

РАЗДЕЛ 1

ТЕХНИЧЕСКИЕ и ПРАКТИЧЕСКИЕ ПЕРСПЕКТИВЫ

Глава 12. Принятие решения о том, что подлежит сохранению

ВВЕДЕНИЕ

12.1 Цели

Материалы этой главы дают читателю возможность понять ключевые вопросы, на которые должен быть найден ответ при принятии решения о том, какие цифровые материалы должны быть отобраны для сохранения, а также они содержат некоторые руководящие принципы. В главе также даются некоторые советы технического и практического характера, включая предложения по идентификации основных элементов, которые должны быть сохранены.

12.2 Краткое содержание главы

Обычно необходимо принимать решение о том, какие цифровые материалы стоит сохранять, так же как и в случае с нецифровыми материалами. Многие из тех же самых подходов — отбор, основанный на критериях, применяемых при создании коллекции, и хорошее знание материалов и их контекста — являются фундаментальными при отборе цифрового наследия. Программы сохранения также должны определить элементы или характеристики выбираемых материалов, которые придают им значимость, и руководствоваться.

12.3 Терминология

Термин *отбор* используется в этой главе как универсальное понятие. Его необходимо уяснить, чтобы обращаться к такому понятию как *экспертиза*, которое имеет частное значение в сообществе архивных записей.

ОБЩИЕ ВОПРОСЫ

12.4 Что должно быть сохранено?

Решение, что должно быть сохранено, кем, на какой период времени, является фундаментальным в процессе управления всеми видами материального наследия. Такие решения необходимы, потому что обычно различных вещей имеется большее количество — информации, записей, изданий, данных — чем мы в состоянии сохранить. Выбор каждого объекта сохранения происходит за счет какого-то другого.

12.5 Опора на существующие концепции

Процесс отбора цифрового наследия концептуально является таким же как и отбор нецифровых материалов. Любая существующая программа с хорошо разработанными процедурами оценки и отбора материала для сохранения уже будет иметь практику, навыки и инструментальные средства, которые могут быть полезными при выборе цифровых материалов, даже при том, что они, возможно, будут нуждаться в некоторой корректировке.

12.6 Проблема цифровых материалов

Однако, при работе с цифровыми материалами действительно возникают некоторые новые проблемы, которые менеджеры программы должны учитывать при принятии самых оптимальных решений по отбору.

- Часто имеется большое количество материала, который необходимо оценить
- Способы создания и распространения цифровых материалов широко доступны, так что зачастую их качество не соответствует предъявляемым требованиям
- В то же время, может оказываться давление с целью обеспечить сохранение всего трафика новых каналов связи, таких как Всемирная паутина, независимо от его качества
- Синхронизация отбора обычно является проблематичной, поскольку цифровые материалы сами по себе быстро снимают проблему выбора становясь непригодными. В этом случае может быть целесообразно принимать решения об отборе до появления доказательств непреходящего значения материалов
- Некоторые цифровые объекты бывает трудно найти. Новые жанры могут не соответствовать принятой классификации; некоторые цифровые ресурсы состоят из скомпилированных или накладывающихся частей; многие имеют также параллельные версии. В процессе отбора эти сложности должны быть преодолены и в результате принято однозначное решение о том, что должно быть сохранено
- Даже когда определены внешние границы может быть сложно сказать, какие элементы необходимо сохранить, если цифровой объект должен выполнять свою основную функцию
- Трудность может представлять даже определение источника появления цифровых материалов, что осложняет поиск ответственных за его сохранение лиц и тех, с кем необходимо договориться о правах, необходимых в соответствии с требованиями программ сохранения.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

12.7 Обоснованные, непротиворечивые и ответственные решения

Процессы отбора часто сопряжены с неопределенностью, и они связаны с решениями, которые имеют субъективный и умозрительный характер; однако они должны быть обоснованными, непротиворечивыми и ответственными:

- Решения в отношении материала должны быть хорошо обоснованными, учитывать его контексте, и потребности участников, интересы которых будут затронуты
- Принимаемые решения должны последовательно опираться на практику отбора, отражающую цели организации, которая берет ответственность за сохранение. Для учреждений типа библиотек, музеев и архивов хорошим ориентиром может служить существующая практика создания коллекции
- С точки зрения обеспечения отчетности, процессы отбора должны быть прозрачными, основанными на доступных для общественности документах, четких и ясных заявлениях о том, что было отобрано, а что исключено.

12.8 База для критериев отбора

Невозможно предложить определенные критерии для отбора материалов цифрового наследия, потому что обоснованность их сохранения оценивается по очень разнообразным показателям. Однако, в принципе:

- Решения должны основываться прежде всего на значимости материала и его соответствии задачам организации, принявшей ответственность за сохранение
- Эта значимость должна быть соотнесена с вероятными затратами и трудностями сохранения, и с ожидаемой доступностью ресурсов. Можно много говорить о том, чтобы начинать процесс с материала, легко поддающегося хранению. Однако, будущая стоимость и возможности цифровых программ сохранения все еще неясны, так что, вероятно, было бы безответственно отказываться от ценного материала только потому, что его сохранение может оказаться трудным
- В случае, если программы в области сохранения не в состоянии управлять материалом, который, по их мнению, должен быть отобран для сохранения, они должны указать это в своих принципах отбора
- Желательно, чтобы в результате всех усилий по сбору и сохранению была подготовлена по меньшей мере выборка всех видов цифровых материалов, включая кратковременные выборки

12.9 Определение элементов, придающих ценность материалу

Одного решения о выборе элемента, или класса элементов для сохранения может быть недостаточно:

- Процесс сохранения включает поддержание элементов и характеристик, которые придают материалу его значение. В процессе отбора необходимо определить, что является такими элементами и характеристиками
- Должны быть задокументированы причины, по которым материал был отобран, чтобы менеджеры могли понимать, что они обязаны сохранять. (Еще некоторые подробные замечания представлены далее в этой главе.)

12.10 Взвешенный подход

Решение не сохранять данные обычно является окончательным для цифровых материалов. Взвешенный подход мог бы заключаться в том, чтобы принять решение, какие материалы определенно должны быть сохранены и на какой период; что точно не надо сохранять; и что должно быть принято на промежуточное сохранение, пока не дано более четкое указание.

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ВОПРОСЫ

12.11 Оценка значимости материала

Есть сомнения, что предполагаемые текущие оценки значения материала являются полностью достоверным ориентиром их будущей значимости. Например, старые данные, собранные в течение предыдущих десятилетий, неожиданно стали важными для оценки экологических изменений. Этот опыт подсказывает, что, вероятно, лучше допустить ошибку, собрав больше материала, если программа в области сохранения может управлять им.

12.12 Документация

В случаях, когда цифровые материалы могут быть поняты только при наличии свода правил, типа системы сохранения записи, базы данных или системы поколения данных, или другой контекстной информации, в процессе отбора должна определяться документация, которая также должна сохраняться.

12.13 Роль создателей

Создатели цифровых материалов могут играть существенную роль в выборе объектов сохранения. Они обычно хорошо понимают причины создания цифровых объектов, их основное «сообщение», и отношения между объектами и их контекстом. Если такая информация не получается от создателя, позже ее воссоздание может оказаться слишком трудной задачей.

12.14 Селективный или всесторонний отбор

Может возникнуть вопрос, какой из этих методов является предпочтительным. (Этот вопрос часто возникает при обсуждениях материалов, ставших доступными благодаря, например, Всемирной паутине.)

В поддержку и всестороннего, и селективного подходов есть веские аргументы. Защитники комплексного подхода утверждают, что любая информация может иметь долгосрочное значение, а затраты на детальный отбор больше, чем на сохранение всех данных. Защитники более выборочного подхода заявляют, что он позволяет им создавать коллекции ресурсов, представляющих большую ценность, с определенной гарантией технического качества и возможностью договориться о правах на использование с его создателями.

Вполне могут быть возможными оба подхода, поскольку они, вероятно, будут способствовать созданию различных коллекций цифрового наследия, которые важны для различных целей.

12.15 Соглашения, касающиеся отбора материалов

Чтобы минимизировать риск пропустить важные материалы и избежать ненужного дублирования усилий, может стать необходимым заключение соглашения с другим учреждением по отбору и сохранению о соответствующих задачах и обязанностях.

12.16 Определение элементов

Практика отбора должна отработать механизм, когда выбираются все элементы и целые коллекции, а когда делается только выборка. Вообще то предпочтительно сохранять все элементы, чтобы сохранить их целостность, но может возникнуть необходимость ограничить отбор репрезентативными выборками, являющимися, по меньшей мере, доказательством существования некоторых видов материалов.

Принципы отбора должны также определять, является ли многократное использование материала новым элементом, который также необходимо сохранить.

12.17 Правовые вопросы

Правовые вопросы могут оказывать влияние на решения, касающиеся отбора информации. Программы в области сохранения часто выбирают материалы, на которые все еще распространяется законодательство об авторском праве, и вообще не могут отбирать материал, если права на него имеют настолько ограничительный характер, что о соглашениях о предоставлении доступа в будущем нельзя договориться. Если материал никогда не будет доступен для использования, или в случае невозможности предпринять необходимые шаги по его сохранению, становится бессмысленным его отбор как материала наследия.

12.18 Периодический отбор

Должны ли решения об отборе быть окончательными? Рассмотрение решений об отборе в соответствии с точно установленными периодами хранения является давнишней практикой в архивном сообществе. Этот подход может быть полезным и по отношению к другим видам цифровых материалов, таких как записи, с целью убедиться, что ценность материала все еще оправдывает расходы на его хранение. С другой стороны, сам процесс отбора является дорогостоящим мероприятием и должен повторяться как можно реже. Еще более важно, чтобы любые решения, являющиеся предметом рассмотрения, не оставляли никаких сомнений относительно постоянной ответственности за сохранение.

12.19 Поддержка процесса отбора

Процесс отбора требует распределения ресурсов: людей, обладающих необходимыми знаниями, времени, средств и оборудования, чтобы исследовать материал. Управление процессом отбора требует также разработки критериев оценки. Где объем материала является настолько большим, что невозможно оценить элементы в отдельности, может возникнуть необходимость установить классы материала, которые можно оценить на основе репрезентативных выборок.

СПЕЦИАЛЬНЫЕ ВОПРОСЫ

12.20 Выбор необходимых элементов и характеристик для сохранения

Программы сохранения часто выполняют роль посредников для других участников процесса: они осуществляют мероприятия по сохранению по поручению других, кто хочет сохранить материал с какой-то целью. «Кто-то еще» может иметь очень широкое и многогранное содержание, такое как «нация» или «общество в целом», которые могут уполномочить программу собирать и сохранять очень широкий диапазон материалов; или оно может ограничиваться членами организации или исследователями, занимающимися специфическим предметом, которые хотят сохранить результаты своей работы для использования в будущем.

Потребности «общества» — тем не менее определенные — для которого сохраняется материал, диктуют необходимость осуществлять управление процессом принятия многих решений, от того, какой материал отобран, до вида и уровня записанной документации, степени его подлинности и используемых стратегий. Например, некоторые программы должны предложить пользователям опцию запрашиваемых старых данных для достижения новых результатов, тогда как другие имеют узкую задачу представлять материал только для чтения в формате, который не может быть изменен. Некоторые программы должны даже гарантировать пользователям возможность запускать старое имитационное моделирование, играть в старые компьютерные игры, или рассматривать цифровое искусство, воспроизводимое в первоначальном виде, а не в усеченном варианте, который предоставляют более поздние технологии.

Определить основные элементы или характеристики (также именуемые как *существенные свойства* в соответствии с некоторыми программами) концептуально не сложно, как показывают приведенные выше примеры. При некоторых обстоятельствах — таких как четко определенные и удовлетворяемые ожидания пользователей и схожести между собой материалов, охарактеризовать которые не сложно — определение и кодирование основных элементов должно быть открытым. Например, программа может решить, что пользователи большой коллекции сообщений электронной почты должны видеть только элементы, которые могут быть охарактеризованы как «контент информация», типа наименования и адреса отправителя, темы, даты и времени, получателей, и сообщения в стандартизиро-

ванной структуре с только наиболее простым форматированием. Правительственный архив при таком подходе может применять этот основной шаблон элементов к очень большому количеству почтовых отчетов.

С другой стороны, некоторые материалы характеризовать значительно труднее, а сложный процесс определения перспектив того, как они будут представлены для использования, особенно открытыми сообществами потенциальных пользователей, становится практически невозможным.

Различные подходы к этой проблеме разрабатываются по мере того как все больше людей сталкивается с проблемами описания, сохранения и планирования включить цифровые объекты в растущие коллекции на длительные периоды времени.

В то время как развиваются более сложные методы определения и описания основных элементов, приведенные далее вопросы могут оказать помощь в процессе отбора. (Будет понятно, что это действительно часть процесса оценки, через который проходят менеджеры записей, чтобы понять материал, которые они рассматривают для отбора.)

- Для кого должен сохраняться этот материал? Есть ли у них определенные ожидания относительно того, что они смогут сделать с материалом когда он будет представлен?
- Почему материалы заслуживают сохранения? Что придает им ценность, оправдывающую трудности, связанные с их сохранением? Связана ли эта ценность с:
 - Доказательством;
 - Информацией;
 - Художественными или эстетическими факторами;
 - Существенным новшеством;
 - Исторической или культурной памятью
 - Тем, что пользователь может заставить материал сделать, или сделать с материалом;
 - Существенными характеристиками культурного характера?
- Связано ли значение со способом рассмотрения материала? (Было бы оно утрачено или значительно уменьшилось, если бы материал рассматривался по-другому?)
- Связана ли ценность с тем, как работает объект? (Будет ли он потерян, если определенные функции будут удалены? Или если они будут выполняться на другой скорости или требовать нажатия различных клавиш?)
- Связано ли значение с контекстом материала? (Будет ли он потерян, если не будут работать ссылки, внедренные в материал? Или если пользователь больше не сможет увидеть доказательства того, что материал связан с его первоначальным контекстом?)
- Имеется возможность определить различия между элементами в рамках каждой из этих областей? Например, можно ли по рекламным заголовкам составить представление об основной части материала? Будут ли необходимы некоторые навигационные элементы или функции дисплея, но не другие?
- Если трудно определить, какие потребности должны поддерживаться, может быть проще рассмотреть значение элемента, который не поддерживается, и найти функции или элементы, в которых определено нет необходимости.

ДЛЯ ПРОГРАММ СОХРАНЕНИЯ, РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ

12.21 Избирательность

Программы в области сохранения, испытывающие недостаток ресурсов, должны все еще принимать решения о материалах, за которые они берут на себя ответственность. Поскольку затраты по сохранению связаны с объемом управляемого материала, деятельность таких программ, возможно, должна иметь очень селективный характер, ограничивая свои амбиции небольшим количеством очень ценного материала.

Затраты по сохранению также связаны с диапазоном проблем и форматов, которые должны управляться, так что может также иметь смысл строгое ограничение видов материалов, выбранных в очень немногих форматах.

12.22 Сотрудничество

Заключение соглашений с другими программами может облегчить бремя отбора материалов. В то время как идет обычный договорной процесс с другими программами в области сохранения, может быть возможность для заключения соглашений с производителями, которые подтолкнули бы их к принятию решений о том, что должно быть отобрано и сохранено. Программам в области сохранения все еще необходимо брать на себя некоторую ответственность за контроль качества, равно как и озвучивать критерии, по которым должен выбираться материал.

12.23 Начинать с простого

Процессы отбора могут совершенствоваться со временем, начинаясь с принятия некоторых простых решений по отбору легко коллекционируемых и сохраняемых материалов («низко висящий плод»), и нацеливаясь через какое-то время на более сложные вопросы, касающиеся более широкого диапазона более сложных материалов.

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

Основные элементы см. также Понимание цифрового сохранения: глава 7, и Поддержание доступности: глава 17

Права см. также Управление правами: глава 15

Дополнительные ссылки (все ссылки были действительны в марте 2003 г.)

- Проект Cedars, (1999). Why Can't We Preserve Everything? Selection Issues for the Preservation of Digital Materials. <http://www.leeds.ac.uk/cedars/documents/ABS01.htm>
- Проект InterPARES, (2002). The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project. <http://www.interpares.org/book/index.htm>
- Marcum Deanna B. (2001). «Scholars as Partners in Digital Preservation», in CLIR Issues, no. 20, Совет по библиотечным и информационным ресурсам. <http://www.clir.org/pubs/issues/issues20.html-scholars>
- Национальная Библиотека Австралии, (updated 2002). Guidelines for the Selection of Online Australian Publications Intended for Preservation by the National Library of Australia. <http://pandora.nla.gov.au/selectionguidelines.html>

- Королевское статистическое общество и Британский архив данных (2002). Preserving and Sharing Statistical Information. <http://www.data-archive.ac.uk/home/PreservingSharing.pdf>

Глава 13. Работа с создателями

ВВЕДЕНИЕ

13.1 Цели

Глава предназначена для того, чтобы подвигнуть менеджеров программы к рассмотрению способов работы с создателями цифрового наследия, а также снабдить первых некоторыми принципами в отношении практических мер и стандартов, которые облегчат выполнение задачи сохранения.

13.2 Краткое содержание главы

Цифровое наследие часто создается без учета обеспечения его непрерывного использования и доступности. Однако есть определенные стандарты и практика, которыми могут пользоваться создатели для того чтобы способствовать или препятствовать осуществлению процесса сохранения. Менеджеры программы должны найти способы извлечь положительные аспекты из самых ранних материалов цифрового наследия. Для этого зачастую требуется готовность работать с создателями.

13.3 Терминология

Термин *создатели* используется в этой главе применительно ко всем, кто участвует в разработке дизайна, написании, создании и распространении цифровых материалов прежде, чем они становятся объектами программы в области сохранения. Программы оцифровки попадают непосредственно в категорию «создателей», к чьей цифровой продукции должен быть обеспечен непрерывный доступ в соответствии с программами в области сохранения.

ОБЩИЕ ВОПРОСЫ

13.4 «Предыстория» цифрового наследия

Цифровые материалы создаются производителями, которые не обязательно заинтересованы в их долгосрочной доступности: создание «цифрового наследия» может не входить в их планы. Даже те, кто надеется сделать что-то, имеющее непреходящее значение, могут не иметь необходимых знаний или средств для этого, или сдерживаться другими препятствиями в их рабочей среде.

Без определенного вмешательства, маловероятно, что материалы цифрового наследия будут создаваться автоматически, что минимизирует затраты и устранил барьеры на пути к сохранению. Многие мероприятия фактически осложняют процесс сохранения.

13.5 Трудности делового общения с создателями

Стремясь работать с создателями для преодоления барьеров на пути к сохранению, программы сталкиваются с проблемами:

- Во многих случаях «создатель» — это многоуровневое понятие, включающее множество действующих лиц, выполняющих совершенно разные функции, таких как разработчики программного обеспечения, создатели (часто многократные), редакторы, издатели и поставщики услуг

- Некоторые создатели могут испытывать неуверенность в себе или даже враждебность к идее, что третье лицо интересуется так или иначе «управлением как цифровым наследием» материалами, которые они создали.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

13.6 Необходимость проводить работу с создателями

Усилия по сохранению, которые предпринимаются в то время, когда проблема уже назрела, вероятно, будут более дорогостоящими, трудными и менее эффективными, чем усилия на ранней стадии.

Организации, которые осуществляют функции и по созданию наследия и по его сохранению, знают из опыта, что внимание, проявленное с самого начала к использованию стандартов, документации, хорошему управлению файлами и другим мероприятиям, приносит позже дивиденды, заключающиеся в более низкой стоимости сохранения и эксплуатационных расходов, равно как и в создании более легкодоступных, многократно используемых и управляемых коллекций.

В то время как все программы в области сохранения не располагают одинаковыми возможностями оказывать влияние на деятельность промышленности, все программы должны стремиться влиять на способы создания и управления материалами на самой ранней стадии.

13.7 Что означает «работа с создателями»?

В широком смысле работа с создателями, вероятно, включает некоторые, или все, приведенные ниже аспекты:

- Доведение до них информации о существовании программ в области сохранения, их задачах и мероприятиях
- Обсуждение путей, с помощью которых промышленный процесс может помочь или воспрепятствовать процессу сохранения
- Определение выгод для обеих сторон при уменьшении любых помех для сохранения
- Поиск взаимоприемлемых путей облегчения процесса сохранения
- Определение проблем создателей и поиск взаимоприемлемые способов обращения к ним
- Если приемлемо, представить подробные советы относительно хороших мероприятий типа использования стандартов, форматов, управления файлами и метаданными
- Ведение переговоров о соглашениях по передаче и управлению правами
- Заключение соглашений о проведении определенных действий, часто основанных на реализации пилотных проектов и проведении совместных оценок.

13.8 Эффективное сотрудничество

Эффективность сотрудничества между программой в области сохранения и создателями может зависеть от ряда факторов, таких как:

- Характер взаимоотношений между ними. Например, рассмотрите разницу в потенциальном влиянии на:
 - Создание архива организационных записей органов правосудия, с созданием записей в пределах их организации;

- Архив данных, признанный в национальном масштабе, ведущий переговоры с независимыми исследователями, создающими наборы данных по широкому аспекту академических дисциплин;
- Правительственный аудиовизуальный архив, стремящийся убедить создателя независимых записей, что их записи «с заднего двора» являются частью национального наследия;
- Небольшую специализированную коллекцию, пытающуюся сохранить созданные на коммерческой основе и продающиеся во всем мире компьютерные игры.

- Готовность создателей к участию

- Техническую экспертизу и понимание, которые могут предложить программы в области сохранения
- Навыки программ сохранения в ведении переговоров о заключении взаимовыгодных соглашений.

Программа в области сохранения должна стремиться максимально повысить эффективность своего влияния в пределах возможного.

