

УДК 771.53

Н.Е. Проскуряков, С.Ю. Борзенкова, Е.Е. Евсеев, О.В. Чечуга

Современные технологии создания страховых фондов документации

Рассмотрены перспективные технологии создания страховых фондов документации в современных условиях. Определены основные технологии записи и воспроизведения бинарных данных с использованием микрофильма и штрихового кодирования. Приведены варианты оценки эффективности гибридных технологий микрофильмирования.

Ключевые слова: исходный электронный документ, микрофильм, методы записи и воспроизведения бинарных данных, сканирование, декодирование, штрих-код.

Задача создания страховых фондов документации. В настоящее время в России стремительными темпами растут объемы сканирования и оцифровки бумажной документации предприятий и организаций, библиотечных и архивных фондов. Утверждены различные государственные документы, концепции и программы, нацеленные на увеличение электронного документооборота. Практически во всех федеральных органах исполнительной власти и органах исполнительной власти субъектов Российской Федерации завершается переход на использование в своей деятельности электронных документов.

Вместе с тем часть документов, относящихся к особо ценным и особо важным, требует обеспечения их длительного и надежного хранения. Эта задача в настоящее время решается системой Единого российского страхового фонда документации (далее – ЕР СФД) с применением традиционного для этих целей носителя информации – микроформы. Ежегодно увеличивающийся объем электронных документов уже сегодня ставит перед системой ЕР СФД решение вопросов по разработке принципов и методов долговременного сохранения электронных массивов информации.

Положение о ЕР СФД, утвержденное Постановлением Правительства Российской Федерации от 26.12.1995 г. №1253-68, допускает фиксацию массивов конструкторской, технологической, проектной, нормативной, научной, историко-культурной и другой документации, относящейся к ЕР СФД, не только на микроформах, но и на других компактных носителях информации.

Как показали информационные исследования, проводимые регулярно на протяжении последних лет ФГУП «НИИ Репрографии» (г. Тула), в настоящее время для долговременного сохранения различных видов информации в ведущих зарубежных странах применяется два основных подхода – микрофильмирование и оцифровка [1].

Между сторонниками и противниками этих направлений ведутся горячие научные споры. Особую актуальность приобретает вопрос долговременного сохранения электронной информации.

Информационное страхование бумажных документов с помощью классических технологий оптического микрофильмирования, несмотря на некоторый спад объемов, по-прежнему продолжает осуществляться практически во всех странах. Но объективное возрастание в жизни общества роли электронного документооборота и стремительное нарастание объема документов, создаваемых, обрабатываемых и хранимых в электронной форме, диктуют необходимость развития новых подходов и технологических решений, таких как гибридные электронно-микрографические технологии.

Внедрение данных технологий в практику создания долговременно хранимых страховых информационных ресурсов происходит практически повсеместно. Преимущества электронного документооборота хорошо известны – это высокая оперативность поиска и доступа к документам, экономия времени и расходных материалов, возможность обмена документами по различным электронным каналам связи, снижение бюрократической волокиты и т.д.

Однако повсеместное внедрение электронного документооборота влечет за собой ряд серьезных проблем, важнейшей из которых является проблема долговременной сохранности электронных документов в целях их информационного страхования и архивирования. Без решения этого вопроса невозможно гарантировать сохранение и доступность для потомков цифрового интеллектуального, научного и культурного наследия цивилизации.

Возможности долгосрочного хранения электронных документов ограничены частой сменой поколений цифровых носителей и поддерживающих их аппаратно-программных платформ, которые склонны к быстрому устареванию и исчезновению. В поисках выхода из сложившейся ситуации мировым научным сообществом предлагаются различные варианты обеспечения длительности существования электронных документов в цифровой среде.

Самыми распространенными решениями являются миграция документов в новые программные среды и форматы, периодическая многократная перезапись на новые носители, а также эмуляция, т.е. имитация старой программной оболочки на новых операционных системах и оборудовании.

Однако оба данных подхода (миграция и эмуляция) принципиально не выходят за рамки цифровой среды, которая по самой своей природе достаточно динамична, изменчива и нестабильна. Для обеспечения постоянной миграции и эмуляции требуются большие финансовые, организационные и трудовые ресурсы. Кроме этого, проведенные эксперименты показали, что указанные процессы не обеспечивают защиты информации от потерь при частой перезаписи и переформатировании, т.е. не дают гарантии того, что она сохранится в неизменном оригинальном виде.

