

Новейшие технологии долговременного хранения электронных информационных ресурсов

Вступление

Объем информации, представленной в цифровом виде, стремительно растет. Задача ее сохранения на всех этапах развития общества была одной из приоритетных, при решении которой необходимо обеспечить для будущих поколений как сохранность знаний, накопленных предыдущими поколениями, так и новой информации [1]. Объемы информации, представленной в цифровой форме, увеличиваются не только за счет новой информации, которая в большинстве случаев уже имеет и электронную форму представления, но и за счет перевода в цифровую форму ранее созданных информационных ресурсов, в первую очередь тех, которые составляют научное, историческое и культурное достояние человечества. Представление информации в цифровом виде позволило решить ряд проблем ее сохранения, создать новые возможности для доступа к информации и к ее обработке. Быстрое внедрение цифровых технологий обработки информации, развитие электронных библиотек обусловили необходимость проведения специальных исследований по созданию технологий долговременного хранения информации, представленной в цифровом виде. Необходимость этих исследований связана с тем, что:

- в цифровую форму переведены значительные объемы информационных ресурсов, которые с течением времени не теряют своей научной, исторической или культурной значимости и ценности;
- оцифровывание изданий минувших лет создало условия для того, чтобы они стали известны широкому кругу исследователей [2, 3];
- постоянно возрастает спрос на "онлайновое" получение информации;
- при создании электронных библиотек израсходованы огромные средства на оцифровывание документов, создание баз данных, организацию доступа к электронным ресурсам.

Перевод документов в цифровую форму позволил решить ряд проблем длительного хранения, а именно: обеспечить возможность контролирования документов без потери качества, устранения дефектов на носителях с аналоговой формой представления (царапины на видеоматериалах, шумы на аудиозаписях и др.), мультимедийного представления информации.

Хранение электронных документов имеет особенности, которые необходимо учитывать при создании электронных библиотек и архивов, а именно:

- электронные документы, их форматы и структура обуславливаются программно-аппаратными средствами воспроизведения информации;
- электронные документы сохраняются на носителях, которые имеют ограниченный срок хранения информации;
- независимо от архивных работников может осуществляться фрагментация документов.

Создание огромных архивов информационных материалов в цифровой форме привело к появлению ряда проблем, которые требуют решения в ближайшее время: создание

надежных носителей для долговременного хранения цифровой информации, разработка и внедрение специальных форматов для записи информации, которая подлежит длительному хранению, разработка специальных устройств, считывающих информацию. Несмотря на важность и сложность задач относительно решения проблемы быстрого морального старения программного обеспечения, а также технического и морального старения систем считывания информации, важнейшей проблемой является создание носителей для ее долговременного хранения.

При использовании цифровой информации следует разрабатывать и использовать специальные меры по исключению возможностей ее изменения, дополнять данные в электронном виде контекстной информацией, сохранять не только сами носители, но и устройства записи информации на них, а также программное обеспечение, которое реализует представление информации в форме, приемлемой для восприятия [4, 5]. Носители с цифровой формой представления информации более чувствительны к появлению локальных дефектов в записях.

Создание условий для долговременного и надежного хранения такой информации решается не только за счет использования специальных носителей информации и соответствующих технологий воспроизведения информации с них, а и за счет использования необходимых форматов представления информации и программного обеспечения. В частности, для долговременного хранения цифровой информации предлагается использовать простейшие цифровые форматы. Так, значительное внимание уделяется созданию универсального формата хранения данных (UPF). UPF рассматривается как файловый механизм данных, который использует контейнерную или свернутую структуру. Его базой являются метаданные, которые определяют содержание с помощью регистра стандартных типов данных и являются источником кода для разметки или перевода двоичных наборов в доступные или полезные формы. Предполагается, что UPF будет создан таким образом, что не будет зависеть от компьютерных дополнений и операционных систем, которые использовались при создании контекста, а также от свойств носителя, на который записан контекст. Разработка универсального формата хранения данных позволит избежать операций их повторного копирования и переформатирования и создать условия для полного использования физического ресурса носителей по надежному хранению информации.