13.9 Выгоды

Есть много потенциальных выгод для программы в области сохранения, работающей с создателями над преодолением барьеров на пути сохранения; существуют также потенциальные выгоды для создателя. Некоторые из них представлены в Таблице 13-1.

Краткосрочные выгоды для программ сохранения	Долгосрочные выгоды для программ сохранения	Выгоды для создателя
Установленные темы контакта могут облегчить общение	Улучшенный выбор форматов и то, как они используются, и возможности заключить соглашения об установке обходных устройств защиты, блокирующих копирование сохраняемых материалов	Улучшенное представление материалов в заархивированных коллекциях
Передачи могут быть более просты, особенно там, где не налажен автоматизированный «сбор» (см. Главу 14)	Улучшенная передача документации	Более эффективные технологические процессы, меньше работы по «восстановлению» с целью выполнения требований по архивированию
Причастность создателя к принятию решения, что должно быть сохранено	Лучшее понимание ролей и обязанностей	Более широкое признание значения их работы
Понимание, какой материал является доступным, и как он воспринимается сообществом создателей	Видение будущих тенденций в производстве материалов цифрового наследия	Представление их работы более широкой аудитории, которая может создать новые рынки и способствовать проявлению более широкого интереса
Идентификация «невидимых» материалов	Основание для идентификации приоритетных проблем с определенными сообществами	Может помочь установить доверие к новым формам создания и распределения информации
Менее дорогостоящие передачи	Менее дорогостоящее долгосрочное сохранение	Возросший интерес к использованию открытых источников программного обеспечения (для целей сохранения) может поощрять новые совместные производственные модели

Таблица 13-1 Некоторые потенциальные выгоды сотрудничества между программами в области сохранения и создателями материалов цифрового наследия

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ВОПРОСЫ

13.10 Осознание различий

Важно осознать, что создатели цифровых материалов работают в различной среде и, вероятно, существенно отличаются друг от друга: тем, как они относятся к своей работе, размер их действий, организационная и техническая поддержка, находящаяся в их распоряжении, и их интерес к проблемам обеспечения долгосрочного доступа. Например, собирающие данные ученые, вероятно, будут проявлять большую заинтересованность в том, как точно и надежно защищены их данные; насколько соответствуют предложенные форматы и стандарты их рабочим потребностям; приемлемы ли соглашения о передаче; и сохранены ли вероятные права и контроль за доступом к их данным. С другой стороны, коммерческие издатели пакетов CD-ROM, вероятно, больше будут интересоваться средствами управления неуполномоченным копированием; затратами и рисками, связанными с представлением «незащищенных» версий программе в области сохранения; потенциальное многократное использование их содержания; и их лицензионные обязательства перед владельцами программного обеспечения, чью продукцию они используют.

13.11 Подходы к работе с создателями

Есть много походов к формам работы ответственных хранителей цифрового наследия с его создателями и распространителями.

- Очевидным первым шагом должно стать определение круга участников. Некоторые мероприятия полезно будет провести с группами, представляющими промышленность, но некоторые шаги могут потребовать личных контактов и проведения переговоров
- Создатели также должны знать с кем имеют дело. Программы в области сохранения должны действительно способствовать осознанию их собственной роли
- Может быть выгодно идентифицировать группы создателей и работать с ними, решая определенные проблемы, вместо того, чтобы пытаться решить все вопросы на общем уровне
- На более широком промышленном уровне, может быть полезно разработать практическое руководство, которое устанавливает согласованные решения о ролях и обязанностях, и определяет сферу и условия осуществляемого сотрудничества
- Во многих секторах есть активные промышленные группы, которые обеспечивают форумы для обсуждения проблем. Наряду с предоставлением возможностей для диалога с промышленными лидерами, такие форумы могут способствовать установлению новых взглядов на долгосрочную перспективу
- Для программ в области сохранения важно оказать поддержку и обеспечить обратную связь мероприятиям, которые хотят осуществить создатели, и обеспечить уровень ответственности в процессе работы с сохраняемым материалом. Доказательство того, что результатом сотрудничества являются эффективные меры по сохранению, вполне может побуждать создателей устанавливать и поддерживать дальнейшее сотрудничество

13.12 «Улица с двусторонним движением»

Во многих ситуациях работа с создателями означает реальный ввод ресурса программой в области сохранения, а не только создателем. Возможные области ввода могут включать:

- Обеспечение письменных рекомендаций и спецификаций

- Обеспечение возможности обучения персонала
- Оказание помощи в проектировании систем и технологических процессов
- Обмен информацией и рабочими консультациями
- Ряд соглашений по материалу, в управляемой создателем программе по сохранению.

СПЕЦИАЛЬНЫЕ ВОПРОСЫ

13.13 Спецификации и практическое руководство

Практическое руководство, вероятно, должно включать советы по следующим аспектам:

- Организационные проблемы, облегчающие управление цифровыми материалами
- Планирование проекта, делая акцент на дизайне системы, предшествующей созданию любых записей или изданий
- Выбор носителя. Создатели должны побуждаться к использованию продуктов «индустриальной силы», которые будут служить достаточно долго, чтобы перенести данные на другие носители, или создателем, или передав программе в области сохранения
- Выбор соответствующих форматов файла и стандартов данных. Если нет очень веских причин сделать по-другому, создателей необходимо побуждать к использованию очень широко принятых, хорошо стандартизированных файловых форматов, которые соответствуют их целям. Вообще говоря, данные в более простых форматах, использующие открытый источник, не частное программное обеспечение более удобны для сохранения (хотя некоторые частные приложения используются настолько широко, что они могут быть приняты как промышленный стандарт, особенно если их спецификации открыто изданы). Интерактивные материалы, изданные для общественного доступа, должны быть читаемы обычно используемыми браузерами. Структурированные по стандарту документы в легко распознаваемом и долговечном формате типа XML (Расширяемый язык разметки) следует рассматривать как материал непреходящего значения
- Проверка правильности форматов. Недостаточно просто выбрать стандартный формат и затем использовать это ненормативными способами: форматы должны эксплуатироваться в соответствии с их стандартом и в случае необходимости проверяться, чтобы удалить любые особенности, которые, вероятно, могут усложнить сохранение. (Существует много интерактивных инструментальных средств, доступных для проверки правильности диапазона файловых форматов.)
- Названия файла должны быть непротиворечивы и однозначны
- Интерактивные файлы должны управляться в режиме уверенного доступа с помощью устойчивого идентификатора и преобразователя услуг, или переадресовывать сообщения, если файлы перемещаются. Множество схем РІ используется интернационально в различных секторах, хотя ни одна — универсально. Схема DOI (Цифровой идентификатор объекта), используемая коммерческими издателями для управления правами, принята наиболее широко

- Создатели должны подготовить метаданные хорошего качества для производимых ими ресурсов, используя широко принятые схемы типа MARK, элементы метаданных the Dublin Core или одного из многих секторных расширений. Метаданные помогут пользователям находить и использовать их ресурсы. Метаданные должны быть также записаны и содержать техническую характеристику цифровых объектов, необходимую информацию для доступа к ним и о любых изменениях в течение срока службы: эта информация будет необходима для управления ими. Метаданные могут быть внедрены в ресурсы или сохранены в сопряженном файле метаданных
- Управление файлами. Основные файлы сохранения должны храниться и управляться отдельно от копий для распространения. Процедуры управления базой данных должны гарантировать, что данные не записываются поверх прежде, чем они зафиксированы
- Системная защита. Файлы и системы должны быть полностью защищены от повреждения или потери с помощью внедрения лучших мер защиты и соответствующих резервных мероприятий даже для краткосрочного хранения
- Подлинность. Все файлы должны быть идентифицированы, а их происхождение и хронология зарегистрированы, чтобы обеспечить доказательство их подлинности
- Обучение. Персонал, подрядчики и другие люди, работающие с цифровыми материалами, должны руководствоваться соответствующими процедурами и наставлениями, быть обученными, активными и готовыми к их использованию
- Если барьеры на пути к доступу или копированию считаются необходимым условием для защиты интеллектуальной собственности, они вполне могут сделать сохранение невозможным. Понадобятся соглашения для осуществления мер по сохранению, таких, например, как копирование
- Начальные шаги по поддержанию доступа могут включать хранение всего программного обеспечения, необходимого для обеспечения доступа, равно как и любых специализированных аппаратных средств. Это не является эффективной долгосрочной стратегией, но вполне может потребоваться на краткосрочный период
- Может возникнуть необходимость оценить цифровые материалы, определить срок их хранения и круг ответственных лиц, в соответствии с одобренной практикой, как, например, полномочия по передаче в архив.

ДЛЯ ПРОГРАММ СОХРАНЕНИЯ, РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ

13.14 Снижение загрузки

Цифровые программы в области сохранения могут оказаться не в состоянии экономить ресурсы для работы с создателями. Однако может представиться возможность осуществить ряд заслуживающих внимания мер ограниченного, целенаправленного характера, которые приведут к снижению уровня расходов в будущем. Например:

- Привлечение одного или двух создателей для изучения, какие из мероприятий могут облегчить некоторые дальнейшие действия, являющиеся предметом переговоров

- Ограничение диапазона материалов, отбираемых в нескольких хорошо стандартизированных форматах, может облегчить обеспечение спецификациями, которым смогут следовать создатели без необходимости индивидуального ввода
- Использование существующих рекомендаций, подготовленных для других программ, способствует достижению аналогичной цели, пока они соответствуют обстановке. (Ко многим подобным документам возможен интерактивный доступ, как, например, к рекомендациям Библиотеки Конгресса, Британской службы архивных данных по искусству и гуманитарной тематике. Различные организации имеют также договорные лицензионные соглашения с коммерческими издателями, которые, например, могут являться хорошими моделями для обсуждения с локальными создателями.)

13.15 Распределение загрузки

Программы сохранения могут также найти партнеров, проявляющих желание совместно с создателями внести свой вклад, как соисполнители:

- Может быть возможно найти партнерское учреждение с лучше обеспеченной ресурсами программой, которая уже установила хорошие рабочие контакты с сообществом создателей. Согласно соглашению о развитии, создатели могут проявить желание включить в договор других партнеров, если есть адекватные меры защиты их интересов
- Возможно также, что множество более мелких программ, работающих в том же самом регионе, сформируют консорциум с целью заключения соглашений с создателями от имени всех.

ОПЫТ ПРИМЕНЕНИЯ

В таблице 13-3 представлены некоторые возможные сценарии действий в различной обстановке.

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

Правовые вопросы см. также Управление правами: глава 15

Другие источники (все ссылки были доступны в марте 2003 г.)

- Arts and Humanities Data Service, (nd). Digitisation: a Project Planning Checklist. <http://ahds.ac.uk/checklist.htm>
- Канадская информационная сеть наследия (CHIN) (2002). Creating and Managing Digital Content. http://www.chin.gc.ca/English/Digital_Content/index.html
- Библиотека Корнуэльского университета (2003). Moving Theory into Practice: Digital Imaging Tutorial. <http://www.library.cornell.edu/preservation/tutorial/contents.html>
- Федерация Цифровых Библиотек (2002). Benchmarks for Digital Reproductions of Monographs and Serials. <http://www.diglib.org/standards/bmarkfin.htm> – [bench](http://www.diglib.org/standards/bmarkfin.htm)
- Институт новейших технологий по гуманитарной тематике и информатике (НАТИ), Университет Глазго, и Национальная инициатива по сетевому культурному наследию

(NINCH) (2002). The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials. <http://www.nyu.edu/its/humanities/ninchguide/>

- Институт музейных и библиотечных услуг (IMLS) (2001). A Framework of Guidance for Building Good Digital Collections. <http://www.imls.gov/pubs/forumframework.htm>
- Библиотека Конгресса (2002). Building Digital Collections: Technical Information and Background Papers [for American Memory Programme]. <http://lcweb2.loc.gov/ammem/ftpfiles.html>
- МАТРИЦА: (Интерактивный центр искусств, литературы и общественных наук в университете штата Мичиган), (nd). Working Paper on Digitizing Audio for the Nation Gallery of the Spoken Word and the African Online Digital Library. <http://africandl.org/bestprac/audio/audio.html>
- Национальная библиотека Австралии (2002). Safeguarding Australia's web resources: guidelines for creators and publishers. <http://www.nla.gov.au/guidelines/webresources.html>
- Совет по изучению окружающей среды (NERC) (Великобритания), (2002). NERC Data Policy Handbook. <http://www.nerc.ac.uk/data/documents/datahandbook.pdf>
- Pockley, Simon (1998). Cinemedia: Good Practice Guide. <http://www.acmi.net.au/FOD/DuckDigital/GoodP.html>
- Rowe, J (2002). 'Developing a 3D Digital Library for Spatial Data: Issues Identified and Description of Prototype' in RLG DigiNews, 6(5). <http://www.rlg.org/preserv/diginews/diginews6-5.html – feature1>
- Townsend, S, Chappell, C, Struijve, O, (1999). Digitising History: a Guide to Creating Digital Resources from Historical Documents. (Практическое руководство AHDS). http://hds.essex.ac.uk/g2gp/digitising_history/index.asp

	Библиотека с правами сохранения	Библиотека с лицензированным доступом	Архив организации	Специализированный аудио-видео архив	Архив данных	Проект общественного архива
Вероятный уровень влияния или управления:	Низкий: может быть сложно найти создателей или войти в контакт с ними, круг создателей может быть слишком широк (например, веб-издатели), проблемы с коммерческими правами, невозможность определить используемые форматы	Хороший с точки зрения управления форматами, но может быть низким с точки зрения уровня управления сохранением	Потенциально хороший, если установленны спецификации и стандартные процедуры, возможна разработка правозащитных мер или организационных мер для взаимовыяснения	Может быть хорошим для пополнения коллекций, но низким для других, возможно использование различных форматов и стандартов	Вероятно будет хорошим, если получит знание со стороны создателей и пользователей	Вероятно, будет хорошим, если подключение к разработке проекта на ранней стадии, но недостаточно, если это не будет следовано
Возможные стратегии воздействия:	Идентификация создателей и работа с ними, пробуйте разработать программу практического обучения, проведения семинаров, рекомендации, подчеркните выгоды для создателей, может потребоваться отдельное согласование для создания правовых основ	Могут потребоваться длительные переговоры для обеспечения непрерывного доступа	Разработка программ обучения и технической поддержки для гибкости процесса	Разработка стандарта форматов для включения в программу	Опубликуйте требования, помогите создавать собственным проектам, поощряйте создание баз кон-текстной информации	Работайте в тесном взаимодействии с создателями, обеспечьте наличие инструментов, которые облегчат включение интегрированных программ сохранения в задачи общественного проекта

Таблица 13-3 Некоторые возможные методы работы с источниками в различных секторах

Глава 14. Контроль за метаданными и их пересылкой

ВВЕДЕНИЕ

14.1 Цели

Эта глава предназначена для того, чтобы дать представление об общих и технических проблемах контроля сохранения цифрового наследия соответствующими программными средствами.

14.2 Краткое содержание главы

Контроль за обработкой цифровых материалов является ключевым вопросом их сохранения. В большинстве случаев это требует обеспечения безопасной пересылки данных и документации на программные модули сохранения информации, где им присваивают уникальные идентификационные атрибуты и описывают с использованием различных видов метаданных. Метаданные дают возможность поиска цифровых материалов, а также их обработки и наиболее точного отображения. Хотя стандарты сохранения метаданных все еще находятся в стадии разработки, программы должны описать технические характеристики, источник и цель ввода хранящихся в них цифровых материалов.

ОБЩИЕ ВОПРОСЫ

14.3 Безопасная пересылка данных

Мы рекомендуем осуществлять пересылку материалов цифрового наследия на безопасные носители с целью сокращения риска повреждений или потерь, связанных с ежедневным использованием цифровых файлов. В большинстве случаев это требует пересылки данных под контролем специальных агентств, ответственных за сохранность информации.

Сам по себе процесс пересылки данных уже содержит элемент риска, так как при этом сохраняются возможности потери, случайных изменений, нераспознавания части информационных массивов или потери контекстных поясняющих сведений.

14.4 Вопросы правообладания

Поскольку источники цифровых материалов уже сами по себе обладают некоторыми авторскими правами на свою продукцию, процесс их пересылки порождает ряд вопросов правообладания в юридической и моральной области.

14.5 Вопросы контроля

По мере осуществления процесса пересылки информации в цифровом виде, необходимо организовать эффективное управление и контроль за информационными массивами. Это включает в себя требования по обеспечению возможности доступа, поиска, использования, управления и сохранения в соответствии с правами пользователя и владельца.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

14.6 Имеющийся опыт

Процессы пересылки и контроля использовались еще при управлении наследием с использованием нецифровых способов сохранения. Сегодня, в век цифровых технологий сохранения, эти процессы должны претерпеть определенные изменения.

- Необходимо формирование соответствующего правового поля, сопровождающего процессы пересылки. В нем должна найти отражение как возможность обеспечения многократного использования материалов, так и случаи копирования данных для сохранения.
- Пересылка должна осуществляться без потери данных, при этом необходимо учитывать, что зачастую используются методы, совершенно отличные от используемых для пересылки нецифровых материалов
- Пересылка сопровождающей документации, как описательной базы, является особенно важной для цифровых данных.

14.7 Два подхода к осуществлению пересылки данных

Большинство стратегий пересылки сводятся к двум основным группам: источники направляют информацию в цифровом виде на программы сохранения, или программы сохранения сами изымают материалы от источника.

Менеджеры программы должны выбрать наиболее приемлемый подход, исходя из вида материалов и имеющейся технологической оснастки.

14.8 Контроль форматов и стандартов

Ряд программ при пересылке осуществляет контроль за форматом получаемого материала. Это предназначено для упрощения процесса сохранения, при котором должно снижаться многообразие и упрощаться доступ к сохраняемым материалам. Не все программы способны ограничивать воспринимаемые форматы, но в каждую из них необходимо закладывать функцию проверки заданных форматов.

14.9 Контроль за материалами при помощи идентификации

Цифровым файлам необходимо присваивать соответствующие файловые идентификаторы для облегчения поиска. Каждый файл, хранящийся в памяти, должен быть идентифицирован с уникальным файловым именем с тем, чтобы его не могли перепутать с другими файлами.

Также для программ в области сохранения важно обеспечить возможность гарантированного поиска материалов вне зависимости от их местонахождения. Для опознания доступных через сеть ресурсов используется унифицированный идентификатор информационного ресурса (URL). Однако он обладает некоторыми недостатками. Так, например, он не позволяет пользователям найти материал после его перемещения. Таким образом, некоторые элементы могут быть потеряны, даже если они все-таки существуют и хорошо защищены. Преодоление этой проблемы требует наличия системы *устойчивой идентификации*, сформированной на базе отдельных идентификаторов, а также способов их поиска или осуществления связи с файлом в его текущей локализации. Существует и разрабатывается ряд схем, в том числе использование Идентификатора цифровых объектов (DOI), применяемого издательствами, а также различные схемы, применяемые в библиотеках и архивных службах, но ни одна из них не нашла универсального применения.

14.10 Управление материалами при помощи описаний

Программы сохранения используют метаданные — структурированную информацию об информационных ресурсах — предназначенные для описания сохраняемой ими информации в цифровом виде. Существуют, по крайней мере, три основных причины для подробного описания материалов цифрового наследия:

- Необходимость их легкого поиска, доступа и распознавания. Эта потребность привела к разработке метаданных *ресурсного поиска*, работающих по схемам от простого поиска файла по названию до обширных контекстных описаний. Схемы метаданных *ресурсного поиска*, вроде Дублинского каталога, MARC, стандартов архивного поиска и музейных каталогов, являются важными инструментами, используемыми для нужд программ в области сохранения
- Технологические процессы должны быть управляемы. Программы сохранения генерируют большое количество информации о создании, пересылке и использовании материалов, о правах на доступ и других элементах процесса управления. Одним из примеров применения метаданных ресурсного управления является Словарь данных Института NISO в США — Технические метаданные для цифровых неподвижных изображений, изданные как проект стандарта в 2002 году (доступно в интерактивном режиме по адресу http://www.niso.org/standards/resources/Z39_87_trial_use.pdf)
- Программы в области сохранения должны определять способ отображения цифровой информации в виде, необходимом для нормального доступа. *Метаданные сохранения*, наряду с управляющими процессом сохранения элементами метаданных ресурсного управления, описывают возможные способы доступа. Это крайне важно для любой программы сохранения. Их разработка и размещение особенно важны для больших массивов информации, которые должны быть обработаны с максимальной степенью автоматизации.

14.11 Метаданные как информационный ресурс

Метаданные представляют собой самостоятельный информационный ресурс, который должен сохраняться наряду с материалом, который они описывают.

14.12 Подход к стандартизации метаданных

Отдельно разработанные схемы метаданных могут успешно применяться для описания массивов цифровых материалов, однако, по опыту других распространенных схем, имеются веские основания для стандартизированного подхода:

- Возможность значительно снизить затраты на разработку отдельных схем
- Использование в своих интересах имеющихся программных инструментов, которые автоматически распознают и вносят стандартные элементы метаданных о цифровых материалах, тем самым значительно снижая расходы на формирование метаданных
- Обеспечение возможности программам в области сохранения совместно использовать информацию, делая их массивы доступными для поиска более широкой аудиторией
- Обеспечение возможности переноса информационных массивов из одного архива в другой без полного изменения метаданных
- Стандартизация процессов сохранения, которые описываются и управляются метаданными.

В рамках настоящих Рекомендаций дополнительная информация о метаданных, кроме метаданных сохранения, которые будут более подробно рассмотрены ниже, не рассматривалась. Как правило, программисты, работающие с программами в области сохранения, должны иметь представление о стандартных схемах метаданных, которые широко используются в этом секторе, и в идеале должны работать с теми из них, которые наилучшим образом отвечают их потребностям. При этом они должны уделять внимание работам различных международных сообществ, заинтересованных в обработке цифровых материалов, по развитию стандартов метаданных.