Поэтому в настоящее время ученые и специалисты обращаются к исследованию и разработке других, более надежных и экономичных стратегий архивирования важнейшей электронной информации с использованием таких технологий долговременного хранения, которые не требуют постоянного обновления и поддержки. И здесь на помощь человечеству снова приходит микрофильм, проверенный и испытанный аналоговый носитель, обладающий огромным потенциалом.

Разработка гибридных способов сохранения информации. В международном стандарте по микрографии долгосрочное сохранение цифровой информации определяется в широком смысле как «действия, необходимые для поддержания доступа к цифровым данным после отказа носителя или смены технологии». По сути, управление хранением цифровых данных состоит в управлении рисками утраты цифровой информации со временем.

Цель управления хранением – обеспечить долговечность цифровой информации в приемлемой форме и гарантировать ее целостность. Для достижения этой цели лучше всего подходит архивный микрофильм как технологически независимый носитель, обеспечивающий гарантированное хранение информации сроком до 500 лет, а также ее неизменность и устойчивость за счет минимального вмешательства в процесс хранения.

Но как совместить аналоговый носитель – микрофильм, и цифровое содержание электронных документов? Для этого в микрографии необходимо осуществить интеграцию цифровых и аналоговых технологий. Принципиальная возможность такой интеграции появилась в начале 70-х годов прошлого века с изобретением СОМ-систем – устройств, позволяющих экспонировать электронную текстовую и графическую цифровую информацию из компьютера на микроформы. Сейчас на современном мировом рынке насчитывается около 20 моделей СОМ-систем ведущих мировых производителей.

Эти системы различаются по принципу записи, типам микроформ, с которыми работают, форматам принимаемых исходных файлов и другим техническим характеристикам, однако все они способны записывать цифровую информацию из компьютера на пленочные носители. Последним достижением в производстве СОМ-систем стала разработка лазерной цветной системы, способной качественно и с высокой скоростью вести запись цифровой информации на цветной микрофильм.

При этом продолжают совершенствоваться и существующие, хорошо зарекомендовавшие себя на рынке СОМ-системы. Так, фирмой Microbox была представлена новая версия изделия Polysom, способная работать с электронными образами документов до формата А0 включительно и в связи с этим являющаяся наиболее пригодным аппаратом для создания СФД для различных отраслей промышленности.

СОМ-системы вместе со сканерами микрофильмов по праву можно назвать ключевым звеном современных электронно-микрографических технологий, своего рода мостом между цифровым и аналоговым мирами. Несколько лет назад несовершенства и недостатки отдельных моделей, а также общая увлеченность стремительным развитием технологий оцифровки дали повод некоторым ученым считать, что микрофильм как носитель безнадежно устаревает, а СОМ-системы необходимы только для локального применения при сохранении специфических видов электронных документов.

Однако неудачи различных стратегий долгосрочного цифрового сохранения заставили исследователей пересмотреть свои взгляды и снова обратиться к традиционному микрофильму, теперь уже

как к носителю для сохранения цифровой информации, долгосрочный и стабильный потенциал которого может быть усилен возможностями современных СОМ-систем.

СОМ-устройства коренным образом изменили способ создания архивных микрофильмов. Вместо использования для создания изображения оптической съемки эта технология считывает бинарные данные оцифрованного изображения и записывает положение каждого пикселя на пленку с помощью лазера (напрямую) или подобных устройств. Вариантом этой технологии являются записывающие устройства, способные переносить на микрофильм изображение с монитора – это стало возможным благодаря разработкам новых графических карт и специальных мониторов с очень высоким разрешением экрана. Современные СОМ-устройства могут принимать большую часть распространенных электронных текстовых и графических форматов, а новые аппараты позволяют улучшить качество вывода при работе с самыми различными оригинальными изображениями.

Важная роль СОМ-систем в современном сохранении цифровых материалов подтверждается официальным принятием и введением в действие в 2009 г. международного стандарта ISO 11506 «Архивирование электронных данных. Компьютерный вывод на микрофильм (СОМ) и запись на оптический диск (СОЛД)». Данный стандарт впервые в мировой практике нормативно закрепляет стратегию долгосрочного архивного сохранения цифровой информации с помощью компьютерной записи на микрофильм для долгосрочного сохранения и на лазерный оптический диск для оперативного использования. Данный стандарт приобретен нашим институтом, переведен на русский язык и используется в работе.

В настоящее время в мире реализуется множество проектов сохранения цифровой информации с использованием СОМ-систем. Известно, что данные устройства широко применяются в библиотеке Конгресса США, различных отраслях Германии, Японии, Швеции, Франции и Великобритании и множестве других инновационных проектах по долгосрочному сохранению цифровой информации в ведущих странах мира.