Важной для носителей архивного хранения является также форма представления информации. Она должна позволять воссоздавать информацию разными физическими методами. Примером удачного выбора формы физического представления может служить использование микрорельефного представления информации на восковых цилиндрах, грампластинках, компакт-дисках.

Записи на восковых цилиндрах, шеллачных грампластинках, выполненные с рельефным представлением информации на довольно несовершенном оборудовании, при использовании современных технологий могут быть воспроизведены с лучшим качеством [6]. В данное время создаются методы рельефной записи с плотностью записи порядка десятков Гбайт/см².

Современные носители для архивного хранения информации

Для архивного хранения информации в цифровой форме сегодня широко используются магнитные ленты [7, 8]. Емкость электронных архивов на магнитных лентах может составлять сотни терабайт. Накоплен большой опыт хранения аудио- и видеоинформации на магнитных носителях. Существуют образцы магнитных лент с аналоговыми

аудиозаписями, на которых информация сохраняется более 50 лет. За последние десятилетия осуществлены значительные усовершенствования систем магнитной записи на ленточные носители, в частности:

- использование основ из полиэтилентерефталата позволило уменьшить толщину лент и увеличить скорость перемещения носителей;
- нанесение алмазоподобных защитных покрытий обеспечило механическую защиту информационного слоя;
- вакуумно-напыленные магнитные покрытия имеют минимальное количество дефектов, связанных с посторонними включениями, высокую стабильность магнитных свойств.

Основное внимание при разработке новых типов магнитных лент уделялось повышению плотности записи информации (емкость современных магнитных лент составляет сотни гигабайт) и скорости записи информации (скорость записи/воспроизведения составляет десятки миллионов байт/с). Однако гарантированный срок хранения информации на лентах практически не изменился — 30-40 лет [9, 10]. Существенные трудности при использовании магнитных лент для архивного хранения данных создает большое число форматов записей (сотни), которые в большинстве случаев являются несовместимыми или мало совместимыми, а также наличие разных устройств записи, которые выпускаются ограниченным числом производителей. Такая ситуация приводит к необходимости периодически проводить перезапись на новые носители. Если учитывать объемы информации, хранящейся на магнитных лентах (только аудиозаписей — десятки миллионов часов), то эта операция становится очень дорогостоящей и труднореализуемой.

В последнее время предлагается организация хранения больших информационных массивов на магнитных дисках. Это связано со значительным усовершенствованием их характеристик, созданием эффективных систем повышения надежности работы (RAID-системы). Исследования показывают, что срок хранения данных на жестких магнитных дисках не превышает 10 лет. В условиях постоянного усовершенствования технических характеристик магнитных дисков, замены винчестеров через 2-3 года, проблеме долговременного хранения информации на магнитных дисках не уделялось должного внимания. Необходимая информация переписывалась на новый носитель значительно большей емкости.

Повышение надежности хранения информации, исключение возможности ее несанкционированного изменения на RAID-массивах может обеспечить использование режима записи WORM, который раньше использовался только на магнитных носителях.

Несмотря на емкость и скорость записи/воспроизведения информации, резервирование записанной информации, магнитные дисковые системы не могут использоваться для долговременного хранения информации. В многоуровневых системах хранения данных жесткие магнитные диски в соответствии с иерархической структурой устройств используются для хранения часто запрашиваемых данных, доступ к которым осуществляется с максимально высокой скоростью.

Отсутствие надежных носителей для цифровой записи определило в последние годы интерес к хранению информации на микрофильмах в аналоговой форме. Предлагается технология архивного хранения, а именно — создается цифровой файл, который записывается в аналоговой форме на микрофильм как резервная копия для долговременного хранения [11]. При потере цифровой копии она может быть восстановлена при сканировании микрофильма. Срок хранения галогенидосеребряных

микрофильмов может составлять десятки лет. Однако эта технология не позволяет записывать мультимедийную информацию, обеспечивать высокие скорости воспроизведения данных.