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ВОПРОСЫ

14.13 Подход к организации пересылки данных

- Возможность воспроизводства цифровых материалов означает, что процесс пересылки более не требует удаления материала из сайта для перемещения его на другой адрес. Сейчас подлинные копии могут быть переданы для сохранения, в то время как «живые» копии остаются у создателя
- Синхронизация пересылки может иметь достаточно важное значение. Даже если материал отбирается для сохранения, на этот процесс не влияют устаревание носителей или технологии. Пересылка произойдет достаточно быстро, чтобы быть подверженной этим угрозам
- Процесс пересылки, возможно, должен включать, в дополнение к выбранным файлам:
 - Пересылку документации (включая обложки и упаковку материальных носителей типа компакт-дисков и дискет), правил обращения с данными и информацию о происхождении и первоначальном значении
 - Пересылку существующих метаданных
 - Информацию о правах, включая любые лицензионные соглашения
 - Информацию о способах доступа, возможно со специальными программами или аппаратными средствами, которые необходимы для правильного доступа.

14.14 Определение форматов файлов и носителей

Не существует универсального способа осуществить пересылку цифровых материалов. Данные могут быть перенесены с помощью широкого диапазона материальных носителей, в том числе различных дискет, компакт-дисков, лент, картриджей и дисководов; либо с использованием систем связи, в том числе почтовых вложений, протокола пересылки файлов (FTP) и веб-узлов. Выбор при этом зависит только от потребностей сторон.

Независимо от выбранного способа должна обеспечиваться безопасность данных. Некоторые способы передачи могут содержать специфические риски для носителей; например, материальные носители легко могут быть потеряны или похищены, в то время как системы связи могут быть ненадежны, и более безопасным способом станет передача материального носителя с посылным.

Способ пересылки должен позволять последующую загрузку и восстановление данных. Когда источник и получатель используют ту же технологию, пересылка становится относительно легким делом. В иных случаях одна или обе стороны должны будут нести затраты из-за использования разных технологий.

Программы в области сохранения могут иметь средства обработки широкого диапазона носителей, или быть более ограничены в их выборе. Материальные носители требуют специализированных аппаратных средств, которые организации, работающие в области сохранения, могут быть не в состоянии обеспечить. В таких случаях они должны будут решить, разумно ли требовать, чтобы пересылка осуществлялась только на тех носителях, которые они могут обработать, или вкладывать капитал в средства обработки более широкого диапазона носителей.

Некоторые соображения по выбору способа пересылки приведены в Таблице 14-1.

Если данные должны оставаться на носителях для среднесрочного хранения	избегать краткосрочных носителей типа лент или дискет
Если данные будут немедленно загружены на другие носители для хранения	краткосрочные носители могут отвечать требованиям пересылки
Если затраты на поддержку широкого диапазона носителей недопустимы	необходимо определить более узкий диапазон носителей
Если технологические процессы сформированы на базе специальных носителей	необходимо определить носители, которые соответствуют технологиям, или определить носители, простые в обслуживании и под которые можно скорректировать технологию

Таблица 14-1 Принятие решения по выбору носителей для пересылки

14.15 Стратегии пересылки

Пересылка данных обычно подразумевает, что программы сохранения получают файлы от источника (пассивный способ), или сами забирают файлы с сайта источника (активный способ).

Используется множество пассивных способов, таких, как передача файлов на материальном носителе почтой или курьером; пересылка файлов в виде вложений электронной почты или их пересылка непосредственно на сервер программы в области сохранения. Такие способы имеют ряд преимуществ, поскольку они позволяют источнику сохранить более простые версии продукта, чем их общедоступный аналог, и предоставить ему большие возможности по влиянию на отбор сохраняемых материалов.

С другой стороны, если программы сохранения базируются только на пассивных способах, то процесс пересылки может быть поставлен в зависимость от факторов, находящихся вне их контроля, включая смену персонала, изменение приоритета или снижение интереса, что может привести к нарушению процесса.

Активные способы дают программам в области сохранения большие возможности по управлению, в том числе и в вопросах синхронизации и содержания материалов. Некоторые источники рассматривают это как нарушение их прав и либо блокируют возможность копирования их файлов, либо устанавливают программы запроса прав на копирование, поэтому активные способы могут не всегда срабатывать. (С другой стороны, множество источников испытывают удовлетворение от того, что их материалы будут сохранены, причем без всякой платы.)

Сбор или автоматизированное накопление материалов с сайтов источников стали возможными в эпоху информационных сетей. Используя программное обеспечение, нацеленное на просмотр сетей в поисках файлов, которые удовлетворяют заданным критериям, программы сохранения могут копировать и загружать файлы на собственные компьютерные системы. Такой подход широко используется поисковыми средствами Интернета и большинством программ в области сохранения, собирающих материалы в сетях. Различ-

ные программы индексации и поиска могут с различной точностью определить, что должно или не должно быть сохранено.

Сбор может быть достаточно эффективным способом накопления данных, однако и он может также столкнуться с определенными проблемами. Некоторые файлы могут быть невидимы для программного обеспечения, будучи доступными только через интерфейс пользователя. Некоторые источники также сохраняют более высококачественные версии их работ, например изображений и звуковых файлов, отдельно от промежуточных или вторичных версий, предназначенных для распространения по сети: в этом случае пропускаются версии, которые должны сохраниться и сохраняются версии, предназначенные только для временного доступа.

Решение этих проблем часто осуществляется при использовании комплексного подхода, когда источники соглашаются разместить соответствующую версию своих работ в доступном для программ в области сохранения месте.

14.16 Контроль качества

Независимо от способа пересылки, программы сохранения должны проверить материалы, поскольку необходимо убедиться, что все заданные файлы были получены и работают, а метаданные и другая документация в наличии.

14.17 Идентификация файлов

Цифровые объекты могут обладать множеством идентификаторов, по-разному используемых в процессе управления, для идентификации всей системы и для обеспечения глобального доступа, (точно так же книга на библиотечной полке может быть идентифицирована в соответствии с ее заголовком, номером по классификации, местонахождением полки, порядковым номером читателя, номером в базе данных каталога, международным серийным номером книги и так далее).

Некоторые виды идентификации необходимы для поиска элементов, даже если они были перемещены. Любые ссылки в пределах объектов будут продолжать работать, только если они связано с идентификаторами положения в системе.

Существуют иные подходы, которые включают:

- Информирование пользователей о любых изменениях местоположения, применяемое в малых системах
- Автоматическая переадресация пользователей на новый адрес
- Управление хранением файлов, сводящее их перемещения к минимуму
- Использование схемы идентификаторов положения (ИП), которая сохраняет уникальное имя и описание файла на специальное устройство регистрирующее ИП и их текущее местоположение, (следует отметить, что существующие схемы ИП в основном находятся или в разработке, либо слишком дорогостоящи.)

14.18 Обеспечение метаданных

Было уже отмечено, что метаданные необходимо не только регистрировать, но также обеспечивать их работу и хранение. Этот процесс включает ряд элементов:

- Структурирование. Перевод метаданных в структуру стандартизированного документа, такого, как шаблон XML, должен облегчить возможность их хранения

- **Взаимосвязь.** Необходимо поддерживать связь между метаданными и цифровыми объектами, которые они описывают. Есть множество мнений о том, где лучше хранить метаданные для наилучшего доступа. В то время, как некоторые метаданные должны быть приложены к материалу, чтобы программные инструменты могли автоматически его обрабатывать, существуют разногласия, должны ли полные метаданные быть сохранены отдельно, приложены к объектам, которые они описывают, или даже становятся их частью. Раздельное хранение позволяет обращаться и модифицировать метаданные без извлечения из памяти сопряженных цифровых объектов, что уже является большим преимуществом. С другой стороны, множество менеджеров программы обеспокоены возможностью обеспечения связи между объектом и записью, которая может быть прервана за длительный промежуток времени хранения. При этом необходимо оценить возможные риски и выбрать наиболее безопасный подход
- **Контроль качества.** Обеспечение доверия к записям метаданных обладает высоким приоритетом. Необходимо принимать меры по контролю качества всякий раз, когда создаются или изменяются записи метаданных
- **Защита.** Целостность отчетов метаданных должна обеспечиваться с теми же требованиями, как и для объектов, которые они описывают.

14.19 Подготовка архивного пакета для хранения

По мере переноса цифрового материала и начала работ по его управлению и описанию, материал должен быть подготовлен к записи в систему памяти. Для этого определяются связи между частями информационного пакета (включая содержание и любые метаданные) и создается поток данных, который может быть сохранен на носителях для последующего использования и может быть найден соответствующим поисковым файлом.

После этого считается, что пакет сохранен в памяти.

Перед внесением на хранение цифрового объекта ряд программ создают дополнительные копии, по крайней мере по двум важным причинам:

- чтобы иметь в наличии копию для постоянного использования без извлечения сохраненного пакета. Пользовательские копии часто оптимизируются для обеспечения возможности доступа с использованием современных технологий связи и отображения (низкое разрешение, сжатые загрузочные модули для быстрого интерактивного доступа). Вторичные копии для доступа вообще не должны сохраняться после изменений технологии, и они часто не имеют детальных отчетов метаданных сохранения.
- чтобы сохранять объекты в более чем одном формате, привлекая различные стратегии для обеспечения доступа в будущем. Как отмечено в главе 16, это удобная практика — сохранять копии цифровых объектов в их первоначальных форматах, независимо от потребности создания новых форматов или копий для обеспечения доступа.

Разумеется, любые параллельные версии должны управляться как отдельные, но взаимосвязанные цифровые объекты.

ОСОБЕННОСТИ

14.20 Метаданные сохранения

Метаданные сохранения — это структурированная информация о цифровом объекте, которая:

- Идентифицирует материал, за который несет ответственность программа в области сохранения
- Сообщает, что необходимо для сохранения и защиты данных
- Сообщает, что необходимо для отображения определенного объекта (или его необходимых элементов) пользователю, независимо от изменений в технологии хранения и доступа
- Записывает историю и результат работы с объектом
- Документирует достоверность и целостность объекта как основание для определения подлинности
- Позволяет пользователю и программе сохранения понять контекст объекта при его хранении и использовании.

Условия записи метаданных сохранения должны учитывать, что одно и то же основное содержание (или концепция объекта) может существовать во многих проявлениях на протяжении срока его существования. Некоторые из этих проявлений совместимы как цифровые объекты, в то время как иные могут следовать друг за другом в ряде отдельных или наслаивающихся поколений. Некоторые программы сохранения учитывают это, создавая запись единой версии, известной как Мастер сохранения, документирующей варианты и изменения как часть истории объекта. Другие программы создают запись для каждого проявления, требующего сохранения, отражая зависимости между проявлениями в записях метаданных.

Информация, заданная для метаданных сохранения часто разделяется на два класса (в соответствии со Справочной моделью Открытых архивных информационных систем, или OAIS, упомянутой в главе 8):

- *Контентная информация*, состоящая главным образом из подробностей о технической природе объекта и указывающая системе способ отображения данных в виде отдельных типов и форматов данных. По мере изменения технологий доступа, эти метаданные отображения также должны дорабатываться
- *Информация, описывающая сохранение*, состоящая из другой длительно необходимой для управления и использования объекта информации, включая идентификаторы и библиографические подробности, информацию о владельцах, правах, происхождении, истории, связях с другим объектами и методах контроля.

Очевидно, часть из этих метаданных может относиться к другим информационным объектам типа программных инструментов и инструкций по форматированию, которые также должны управляться. Взаимозависимая природа цифровых материалов означает, что программам зачастую приходится иметь дело с целым клубком взаимосвязанных объектов и их метаданных.

Все еще не существует никаких общепринятых стандартов для схем метаданных сохранения, так что программам, вероятно, придется выбирать между принятием (и, возможно, адаптацией) одной из используемых моделей, или разработкой собственной схемы (в окончательном виде или в виде промежуточного решения до разработки стандарта).

Ряд национальных архивных организаций используют описания метаданных для систем хранения записей, используемых для сохранения., Международная рабочая группа, со-званная OCLC и Исследовательской группой библиотек (RLG) в середине 2002 г. разрабо-тала рекомендации по структуре метаданных сохранения в библиотечной области (доступ-ны на адресе <http://www.oclc.org/research/pmwg/pm_framework.pdf>). Их отчет может стать отправной точкой при разработке необходимых метаданных.

Интересен подход Национальной библиотеки Новой Зеландии к использованию этих ре-комендаций для конкретной программы (доступно на адресе <http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf>). Эта схема предполагает сле-дующие элементы:

<u>Описание цифрового объекта</u>	Характеристики отдельных типов файлов
Название объекта	(для файлов изображений: разрешающая способность, размер, тональная разре-шающая способность, цветовая гамма, управление цветом, цветовая таблица, ориентация, сжатие)
Локальные идентификаторы	(для текстовых файлов: сжатие, набор сим-волов, DTD для структурного текста, структурные разделы)
Глобальный устойчивый идентификатор	(для аудио: разрешающая способность, продолжительность, скорость пересыл-ки информации в битах, сжатие, форми-рование пакета, номер и тип дорожки)
Место файла в системе памяти	(для видео: размер окна, продолжитель-ность, скорость пересылки кадров, сжа-тие, структура кодирования, звук)
Дата создания в качестве Мастера сохране-ния	(для наборов данных: все вышеуказанные элементы)
Оценка технического решения (кроме файлов стандарта MIME)	(для исполняемых файлов: все вышеука-занные элементы)
Структурный тип (текст, изображение...)	<u>Описание обновления метаданных</u>
Аппаратные средства, необходимые для работы объекта	Средства изменения метаданных
Программное обеспечение, необходимое для работы объекта	Когда изменялись
Специальные инструкции по установке	Область внесения изменений
Встроенные средства доступа	<u>Описание технических характеристик лю-бого файла в объекте</u>
Недостатки (встроенные аномалии)	Идентификаторы файла
Идентификация или ключи оценки досто-верности	Зависимость от других файлов объекта
Кто и когда проверял правильность созда-ния метаданных	Размер файла
<u>Описание любого действия с объектом (включая создание)</u>	
Название и цель действия	
Исполнитель	
Кто и когда разрешил действие	
Использованные аппаратные средства и программное обеспечение	
Этапы действия	
Результаты	
Использованные стандарты или специфи-кации	
Когда закончено действие	
<u>При создании типа/формата MIME (image/tif)</u>	
Версия	
Файл, обеспечивающий доступ	

ДЛЯ ПРОГРАММ СОХРАНЕНИЯ, РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ

14.21 Пересылка

Программы, работающие с небольшими ресурсами, возможно, должны искать способы снижения затрат на пересылку:

- Подход «пассивного» накопления может потребовать меньше усилий от програм-мы в области сохранения и возложить основные затраты на пересылку на источни-ка. Однако, без соглашений о носителях, форматах и контроле качества при пере-сылке материалов, тактический выигрыш может привести к большим стратегичес-ким потерям
- Обоснованные ограничения на диапазон носителей и форматов, принимаемых программой, может привести к некоторой экономии средств
- Программы могут сохранять перенесенный материал на тех же носителях, при ус-ловии их устойчивости и наличия резервных копий.

Некоторым сообществам, которым недоступны агентства по сохранению, вероятно, при-дется следовать «непересылочной» модели, устанавливая наилучшие возможные условия сохранения. Даже при этих обстоятельствах применяются многие обычные правила: обес-печить доступность более вероятно с некоторой внутренней пересылкой, даже в неболь-шой «архив копирования», где файлы могут управляться без обычных рисков использова-ния. Файлы все еще должны быть достаточно хорошо описаны и защищены, чтобы позже с ними смогла работать более безопасная программа сохранения.

14.22 Метаданные

Затраты на запись метаданных могут быть существенной частью от полных затрат на со-хранение. Может появиться возможность экономии за счет сокращения количества мини-мально зарегистрированной информации (созная, что доступ и сохранение будут за-труднены); или вложений в программное обеспечение, которое будет автоматически запи-сывать метаданные (этот процесс упрощается с развитием стандартов метаданных).

При выборе минимального набора метаданных, программные менеджеры могут рассмот-реть запросы пользователей по поиску материалов и какие запросы должны быть удовле-творены в процессе сохранения.

ОПЫТ ИСПОЛЬЗОВАНИЯ

14.23 Пример 1

Национальная библиотека, собирающая интерактивные цифровые издания, собирает файлы с сайтов издателей, используя программируемый поиск, копирование и загрузку программного обеспечения типа HTTrack, в соответствии с соглашениями, о которых до-говариваются с каждым владельцем сайта. Процесс накопления предусматривает привле-чение персонала для поиска потенциальных сайтов, которые могли бы отвечать рекомен-дациям библиотеки, решая, что должно быть сохранено и как долго сайт должен отслежи-ваться, (политика отбора предполагает, что должны быть сохранены связанные докумен-ты с того же сайта, но не другие ссылки). Когда файлы загружены программным обеспече-нием, необходимо убедиться, что весь нужный материал загружен и все файлы работают. Запись метаданных производится с использованием информации, сгенерированной про-

граммным обеспечением и введенной вручную. Индивидуальная страница создается для каждого заголовка, используя сгенерированный системой шаблон так, чтобы пользователи могли определить, что они получают и как это связано с веб-сайтом издателя и с другим материалом, находящимся в архиве. По окончании метаданные, которые включают ссылку на сохраненные объекты, записываются в архив метаданных, и объекты сохраняются в системе памяти массива.

14.24 Пример 2

Малый архив этнической музыки получает полевую запись от коллекционеров на ленте звукозаписи, дешевой и удобной для коллекции, но не подходящий для хранения. После проверки соответствия материала политике сбора архива и качества записи, материал принимается с вводом информации об источнике в отдельную базу данных вручную. Материал сохраняется и ему присваивается уникальный номер в коллекции. Данные на ленте копируются на два компакт-диска: один для хранения и один как резервная копия. Оригинальная лента звукозаписи хранится как копия для краткосрочного доступа, копии на компакт-диске хранятся отдельно. В метаданные дописывается местонахождение всех копий.

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

Связь с источниками см. также Работа с источниками: Глава 13

Метаданные и средства доступа см. также Обеспечение доступа: Глава 17

Другие ссылки (все ссылки были действительны в марте 2003 г.)

1. Пересылка

Служба данных искусствоведческой и гуманитарной тематики Великобритании (AHDS) и ее объединенный архив данных (в области литературы, археологии, искусства и истории) выпустили ряд отличных «руководств для вкладчиков». Они включают техническую информацию о предпочтительных форматах для широкого диапазона материалов и могут служить хорошими моделями для программ на основе данных. Среди них:

- History Data Service (nd). Guidelines for Depositors. <http://hds.essex.ac.uk/depguide.asp>
- Oxford Text Archive (1999). Depositing with the OTA: the Depositors Guidelines. http://ota.ahds.ac.uk/publications/ID_Depositing-Introduction.html
- Visual Arts Data Service (nd). Guidelines for Depositors. http://vads.ahds.ac.uk/depositing/depositor_guidelines.pdf

2. Устойчивая идентификация

- Corporation for National Research Initiatives (CNRI), (nd). The Handle System. <http://www.handle.net/index.html>
- Dack Diana (2001). Persistent Identification Systems (Report on a consultancy for the National Library of Australia). <http://www.nla.gov.au/initiatives/persistence/PIcontents.html>
- International DOI Foundation (nd). The Digital Object Identifier System. <http://www.doi.org/>

- Internet Engineering Task Force (IETF), (2001). Uniform Resource Names (URN). <http://www.ietf.org/html.charters/urn-charter.html>
- National Library of Australia (2001). Managing Web Resources for Persistent Access. <http://www.nla.gov.au/guidelines/2000/persistence.html>
- The PURL Team (nd). PURL – Persistent URL Homepage. <http://purl.oclc.org/>

2. Метаданные

Большое разнообразие стандартов и инициатив метаданных создано для различных областей наследия, с расширением или адаптацией для информации в цифровом виде. Для примеров, см.:

- Dublin Core Metadata Initiative. <http://dublincore.org/>
- IFLA Universal Bibliographic Control and International MARC Core Programme (UBCIM) (2000). *UNIMARC Guidelines no 6: Electronic Resources*. <http://ifla.org/VI/3/p1996-1/guid6.htm>
- International Council on Archives, (1999). *General International Standard Archival Description*, 2nd edition. http://www.ica.org/biblio/cds/isad_g_2e.pdf
- Consortium for the Computer Interchange of Museum Information, (1999). *CIMI Dublin Core Metadata Testbed Project*. http://www.cimi.org/old_site/documents/meta_webliography.html
- International Association of Sound and Audiovisual Archives, (1998). *The IASA Cataloguing Rules*. <http://www.iasa-web.org/icat/>

Некоторые источники метаданных сохранения:

- Colorado Digitization Project Metadata Workgroup, Audio Taskforce (2002). *Metadata for Digital Audio (draft)*. <http://coloradodigital.coalliance.org/digaudio meta.pdf>
- National Library of New Zealand, (2002). *Metadata standards framework – preservation metadata*. <http://www.natlib.govt.nz/files/4initiatives metaschema.pdf>
- NISO/AIIM, (2002). *Data dictionary – technical metadata for digital still images*, проект стандарта для NISO Z39.87 – 2002. http://www.niso.org/standards/resources/Z39_87 trial use.pdf
- *Preservation metadata and the OAIS Information Model: a metadata framework to support the preservation of digital objects: a report by the OCLC/RLG Working Group on Preservation Metadata*, (2002). <http://www.oclc.org/research/pmwg/pm framework.pdf>
- Public Record Office (UK) (nd). *PRONOM* (о системных базах данных, которые сохраняют и предоставляю информацию о форматах и прикладном программном обеспечении, необходимом для работы с ними.) <http://www.pro.gov.uk/about/preservation/digital/pronom.htm>
- The British Library, (nd). *Code of Practice for the Voluntary Deposit of Non-Print Publications*. <http://www.bl.uk/about/policies/codeprac.html>

Глава 15. Управление правами

ВВЕДЕНИЕ

15.1 Предостережение

Настоящие рекомендации не должны рассматриваться в качестве компетентных юридических советов по вопросам использования прав.