Что касается России, то, по приблизительным подсчетам, в настоящее время в нашей стране находится в эксплуатации около 50 СОМ-систем различных типов и производителей. Основными потребителями этих устройств являются организации и учреждения, участвующие в создании и наполнении единого российского страхового фонда документации, а также другие организации, осознающие важность долгосрочного страхового сохранения своих информационных активов.

Российский рынок такого рода оборудования представляется достаточно развитым. На нем представлены практически все основные мировые производители СОМ-оборудования, включая «большую тройку» ведущих немецких компаний – SMA, Zeuschel и Microbox.

Российская наука не стоит в стороне от указанных проблем. Так, в нашей стране именно ФГУП «НИИ Репрографии» на протяжении последних лет в интересах национальной безопасности государства теоретически обосновывает, нормативно и методически закрепляет, а также внедряет современные гибридные электронно-микрографические технологии создания, сохранения и использования ЕР СФД, которые позволяют интегрировать традиционные (микрографические) и современные (электронные) способы создания страховых фондов документации различного назначения.

Данные гибридные технологии позволяют долгосрочно сохранять на микрофильме определенные виды цифровой информации, в частности текстовую, фотографическую и чертежно-графическую документацию, созданную как путем оцифровки бумажных оригиналов, так и непосредственно в ЭВМ. Исследования, проводимые в данной области, опираются на твердую государственную поддержку, высокую научную квалификацию сотрудников НИИ Репрографии, передовой зарубежный опыт и парк современного электронно-микрографического оборудования (СОМ-системы, сканеры микроформ), позволяющего проводить различные эксперименты, отрабатывать технологические схемы и моделировать цепочки взаимодействия новых устройств в условиях функционирования системы СФД. При этом сотрудниками НИИ Репрографии осуществляется регулярный мониторинг зарубежной информации по проблеме исследований, осуществляется ее сбор, накопление и анализ.

Благодаря СОМ-системам открываются новые возможности в области долгосрочного сохранения цифровой информации. Современные инновации в сфере СОМ-систем существенно расширяют сферу их применения.

Так, по результатам последних зарубежных исследований теоретически обоснован и экспериментально подтвержден новый подход к сохранению цифровой информации на микрофильмах. Идея такого подхода заключается в следующем.

Любой цифровой документ состоит из набора двоичных данных – битовой информации. Эта битовая информация может быть закодирована в виде двухмерного штрих-кода, состоящего из информационных точек, а далее представлена в виде двухмерного растрового изображения.

Изображение при помощи СОМ-системы сохраняется на микрофильме. При необходимости восстановления информации штрих-кодовые данные считываются с микрофильма сканирующим устройством, а затем декодируются, в результате чего происходит восстановление оригинального электронного документа.

Значение этой технологии заключается в том, что впервые появилась теоретически обоснованная и технологически реализуемая возможность долгосрочно сохранять на микрофильме любую цифровую информацию и документацию.

При этом тип электронного документа не имеет значения, так как все цифровые файлы состоят из набора двоичных данных и соответственно могут быть представлены в виде двухмерных графических штрих-кодов.

Помимо уже осуществляемого сохранения цифровой цветной и черно-белой чертежно-графической, текстовой и фотографической документации, применение данного метода открывает казавшиеся ранее невозможными перспективы сохранения на микрофильмах цифровой аудиовизуальной документации, программных продуктов, трехмерной документации САД-приложений и др., т.е. любого типа цифровых данных.

Сейчас предлагаются различные варианты этого подхода, такие как гибридное хранение, т.е. совместная запись на микрофильм как самого оригинала изображения документа, так и его цифрового штрих-кода, использование цветного микрофильма, что позволит повысить объем записываемых кодированных данных благодаря использованию трех цветных слоев и т.д. Однако принципиальная схема технологии остается такой, как на рис. 1.



Рис. 1. Схема сохранения бинарной информации на микрофильме

Исходный цифровой документ любого типа с помощью программных алгоритмов представляется в виде двухмерного штрих-кодового растрового изображения, которое может восприниматься СОМ-системой. Затем данное изображение экспонируется СОМ-системой на микрофильм, который направляется на хранение. Далее с использованием сканера микрофильмов микрофильм сканируется, отсканированное штрих-кодовое изображение декодируется и происходит восстановление оригинального электронного документа (файла).

Необходимо заметить, что алгоритм кодирования / декодирования снабжен механизмом коррекции ошибок Рида–Соломона, аналогичным тому, который используется при записи / считывании оптических дисков, что повышает надежность считывания и декодирования штрих-кодовой информации.