Анализ возможностей использования оптических дисков для долговременного хранения информации

Среди разработанных до настоящего времени технологий наиболее полно удовлетворяют требованиям по созданию носителей для долговременного хранения оптические методы записи и хранения информации.

Оптические носители имеют ряд особенностей, которые позволяют рассматривать их как перспективные носители для долговременного хранения информации, а именно:

- бесконтактное считывание информации, что обеспечивает доступ к содержанию документа без нарушения оригинала и возможность долговременного хранения информации;
- использование физических методов защиты записанной информации от механических повреждений;
- реализация обратной совместимости на новых типах устройств воспроизведения информации;
- высокая плотность записи, возможность увеличения плотности и скорости записи информации;
- использование режима однократной записи и многократного считывания, при котором сделанная на таком диске запись не может быть стерта или заменена на новую (архивные документы не подлежат любому восстановлению или корректированию);
- использование надежного рельефного представления информации;
- возможность применения высокостабильных материалов для изготовления оптических дисков;
- использование универсальных защитных контейнеров для всех типов оптических дисков.

Наиболее длительный срок хранения имеют оптические диски на стеклянных подложках с однослойным покрытием типа WORM [12]. Срок хранения записанной информации на них может составлять, по мнению разработчиков, 100 лет, тогда как срок хранения информации на стандартных компакт-дисках составляет не более 20-30 лет. Такой срок хранения определяется особенностями технологии изготовления носителей, предназначенных для массового использования. Срок хранения информации на стандартных компакт-дисках ограничен по следующим причинам:

- нестабильность свойств и недостаточная механическая прочность подложек из поликарбоната;
- использование оптических интерференционных структур, выполненных из разнородных материалов (в некоторых типах компакт-дисков (CD-R) имеет место малая адгезия между слоями), что приводит к изменению оптических свойств при колебаниях значений температуры и влажности;
- сильная зависимость свойств носителей от условий изготовления компакт-дисков и режимов записи информации на них.

Опыт использования компакт-дисков CD-R (DVD-R), CD-RW (DVD-RW) в электронных архивах и библиотеках позволил разработать рекомендации, согласно которым

информация, хранящаяся на таких компакт-дисках, должна перезаписываться каждые 2-3 года [13].

Повысить надежность хранения информации можно за счет изготовления носителей с более стабильными во времени характеристиками, способными выдерживать колебание температур и влажности в более широком диапазоне, чем носители с поликарбонатными слоями. Кроме того, целесообразно выполнять носители информации из однородных материалов [14, 15].

Эти носители изготавливаются с использованием технологий, которые применяются в производстве компакт-дисков. В них используются форматы представления данных на стандартных компакт-дисках, они имеют геометрические размеры стандартных компакт-дисков и воспроизводятся на плеерах стандартных компакт-дисков. Они могут рассматриваться как расширение "семейства" компакт-дисков со специфической областью применения. В металлических носителях информации используется микрорельефное представление информации на никелевых слоях толщиной 0,1-0,3 мм, которые герметизированы прозрачными защитными слоями из полимерных материалов или из силикатного стекла. При необходимости может производиться замена защитных слоев (реставрация носителя). Срок хранения информации на металлических носителях может составлять сотни лет. Так, медно-никелевые штампы для тиражирования грампластинок, изготовленные 60-70 лет тому назад, прекрасно сохранились, аудиоинформация с них воспроизводится с высоким качеством звучания. Проведенный анализ химического состава поверхностного никелевого пласта показал, что проникновение кислорода в штамп незначительно и окислы локализованы в основном на примесях. Сплошная окисная пленка на поверхности никеля отсутствует. Современные методы гальванопластики позволяют получать никелевые носители с количеством примесей не более 0,005-0,01 %, что меньше чем на порядок количества примесей в металлах для тиражирования грампластинок. Это создает дополнительные возможности для обеспечения долговременного хранения металлических носителей. Окисные слои на поверхности никелевых носителей, которые приводят к изменению геометрических размеров питов, могут образовываться за 250-300 лет. Главная задача при использовании металлических носителей состоит в защите поверхности носителей от загрязнения. Широкие возможности может иметь технология изготовления компакт-дисков со стеклянными подложками. Они будут иметь срок хранения информации сотни лет.