15.2 Цели

Глава предназначена для освещения ответственности программ в области сохранения в правовой сфере и формирования общих предложений о путях решения подобных проблем.

15.3 Краткое содержание главы

Существует ряд прав и ожиданий владельцев, о которых программы в области сохранения должны знать и, в случае необходимости, учитывать их при планировании управления. Многие из этих прав имеют юридический подтекст, включая права конфиденциальности и права интеллектуальной собственности. Поскольку программы сохранения должны копировать информацию в цифровом виде для сохранения, и большинство из них предназначены для некоторого уровня доступа, необходимо активное решение правовых вопросов.

ОБЩИЕ ВОПРОСЫ

15.4 Цифровое наследие и права

Материалы цифрового наследия являются предметом реализации ряда прав и ожиданий, некоторые из которых имеют юридическую силу. Многие из них, как, например, авторские права, проистекают из интеллектуальной собственности, которая вложена в материал. Однако могут существовать другие права и ожидания, что также должно быть принято во внимание.

15.5 Перечень прав и ожиданий

Перечень прав и ожиданий, с которыми могут столкнуться программы в области сохранения, включает:

- Права интеллектуальной собственности источников, включая авторское право, которое может существовать в различных видах, связанных с различными видами материала; право на установление условий доступа и использования; выяснение моральных прав создателя
- Законные права некоторых учреждений по сбору, сохранению и предоставлению доступа к некоторым материалам

- Права и ожидания в вопросах частной собственности, конфиденциальности и разрешения использования, связанные с некоторыми материалами, например, организационные отчеты, исторически значимые записи устной информации, персональные данные и частная переписка
- Ожидания пользователей в вопросах доступа и использования
- Ожидания широкого круга заинтересованных лиц, что значение материала наследия сохранится и будет доступным в рамках законодательно установленного правового режима.

15.6 Основные права, необходимые для сохранения

Сохранение включает ряд процессов, затрагивающих правовые вопросы. Чтобы достигнуть непрерывности цифрового наследия, программы сохранения должны:

- Получать и содержать материал, обычно используя копирование
- Производить дополнительные копии для сохранения
- В случае необходимости, обходить меры источников по ограничению доступа и запрету копирования
- Определить, какие материалы и какие виды материалов должны сохраняться
- Создавать в материалах метаданные
- Изменять структуру и имена файлов в случае необходимости
- Использовать любой способ для сохранения доступа
- Предоставлять управляемый доступ для зарегистрированных пользователей.

15.7 Вызовы и угрозы

Получение разрешений на подобные действия может быть затруднено:

- Источники и другие владельцы прав могут не давать разрешений
- Права доступа, частной собственности и конфиденциальности часто противоречат
- При отдельном или совместном создании цифровых материалов, может быть тяжело вести переговоры с владельцами всех прав или вообще найти их
- Юридическая позиция может быть неоднозначной, поскольку много вопросов юрисдикции еще в процессе определения правовых рамок и порядка обращения с ними
- При обращении с материалами в мировом масштабе может отсутствовать четкое понятие о том, чья юрисдикция применяется: места производства материала, его издания или сохранения, или проживания пользователя — все они могут иметь разную юрисдикцию.

Затраты на организацию управления правами могут быть высоки, особенно если требуются индивидуальные переговоры. С другой стороны, затраты, связанные с обращением материалов без наличия соответствующих прав, вероятно, будут высоки.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

15.8 Правосознание

Программы в области сохранения должны знать о правовых рамках, в которых они работают, включая юридические права, ограничения и обязательства. Это может потребовать обращения за юридической консультацией к компетентному источнику. Даже руководствуясь благородными принципами сохранения важных материалов наследия, программы в области сохранения ответственны за поиск способов решения своих задач, не нарушая законных прав других участников.

15.9 Защита прав

Программы сохранения должны определить степень своего участия в защите прав при представлении данных для сохранения широкого диапазона цифровых материалов в рамках законодательства.

Как минимум, программы сохранения должны гарантировать, что заинтересованные стороны знают о правах, необходимых для эффективного сохранения.

15.10 Поиск приемлемых решений

Решение правовых вопросов может быть непростым, хотя обычно проблемы преодолимы. Их решение действительно требует уважения законных интересов других. Решения могут обычно разрабатываться при совместном подходе, который выявляет взаимные потребности и выгоды. Программы сохранения могут сделать большой вклад, показав, что:

- Управление правами возможно
- Есть способы достижения целей сохранения без создания опасности для коммерческих интересов
- При помощи документации и метаданных, программы в области сохранения могут способствовать просвещению общества с соблюдением прав владельцев
- Выбирая материал для сохранения, программы сохранения могут подтвердить значение записей, результатов исследований и других неизданных материалов.

Ряд программ сохранения нашли способ удовлетворения правовых вопросов, часто во взаимодействии с правообладателями. Это организуется на основе как простых соглашений с отдельными правообладателями (общими для архивов данных и выборочных архивов веб-публикаций), так и долгосрочных соглашений с крупными коммерческими издателями и национальными библиотеками.

Все модели обычно основаны на комплексе прав переноса, управления и сохранения. Например, право хранить и сохранять материал может быть передано полностью, в то время как программа сохранения обязана управлять доступом, а источник сохраняет авторское право.

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ПРОБЛЕМЫ

15.11 Правовые рамки

Может существовать ряд положений, которые позволяют программам сохранения обрести право собирать и сохранять информацию в цифровом виде. Наиболее общий перечень включает:

- Юридическое право или законодательство, регламентирующее хранение или запись
- Организационные правила управления общей информацией
- Оговоренные требования по хранению данных
- Условия грантов, решений, занятости или членства организаций
- Права, унаследованные одной организацией от другой
- Предоставленные или приобретенные лицензионные соглашения
- Права, подразумевающие добровольное представление материала программам сохранения
- Некоторые агентства сохранения отбирают и сохраняют материалы, вроде общественно доступных по веб-сайтам, не уточняя наличия разрешений. Некоторые делают это на основании «общего использования» материала в общественном домене; другие полагаются на «отказ от участия», из-за чего владельцы прав часто возражают.

Программы сохранения ответственны за определение, на основе компетентного юридического заключения, является ли любой из этих или других подходов применимым, и что требуется для адекватной юридической защиты прав.

15.12 Некоторые общие подходы

Каждая ситуация требует отдельного рассмотрения, но программы сохранения должны учитывать некоторые общие подходы, в том числе:

- Определение правовой ситуации, ее точное установление в соответствии с законодательством, существующими административными правилами или лицензионными соглашениями
- Определение прав, которые будут необходимы для сохранения
- Определение владельцев прав и других депозитариев для решения правовых вопросов
- Определение того, что необходимо и как оно будет управляться
- Поиск правообладателей и определение взаимоприемлемого правового режима
- Запись правовых вопросов в метаданных, которые надежно связаны с соответствующими материалами
- Обеспечение понимания прав персоналом
- Установка систем, процедур и инструментов безопасного управления доступом и копированием, при осуществлении контроля за процессом
- В случае необходимости, ограничение хранения от других видов доступа и использования
- Доведение прав и обязательств до пользователя
- Регулярная оценка систем и процедур с точки зрения их адекватности

- Контроль вопросов изменения прав, к примеру, прохождение установленного периода времени.

15.13 Ведение переговоров об условиях доступа

Уровень доступа, которого должны добиваться программы в области сохранения, будет зависеть от их задачи: некоторые материалы цифрового наследия могут иметь ограниченный доступ по причинам секретности, защиты или другим причинам, тогда как разумно ожидать, что изданные материалы будут предоставлены для долгосрочного доступа через управляемую программу сохранения.

Возможные предложения, которые могут быть сделаны при ведении переговоров об условиях доступа, включают:

- Географические ограничения, в виде ограничения доступа отдельным пользователям
- Ограничения на возможность копирования, к примеру, использование автономного компьютера без доступа к внешним сетям или дисковым
- Ограничение числа пользователей, которые могут обращаться к материалу в любое время
- Временные рамки, позволяющие неограниченный доступ после разумного периода коммерческой эксплуатации
- Взаимно принятые условия передачи прав доступа, к примеру, когда материал больше не находится на сайте издателя
- Разрешение доступа только зарегистрированным пользователям, которые обязаны отвечать точно установленным условиям.

15.14 Управление правовыми вопросами

По мере решения правовых вопросов, необходимо обеспечить их управляемость, как основу ответственности программ в области сохранения.

- Программы в области сохранения могут ожидать больших массивов материалов, поэтому использование стандартных лицензионных соглашений, охватывающих целые виды материала, поможет избежать необходимости ведения отдельных переговоров и управления разными правами
- Существуют и продолжают развиваться служебные программы, управляющие правами. Такие инструменты записывают условия доступа к отдельным элементам, делают записи и запросы о возможности использования и формируют отчеты. При выборе инструментов управления правами важно решить, какие из них обеспечивают уравновешенный подход к управлению правами
- Следует упростить для пользователей контакт с владельцами прав, чтобы договориться о собственных разрешениях, к примеру, права копирования под ответственность пользователя
- Создание разрешенного доступа простым насколько возможно может стать препятствием для неуполномоченного доступа и использования
- Поощрение источников к использованию открытого исходного программного обеспечения должно помочь снизить осложнения и затраты, возникающие при ведении переговоров о правах с частными программными разработчиками.

ДЛЯ ПРОГРАММ СОХРАНЕНИЯ, РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ

15.15 Поиск полезных решений

Вопросы прав могут резко увеличить затраты программ в области сохранения, поэтому все программы заинтересованы в поиске полезных решений и уходе от тяжб. Программы, работающие с небольшими ресурсами, могут быть особенно заинтересованы в наличии стандартных соглашений, которые снижают затраты на ведение переговоров о правах. Им, вероятно, также придется уяснить, что управление правами является фактором ограничения размаха деятельности.

С другой стороны, они, возможно, должны ограничить свою активность материалами, для которых правовой режим уже облегчен, например, потому что:

- Они уже имеют разрешение
- Права истекли (хотя вряд ли для цифровых материалов пока применим срок в несколько десятилетий)
- Источник заинтересован в программе сохранения
- Существует достоверное юридическое уведомление, что «справедливое использование» или другие условия будут успешно защищены.

ОПЫТ ИСПОЛЬЗОВАНИЯ

15.16 Пример 1

Архив данных, работающий по научной дисциплине, использует стандартное соглашение для вкладчиков, разрешающее архиву делать копию данных и предпринять любое необходимое действие по сохранению, включая создание копий в любых форматах, которые необходимы для обеспечения надежного доступа. Вкладчики должны указать, есть ли ограничения объемного, пользовательского или временного характера. Максимальный период для закрытия доступа — 10 лет. Архив управляет правами в отдельном режиме, поскольку данные не доступны в интерактивном режиме: пользовательские запросы сверяются с записями метаданных прежде, чем позволяется доступ.

15.17 Пример 2

Государственная библиотека работает на основании законодательства в области сохранения, которое позволяет производить и сохранять копии для хранения. Условия авторского права все еще действуют, так что библиотека сообщает пользователям о необходимости получить разрешение от владельца авторского права перед созданием копий. Библиотека договаривается об ограничениях доступа с владельцами коммерческих изданий, чтобы соблюсти их коммерческие интересы в течение согласованного периода времени, обычно составляющего 5 лет, в течение которого допускается только локальное использование. Ряд владельцев снижают ограничения доступа, поскольку это расширяет аудиторию, в то время как некоторые требуют более длительного периода ограничений. Система метаданных управления правами используется для записи ограничений и автоматического исполнения или отклонения запросов.

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

Работа с создателями см. также главу 13

Метаданные см. также главу 14

Другие ссылки (все ссылки были действительны в марте 2003 г.)

Ряд архивов данных используют стандартные лицензионные соглашения с вкладчиками для формализации передачи прав. Например, см.:

- Oxford Text Archive, (2003). Licence for depositors. <http://ota.ahds.ac.uk/>, издан «OTA Publications»

В качестве примеров переговоров по правовым вопросам с положительным исходом см.:

- IFLA and the International Publishers Association, (June 2002). Preserving the Memory of the World in Perpetuity: a Joint Statement on the Archiving and Preserving of Digital Information. <http://www.ifla.org/V/press/ifla-ipa02.htm>
- Koninklijke Bibliotheek (August 2002). National Library of the Netherlands and Elsevier Science Make Digital Preservation History. <http://www.kb.nl/kb/resources/frame-setkb.html?/kb/ict/dea/ltp/ltp-en.html>

Некоторые другие ресурсы:

- CEDARS Project (2002). CEDARS Guide to Intellectual Property Rights. <http://www.leeds.ac.uk/cedars/guideto/ipr/guidetoipr.pdf>
- Kavcic-Colic, Alenka (2002). Archiving the Web: Some Legal Aspects, 68th IFLA Council and General Conference, Glasgow. <http://www.ifla.org/IV/ifla68/papers/116163e.pdf>

Глава 16. Защита данных

ВВЕДЕНИЕ

16.1 Цели

Из данной главы менеджер программы должен уяснить важность обеспечения сохранности целостных данных, лежащих в основе цифровых объектов. Лица, занятые в данной области должны использовать материалы главы как основу для обсуждения специальных требований со специалистами в области компьютерных технологий или поставщиками услуг.

16.2 Краткое содержание главы

Защита данных является основным вопросом всех программ в области сохранения. Для многих программ немаловажной является также проблема достоверности данных. Это касается целостности данных на всем протяжении обработки, а также их идентификации. Стратегия защиты данных включает проработку вопросов ответственности, технической инфраструктуры, обеспечения, пересылки данных, соответствующих способов хранения на носителях, резервного копирования, системной защиты и планирования действий в условиях чрезвычайных ситуаций. Достоверность также опирается на исходную документацию об оригинале и хронологию цифровых материалов.

КЛЮЧЕВЫЕ ПРОБЛЕМЫ ГЛАВЫ

16.3 Сохранение и защита данных

Данные должны быть сохранены. Несмотря на то, что основной упор в процессе хранения направлен на представление цифровых объектов в максимальном приближении к первоначальному виду, нельзя забывать о том, что цифровые объекты состоят из данных. Именно данные являются объектом сохранения, управления и защиты с тем, чтобы любой цифровой объект мог быть представлен пользователю.

16.4 Достоверность

Часто материалы наследия оцениваются, по крайней мере, частично, по степени их достоверности, то есть степени, с которой пользователь может быть уверен в их подлинности. Для архивных отчетов, научных данных и многих других видов цифровых материалов уверенность в их достоверности является основным показателем, без чего они фактически не имеют никакого значения.

Достоверность проистекает как из уверенности в тождественности объекта — то есть того, что он является именно тем, о чем говорится в описании и не спутан с каким-либо другим объектом — так и из уверенности в целостности объекта — то есть уверенности в том, что объект не был изменен каким-либо способом искажающим его содержание.

Обеспечение тождественности и целостности подразумевает устойчивые и зарегистрированные связи и взаимные ссылки между объектом как в первоначальном его виде, так и в виде, представленном в настоящее время.

Оценка, поддержание и обеспечение доказательств подлинности являются ключевыми задачами большинства программ сохранения.

16.5 Угрозы достоверности

Достоверность может быть подвергнута сомнению в следующих случаях:

- Угроза тождественности. Потеря уверенности в том, что объект отличен от других объектов, снижает степень достоверности. К этому могут привести ошибки идентификации данных, замена идентификаторов или сбой в отображении связей между различными версиями или копиями
- Угроза целостности. Изменения содержания непосредственно объекта также могут привести к снижению достоверности. Большинство таких изменений несет угрозу содержанию объекта на уровне данных.
- Как природа цифровых материалов, так и вопросы управления ими с целью обеспечения сохранности и доступности, несут в себе следующие проблемы:
- Информация в цифровом виде легко может быть изменена, с мошенническими намерениями или без таковых, или даже вообще без какого бы то ни было на то желания
- Внесенные изменения могут быть незаметны
- Процессы сохранения почти всегда содержат внесение изменений, связанных с пересылкой данных от системы к системе, от носителя к носителю, наращиванием или модификацией метаданных, созданием новых копий, для которых необходимы новые имена файлов, изменением способа отображения с заменой его технологии и так далее.

16.6 Угрозы целостности данных

Общие угрозы целостности данных, с которыми могут столкнуться программы сохранения, включают:

- «Естественные» ошибки, которые возникают в цифровых накопителях
- Сбой носителей. Большинство носителей имеют достаточно короткий срок службы до снижения надежности хранения данных
- Атаки системных хакеров и вирусов, которые могут исходить от персонала или внешних злоумышленников
- Ущерб, нанесенный в результате террористических атак, боевых действий или общественных беспорядков зданиям или источникам электропитания
- Небрежные действия персонала или посетителей (выключение энергии, случайное уничтожение дисков или лент, переформатирование запоминающих устройств и т.д.)
- Природные катастрофы (пожары, наводнения или разрушение зданий)
- Банкротство.

Риск этих и других угроз изменяется в зависимости от конкретной ситуации, однако их вероятность непременно должна быть учтена.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

16.7 Какова должна быть степень достоверности?

Поскольку информация в цифровом виде является описанием реально существующего объекта, ее достоверность приобретает чрезвычайную важность. Не вся информация в цифровом виде может быть описанием: например, она может содержать творческую идею, дебаты по некоторой проблеме или вообще содержать информацию развлекательного характера. Даже для таких материалов может существовать проблема достоверности, поскольку должна быть защищена целостность самого продукта или содержащихся в нем идей.

В конечном счете, программы сохранения должны определить, какова должна быть степень достоверности для обеспечения подлинности обрабатываемых ими материалов, имея в виду, что тождество объектов и целостность данных являются фундаментальными принципами.

16.8 Значение защиты данных

Защита данных должна играть ключевую роль в любой программе сохранения по двум причинам:

- Возможность обеспечения доступа пользователя к цифровому объекту. Это фундаментальное требование: если данные потеряны или серьезно нарушены, сам доступ к цифровому объекту может быть невозможен, и сам процесс его хранения станет бессмысленным
- Целостность данных должна поддерживаться без вмешательства или внесения искажений с тем, чтобы пользователи могли доверять подлинности отображаемого объекта.

16.9 Значение документации

Документация также играет ключевую роль по двум причинам:

Поясняя связи между объектами и описывая различия между ними, она является доказательством тождества

Показывая, какие изменения, если таковые вообще имеются, имели место, с чьей властью и с каким результатом они были внесены, она обеспечивает наличие ссылок для проверки достоверности.

16.10 Обязанности по поддержанию достоверности

Практически невозможно рассчитывать на наличие полностью объективных гарантий достоверности — в них всегда может быть привнесен элемент общего или субъективного суждения при оценке достоверности — однако, как кажется, разумно возложить на программы сохранения цифрового наследия три обязанности:

- Они должны оценить, является ли демонстрируемая достоверность достаточной по отношению к текущей значимости материала
- Они должны предохранить вверенный их попечению материал от изменений в его содержании. (Это допускает внесение внешних изменений типа новых интерпретаций, не позволяя внутренние изменения, которые изменили бы содержание исходного материала)

- Они должны отражать все связи, на которые распространяется заданный уровень достоверности. Это включает связи между объектом и его идентификаторами; между объектом и его автором; между различными объектами; между объектом и системой управления.

16.11 Стратегии защиты данных

Возможно, с материалами наследия в каком-либо ином виде и поступали небрежно, однако после перевода их в цифровую форму такого случиться не должно. Работа с цифровыми объектами требует наличия хорошо спланированных, обеспеченных и долгосрочных стратегий защиты данных как необходимого условия преемственности. При этом они должны предусматривать:

- Четко очерченный круг ответственности
- Наличие соответствующей технической инфраструктуры, включая системы памяти
- Обеспечение наличия приемлемых устройств и накопителей
- Обслуживание, поддержку и своевременную замену программ
- Регулярную перезапись данных на новые носители для снижения зависимости от замены аппаратных средств доступа
- Наличие необходимого объема памяти и средств обработки информации на носителях
- Высокий уровень дублирования как страховку от отказа узлов или носителей; наличие соответствующих резервных режимов
- Высокий уровень системной защиты, в том числе средств управления доступа к данным
- Планирование защиты от стихийных бедствий.

Более подробно эти вопросы рассмотрены в следующем разделе.

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ВОПРОСЫ

16.12 Использование услуг провайдеров

Из всех задач программ сохранения, накопление и защита данных являются наиболее приемлемыми и привлекательными для сторонних поставщиков услуг из-за значительных инвестиций в закупки оборудования и производство экспертиз. Однако важность процесса сохранения данных предъявляет требования к надежности решения этих задач с точки зрения безопасности и управляемости процессом.

16.13 Практические аспекты стратегий защиты данных

Существует стандартный набор стратегий управления данными при длительном хранении. Большинство из них основаны на том, что в хранении данных нуждается информация как таковая, а не ее носители.