Предлагается использовать для этих целей следующий вариант такого подхода. Хранение должно осуществляться гибридным способом, т.е. на микрофильм записываются как само аналоговое изображение, так и его цифровой код. По своей природе микрофильм позволяет считывать информацию и человеку, и машине, поэтому он может использоваться как гибридный носитель, сочетая аналоговую и цифровую информацию.

В качестве конкретного носителя предлагается цветной микрофильм производства Pfochrome Micrographic [2]. Для хранения данных на цветной пленке есть свои основания, главное из которых заключается в том, что при хранении можно использовать все три цветовых слоя, благодаря чему увеличится объем сохраняемых данных. Двухмерный штрих-код, в который преобразовываются оригинальные документы, – это растровое изображение, в котором каждая растровая точка представляет собой состояние. Одна растровая точка служит бинарным описанием состояния (максимальная или минимальная оптическая плотность) или описанием состояния более высокого порядка (несколько уровней плотности).

По данным экспериментальных исследований, в которых для записи цветного микрофильма использовалась цветная лазерная СОМ-система нового поколения Archive Laser Recorder, была достигнута достаточно высокая плотность записи информации. Так, при размере точек 15 мкм на шестисотметровом рулоне цветной пленки 35 мм можно сохранить 22 гигабайта данных. При размере точки 12 мкм – 38 гигабайт. При 9 мкм – примерно 70 гигабайт на одном рулоне.

Кажется, что такой объем не составляет конкуренции таким носителям, как, например, жесткий диск. Но не стоит забывать, что при хранении цифровой информации вместимость не всегда является определяющим фактором, особенно по сравнению с долговечностью и стабильностью.

Однако такая технология является достаточно затратной, так как для записи требуются цветная пленка, цветной лазер (цветные СОМ-устройства) и химико-фотографическая обработка цветной пленки, что достаточно дорого. Сканирующее оборудование, необходимое для считывания цветной пленки, также является более сложным и дорогим, чем аналогичное оборудование для черно-белых материалов.

Соответственно если цвет решающего значения не имеет, рациональнее использовать черно-белый микрофильм. Тогда в качестве носителя используется обычный черно-белый микрофильм, а исходные электронные документы (их бинарные данные) кодируются с помощью двухмерного черно-белого графического штрих-кода. Затем эти данные трансформируются в изображение и сохраняются (экспонируются) на микропленку (рис. 2). При воспроизведении бинарных данных микрофильм сканируется, а изображение декодируется с помощью расшифровки отсканированного штрих-кода. В результате снова получается поток бинарных данных, из которых восстанавливается исходный электронный документ.

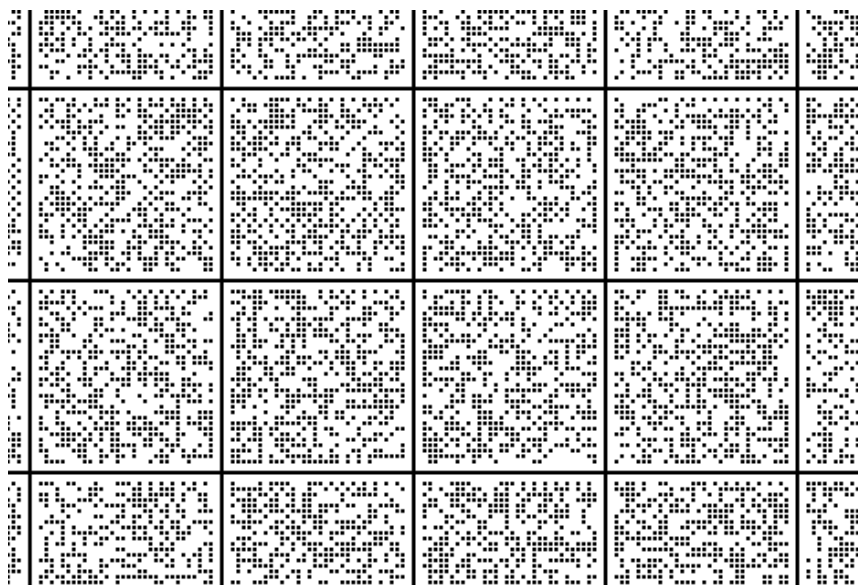


Рис. 2. Увеличенный фрагмент черно-белого штрих-кодowego изображения

Черно-белые штрих-коды позволяют добиться относительно высокой плотности записи информации. Выяснилось, что на одном 16-мм микрофильме длиной 30,5 м в штрих-кодах можно сохранить 7200 изображений формата А4 или 45,32 мегабайт информации (на 35-мм микрофильме соответственно в 2 раза больше). В данном случае стоимость хранения 1 мегабайта составит \$0,28, что в долгосрочной перспективе хранения представляется наиболее оптимальной по сравнению с дру-

гими системами, особенно в сравнении ее со стоимостью миграции каждые 5–7 лет, необходимой для других форматов, и стоимостью их технической поддержки.

