный процесс). Например, наиболее популярные процессоры сегодня производят фирмы Intel, AMD и IBM. Intel: 8086, i286, i386, i486, Pentium, Pentium II, Pentium III, Celeron (упрощённый вариант Pentium), Pentium 4, Core 2 Duo, Xeon (серия процессоров для серверов), Itanium и др. AMD имеет в своей линейке процессоры Amx86, Duron, Sempron, Athlon, Athlon 64, Athlon 64 X2, Opteron и др.

Поле «Частота процессора» – указание тактовой частоты процессора в герцах (Один герц соответствует одному периоду колебаний в секунду. Обозначается Гц или Hz.) Например, 900 МГц, 2 ГГц и т. д.

Поле «Разрядность операционной системы» – указание разрядность операционной системы. Например, 16,32,64.

Поле «Семейство операционной системы» – наименование семейства операционной системы (базовый комплекс компьютерных программ, обеспечивающий управление аппаратными средствами компьютера, работу с файлами, ввод и вывод данных, а также выполнение прикладных программ и утилит.) Например, семейства Windows, UNIX, MacOS, BeOS и другие.

Поле «Версия операционной системы» – указание конкретной версии выбранного семейства. Например, Windows XP, GNU, Linux и другое.

Поле «Объем оперативной памяти» – объем оперативной памяти (память, предназначенная для временного хранения данных и команд, необходимых процессору для выполнения им операций) Например, 64 Мб, 128 Мб и т. д.

Поле «Объем памяти на диске» – объем жесткого диска (энергонезависимое, перезаписываемое компьютерное запоминающее устройство) Например, 500 Мб, 1 Гб и т. д.

Поле «Модель видеокарты» – указание конкретной модели видеокарты (плата расширения, предназначенная для работы IBM-совместимыми ПК с видеосигналами, поступающими, например, из кабельной телевизионной сети, ТВ антенны или видеомагнитофона, а также для вывода компьютерной графики на телевизионные приемники). Например, ASUS GeForce 8600 GT, Leadtek GeForce 8800 GTS и другие.

Поле «Аппаратная поддержка видеотехнологий» – поддерживаемые картой стандарты и программы. Например, поддержка API, памяти DDR3 и др.

Поле «Объем видеопамати» – часть оперативной памяти, отведённая для хранения данных, которые используются для формирования изображения на экране монитора. Или отдельно выделенная микросхема памяти на видеокarte. Например, 32 Мб, 1 Гб и др.

Поле «Разрешение экрана» – размеры изображения в пикселах. Например, 800 Ч 600,1024 Ч 768,1280 Ч 1024.

Поле «Необходимость аудиокарты» – необходимость наличия платы для работы со звуковой информацией.

Поле «Наличие специальных требований» – необходимость в специальных требованиях к звуковой карте.

Поле «Аппаратная поддержка аудиотехнологий» – стандарты, надстройки и технологии, поддерживаемые аудиокартой. Например, EAX, EAX 5. 0, OpenAL и другие.

Поле «Требования к дисководу» – требования к устройству, управляющему считыванием и записыванием информации. Например, CD-ROM, CD-RW, DVD/CD-RW, BD-RE и другие.

Поле «Необходимое дополнительное программное обеспечение» – дополнительные программы, управляющие работой компьютера или автоматизированной системы. Например, MS Office, различные версии Интернет-браузеров, Adobe Acrobat и другие.

Поле «Необходимость доступа к сети Интернет» – возможность доступа к всемирной системе добровольно объединённых компьютерных сетей, построенной на использовании протокола IP и маршрутизации пакетов данных.

Поле «Носитель для запуска» – несколько носителей одного электронного издания, возможность выбора воспроизведения нужного носителя.

Поле «Родительское электронное издание» – наличие «родителя» у электронного издания и необходимость запуска «родителя» для воспроизведения «дочернего» электронного издания.³

Эффективным способом решения проблемы сохранения электронных издания является использование технологии с применением виртуальных машин, что позволяет реализовывать запуск практически любой необходимой для воспроизведения электронного издания системы. Важной составляющей частью технологии виртуализации следует рассматривать извлечение определенного набора метаданных для библиографического описания с последующим их использованием для определения типа загрузки нужной системы.

³ Исследование вопросов сохранности электронных документов для ведения хранилища национального библиотечно-информационного фонда на базе ФИЦ: отчет по научно-исследовательской работе: № 12-1-2007 / № ГР 01. 2. 00614717 Федеральное агентство по информационным технологиям, Федеральное государственное унитарное предприятие Научно-технический центр «Информрегистр». – М., 2007. – 212 с.