- Распределение ответственности. Процессы хранения и защиты данных требуют наличия конкретных ответственных лиц. В их круг обязанностей должна быть

включена техническая ответственность, требующая определенного набора знаний и навыков, а также практики управления. За исключением небольших массивов, хранение и защита данных требует наличия специализированных ресурсов, работающих по соответствующей схеме и ответственных за данные вопросы

- Необходимая техническая инфраструктура. Хранение данных и управление ими должно осуществляться специализированными системами и на специализированных носителях. Существуют системы управления цифровыми массивами или системы хранения цифровых объектов, отвечающие требованиям программ сохранения. По мере определения требований их необходимо обсуждать с возможными поставщиками. Различные системы и носители созданы для удовлетворения различных потребностей, и их отбор должен делаться на основании соответствия требованиям программ сохранения. В целом система должна обладать всеми необходимыми возможностями, в том числе:

— Достаточный объем хранения. При этом он может наращиваться через какое-то время, однако система должна быть способна хранить необходимый объем данных на протяжении всего срока службы

— Главной является способность системы дублировать, а также переносить данные на новые или «обновленные» носители без потерь информации

— Наличие возможности и технического обеспечения для надежного и быстрого решения возникающих проблем

— Способность отображать имена файлов в схеме, соответствующей архитектуре памяти. Системы памяти строятся из поименованных объектов. Различные системы используют различную архитектуру организации объектов. Это может оказать влияние на способ отображения и запоминания имен объектов, хранящихся в памяти; например, дисковые системы могут использовать иерархическую структуру каталога имен файлов, отличную от используемой ленточными системами. Система должна позволять, а лучше обеспечивать самостоятельно, отображать имена файлов, существующих в системе, совместно с идентификаторами

— Способность управлять избыточной памятью

— Возможность проверки ошибок. Обычно запоминающие устройства позволяют проводить автоматизированную проверку на наличие ошибок. Так как материалы наследия должны храниться долгое время, часто не будучи востребованными продолжительное время, система должна быть способна обнаружить изменения или потерю данных и предпринять соответствующие меры

- Техническая инфраструктура должна также включать способы сохранения метаданных и обеспечения надежной связи метаданных с цифровыми объектами. Масштабная обработка данных зачастую выявляет необходимость установки систем управления цифровыми объектами, связанных с системами хранения цифровых массивов. При этом такие системы должны быть отдельными, чтобы справляться с потоком решаемых задач и позволять замену метаданных и интерфейса без замены сохраняемых массивов.

Ниже, в Таблице 16-1, приведен перечень и возможности носителей для хранения данных:

Носитель	Доступ к данным	Возможность внесения изменений	Возможность хранения информации	Возможности по наращиванию емкости	Срок жизни носителя	Примечание
Магнитный (или жесткий) диск	быстрый	есть	до 200 гигабайт	в 2 раза каждые 12-18 месяцев	около 5 лет	в основном хранятся стационарно
Магнитная лента	линейная запись замедляет процесс поиска и доступа	в основном отсутствует, процесс требует полной перезаписи информации	до 200 гигабайт	в 2 раза каждые 12-18 месяцев	около 5 лет	нестационарное хранение, удобны для копирования
Оптический (CD,DVD)	диск быстрый, но медленнее, чем у магнитных дисков	в основном есть	до 4 гигабайт	низкие из-за неприменимости для хранения больших массивов	в диапазоне от 5 до десяти лет в зависимости от качества	нестационарное хранение, дешевые носители, доступное оборудование

Таблица 16-1 Сравнительные свойства носителей для хранения данных

- **Обслуживание, поддержка и замена программ.** Системные компоненты должны меняться раз в несколько лет. Аппаратные средства в большинстве своем имеют срок жизни продолжительностью пять лет до прихода в неремонтопригодное состояние. Носители также нуждаются в регулярном *обновлении* (перезаписи данных) и периодической замене на новые носители.
- Потребность замены систем памяти влечет за собой существенные затраты как на приобретение оборудования, так и на процесс промежуточного сохранения данных при установке нового оборудования. Эти затраты должны быть предусмотрены при долгосрочном планировании бюджета.
- Наряду со стоимостью замены носителей необходимо учесть и возможности по увеличению объема памяти. К сожалению, любое сохранение обычно ведет к росту количества сохраняемых данных.
- Рынок систем сохранения и управления данными развивается параллельно программам сохранения, поэтому на нем представлено достаточное количество раскрученных продуктов. Их использование обеспечивает возможность легко управлять данными, снизить уровень рисков и существенно снизить стоимость технического обеспечения. Кроме того, такой подход приводит к росту конкурентоспособности рынка технической поддержки и обновлений. В этом сегменте рынка достаточно широко используются стандарты, что позволяет совместно использовать различные программы при широком выборе способов обновления и замены.
- **Регулярная перезапись данных.** Системы хранения строятся на принципе безопасного и полного дублирования данных, вместо обеспечения безопасности носителей. Данные должны быть скопированы на другие носители, чтобы избежать опасности их потери из-за поломки носителя. Если новые виды носителей покажут бо-

лее высокую эффективность в системах хранения, данные должны быть перезаписаны на новые носители. Это должно случиться до того, как поисковые аппаратные средства или программное обеспечение потребуют ремонта или замены.

Планирование пересылки данных является ключевым вопросом управления безотносительно места использования системы. Так, при использовании малых архивов на компакт-дисках необходимо следить за сроками их использования и состоянием, а также учитывать изменения в технологии компакт-дисков. В более сложных системах вообще решения о регулярной пересылке данных на новые носители автоматизированы, хотя при этом персонал все равно должен самостоятельно принимать решения о замене носителей и технологий.

- **Обеспечение хранения и условий работы носителей.** Условия хранения цифровых носителей не должны снижать сроки их использования.

Основной опасностью для носителей являются несоблюдение режима температуры и влажности; пыль или механическое засорение, которое может затруднить доступ к данным; в случае оптически кодированного материала – свет может повредить данные. При этом современные ленточные носители более устойчивы к магнитным полям.

Магнитные ленты-носители могут быть интегрированы в систему цифрового накопителя. Оптимальными условиями при их хранении вместе с компьютерными системами является чистое помещение с постоянной температурой 18° С и относительной влажностью 40 %, непрерывным притоком чистого, без пыли, воздуха с ежедневной уборкой. Граничные перепады должны не превышать 2° С и 10 % влажности.

Максимальный срок службы магнитных лент-носителей (используемых не совместно с компьютерами) обеспечивается при более строгих условиях хранения (температура от 18 до 10°С ±1°С, и относительная влажность от 30 до 40 % ±3 %). Подобные условия, без доступа света из-за светочувствительности носителя, должны обеспечиваться и для оптических носителей (CD, CD-RW и т.д.).

Существует мнение, что низкие температуры (около 0°С или ниже), могут снижать срок службы некоторых носителей, однако это пока не доказано.

- **Режимы резервирования и дублирования.** Значение резервирования и дублирования трудно переоценить: для всех программ сохранения это основной способ страховки от повреждения или потери любой единственной копии.

Действительно, хранение нескольких копий предполагает некоторую гарантию защиты от сбоя, однако для программ сохранения также необходимо учесть угрозу повреждения всех копий, сохраненных на одном сайте. Поэтому основным требованием становится сохранение копий на разных сайтах; чтобы предупредить возможность масштабных бедствий – наводнений, землетрясений, пожаров и войн, программы должны учесть необходимость хранения дополнительных резервных копий важных данных на несвязанных адресах.

Программы в области сохранения также должны корректировать графики резервирования с тем, чтобы сохраняемые данные обновлялись (то есть перезаписывались) не поверх новых данных.

- **Системная защита.** Средства защиты должны следить за тем, чтобы сохраняемые данные обрабатывались только управляемыми и разрешенными процессами. Стандартные меры защиты вычислительной техники и информационных ресурсов полностью применимы и востребованы.

Обеспечение защиты от чрезвычайных ситуаций и стихийных бедствий. Стандарт-

ные схемы действий в условиях чрезвычайных ситуаций и стихийных бедствий должны быть в наличии и регулярно проверяться. Они могут включать планы восстановления данных с поврежденных носителей, однако этот процесс представляется дорогим и сомнительным, поэтому лучшим вариантом является способ восстановления с дубликатных записей.

16.12 Управление рисками

Таблица 16-2 представляет общий анализ некоторых угроз процессу сохранения данных.

Угроза	Объект угрозы	Вероятность события	Скорость нарастания угрозы	Результаты	Способ предупреждения
«Естественные» ошибки	целостность данных	почти 100 %	постепенная	сбои в работе; данные могут не восстанавливаться	проверка наличия и исправление ошибок, обновление и перенос данных
Разрушение носителя	целостность данных	почти 100 %	постепенная	тяжелые; невозможность чтения и восстановления данных	использование качественных продуктов и носителей; частая проверка условий хранения; перенос данных до истечения срока службы носителей
Целевые атаки: хакеры, вирусы, преступники	целостность данных, целостность файлов	почти 100 % для сетевых архивов	обычно внешняя	обычно тяжелые; могут вести к подмене или исчезновению данных	решительные физические и программные меры; защитные системы, контроль доступа; автономная работа с данными
Иные, не целенаправленные атаки	целостность данных, целостность файлов, повреждение оборудования	низкая, в зависимости от ситуации	обычно внешняя	обычно тяжелые, непредусмотренные обычными средствами защиты	дублирование данных; ограничение доступа к дубликатам
Небрежность — отключение энергии, выемка носителей переформатированием	целостность данных, целостность файлов	непредвиденная	обычно внешняя	от незначительных до катастрофических	дублирование данных; подготовка персонала и физическая защита
Стихийные бедствия — пожар, наводнение	целостность данных, целостность файлов, повреждение оборудования	обычно известно заблаговременно	обычно внешняя, однако может предваряться предупреждением	от минимальной до полной потери данных	подготовка к возможным стихийным бедствиям; учет при выборе зданий и сооружений; дублирование данных с расположением на других территориях
Банкротства и финансовые неурядицы	доступ к данным	низкая	может быть постепенной или внезапной	обычно отмена доступа	носители утилизируются или используются в других целях планирование и управление бизнесом; договоренности по вопросам правопередачи; четкая их идентификация

Таблица 16-2 Общий анализ угроз процессу хранения данных

ДЛЯ ПРОГРАММ СОХРАНЕНИЯ, РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ

16.13 Основные вопросы

Защита данных является настолько важной задачей, что даже программы в области сохранения, работающие с небольшими ресурсами должны уделять ей большое внимание. Общий анализ угроз процессу хранения данных, проведенный выше, позволяет определить возможности по снижению их уровня для некоторых программ. Также можно предположить, что некоторые угрозы можно снизить за счет сокращения объема или скорости доступа. Это может найти применение в некоторых массивах.

16.14 Установка приоритета

Одним из путей повышения уровня защищенности является градуирование части массивов по приоритетам для введения дополнительной защиты, со снижением ее уровня (менее частое резервное копирование и пересылка данных на новые носители, использование носителей более низкого качества) для данных с меньшим приоритетом важности.

ОПЫТ ПРИМЕНЕНИЯ

16.15 Пример 1

Типичным решением является сохранение данных на магнитной ленте в трех копиях: одна — общедоступная — в библиотеке магнитных лент системы, одна хранится отдельно, но в пределах основного хранилища, и одна — за его пределами. Все копии в системе идентичны. Для постоянной работы создана временная копия на диске, находящемся в РФНД (резервный фонд оборотных дисков), в котором отказ одного диска компенсируется копиями на других дисках. Резервирование носителей может также достигаться отдельными копиями, сохраняемые на оптических носителях за пределами основного хранилища.

16.16 Пример 2

Документирование архива правительственных деловых транзакций включает в себя большой объем работы по удостоверению подлинности каждой сохраняемой записи. Все записи, сделанные при сканировании нецифровых оригиналов содержат подписанное удостоверение подлинности копий; цифровые отчеты, записанные с электронных носителей, содержат сгенерированные системой подтверждающие знаки. Все процессы, которые могли вызвать непреднамеренные или неразрешенные изменения, зарегистрированы в метаданных, приложенных к записи.

В региональных библиотеках цифровых изданий используются системы проверки и контроля качества, что позволяет гарантировать соответствие файлов копиям, остающимся на сайте издателя. Они документируют все случаи обращений к материалу и предотвращают возникновение большей части угроз целостности данных. При этом такие системы учитывают возможность того, что некоторые процессы могут привести к созданию новых объектов, которые в будущем будут отличаться от их первоначального вида. Конечно, невозможно полностью удостовериться в подлинности копий, однако наличие таких систем позволяет рассматривать библиотечные копии как аутентичные объекты для исследовательских целей.

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

- *Связь между данными и цифровыми объектами* см. также в главе 7 Понятие о сохранении цифровой информации

Другие источники (все ссылки были доступны в марте 2003 г.)

Некоторые заслуживающие внимания взгляды по вопросу подлинности могут быть найдены в:

- Gladney Henry M, *Digital Document Quarterly*. <http://home.pacbell.net/hgladney/ddq.htm>
- Graham Peter S, (2000 г.). *Authenticity in a Digital Environment*, Совет библиотечных и информационных ресурсов. <http://www.clir.org/pubs/reports/graham/intpres.html>
- InterPARES Project (2002). *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*. <http://www.interpares.org/book/index.htm>

Техническая информация об аппаратных устройствах, применяемых для сохранения данных, может быть найдена по адресам:

- Bogart, John Van. (1995). *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives* Совет библиотечных и информационных ресурсов, Вашингтон, О.К. <http://www.clir.org/pubs/reports/pub54/index.html>
- *CoOL [Conservation OnLine]: electronic storage media*. <http://palimpsest.stanford.edu/bytopic/electronic-records/electronic-storage-media/>
- Kodak Professional (nd). *Permanence and Handling of CDs*. <http://kodak.com/global/en/professional/products/storage/pcd/techInfo/permanence.jhtml>
- Библиотека конгресса (2002 г.). *Cylinder, Disc and Tape Care in a Nutshell*. <http://www.loc.gov/preserv/care/record.html>

Глава 17. Обеспечение доступа

ВВЕДЕНИЕ

17.1 Цели

Глава предназначена для рассмотрения общих и частных вопросов обеспечения доступа, а также рассматривает ряд используемых при этом подходов.

17.2 Краткое содержание главы

Замена программного обеспечения и оборудования часто приводит к потере доступа к материалам цифрового наследия. Скорее всего, это станет основной проблемой для большинства программ сохранения. Понимая зависимость между цифровыми объектами и способом доступа к ним, и принимая во внимание способ их отображения пользователю при обеспечении доступа, менеджеры программы должны обеспечивать принятие таких стратегий, которые будут обеспечивать доступ в любой момент времени. Процесс разработки этих стратегий ведется постоянно, в соответствии с требованиями времени и запросами пользователей. Рассматриваемые стратегии, сгруппированные по признаку ресурсоемкости этапов обработки цифровых материалов, обладают кратко-, средне- и долгосрочной эффективностью. Кроме того, рассмотрены некоторые альтернативные, «нецифровые» и «не сохраняющие» стратегии.

ОБЩИЕ ВОПРОСЫ

17.3 Необходимость наличия путей доступа

Сохранение возможности доступа к цифровым материалам — одна из основных целей программ сохранения. Основываясь на сохранении данных и метаданных, и используя программные и инструментальные средства доступа, цифровые объекты должны отображаться в понятной пользователям форме. Это должно обеспечиваться при первой необходимости, с использованием существующих на момент обращения технологий доступа.

Поскольку цифровые объекты строятся на различных технологиях отображения, возможность доступа к ним меняется в зависимости от замены технологий. Явление изменения технологий доступа становится настолько общим, что оно превратилось чуть ли не в основную характеристику сохраненных цифровых материалов.

17.4 Время хранения

Скорость смены технологий в настоящее время повышает сам риск потери существующей в цифровом виде информации. Некоторые материалы создавались на технологиях десятилетней давности и не могут использоваться с применением современных технологий.

В то время, как основная цель состоит в том, чтобы найти способ обеспечить доступ в долгосрочной перспективе, существует потребность в обеспечении доступа даже в ближайшем будущем.

17.5 Определение приемлемых потерь

Программы сохранения, вероятно, будут нуждаться в подходах к определению приемлемых и недопустимых уровней потерь.

Полная точность первоначальной индикации цифровых материалов в любом случае затруднена; многие современные стратегии могут приводить к потерям, включая возможную потерю содержания, изначального восприятия или некоторых функций оригинала.

Эти потери могут быть непреднамеренными побочными продуктами выбранной стратегии (обычными при преобразовании файлов в новые форматы), или результатом упрощения процессов, принятых с целью удешевления процессов сохранения (например, уничтожение ссылок или динамических элементов веб-страниц). Они могут даже быть встроенны в программы сохранения (например, удаление редакторских функций из статических отчетов).

При этих и подобных подходах для программ может потребоваться разработка способа определения приемлемых потерь.

ОСНОВОПОЛАГАЮЩИЕ ПРИНЦИПЫ

17.6 Ответственность программ сохранения

Программы сохранения должны найти пути по преодолению угрозы изменения и устаревания технологий в целях сохранения постоянного доступа.

17.7 Выбор сохраняемых версий

Ряд массивов содержит несколько версий одних и тех же материалов, например цифровых изображений высокого и более низкого качества, вторичные версии для обычного сетевого доступа. Программы сохранения должны определить, какая версия (или версии) должны оставаться в их ведении и какие могут быть сгенерированы позже.

17.8 Выбор обслуживаемых элементов

Для определения как приемлемых, так и недопустимых уровней потерь, программы в области сохранения должны определить основные элементы сохранения. Как уже говорилось ранее (см. главу 12), это необходимо для:

- Выбора необходимых стратегий обеспечения
- Выбора самой рентабельной стратегии
- Оценки успешности стратегий.

Определение объектов сохранения на этом уровне требует внимательного изучения материала с тем, чтобы уяснить принципы его существования, работы и необходимого уровня и качества отображения копии для пользователя

По мере определения необходимых элементов программы в области сохранения должны определять необходимый набор данных, программного обеспечения и оборудования, которые обеспечат необходимый уровень их отображения.

17.9 Связь между данными и программным обеспечением

Существует четкая связь между данными и программным обеспечением: все данные требуют наличия программного обеспечения, чтобы быть отображенными в понятной пользователю форме. Важное значение имеет степень связи:

- Некоторые объекты относительно независимы от специального программного обеспечения; так, базовые данные, простые или размеченные тексты типа ASCII могут отображаться с использованием основного набора программных средств

- Некоторые объекты зависят от более специального, но универсального или широко распространенного программного обеспечения; так, HTML, стандартные форматы отображения типа TIFF и другие форматы разработаны с учетом взаимозаменяемости платформ

- Некоторые объекты зависят от специального прикладного программного обеспечения и не предназначены для работы в других средах — несмотря на то, что изготовители зачастую обеспечивают их инструментами для чтения или использования в более широком формате; например, текстовые редакторы, электронные таблицы, некоторые базы данных, средства рисования и географические информационные системы

- Некоторые объекты сами являются программами; например — самораспаковывающиеся или исполнительные файлы и программы

- Ряд комплексных материалов содержит комбинации объектов с различным уровнем зависимости от программного обеспечения.

Степень зависимости от программного обеспечения может ограничить приемлемый диапазон стратегий. Например, объекты, содержащие информацию или документы, могут эффективно отображаться многими видами программного обеспечения, в то время как «программные» объекты могут иметь гораздо меньший выбор при отсутствии начальных сред.

17.10 Выбор стратегий

Пока еще не существует универсально применимого и практического решения проблемы технологического устаревания для цифровых материалов. Было предложено несколько подходов, однако маловероятно, что будет найдено единственное решение, которое предложит рентабельный способ обеспечения постоянного доступа для всех материалов и объектов. В этой ситуации, разумно применять стратегии, учитывающие возможность работы с широким ассортиментом материалов на протяжении длительного периода времени.

Сейчас важно предпринимать активные шаги, пусть даже небольшие, которые обеспечат сохранение доступа в обозримом будущем, при этом продолжая планировать долгосрочные перспективы, как наиболее приемлемые.

Нынешними прообразами долгосрочных стратегий являются: использование *стандартов* для кодирования, структурирования и описания данных, что может позволить им оставаться распознаваемыми на протяжении длительного периода времени; *эмуляция* устаревшего программного обеспечения или аппаратных средств в новой среде; перевод данных на новые технологии. Весь этот перечень подходов удовлетворял требованиям в течение ограниченного периода времени. Конечно, они еще не смогли зарекомендовать себя, «отвечая все новым и новым вызовам на протяжении нескольких веков». Тем не менее, они нашли свое применение в процессе управления данными и, вероятно, ряд из них ляжет в основу исследований и предложений в области длительного хранения.

17.11 Принципы, лежащие в основе современных подходов

При оценке способов преодоления влияния технологических изменений было отмечено, что большинство предложенных подходов основано на одном или нескольких из следующих принципов:

- Преобразование данных в удобочитаемую форму на простых в обслуживании носителях (документы, фильмы, металлические конструкции и т.д.)
- Создание данных или их конвертация с использованием стандартизированных форм кодирования и/или структур документа (или формата файла), которые будут широко распространены в компьютерных системах на протяжении длительного периода времени
- Создание данных в «самоописывающей» и «автономной» форме, упаковывая в них метаданные и со ссылками на программы, которые смогут обеспечить доступ (возможно, даже упаковка программного обеспечения с данными)
- Преобразование данных в формат с более легкодоступным способом доступа
- Сохранение данных в первоначальной форме (или упрощенной версии), и обеспечение их инструментами отображения в оригинальном виде, будет ли при этом использоваться оригинальное программное обеспечение и оборудование (которые также сохраняются), или новое программное обеспечение с эмуляцией работы оригинального программного обеспечения и/или аппаратных средств
- Обеспечение условий эмуляции первоначального способа доступа на виртуальной промежуточной компьютерной платформе, как базе для эмуляции в будущих операционных средах
- Преобразование (перемещение) данных в новые форматы, доступные новым операционным технологиям
- Задержка в пересылке для сохранения в данных достаточного количества информации, позволяющей будущему пользователю или менеджеру преобразовать их в удобочитаемую форму
- Обеспечение данных и поставка новых программ отображения (средств просмотра), которые обеспечат их отображение в новых операционных средах.