Выводы

1. Использование оптических носителей позволяет решать основные проблемы по обеспечению долговременного хранения цифровой информации.
2. Срок хранения информации на оптических носителях может быть существенно увеличен (до 200-300) лет при использовании высокостабильных материалов для подложек носителей и однородных структур для представления информации.
3. Одним из основных факторов, которые определяют длительный срок хранения информации на металлическом и стеклянном оптическом дисках, является возможность считывать информацию с них различными методами, например методом сканирующей туннельной микроскопии.

Литература

1. *Бродолин Л. И., Воргачев В. Н.* Проблемы сохранности научной и технической информации // Тр. Междунар. конф. "Информационное общество и

- информационные ресурсы и технологии телекоммуникаций", Москва, 22-24 ноября 2000 г. — НТИ. — 2000. — С. 61-62.
2. *Смит Б.* Исследовательская деятельность Европейского Союза в области культурного наследия // Тр. конф. "Информация для всех: культура и технологии информационного общества", Москва, 1-5 декабря 2003 г. — С. П1-3-1(П1-3-18).
 3. *Афиногенов Л. П.* Длительное хранение информации в технических системах. — Л.: Гидрометеиздат, 1983. — 256 с.
 4. *Hedstrom M.* Digital preservation: a time bomb for digital libraries [Electronic resource]. — Way of access: <http://www.uky/kierman/DL/hedstrom.html>. — Title from the screen.
 5. *Woodyard D.* Digital preservation: the Australian experience // Proc. 3rd Digital Library Conference: Positioning the fortain of knowledge. State Library of Sarawak, Malaysia, 2 nd-4th of October, 2000, p. 1-10.
 6. Optomechanical method of Edison cyliders sound reproduction / V. V.Petrov, O. S. Onyshchenko, A. A. Kryuchin, et. // Proc. of the 102 nd AES Convention 1997, March 22-25. — Munich, Germany. — Preprint 4491 (M4).
 7. *Устинов В. А.* Обеспечение физической сохранности архивных документов // Техника кино и телевидения. — 2000. — № 6. — С. 44-48.
 8. *Фрадкин В.* Прошлое, настоящее и будущее носителей информации [Электронный ресурс]. — Режим доступа: <http://www.dw-world.de>. — Заглавие с экрана.
 9. Memory of the World. Safeguarding the Documentary Heritage. A guide to Standards, recommended practices and reference literature related to the preservation of documents of all kinds [Electronic resource]. — Way of access: <http://www.unesco.org/webworld/mdm/administ/en/guide/guidetoc.htm>. — Title from the screen.
 10. Digital preservation management: implementing short-term strategies for long-term problems [Electronic resource]. — Way of access: <http://www.library.cornell.edu/iris/tutorial/dpm/>. — Title from the screen.
 11. *Михайлов О. А.* Электронные документы в архивах. — Г.: Диалог — МГУ, 2000. — 325 с.
 12. Phase change recording in 12-inch. True WORM technology [Electronic resource]. — Way of access: <http://www.plasmon.com>. — Title from the screen.
 13. Good ractices handbook [Electronic resource]. — Way of access: http://www.minervaeurope.org/publications/goodpractices1_3.htm. — Title from the screen.
 14. *Петров В. В.* Способы решения проблемы долговременного хранения информации, записанной в цифровом виде / В. В. Петров, А. А. Крючин, С. М. Шанойло, И. О. Косско, В. Г. Портной // Доклады Национальной академии наук Украины. — 2003. — № 4 — С. 52-58 (укр.яз.).
 15. HD-Rosetta archival preservation services [Electronic resource]. — Way of access: <http://www.norsam.com>. — Title from the screen.