Так, например, хранение на современных жестких дисках обходится 0,1...0,3 доллара за 1 гигабайт, но эти технологии требуют значительных затрат в процессе хранения, так как большое количество жестких дисков должно постоянно функционировать, чтобы поддерживать систему в рабочем состоянии. Это требует значительных затрат на электроэнергию, инфраструктуру и техобслуживание на протяжении относительно короткого срока службы.

К тому же в отличие от других носителей, таких как жесткие диски, флеш-карты, CD- или DVD-диски, технологии считывания микрофильма очень просты и универсальны. Тогда как для воспроизведения данных с популярных электронных носителей необходимы специализированные интерфейсы и сложные технологии (оптические диски с лазерной технологией, высокоточное расположение считывающих устройств для магнитных носителей, контролирующие программы и оборудование и т.д.), для считывания данных с микрофильма необходимы только простые оптические устройства. Это выгодно отличает данный носитель от IT-систем. Если найти в будущем устаревший привод для DVD или лент, или USB-порт, совместимый с новыми компьютерными системами, будет очень сложно, то для микрофильма будет достаточно любого современного оптического устройства для формирования изображения – это может быть сканер, камера или другой аппарат.

Заключение. На основе проведенных исследований можно сделать следующие выводы:

1) В свете последних достижений науки технологический потенциал микрофильма и СОМ-систем в деле долгосрочного сохранения цифровой информации представляется очень существенным. Разумеется, что новые технологии требуют совершенствования, исследований и экспериментов по подбору параметров записи, отработке режимов, синхронизации оборудования, оптимизации настроек элементов системы, технико-экономические расчетов и т.д. Однако первые шаги уже сделаны, и дальнейшие исследования возможности применения данного перспективного метода обязательно будут продолжены как за рубежом, так и в нашей стране.

2) Основными моментами, определяющими направления развития работ по информационному страхованию различных видов информации за рубежом, являются следующие:

- Рост тенденции архивирования цифровой информации на микрофильме.
- Снижение доли классического оптического микрофильмирования.
- Развитие технологий цветного микрофильмирования.
- Совершенствование возможностей и улучшение технических характеристик современного микрографического оборудования, такого как СОМ-системы и сканеры микроформ.

3) Сегодня специалисты ведущих стран мира опять обратились к апробированной технологии обработки и сохранения информации – микрографии; правда, это теперь существенно усовершенствованная и обогащенная новыми возможностями технология.

Литература

1. Мировой опыт создания и хранения информационных ресурсов в современных условиях / А.К. Талалаев, Е.Е. Евсеев, П.Е. Завалишин, Н.Е. Проскуряков // Изв. Тул. гос. ун-та. Технические науки. – 2013. – № 3. – С. 408–421.

2. Ilfochrome Micrographic Film [Электронный ресурс]. – Режим доступа: <http://www.yumpu.com/et/document/view/549624/ilfochrome-micrographic-film>, свободный (дата обращения: 07.04.2013).

Проскуряков Николай Евгеньевич

Д-р техн. наук, профессор каф. технологии полиграфического производства и защиты информации
Тульского государственного университета (ТулГУ)
Тел.: 8 (487-2) 35-24-93
Эл. почта: tpzzi@tsu.tula.ru

Борзенкова Светлана Юрьевна

Канд. техн. наук, доцент каф. технологии полиграфического производства и защиты информации ТулГУ
Тел.: 8 (487-2) 35-24-93
Эл. почта: tehnol_sb@tsu.tula.ru

Евсеев Евгений Евгеньевич

Канд. техн. наук, доцент, директор ФГУП НИИ Репрографии, Тула
Тел.: 8 (487-2) 56-97-27
Эл. почта: info@reprograf.ru

Чечуга Ольга Владимировна

Канд. техн. наук, доцент каф. технологии полиграфического производства и защиты информации ТулГУ
Тел.: 8 (487-2) 35-24-93
Эл. почта: sourie_1@mail.ru

Proskuryakov N.E., Borzenkova S.Y., Evseev E.E., Chechuga O.V.

The modern technology of insurance funds documentation

Promising technologies for creating documentation of insurance funds in modern conditions. The basic technology of recording and playback of binary data using microfilm and barcoding. Variants of evaluating the effectiveness of hybrid technology microfilming.

Keywords: original electronic document, microfilm, recording and playback methods of binary data, scanning, decoding, barcode.