17.12 Требования по обеспечению стратегий сохранения

Независимо от выбранных стратегий, они должны быть обеспечены:

- Соответствующими организационными рекомендациями об ответственности, политике применения, процедурам и ресурсам
- Соответствующими юридическими документами
- Защитой данных
- Доступом к спецификациям стандартов и форматов файла для справок
- Метаданными, которые устанавливают достоверность, целостность и технические требования к материалу на весь период службы
- Вниманием к вопросам контроля качества на всех стадиях
- Контроль таких угроз, как изменений в технологии, которые повлекут возобновление проблемы разработки стратегии.

17.13 Планирование на случай непредвиденных ситуаций

Для всех стратегий целесообразно сохранять и предохранять первоначальный объектный поток данных, наряду с измененными в результате работы. Сохранение первоначального

потока данных должно рассматриваться, как планирование на случай непредвиденных ситуаций, обеспечивая возможность перейти на другие стратегии при возможном сбое выбранной. Такой подход подразумевает увеличение затрат на управление дополнительными данными и обеспечение связи между параллельными потоками данных. Несмотря на затраты, сомнения по поводу большинства стратегий сохранения делают этот подход достаточно целесообразным.

ТЕХНИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ПРОБЛЕМЫ

17.14 Введение

Настоящий раздел посвящен обсуждению некоторых обычно используемых стратегий. Они распределяются следующим образом:

- «инвестиционные» стратегии (предусматривающие затраты в основном на начальном этапе):
 - Использование стандартов
 - Извлечение и структурирование данных
 - Формирование пакета
 - Ограничение диапазона форматов данных
 - Подход «Универсальный виртуальный компьютер» (УВК)
- краткосрочные стратегии (обычно работающие наилучшим образом в течение короткого периода):
 - Сохранение технологий
 - Использование совместимости и преемственности версий
 - Перевод (который может действовать также и более длительный период)
- средне- и долгосрочные стратегии (возможно применимые для более длительного хранения):
 - (Перевод)
 - Средства просмотра
 - Эмуляция
 - Подход УВК
- альтернативные стратегии:
 - Нецифровые подходы
 - Восстановление данных
- смешанные.

1. Использование стандартов

Описание:

Эта стратегия предусматривает использование предпочтительно открытых, широко распространенных, поддерживаемых или согласованных стандартов и форматов файла, для которых существует большая вероятность стабильности и более длительного срока использования. Такие стандарты или форматы могут быть или формально согласованными, или быть действительно стандартными форматами, широко принятыми производителями. Опора на стандарты может также упростить приложение или повысить эффективность более поздних стратегий сохранения. Эта стратегия может быть связана со стратегией 4 — Ограничение диапазона форматов.

Некоторое усовершенствование этого подхода предложено при совместном использовании с подходом УВК (см. ниже, № 5), известное, как *долговременное кодирование* (Gladney и Lorie, 2002). При этом рекомендуется кодировать данные соотносить с известными стандартами обработки данных, в том числе кодирования битов (ASCII, Unicode UTF-8 и т.д.) и объектов (например, XML). Для объектов, которые не могут быть закодированы этим способом, программы отображения могут быть закодированы указанным образом и упакованы в комплекте с ними.

Примеры:

- Большинство программ оцифровки типа TIFF (тегированный формат файлов изображений) как доступный, устойчивый и широко используемый стандарт для создания сохраняемых изображений, с ожиданиями долговечности формата
- Стратегия электронных отчетов Виктории (VERS) в основном сохраняет цифровые документы в формате Adobe Portable Document (формат PDF) и форматирует их в упаковщике метаданных XML. Формат PDF был выбран, частично, из-за доступности стандарта, независимо от которого были созданы инструменты отображения и просмотра.

Более подробная информация: (все ссылки действительны на март 2003)

Gibbs R, Heazlewood J (2000). ‘Electronic Records — Problem Solved?: the Victorian Electronic Records Strategy and the future of electronic record keeping in Victoria’. Опубликовано в издании: *Books and Bytes : Technologies for the Hybrid Library : Proceedings, 10th Biennial Conference and Exhibition, 16-18 February, 2000, Melbourne Convention Centre*. Victorian Association for Library Automation, Inc., Melbourne, 2000.

Gladney H, Lorie R (2002). *Trustworthy 100-Year Digital Objects: Durable Encoding for When It’s Too Late to Ask*. Saratoga CA, HMG Consulting, 2002. Доступно, совместно с более поздними сопутствующими документами при обращении к HMG Consulting и, в электронном виде, на сайте Quarterly, http://home.pacbell.net/hgladney/ddq_14.htm

Некоторые имеющиеся стандарты обмена данных для различных областей деятельности перечислены в издании: The Diffuse Project (2002). *Diffuse Standards and Specification List*. The Diffuse Project Consortium, 2002. <http://www.diffuse.org/standards.html>

Преимущества использования стандартов

- Упрощение процесса сохранения, замедляя скорость замены технологий доступа; кодирование данных в базовых стандартах, подобно ASCII, что должно сделать их распознаваемыми в течение долгого промежутка времени
- Широко используемые форматы могут иметь широкий диапазон инструментов отображения
- Использование принятых и распространенных стандартов с большей вероятностью позволит заново интерпретировать данные или восстановить, при необходимости, инструментарий в будущем.

Трудности, недостатки и риски

- Может потребовать некоторых затрат на конвертирование материалов в стандартный вид; может привести к несогласованности некоторых элементов при конвертировании; для некоторых типов объектов не существует стандартизированных форматов.

Особые требования

- Знание соответствующих стандартов и постоянный контроль их разработки
- Стандарт форматов файлов должен выбираться как с точки зрения возможности сохранения основных характеристик объектов, так и с точки зрения ожидаемой долговечности инструментария.

Показания к применению

- Использование стандартов вообще должно поощряться, особенно когда заинтересованная организация обладает возможностью влиять на создание материалов или форматов, в которых могут быть представлены материалы
- Целесообразно в случае, когда доступные и стандартные форматы могут кодировать первоначальные объекты с заданной сложностью и без недопустимой потери основных характеристик.

2. Извлечение и структурирование данных

Описание:

Выделение данных, которое иногда называют *нормализацией*, предусматривает анализ и маркировку данных таким образом, чтобы могли быть описаны функции, связи и структура отдельных элементов. Отображение содержания может вестись без привлечения специальных приложений и может быть достигнуто за счет использования различных приложений в качестве замены технологии.

Примеры:

- Суперкомпьютерный центр Сан-Диего использовал заказные алгоритмы в применении к тегам XML в массиве, насчитывающем один миллион единиц электронной почты (Мооре и другие, 2000 [2]). Этот подход был также использован в приложении к обработке текстов и картографических данных (Мооре, 2001). Национальный архив Австралии также исследует этот подход в применении к электронной почте, с расширением действия на другие форматы (Heslop and Davis, 2001)
- Программа VERS Общественного центра сохранения данных штата Виктория (Австралия) исследует использование XML в таблицах баз данных
- Подход «Универсальный виртуальный компьютер» (Lorie, 2000) предполагает включение тэгов в начальные потоки данных с тем, чтобы отметить разделы для их отображения с использованием установленного свода правил для каждого типа данных.

Дополнительная информация:

Heslop H, Davis S (2002) (не опубликовано). *An Approach to the Preservation of Digital Records*. Национальный Архив Австралии, Канберра

Lorie RA (2000). *Long-Term Archiving of Digital Information, IBM Research Report RJ10185*. Исследовательское подразделение IBM, Сан-Хосе, Калифорния.

<http://domino.watson.ibm.com/library/CyberDig.nsf/7d11afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

Moore R, Baru C, Rajasekar A, Ludaescher B, Marciano R, Wan M *et al.* (2000). Collection-Based Persistent Digital Archives – Part 1. *Журнал D-Lib № 6(3)*.
<http://www.dlib.org/dlib/march00/moore/03moore-pt1.html>

Moore R, Baru C, Rajasekar A, Ludaescher B, Marciano R, Wan M *et al.* (2000). Collection-Based Persistent Digital Archives – Part 2. *Журнал D-Lib № 6(4)*.
<http://www.dlib.org/dlib/april00/moore/04moore-pt2.html>

Moore R (2001). Final Report for the Research Project on Application of Distributed Object Computation Testbed Technologies to Archival Preservation and Access Requirements, SDSC TR-2001-8. Суперкомпьютерный центр Сан-Диего. <http://www.sdsc.edu/TR/TR-2001-08.doc.pdf>

Преимущества извлечения данных

- независимость инфраструктуры упрощает перенос данных на новые платформы и технологии.

Трудности, недостатки и риски

- Не все виды объектов могут интерпретироваться таким способом
- Требуется разработка инструментов и методов анализа и обработки с тем, чтобы правильно отображать и размечать каждый вид данных

- Технология, использованная для отображения, может иметь ограничения при повторном отображении функций

Особые требования

- Наличие инструментов разметки и преобразования данных
- Контроль качества разработки методов для обеспечения отображения всех семантических зависимостей и аномалий.

Показания к применению

- структурные или слабоструктурированные данные или документы, для которых содержание, семантика и зависимости более важны, нежели иные характеристики отображения.

3. Формирование пакета

Описание:

Формирование пакета является распространенным способом связывания данных и способов обеспечения доступа к ним, предпочтительно в «упаковщике», описывающем пакет в формате, доступном широкому диапазону технологий (например, документ XML). Поскольку зачастую не практично и не нужно вводить в пакет сами способы доступа – программное обеспечение и оборудование – в пакет обычно включаются метаданные, описывающие или содержащие ссылки на необходимые инструментальные средствами. Иной подход предусматривает упаковку или ссылку на описание программного обеспечения или аппаратных средств с тем, чтобы они могли быть восстановлены в будущем, в случае необходимости.

Примеры:

- Ссылочный модуль Общей Системы Архивной Информации (ОСАИ) описывает связанные объекты, содержащие данные, с их метаданными в Архивных информационных пакетах (АИПах). Метаданные могут быть связаны с заархивированным объектом или прямо, или логически, в рамках системы
- Стратегия VERS предусматривает создание «луковичных отчетов», в которых объекты упакованы непосредственно закодированными в XML метаданными, делая их независимыми от систем управления
- Стратегия Универсального формата сохранения (УФС) стремится сделать объекты независимыми от приложений и операционных систем, упаковывая содержание в «самоописывающих» метаданных, которые содержат технические инструкции и инструменты по работе с содержимым.

Дополнительная информация:

Консультативный комитет по вопросам систем информационного пространства (2002). *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-B-1. Голубая книга. Выпуск 1. Январь 2002 г. Вашингтон, О.К., Секретариат ККСИП, 2002.
<http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>

Gibbs R, Heazlewood J (2000). Electronic Records – Problem Solved?: the Victorian Electronic Records Strategy and the future of electronic record keeping in Victoria. In: *Books and Bytes : Technologies for the Hybrid Library : Proceedings, 10th Biennial Conference and Exhibition, 16-18 February, 2000, Melbourne Convention Centre*. Ассоциация автоматизации библиотек штата Виктория Инкорпорейтед, Мельбурн, 2000.

Shepard T, MacCarn D (1999). *The Universal Preservation Format: A Recommended Practice for Archiving Media and Electronic Records*. Фонд образования WGBH, Бостон.
http://info.wgbh.org/upf/pdfs/991231_UPF_RP.pdf

Преимущества формирования пакета

- Обеспечивает наличие информации, облегчающей как поиск существующих способов доступа, так и их разработку.

Трудности, недостатки и риски

- Обеспечение ссылок на существующие способы доступа все-таки не снимает проблемы изменения технологий
- Поиск или восстановление устаревших способов доступа к упакованной информации могут быть затруднены.

Особые требования

- Детальное знание технических требований процесса доступа
- Безопасное и комплексное создание пакета из данных и метаданных
- Метаданные, описывающие способ доступа должны быть актуальны и современны
- Желательно наличие подпрограмм самоописания, таких, как упаковка XML.

Показания к применению

- Может быть рекомендована в качестве вспомогательного объекта для иных стратегий.

4. Ограничение диапазона форматов

Описание:

Программы сохранения могут хранить данные лишь в ограниченном диапазоне форматов. Это может достигаться или приемом материала, уже созданного в нужном формате, или конвертацией материалов из других форматов перед сохранением.

Примеры:

- Археологическая служба информации Великобритании (АСИ) определяет предпочтительный (но не единственный) диапазон форматов хранения и предоставляет рекомендации по созданию или подготовке материалов для отображения
- При определении видов отчетов от подчиненных учреждений, правительственные архивы могут также установить форматы, в которых может приниматься информация.

Дополнительная информация:

Archaeology Data Service (2001). *Guidelines for Depositors, Version 1.1*. Археологическая служба информации, Йорк <http://ads.ahds.ac.uk/project/userinfo/deposit.html>

Преимущества ограничения форматов

- Снижает диапазон проблем управления
- Может использоваться как усовершенствование существующего стандарта.

Трудности, недостатки и риски

- Не обязательно снимает проблему доступа, даже если используемые форматы оправдали себя в рамках иных подходов

- Может ограничить диапазон принимаемых материалов
- Преобразование может вызвать потерю некоторых необходимых элементов.

Особые требования

- База для принятия решения о приемлемых форматах и порядке действий с неприемлемыми
- Наличие четких правил доработки программного обеспечения для пересылки данных
- Строгий контроль качества.

Показания к применению

- Подход целесообразен для легко стандартизуемых материалов
- Массивы с большим числом однородных элементов.

5. Подход «Универсальный виртуальный компьютер»

Описание:

Подход «Универсальный виртуальный компьютер» (УВК) предназначен для формирования промежуточной платформы, виртуального компьютера, имеющего общие, полностью и точно определенные характеристики. Операции УВК достаточно просто могут быть перестроены для работы в любое время, на любой перспективной платформе.

При сохранении объектов во время архивирования разработанной логической схемы данных объект заносится в нее вместе с программой декодирования, способной к его отображению. Программа декодирования пишется под базу УВК.

Во время восстановления объекта, эмулятор УВК устанавливается на используемую платформу. УВК, в свою очередь, выполняет заархивированную программу декодирования, которая отображает заархивированный объект и передает управление программе восстановления. Последняя восстанавливает представление объекта согласно заархивированной логической схеме.

Примеры:

- Концептуальный прототип подхода УВК (Lorie, 2002) использовался для создания логической схемы, программы декодирования и механизма отображения для документов формата PDF, с тем, чтобы содержание документа могло быть представлено с использованием интерпретатора УВК и программы восстановления.

Дополнительная информация:

Gladney H, Lorie R (2002). *Trustworthy 100-Year Digital Objects: Durable Encoding for When It's Too Late to Ask*. Саратога, Калифорния, HMG Consulting, 2002. Доступно, совместно с более поздними сопутствующими документами при обращении к HMG Consulting и, в электронном виде, на сайте *Digital Document Quarterly*, <http://home.pacbell.net/hgladney/ddq 1 4.htm>

Lorie R (2002). *The UVC: a Method for Preserving Digital Documents – Proof of Concept*. Амстердам, IBM-Нидерланды, 2002. <http://www.kb.nl/kb/ict/dea/ltp/reports/4-uvc.pdf>

Lorie RA (2000). *Long-Term Archiving of Digital Information, IBM Research Report RJ10185*. Исследовательское подразделение IBM, Сан-Хосе, Калифорния. <http://domino.watson.ibm.com/library/CyberDig.nsf/7d11afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

Преимущества подхода УВК

- Возможно сохранение режима работы как материалов в документальной форме, так и программ
- Единая промежуточная платформа может снизить необходимость разработок, вызванных совершенствованием специального программного обеспечения и платформ
- Концепция УВК изначально задумана в упрощенном виде, позволяя ее использование программистами средней квалификации и предусматривая универсализацию интерпретаторов или эмуляторов УВК в будущем
- Может быть предназначен для отображения потока данных объекта как в первоначальном, так и в преобразованном или абстрактном виде
- Программы кодирования и декодирования данных могут быть протестированы в ходе создания на современной версии УВК. Можно с уверенностью ожидать, что подобным образом они будут работать и в среде перспективных версий УВК.

Трудности, недостатки и риски

- Подход в настоящее время находится в разработке и был смоделирован для преобразованных отображений оригинальных документов. Сейчас ведется дальнейшая работа по применению этого подхода к программам. Как и в случае с эмуляцией (см. № 11), его применение может вызвать проблемы в работе программ
- Затраты во время архивирования на разработку методов кодирования или УВК-совместимых программ отображения для каждого типа данных
- Может потребовать от источников информации создания УВК-совместимых версий их продукта (Gladney и Lorie, 2002), в чем они могут быть не заинтересованы или что может быть для них невыгодным
- В ходе восстановления необходимы временные затраты на разработку эмулятора УВК и программ восстановления
- Если первоначальные объекты находятся в абстрактном виде, или преобразованы для кодирования, дополнительное преобразование может вызвать сбой в основных характеристиках.

Особые требования

- Разработка логической схемы или способа отображения для каждого типа данных или программ при кодировании.
- Разработка на стадии кодирования программного декодера, который сможет отображать все типы данных или программ, написанных для выполнения интерпретатором или эмулятором УВК.
- Разработка интерпретатора или эмулятора УВК во время восстановления объекта для его соответствия существующей платформе.
- Разработка программ восстановления для обеспечения отображения объекта в первоначальном виде на основе логической схемы и данных, восстановленных УВК при выполнении программы-декодера.
- Архивирование данных или программы (или ее представления), связанных логических схем, UVC-совместимой программы-декодера и инструкций УВК.
- Проведение экспертиз для разработки логических схем, способов кодирования, программ-декодеров, эмуляторов УВК на базе программ восстановления и детализации.

Показания к применению

- Может применяться в условиях, когда объекты могут быть в достаточной степени совместимы, отображены, закодированы и восстановлены инструментами, разработанными на основе технологии УВК. Во время написания книги технология УВК все еще находилась в стадии разработки.

17.16 Краткосрочные стратегии

6. <u>Сохранение технологий</u> <i>Описание:</i> Эта стратегия предусматривает сохранение и поддержание исходного программного обеспечения и оборудования, с использованием которого создавались цифровые объекты. Она становится основным и, до некоторой степени, важнейшим первым шагом сохранения возможности доступа в условиях неприемлемости иных подходов. Если аппаратные средства и программное обеспечение, обеспечивающие доступ, вышли из употребления до применения других подходов, то возобновление доступа может стать фактически невозможным без проведения дорогих работ по восстановлению данных с сомнительным результатом. <i>Примеры:</i> • Сохранение старых дисководов для дискет более не применяемого размера на существующих компьютерах • Сохранение устаревшего программного обеспечения для обработки давних материалов • Сохранение старых операционных систем для работы с программным обеспечением, которое не работает на современных платформах. <i>Дополнительная информация:</i> Jones M, Beagrie N (2001). <i>Preservation Management of Digital Materials: A Handbook</i>. Британская Библиотека, Лондон, 2001.
--

Преимущества сохранения технологий

- Представление цифровых объектов на исходных аппаратных средствах и программном обеспечении гарантирует полное отображение из элементов и функций
- Дает время на разработку или внедрение иных стратегий
- В качестве дополнительной выгоды, накопление аппаратных средств и программного обеспечения может привести к улучшенному пониманию поддерживаемых ими архивов и их внутренних зависимостей, что, вероятно, будет полезно при планировании и осуществлении других стратегий.

Трудности, недостатки и риски

- Длительное обслуживание оборудования, с невозможностью обеспечить поставку запчастей и ремонт, маловероятно
- Даже с приложением активных усилий, период обеспечения доступа при использовании такого подхода может составить не более пяти — десяти лет со времени замены исходного формата. (Однако это все же лучше, чем потерять возможность доступа сразу)
- Необходимость обеспечения и обслуживания разнообразного оборудования и программного обеспечения, наряду со вспомогательными материалами вроде справочников и лицензий, что может оказаться трудно и дорого
- Необходимая экспертиза и техническая поддержка могут быть просто не обеспечены.

Особые требования

- Необходима четкая идентификация аппаратных средств и программного обеспечения, необходимых для доступа
- Необходимо активное и постоянное обслуживание оборудования и сохранение лицензий на программное обеспечение
- Необходимо обеспечение раздельности экспертиз и независимость их от одного лица
- Возможно, нескольким организациям можно объединить в единый банк оборудования и использовать общий архив программного обеспечения, принадлежащий им или третьей стороне
- Как правило, если необходимые программы доступа существуют, они должны быть найдены и сохранены до тех пор, пока не будет задействована другая стратегия.
- Подход к сохраненному программному обеспечению должен быть подобен подходу к другим цифровым объектам, включая обеспечение, наличие документации, восстановление и обслуживание носителей, а также соблюдение авторских прав.

Показания к применению

- Рекомендуются в качестве начальной стратегии для всех программ сохранения, в отсутствие более долгосрочных стратегий или в период их разработки
- Может оказаться единственной возможностью продления срока хранения таких цифровых комплексных объектов, как программное обеспечение и объекты мультимедиа
- Рекомендуются для программного обеспечения, необходимого для обеспечения других стратегий.

7. Использование совместимости и преемственности версий

Описание:

Эта стратегия опирается на способность некоторых видов программного обеспечения интерпретировать и представить объекты, созданные на предыдущих версиях. В случае совместимости, отображение может быть ограничено лишь временным просмотром, тогда как в случае преемственности новая версия просто конвертирует документы в собственный формат.

Примеры:

- Web-браузеры обычно способны к интерпретации и отображению материалов, созданных на более ранних версиях стандарта HTML
- Офисные приложения, такие, как текстовые редакторы, электронные таблицы и базы данных, обычно позволяют преобразование и повторное сохранение в новой версии файлов форматов предыдущих версий
- Нидерландский проект «Цифровой тестер сохранения» («Сигнальный цифровой тестер») использует перевод документов на новые прикладные версии.

Подробная информация:

Potter M (2002). Researching Long Term Digital Preservation Approaches in the Dutch Digital Preservation Testbed (Testbed Digitale Bewaring). *RLG DigiNews* 6(3). <http://www.rlg.org/preserv/diginews/v6-n3-a2.html>

Преимущества сохранения совместимости

- Применимость: разработчики программ часто встраивают возможность совместимости или конвертирования для документов
- Может увеличить период до более обширного преобразования или переработки
- В некоторых случаях обеспечивает возможность функционирования, как и при начальном отображении.

Трудности, недостатки и риски

- Маловероятно, что совместимость будет сохраняться на протяжении многих поколений программного обеспечения
- Могут появляться нежелательные замены при совместимости со многими поколениями программ
- Такой путь невозможно обеспечить для всех типов объектов
- Из-за возможных упущений разработчиков программ нового поколения надежность метода непредсказуема
- Даже при самых совместимых версиях могут появляться нежелательные изменения в материале.

Особые требования

- Как и при любом перемещении, необходима проверка качества перенесенных документов для обнаружения недопустимых замен

Показания к применению

- Может обеспечить простое и недолгое решение задачи сохранения документальных объектов в форматах, которые предполагают преемственность версий, пока конвертирование не приводит к нежелательным заменам
- Может стать альтернативой технологии сохранения таких объектов, как электронные таблицы и базы данных, для которых иные стратегии все еще не разработаны

8. Перевод

Описание:

Перевод предусматривает перенос цифровых материалов с одного поколения аппаратных средств или программного обеспечения на другое. В отличие от обновления, при котором данные переносятся на другие носители, при перемещении преобразовывается логическая форма цифрового объекта так, чтобы он мог быть концептуально представлен на новых аппаратных средствах или программном обеспечении.

Существует ряд стратегий, которые можно рассматривать в качестве форм перевода, различающихся лишь по времени преобразование и типам преобразуемых объектов. Обычный метод перевода предусматривает постоянное преобразование одного логического формата в другой в соответствии с изменением технологий так, чтобы все обработанные объекты соответствовали преобладающей технологии.

Также можно предложить модели «перевода по требованию» или «перевода при доступе». Этот подход рассматривается ниже, в № 10 («Средства просмотра»).

Примечание. Из-за вероятного кумулятивного эффекта повторных переводов, этот подход может использоваться только как краткосрочная стратегия. Однако для некоторых типов данных и форматов такой подход может использоваться в качестве долгосрочной стратегии.

Примеры:

- В массивах разнородных материалов распространенных форматов, таких, как собрания изображений, вероятно, может использоваться преобразование формата.

Дополнительная информация:

Lawrence GW, Kehoe WR, Rieger OY, Walters WH, Kenney AR (2000). *Risk Management of Digital Information: A File Format Investigation*. Совет по Библиотечным и Информационным Ресурсам, Вашингтон, округ Колумбия
<http://www.clir.org/pubs/reports/pub93/contents.html>

National Archives of Australia. *Managing Electronic Records — Appendix 3: Preserving Electronic Records through Migration*. Национальный Архив Австралии, Канберра.
http://www.naa.gov.au/recordkeeping/er/manage_er/append 3.html

Task Force on Archiving of Digital Information (1996). *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*. Группа Комиссии по сохранению, доступу и исследовательским библиотекам. <ftp://ftp.rlg.org/pub/archtf/final-report.pdf>

Преимущества перевода

- Существуют простые и наработанные механизмы перевода для некоторых форматов
- Переводы, выполненные при изменении в технологии, позволяют сверить переведенные объекты с неперевоенными копиями для проверки сохранности необходимых элементов
- Если перевод осуществлен, пользователь может с уверенностью ожидать, что материал будет отображен с использованием преобладающей технологии, без потребности в специальных аппаратных средствах или программном обеспечении.

Трудности, недостатки и риски

- Перевод формата может сделать невозможным обеспечение доступа к некоторым материалам типа комплексных объектов, потому что в новом формате может не быть способа отображения комплексных функций
- Несмотря на действенность, преобразование логического кодирования может поставить под угрозу целостность или основные элементы материала
- Объекты должны будут регулярно преобразовываться, чтобы идти в ногу с технологией, чем будут вызваны постоянные расходы. Масштабные переводы включают детальный анализ структуры данных, разработку правил преобразования, написание программ для замены кодирования данных, контроль качества и «зачистку». Это может быть легко выполнено для больших, важных для бизнеса баз данных, но такой суровый подход может быть невыполним для менее важных материалов с разнообразным диапазоном форматов файла
- Малые изменения с поколениями могут накопиться в большие изменения или потери в результате повторных переводов.

Особые требования

- Требуется программ и инструментов для преобразования
- Строгая проверка контроля качества как во время разработки метода, так и после перевода
- Документация метода перевода должна быть сохранена в метаданных, как часть хронологии и свидетельства подлинности объекта
- Если возможно, процесс перевода должен быть обратимым, что может быть достигнуто документированием природы и места изменений
- С другой стороны, если преобразование необратимо или могут быть потеряны необходимые элементы, должна быть сохранена копия исходных цифровых объектов. (Сохранение копии в первоначальном формате хорошо в любом случае)
- Процесс перевода должен проверяться перед выполнением и делаться до уничтожения любых промежуточных версий.

Показания к применению

- Перевод, вероятно, будет целесообразен для ряда цифровых объектов, особенно объектов документального типа или содержащих набор данных
- Когда сохранение необходимых элементов отработано и не зависит от вида и содержания материала и не затрагивает исполняемые файлы
- Может быть наиболее рентабелен для однородных коллекций типа цифровых изображений и аудиокolleкций, которые записываются в широко используемых, стандартизованных и общедоступных форматах
- Некоторые широко используемые частные форматы также могут применяться, если владельцы патентов или лицензий или сами разрабатывают новые форматы или инструменты перевода, или позволяют это делать другим.

9. Повторная разработка

Описание:

- Будучи высоко зависимыми от специальных систем или платформ, программные объекты могут повреждаться изменениями технологии и не соответствовать требованиям многих стратегий сохранения, включая регулярный перевод. Повторная разработка программного обеспечения может предложить нескольким стратегиям способ преобразования программ и обеспечения изменения технологий, подобно преобразованию форматов данных. При этом появляются новые возможности, которые включают:
- Настройку и преобразование исходного кода на новую платформу
- Повторную разработку преобразованного кода на более высоком уровне для его переноса на новую платформу
- Повторное кодирование программного обеспечения из временного хранения, или перекодирование на другом языке программирования (Wheatley, 2001)
- Перевод скомпилированных двоичных команд одной платформы, непосредственно в двоичные команды другой платформы. (Исследователи университета в Квинсленде (Cifuentes и др., 1999) исследуют эту концепцию.)

Дополнительная информация:

Cifuentes C, Van Emmerik M, Ramsey N(1999). The Design of a Resourceable and Retargetable Binary Translator. В сборнике: *Proceedings: Sixth Working Conference on Reverse Engineering, October 6-8, 1999, Atlanta, Georgia, USA*. Компьютерное сообщество ШЕЕ, Нью-Джерси, 1999, страницы 280-291

Wheatley P (2001). Migration- a CAMiLEON discussion paper. *Ariadne* 29. <http://www.ariadne.ac.uk/issue29/camileon/>

Преимущества повторной разработки

- Возможен перенос программных объектов с одной платформы на другую.

Трудности, недостатки и риски

- За исключением разработанных самостоятельно открытых исходных программ и программного обеспечения, доступ к исходному коду зачастую невозможен или на это нет прав
- Даже когда доступ к исходному коду возможен, перенос на другие платформы еще не гарантирован, и вообще, необходимы компиляторы или интерпретаторы кодового языка для новой платформы
- Требуется больших затрат времени и усилий на каждый объект
- Повторная разработка в любой форме конечным пользователем обычно строго запрещается лицензионными соглашениями и серьезно нарушает права интеллектуальной собственности. Это может произойти и при других видах внесения изменений.

Особые требования

- Проведение экспертизы на высоком уровне
- Наличие инструментов для преобразования обычно читаемого кода в машинно-читаемую форму
- Наличие разрешения для повторной разработки.

Показания к применению

- Применимо только в случае предоставления соответствующих прав и возможно наличия экспертизы, инструментов и, желательно, исходного кода.

17.17 Средне- и долгосрочные стратегии

10. Средства просмотра и перевода в момент доступа

Описание:

Существует множество вариантов многократного перевода, в том числе использование средств просмотра, программных инструментов или методов преобразования, которые обеспечивают доступ в момент обращения, используя первоначальный поток данных.

Примеры:

- Подход «Перевод по запросу», совместно разработанный проектами CEDARS и CAMiLEON, включают программный инструмент и цифровой объект и предусматривают использование метаданных объекта для описания метода доступа к объекту с использованием данного инструмента. По мере изменения технологии, в метаданные вносятся изменения метода доступа (CEDARS, 2002; Mellor, Sergeant and Wheatley, 2003).
- Подход TOMS (типичный объект для модельного сервера), обеспечивает предоставление методов преобразования для обычных типов документов и данных, позволяя серверу выбрать нужный способ преобразования для стандартных типов объекта. (Thibodeau, 2002)
- Стратегия VERS конвертирует документы в формат PDF, основываясь на том, что средства просмотра для формата PDF могут быть созданы на основе его описания.
- Подход Розетты Стоунс включает методы интерпретации формата данных и просто файлы, как в первоначальном, так и в справочном формате, показывающем, как файлы должны выглядеть при правильной интерпретации. В этом случае могут быть созданы программные инструменты для соответствующего отображения и его проверки, сравнивая полученный результат со справочным. (Thibodeau, 2002)

Подробная информация:

Проект CEDARS (2001). *The Cedars Project Report, April 1998-March 2001*. Cedars, Лидский университет.

<http://www.leeds.ac.uk/cedars/pubconf/papers/projectReports/CedarsProjectReportToMar01.pdf>

Проект CEDARS, (2002). *Cedars Guide to: Digital Preservation Strategies*. Cedars, Лидский, университет. <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html>

Mellor P, Sergeant D, Wheatley P (2002). *Migration on Request: A Practical Technique for Preservation*. Проект CAMiLEON, Мичиганский университет. <http://www.si.umich.edu/CAMiLEON/reports/migreq.pdf>

Thibodeau K (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. В сборнике: *The State of Digital Preservation: An International Perspective – Conference Proceedings, Documentation Abstracts, Inc., Institutes for Information Science, Washington, D.C., April 24-25, 2002*. Совет по Библиотечным и Информационным Ресурсам, Вашингтон, округ Колумбия. <http://www.clir.org/pubs/reports/pub107/thibodeau.html>

Преимущества использования средств просмотра и т.д.

- Первоначальный поток данных интерпретируется и отображается средствами просмотра, инструментами или методами преобразования, вместо использования измененного и переведенного потока данных, что позволит избежать риска кумулятивных искажений объема или зависимости от числа переводов
- Объекты интерпретируются или преобразовываются только когда к ним обращаются, поэтому можно избежать затрат на постоянный перевод объектов независимо от их востребованности.

Трудности, недостатки и риски

- Не существует средств или инструментов просмотра для комплексных материалов, в том числе исполняемых файлов
- Средства просмотра могут отобразить некоторые, но не все элементы некоторых материалов (хотя это может быть преимуществом доступа «только для просмотра»)
- Разница между начальным форматом и современными технологиями в момент доступа может быть слишком большой, чтобы с ней могли справиться имеющиеся инструменты или методы
- Средства просмотра, инструменты или методы, а также соответствующие метаданные также должны обслуживаться и корректироваться в соответствии с развитием технологии
- Даже если его и не было ранее, существует риск, что средства просмотра, инструменты или методы не смогут удовлетворительно отобразить смысл объектов.

Особые требования

- Должно быть обеспечено постоянное документирование форматов файла и методов преобразования
- Масштабное наполнение технических метаданных в соответствии с развитием технологии
- Технические метаданные и методы доступа должны быть связаны, но сохраняться отдельно от цифровых объектов с тем, чтобы метаданные или методы могли быть модифицированы централизованно.

Показания к применению

- Может превалировать над постоянным переводом в случае, если стоимость последнего высока или если существуют долгие промежутки между запросами
- Целесообразно при наличии уверенности в возможности создания инструментов или средств просмотра, которые интерпретируют формат файла исходя из внутренних команд, описаний и методов.

II. Эмуляция

Описание:

Эмуляция предусматривает использование программного обеспечения, которое заставляет одну технологию вести себя подобно другой. В случае долгосрочного цифрового сохранения, это повлекло бы за собой создание технологии программ поведения подобно первоначальной среде сохраняемого цифрового объекта с тем, чтобы первоначальный объект мог быть представлен в его первоначальной форме с использованием первоначального потока данных.

Аппаратная эмуляция часто предлагается как широко применимая стратегия, поскольку аппаратные описания, вероятно, будут более полными или легко определяемыми, чем программные описания. Эмуляция аппаратной платформы также предполагает хорошие методы решения проблемы, при которых можно было бы работать целому ряду систем и цифровых объектов. С другой стороны можно предусмотреть эмуляцию отдельных приложений или функций. Единственным отрицательным моментом при этом становятся усилия по эмуляции отдельно для каждого приложения; с другой стороны, если потребность в эмуляции мала, то усилия по эмуляции всей платформы или системы для малого числа или диапазона объектов может уподобиться стрельбе из пушки по воробьям.

Примеры:

- Специалисты проекта CAMiLEON исследовали эмуляцию в качестве подхода к сохранению цифрового наследия, включая эксперименты по использованию существующих эмуляторов (Hedstrom и Lampe, 2001) и создание эмулятора для системы 1970-х, George3 (Holdsworth и Wheatley, 2001)
- Универсальный Виртуальный Компьютер (UVC) был предложен как промежуточная виртуальная платформа для использования на базе будущих систем, чтобы эмуляция программного поведения обеспечивалась на единой устойчивой платформе, минимизируя потребность в дополнительных уровнях эмуляции (Lorie, 2000)
- Было предложено отложить создание эмуляторов, сохраняя вместо этого детальные инструкции для их создания с возможностью генерации по мере необходимости (Rothenburg, 2000).

Дополнительная информация:

Hedstrom M, Lampe C (2001). Emulation vs. Migration: Do Users Care? *RLG DigiNews* 5(6). <http://www.rlg.org/preserv/diginews/diginews5-6.html#feature1>

Holdsworth D, Wheatley P (2001). Emulation, Preservation and Abstraction. *RLG DigiNews* 5(4). <http://www.rlg.org/preserv/diginews/diginews5-4.html#feature2>

Lorie RA (2000). *Long-Term Archiving of Digital Information, IBM Research Report RJ10185*. Исследовательское подразделение IBM, Сан-Хосе, Калифорния. <http://domino.watson.ibm.com/library/CyberDig.nsf/7d11afdf5c7cda94852566de006b4127/be2a2b188544df2c8525690d00517082>

Rothenberg J (2000). Using Emulation to Preserve Digital Documents. Королевская библиотека, Гаара. <http://www.kb.nl/kb/pr/fonds/emulation/usingemulation.pdf>

Преимущества эмуляции

- Эмуляция достаточно развита и часто используется в информатике, чтобы разрабатывать и проверять новое программное обеспечение перед запуском в серию

- В настоящее время существуют эмуляторы различных платформ и систем, от эмуляторов отдельных систем, созданных энтузиастами, до коммерческих систем для межплатформенного использования или тестирования программного обеспечения
- При широком использовании эмуляция может позволить обновлять функциональные возможности целого ряда цифровых объектов, включая программные объекты, использующие первоначальный, не преобразованный поток данных в сочетании с оригинальным сохраненным программным обеспечением.

Трудности, недостатки и риски

- Технически эмуляция является комплексным, трудоемким и специализированным и, как следствие, чрезвычайно дорогостоящим процессом
- Как широко распространенная цифровая методика сохранения, эмуляция все еще находится в стадии исследования
- Сбой в эффективной эмуляции может быть вызван некорректной документацией программного обеспечения или ненормативным использованием форматов файла, наподобие методов «обходных маневров»
- Поскольку системы становятся более комплексными, к эмуляции могут появиться требования по включению большого количества компонентов. Эмуляция же всех возможностей системы или приложения может быть невозможна
- Для будущих пользователей может быть трудно знать порядок работы с эмуляциями широкого диапазона древних приложений, поэтому может появиться потребность в современных инструментах отображения, что, в свою очередь, увеличит количество изменений в программных инструментах, необходимых для доступа к эмулируемому объекту
- Поскольку технология и платформы постоянно развиваются, сами эмуляторы должны будут или переводиться, или подражать современным системам на новых платформах, потенциально ведя к наслоению и усложнению эмуляторов.

Особые требования

- Достаточный уровень компетентности при разработке эмуляторов, или обращение к уже разработанным эмуляторам
- Полная и точная документация на объекты эмуляции
- Определение уровня эмуляции, например, наличие полной аппаратной эмуляции или имитация отдельных программных функций
- Код эмулятора должен разрабатываться с использованием стандартных методов разработки программного обеспечения, включая структуру кода, а также полные комментарии и документацию
- Код программы эмуляции должен быть написан на базе открытых источников, на стандартном языке программирования с достаточными возможностями долговечности и совместимости
- Любой необходимый нестандартный код (например, для отдельных периферийных функций) должен быть написан как отдельный, но связанный модуль, и хорошо прописан.

Показания к применению

- При наличии соответствующих эмуляторов для заданных платформ
- При наличии достаточной базы для создания эмулятора
- Для объектов высокого уровня сложности или исполняемого программного обеспечения такого же типа, которое может работать только с отдельными системами или аппаратными средствами
- Для объектов, к которым иные пути, например перевод, не подходят
- Для объектов, значение которых зависит от отображения в оригинальных средах.

17.18 Альтернативные стратегии

<u>12. Нецифровые подходы</u> <i>Описание:</i> Альтернативой цифровых методов сохранения является «распечатка» объектов на относительно долговечных аналоговых носителях, например в виде документов, микрофильмов или даже никелевых пластины (как в случае технологии HD-Rosetta, при которой изображения и документы микрогравированы на никеле ионным лучом и просматриваются оптическими лупами), смещая вопросы сохранения на аналоговые копии вместо цифровых объектов. <i>Примеры:</i> • Учреждение, озабоченное сохранностью старой базы данных в устаревшем частном формате, который будет нечитабелен при следующем системном обновлении и для которого уже не существует способа взять или перевести хранимые данные. Учреждение при этом принимает решение напечатать полное содержание баз данных на бумаге в качестве отдельных записей, чтобы сохранить хоть какой-нибудь доступ к ним, хотя это и неэффективно. Также сохраняются и базы данных в цифровой форме, когда станет возможным создать механизм доступа. • Программа оцифровки создает компьютерный микрофильм с цифрового изображения, как материальный дубликат собрания и альтернативный источник сохранения, распространения и доступа. <i>Дополнительная информация:</i> Hedstrom M, Lampe C (2001). Emulation vs. Migration: Do Users Care? <i>RLG DigiNews</i> 5(6). http://www.rlg.org/preserv/diginews/diginews5-6.html#feature1 Norsam Technologies, (2001), <i>HD-Rosetta Archival Preservation Services</i>. http://www.norsam.com/hdrosetta.htm

Преимущества нецифровых методов

- Объекты сохранены в удобочитаемой форме и не подвержены воздействию угрозы технологического устаревания и постоянного процесса цифрового пересохранения
- Обеспечивает более простой вариант сохранения, поскольку аналоговые материалы могут храниться для долгосрочного использования традиционными методами сохранения

- Возможность привлечения единичных затрат на переработку.

Трудности, недостатки и риски

- Вероятно потери преимуществ цифровых технологий, например, удобств использования или эффективности хранения
- Потеря обычных возможностей, обеспечиваемых цифровыми технологиями, например, обработка электронных таблиц, озвучивание или двигающиеся изображения (хотя некоторые из них также могут быть сохранены в отдельной аналоговой форме), или функции поиска и управления. (Это не является недостатком, если данные элементы не должны сохраняться)
- Нельзя полностью преодолеть угрозу технологического устаревания, так как может измениться технология считывателя для некоторых форматов. Даже при том, что теоретически всегда можно использовать оптические лупы для считывания информации, это может быть непрактично, делая материал фактически непригодным для большинства пользователей. Для звукозаписей аналоговый доступ может быть более ненадежным, чем цифровой
- Длительное хранение аналоговых носителей может зависеть от дорогих условий, что может быть менее рентабельным, чем хорошо обслуживаемые компьютерные системы, построенные с высокой степенью дублирования.

Особые требования

- Носители, использованные для аналогового преобразования, должны иметь высокие возможности по сохранению информации и храниться в специальных условиях
- Сохранение цифровых объектов все-таки необходимо, по возможности, проводить, так как в будущем могут быть разработаны необходимые технологии доступа.

Показания к применению

- Целесообразно только для объектов, которые не требуют цифровых технологий отображения, например, содержащих текстовую, визуальную или иную информацию, которая не требует возможностей показа выше простого отображения
- Может применяться, пока разрабатываются другие стратегии
- Может применяться, в крайнем случае, когда не существует иных стратегий и такой ограниченный доступ лучше, чем его отсутствие.

13. Восстановление данных

Описание:

Восстановление данных (иногда это называют *археологией данных*) обычно предусматривает восстановление данных по битам с материальных носителей, сопровождаемое усилиями по восстановлению их читабельности. Этот метод наиболее часто используется при восстановлении данных с неудачных, поврежденных или разрушенных носителей, однако методы по восстановлению читабельности использовались для спасения документов в устаревших форматах. Однако уверенность, что можно будет постоянно предпринимать такие «спасательные операции» в будущем крайне мала и влечет высокий риск нынешней деятельности программ в области сохранения.

Примеры:

- Археологическая информационная служба Великобритании восстанавливала данные с дисков Археологической службы Ньюгемпского музея (Dunning, 2001). Множество файлов было признано разрушенными и не подлежащими восстановлению. Из тех, которые были восстановлены, многие были в устаревших форматах и требовали специального программного обеспечения для отображения, или имели несоответствующую документацию, так что содержание данных не могло быть восстановлено с приемлемой точностью.

Дополнительная информация:

Ross S, Gow A (1999). *Digital Archaeology: Rescuing Neglected and Damaged Data Resources*. Библиотека Центра информационных технологий, Университет южного побережья, Лондон. <http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2con.pdf>

Woodyard D (2001). Data Recovery and Providing Access to Digital Manuscripts. Доклад на конференции Digital Dancing: New Steps, New Partners – Information Online 2001, Tenth Exhibition and Conference, 16-18th January, 2001, Сиднейский выставочный центр, Дарлинг Харбор, Сидней. <http://www.nla.gov.au/nla.arc-14099-20020211-www.csu.edu.au/special/online2001/papers/digital issues iia.htm>

Dunning A (2001) *Excavating Data – The Retrieval of the Newham Archive*. Служба информации гуманитарной и искусствоведческой тематики. <http://ahds.ac.uk/newham.pdf>

Преимущества восстановления данных

- Может позволить восстановление данных, которые в ином случае постоянно терялись бы.

Трудности, недостатки и риски

- Нет гарантии восстановления с носителей или восстановления читабельности данных
- Без достаточной документации, интерпретация данных часто делается по принципу «наилучшего приближения», причем установить достоверность, целостность и контекст крайне трудно
- Способ зачастую дорог, требует затраты значительных усилий на каждый элемент
- Без достаточной документации невозможно заранее судить об оправданности усилий и расходов.

Особые требования

- Наличие подробной документации о типе и содержании файлов
- Может потребовать услуг специалиста по восстановлению данных или программ распознавания

Показания к применению

- Рекомендуются к использованию только для восстановления данных и стратегии в случае повреждения носителей, или в случае нахождения неиспользуемых носителей или форматов файлов, причем значение или важность данных оправдывает понесенные траты.

17.19 Смешанные стратегии

Как отмечено ранее, для собрания разнообразных материалов может подходить не одна, а целый ряд стратегий сохранения; различные подходы могут также хорошо взаимодействовать друг с другом по прошествии какого-то времени. Программы в области сохранения должны также учесть возможные выгоды от следования нескольким стратегиям: даже при хорошем планировании, единственная стратегия может все-таки привести к потере доступа. Вот несколько примеров использования нескольких стратегий:

- Стандарты типа TIFF часто используются для коллекций изображений при подготовке к возможному их переводу на другие стандартные форматы для долгосрочного хранения
- Стратегия VERS объединяет использование стандартов (формат PDF, XML) для последующего использования средств просмотра с переводом в дальнейшем метаданных в стандарт XML
- Постоянные архивы (Moore, 2001) используют отделение данных с возможным переводом в дальнейшем — перевод данных — для обновления систем, обеспечивающих программ и аппаратных средств
- Подход «Универсальный виртуальный компьютер» (УВК) объединяет отделение данных, с правилами их перевода в момент доступа, с эмуляцией программных объектов. Подход «долгосрочного кодирования» дополнительно использует основные стандарты для кодирования данных, включая кодирование в формате УВК.

**ДЛЯ ПРОГРАММ СОХРАНЕНИЯ,
РАБОТАЮЩИХ С НЕБОЛЬШИМИ РЕСУРСАМИ**

17.20 Выбор по принципу снижения затрат

Программам в области сохранения, работающим с небольшими ресурсами, вероятно, придется ограничить количество обеспечиваемых материалов. По отношению к стратегиям обеспечения доступа это означает:

- Возможность принятия подхода «минимального доступа», надежно сохраняя данные с четкой документацией о первоначальном способе доступа, планируя в будущем использовать эту информацию для разработки способа доступа

- Возможность определения приоритетных материалов, которые можно перевести на формат, обеспечивающий некоторый уровень доступа, с одновременным сохранением оригинала для последующего обеспечения более полного доступа
- Принимать материал в немногих, хорошо стандартизированных и широко распространенных форматах, что понизит затраты и ограничит количество специальных инструментов обеспечения доступа. Это может даже позволить перевести ряд материалов под нужды инструментов потребителя.

ОПЫТ ПРИМЕНЕНИЯ

17.21 Некоторые возможные стратегии для различных типов данных (для обсуждения)

- Набор данных — стандартизированное кодирование; метаданные, описывающие структуру; возможен перевод, однако данные не останутся распознаваемыми в течение длительного периода без дальнейших усилий
- Базы данных — сохранение данных и документирование структуры извлечения данных; сохранение программного обеспечения доступа к данным; сохранение интерфейса и истории обращений; перевод данных на новые базы данных и интерфейс пользователя
- Звуковые файлы и файлы изображений — использование стандартов, включая внимание к вероятным осложнениям, например, проблеме сжатия; перевод на новые стандарты и форматы
- Текстовые файлы — кодировка (т.е. перевод на стандартные коды и стандарт XML); возможный перевод на аналоговые носители; перевод на новые форматы
- Файлы HTML — сортировка по форматам и перевод на новые стандарты браузеров
- Программы и материалы на основе программ — сохранение технологии; эмуляция; повторная разработка
- Материалы электронной почты — извлечение данных и стандартизация структуры; перевод по необходимости
- Офисные записи — средства просмотра; извлечение данных и нормализация формата

ССЫЛКИ ДЛЯ ПОЛУЧЕНИЯ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Перекрестные ссылки

Вопрос необходимых элементов см. также в главе 12 Определение того, что должно быть сохранено

Зависимость между данными и программным обеспечением см. также в главе 8 Понятие сохранения цифрового наследия

Другие источники

Веб-узел «Подход к сохранению информации в цифровом виде» (PADI) предоставляет широкий и обновляемый набор ссылок для оценки стратегий по сохранению доступа. Доступно по адресу <http://www.nla.gov.au/padi/>

Глава 18. Некоторые отправные точки

ВВЕДЕНИЕ

18.1 Цели

Трудно создать качественные программы в области сохранения на пустом месте, даже при наличии достаточных ресурсов. В условиях ограниченных ресурсов это становится еще более трудной задачей. В настоящей главе, кратко и с использованием ряда гипотетических сценариев, описаны некоторые отправные точки для этого процесса. Настоящая информация предлагается в качестве стимула для обсуждения, поскольку ситуация для каждой программы зачастую отлична и требует отдельного подхода.

ОБЩИЕ ВОПРОСЫ

18.2 Некоторые начальные шаги

Следующие пункты могут быть использованы при организации работы программ в области сохранения:

1. Определите, за какие материалы Вы ответственны, или какие материалы Вам нужно сохранить.
2. Выясните, кто имеет подобные интересы или обязанности, чтобы оценить возможность совместной работы.
3. Выясните, кто имеет опыт в сохранении или, по крайней мере, обеспечении того вида материалов, которым Вы интересуетесь, для обмена советами и опытом.
4. Пробуйте выяснить, кто создает материал, которым Вы интересуетесь; кто издает, распространяет или содержит материал, и какими возможностями или заинтересованностью по его сохранению они располагают и на какой период.
5. Пробуйте определить потенциальных и нынешних пользователей материала и как они хотят его использовать.
6. Может быть слишком трудно определить на начальной стадии необходимые характеристики для сохранения, однако важно определить уровень функциональных возможностей, которые Вы хотите сохранить, т.е. должны ли пользователи быть способны взаимодействовать с материалом и изменять его, или просто просматривать его в форме «только для чтения».
7. Возможны две широко распространенных модели для того, чтобы предпринять первые практические шаги:
 - Начав с небольшого объема материалов, возможно ограниченного относительно понятной и «простой» информацией для обеспечения наилучшего сохранения, Вы можете по мере накопления опыта определить политику, цели, критерии и инфраструктуру своей работы.

- Стремиться осмыслить программу целиком и решить все проблемы перед ее началом.

Оба подхода имеют свои достоинства и недостатки, однако в настоящих рекомендациях предлагается первый вариант для случая ограниченных ресурсов и необходимости приступить к работе немедленно. Такой подход едва ли может сразу обеспечить всестороннее и достоверное сохранение, однако в его рамках можно обеспечить всестороннюю и надежную деятельность в этой области через некоторое время.

8. Разработайте, хотя бы в первом приближении, план Ваших действий на ближайший период.
9. Определите первичные угрозы, которые требуют немедленного внимания, чтобы предотвратить потери ценных материалов. (Может быть необходимо смириться с тем фактом, что некоторые материалы будут потеряны и сфокусировать внимание на сохранении хотя бы части важного материала.)
10. Определите немедленные шаги, которые Вам нужно предпринять в ответ на угрозы, например, обеспечение восприятия материала, с которым Вы должны будете иметь дело, или установление контакта с создателями.
11. Определите ресурсы (персонал, экспертизу, фонды, оборудование, время), которые могут быть нацелены на преодоление угроз.
12. Определите действия, которые Вам необходимо предпринять, чтобы выиграть некоторое время; или, если Вы располагаете достаточными ресурсами и поддержкой, разработать более качественную программу сохранения. (Некоторые примеры основных угроз и возможных ответов на них приведены ниже в данной главе.)
13. |Определите, какие права или разрешения нужны вам для работы,
14. Определитесь с разрешениями, в том числе проясняя наличие соответствующих прав, или испрашивая права на работу с материалами.
15. Планируйте работу и претворяйте в жизнь свои планы, оценивая их результативность при каждом шаге.
16. Обсудите с создателями способ создания материалов, предлагая им советы и информацию о новациях в области упрощения сохранения.
17. Оцените результаты своей деятельности и определите, является ли ваш подход жизнеспособным, приемлемым или невыполнимым; нуждается ли он в дальнейшем развитии или серьезном изменении.

ТИПОВЫЕ СЦЕНАРИИ

18.3 Сценарий 1

Для веб-публикаций существует высокая вероятность, что они будут изменены или удалены из Сети до сохранения.

Возможные ответные действия:

- Разработайте критерии по отбору сохраняемого материала
- Войдите в контакт с издателями и обсудите совместные планы

- Если материал находится под угрозой утраты и Вы получили предупреждение, скопируйте необходимые файлы на компьютер, проверьте качество пересылки, сам документ и способ копирования и приступайте к обслуживанию данных, изготовив резервные копии и сохранив их в безопасном, удаленном месте
- На основе консультаций с издателем определите возможность размещения материала для общего доступа. В случае согласия Вам понадобится какой-либо интерфейс для обеспечения поиска и отображения материала пользователям
- Это — меры краткосрочного действия, которые будут, в конечном счете, нуждаться в разработке системы поиска, обеспечения договоренностей с издателями, отслеживания и загрузки различных материалов, учета и управления соответствующими метаданными, управления доступом, обслуживания потоков данных, распознавания необходимых для сохранения при отсутствии технологий доступа характеристик материала и поиска способов его отображения, с учетом того, что он может содержать комплект мультимедиа и выполнимые объекты. Такая система и ее разработка требуют существенных затрат времени и денег и может потребовать нескольких лет на разработку и внедрение
- В мире существует множество программ в области сохранения, в рамках которых уже решены эти проблемы: надо только обратиться к их техническому описанию, подходам и описаниям деятельности.

18.4 Сценарий 2

Для объектов Системы хранения записей (СХЗ), может существовать проблема изменений в административных ресурсах, например, смена правительства, что может повлечь крупномасштабную пересылку или уничтожение отчетов; или плановая замена систем.

Возможные ответные действия:

- При наличии времени, определить порядок работы СХЗ, какие записи деловой активности существуют, какие отчеты важны, позволяет ли СХЗ оценить подлинность. Отчеты должны быть расставлены в соответствии с графиком размещения, разработанным в ходе оценки. Отчеты, выбранные для сохранения с метаданными, должны быть перенесены в архив, где они могут быть проверены, преобразованы в стандартный формат, выбранный для сохранения их необходимых характеристик, и помещены в память с соответствующим ссылками
- При отсутствии времени на проведение оценки перед пересылкой, может быть необходимо провести экспресс-оценку с целью определения достаточной для последующей обработки документации, которая позволит обработать материалы позже, когда они уже будут удалены из рабочего оборота. Отчеты и любые инструменты для работы с ними — возможно СХЗ полностью — могут быть также перенесены в архив
- По мере переноса определяются условия для безопасного хранения отчетов, документации на них и любых инструментов доступа до проведения оценки. Поскольку ряд СХЗ привязан к отдельным операционным платформам, может быть потребуется сохранить оборудование и программное обеспечение, которыми оно изначально поддерживалось
- Такой материал может иметь большое значение для оценки и для определения соответствующих стратегий сохранения оригинальности и доступа, определения способа извлечения данных и перевода на стандартные форматы прежде, чем доступ к оригинальным СХЗ будет утрачен.

18.5 Сценарий 3

Для коммерческих аудиовизуальных материалов может существовать риск прекращения деловой активности малых производителей.

Возможные ответные действия:

- Разработка критериев выбора сохраняемого материала
- Определить и войти в контакт с производителями или распространителями для получения копий исходных файлов самого высокого качества, свободных от средств защиты от копирования
- Если материал находится в нестандартном формате, запросить копии в стандартном формате
- Переслать выбранный материал в архив и скопировать или на более устойчивые носители, или на удобные в обслуживании компьютерные системы хранения с созданием резервных копий
- Проверить качество пересылки данных
- Создать документацию на материал, включая информацию о правах, необходимую при управлении правами интеллектуальной собственности, что часто усложнено для материалов данного вида
- Поскольку цифровой аудиовизуальный материал содержит большой объем данных, обычный архив может не располагать средствами сохранения больших массивов, что может поставить их перед проблемой емкости. Выигрышным решением при этом может стать сохранение элементов на достаточно устойчивых автономных носителях, например, компакт-дисках
- Более мощные программы должны обсуждать вопросы размещения и правообладания с источниками или распространителями и устанавливать сложные системы для обработки и хранения больших массивов данных, метаданных, а также, возможно, накапливать оборудование и программное обеспечение для воспроизведения, если нет возможности преобразовать материал для стандартных систем воспроизведения без недопустимых потерь.

18.6 Сценарий 4

Для материала, выпущенного на недолговечных магнитных носителях, например, гибких дисках или лентах, существует большая вероятность ухудшения качества носителей.

Возможные ответные действия

- Постараться определить, какой материал стоит сохранить
- Копировать материалы с гибких дисков или лент на любой более устойчивый носитель, например, компакт-диск, или на другие нестабильные носители, например ленты, но только если они удобны в обслуживании для имеющихся систем
- Проверить качество копирования
- Записать метаданные о материале и порядке пересылки
- По возможности, использовать пересылку как возможность отражения зависимостей материалов от программного обеспечения
- Планировать работу по проблеме программной зависимости.

18.7 Сценарий 5

Для сохраняемого материала может нарастать угроза утраты оборудования и программного обеспечения, необходимого для доступа.

Возможные ответные действия:

- Постараться определить, какой материал стоит сохранить
- По возможности, принять меры по переносу материала с текущего носителя на носитель и формат, которые могут быть обработаны наличным оборудованием и программным обеспечением, если это может быть сделано без существенных потерь
- При необходимости принять меры к сохранению материала в его первоначальном виде наряду с оборудованием и программным обеспечением, использовавшимся в настоящее время для доступа
- Одновременно искать способ максимально долгого хранения оборудования, в противном случае немедленно копировать материал на иной носитель и/или формат
- Проверить материал на нежелательные изменения
- Задokumentировать сам материал, его происхождение и любые изменения в нем
- Надежно хранить материал с возможностью восстановления
- Планировать работу по проблеме программной зависимости материала, особенно если не было возможности конвертировать его в используемый формат файла
- Для материала, который уже утратил аппаратные средства и программное обеспечение, необходимые для доступа, необходимо будет найти владельцев уцелевших аналогов в рабочем состоянии, позволивших бы их использование. Это может потребовать услуг экспертов по восстановлению данных или специального программного обеспечения для восстановления данных.

18.8 Сценарий 6

Для комплексных данных может нарастать угроза утраты персонала, компетентного в кодировке данных и порядке их работы.

Возможные ответные действия:

- Определить, стоит ли сохранять данные в комплексе
- При помощи имеющегося персонала, знакомого с комплексными данными, полностью их документировать с целью обучения других сотрудников или подрядчиков
- Убедиться, что комплексные данные скопированы и адекватно восстанавливаются в защищенной системе хранения
- Планировать порядок обеспечения доступа в случае замены операционной среды.

МИНИМАЛЬНЫЕ ПРОГРАММЫ СОХРАНЕНИЯ

Настоящий раздел определяет возможные шаги для разработки минимальной программы в области сохранения, которая может применяться в случае необходимости сохранения данных организацией, не способной развернуть более масштабную программу сохранения. Раздел содержит также некоторые сценарии.

Определение необходимости сохранения, а также потребностей и ресурсов

- Существует ли в цифровом виде информация, которую Вы должны сохранить? Существует ли еще кто-либо, кто вероятно должен сохранить ее? Какие разрешения необходимы? На какие риски или угрозы нужно обратить внимание? Определите, какие ресурсы вы можете привлечь для решения задачи.

Влияние на процесс сохранения

- По крайней мере, определите принятые форматы. Если возможно, договоритесь с источниками об использовании широко распространенных стандартов и обеспечении соответствующей документации.

Защита данных

- Храните носители в соответствующих условиях
- Копируйте данные на более устойчивые носители и делайте резервные копии, используя носители хорошего качества
- Организуйте надежное хранение данных, включая возможное внешнее хранение резервных копий
- Регулярно проверяйте данные на наличие ошибок
- Определите режим обновления данных, исходя из срока службы носителей.

Работа над способами обеспечения доступа

- Сохраняйте информацию, необходимую для обеспечения быстрого доступа — материал в первоначальном виде, требования к доступу, пароли и т. д.
- Сохраняйте необходимое оборудование и программное обеспечение доступа, поддерживая аппаратные средства и защитное программное обеспечение в рамках существующей лицензии
- Будут необходимы и другие действия по сохранению доступа, поскольку сохранение аппаратных средств в рабочем состоянии непрактично
- Если дальнейшая деятельность с минимальными программными ресурсами невозможна, планируйте схему передачи материала другому исполнителю для дальнейшего сопровождения
- В качестве альтернативы, найдите возможность отобразить материал в устойчивой, нецифровой форме (например, распечатка). Вероятно, такой подход применим только для файлов, содержащих текст или изображения.
- Минимальные программы могут играть положительную, но тем не менее объективно ограниченную роль в сохранении материалов цифрового наследия.

Минимальная программа — Сценарий 1

Небольшая промышленная фирма сохраняет архив файлов для каждого издания. Архив используется для хранения файлов (в качестве контекстных и ссылочных источников), используемых при повторных изданиях. Потребности компании в доступе к заархивированным материалам — не менее пяти лет. Небольшое количество изданий отвечает возможностям по хранению и должно быть сохранено в цифровой форме в соответствующем подразделении для долговременного хранения. Компания не берет на себя никакой ответственности за сохранение материала.

Архив компании сохранен на компакт-диске в защищенном внутреннем помещении, с дополнительным копированием каждого сохраняемого компакт-диска и хранением в ином месте. Старые архивные файлы сохранены на магнитных и магнитооптических носителях. Материалы с них перенесены на компакт-диски, поскольку магнитные носители недостаточно устойчивы в течение заданного периода, и специализированные дисководы для них постепенно выводятся из обращения.

Архив включает как предварительные, так и окончательные варианты файлов в форматах специализированных частных издательских программ. Компания полагается на совместимость или обновление программных версий, чтобы сохранить доступ к более ранним файлам. Для разработки изображений или справочных файлов, созданы дополнительные версии в форматах типа TIFF или PDF для расширенного доступа. Эти версии предназначены для справочной работы в будущем для того, чтобы обновить рабочие файлы, если более поздние программные версии не смогут и дальше поддерживать более ранние версии.

Минимальная программа — Сценарий 2

Некое исследовательское учреждение обеспечивает программу наследия некоего местного сообщества. Члены сообщества создают некоторый объем фольклора, художественных работ, генеалогических записей и интервью, который размещается в Интернете. Материал отслеживается исследовательским учреждением и имеет огромное значение как история традиций и культуры местного сообщества. Вопрос о правах решен с сообществом, в результате чего материал сохраняется неопределенно долго.

Сайт использует текущий стандарт сети и медиа-файлов, по требованию ведущего учреждения, которое берет на себя первичную ответственность за сохранность сайта и хранит некоторую коллекцию программных инструментов для правильного отображения содержащегося материала.

Учреждение регулярно сохраняет резервные копии всего материала, размещаемого на серверах, включая внешнее хранение не менее одной резервной копии. Отдельные сайты также могут регулярно архивироваться на компакт-дисках, особенно при их обновлениях. Для защиты интерактивного содержания используются стандартные меры защиты информации.

Учреждение располагает фондами для работы программы в течение следующих двух лет и ведет переговоры о пересылке материала другому учреждению для дальнейшего хранения.

Пост-минимальная программа — Сценарий 3

Образовательное и научно-исследовательское учреждение использует программу для накопления диссертаций, опытных данных, исследований и моделей и намеревается хранить эти материалы и обеспечивать доступ к ним неопределенно долго.

Учреждение ограничивает диапазон форматов, в которых принимает материалы, чтобы снизить диапазон способов долгосрочного хранения. Принятые форматы соответствуют открытым, широко используемым стандартам, которые, как ожидается, будут доступны в течение длительного периода. Учреждение сохраняет программное обеспечение, способное в настоящее время отображать или исполнять файлы данных форматов. Источников просят снабжать материалы детальными метаданными.

Учреждение обеспечивает безопасное хранение на дисках и лентах и располагает планом резервирования копий и планом восстановления данных в случае бедствия, включая наличие копий в отдельных помещениях и на иных носителях.

Разрабатываются планы обновления систем и оборудования и проводятся исследования методов перевода или интерпретации устаревших форматов в будущем. Технологическая среда отслеживается на предмет признаков устаревания используемых технологий.

РАЗДЕЛ 4

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