

Еремеев Михаил Сергеевич

Доклад
семинаре «Инновационные технологии и практика в сфере архивного дела. Обмен опытом»
(29 апреля 2011 года, г. Тобольск)

Открытые разработки в IT-сфере как один из методов обеспечения сохранности архивной информации

"IT" или "IT-сфера" - это информационные технологии, которые предоставляют людям много средств для работы со столь актуальной сегодня информацией в электронном виде. Одно из таких средств - программное обеспечение, которое можно разделить на две группы: закрытое, ограниченное патентами (оно же собственническое, проприетарное), и открытое, обычно доступное для использования на условиях свободных лицензий.

Самый известный пример закрытых разработок - операционная система «Windows», условия лицензии которой позволяют устанавливать одну копию только на один компьютер, определяют лимит на количество установок, вводят список запрещенных вариантов использования, в том числе и запрет на изменение программы.

Такие ограничения позволяют корпорации «Microsoft», разработчику «Windows», сохранять полный контроль на том, как именно будет работать программа на оборудовании клиента и на каких условиях это будет происходить, что ставит любого пользователя, будь то человек или организация, в неприятную зависимость. А если завтра «Microsoft» обанкротится и исчезнет, вместе с ней пропадет и вся система поддержки, обновления и лицензирования. Уже имеющиеся копии программ будут медленно устаревать и останутся привязанными к устаревшему оборудованию.

Один из способов избежать подобной зависимости - использование открытого программного обеспечения, которое распространяется на условиях свободных лицензий. Лицензия «GNU GPL» («GNU General Public License» - Универсальная общественная лицензия) - одна из самых распространенных. Она предоставляет право использовать программу с любой целью, изучать ее, изменять, свободно распространять.

Именно на условиях этой лицензии в 1991 году была опубликована операционная система «Linux». Сейчас она - альтернатива «Windows» и существует в различных вариантах (дистрибутивах) от самых разных некоммерческих и коммерческих организаций со всего мира: «Debian», «Fedora», «Gentoo», «Mandriva», «openSUSE», «Red Hat», «Ubuntu». Российские разработчики также есть в этом списке: «ALT Linux», «ASPLinux», «Runtu», «Linux XP».

Это разнообразие разработок дает возможность выбора, предлагает альтернативы, а также снижает ущерб в том случае, если организация-поставщик ОС «Linux» по какой-либо причине уйдет с рынка или прекратит поддержку своего продукта. Кроме того условия свободной лицензии позволяют любому человеку или организации изменять существующие программные продукты, приспособлять их под свои частные задачи.

Так поступило, например, Министерство обороны РФ, которое с 2002 года использует внутри своего ведомства операционную систему «МСВС» (Мобильная система Вооруженных сил), созданную на основе американской «Linux Red Hat». Это дает Министерству обороны возможность самостоятельно управлять своим программным обеспечением, не беспокоясь о навязанных извне лицензионных ограничениях.

Архивная служба России такой самостоятельностью похвастаться не может, даже несмотря на то, что владеет исходным кодом своего главного программного продукта - программы «Архивный фонд».

3-я версия «Архивного фонда» создана на языке программирования «Visual Foxpro», принадлежащего «Microsoft», и поэтому на нее распространяются лицензионные ограничения, предписывающие использовать Foxpro-программы только внутри операционной системы «Windows», хотя технически их запуск возможен и в «Linux».

В 4-ой версии «Архивного фонда» Foxpro больше не используется, но присутствуют иные проприетарные компоненты, также привязанные к «Windows»: система управления базами данных «Microsoft SQL Server», набор средств для работы с языком разметки «Microsoft XML Core Services» и программная платформа «.NET Framework». На все перечисленные компоненты при желании можно найти свободные альтернативы, которые могли бы значительно удешевить разработку и сделать работу программы независимой от операционной системы.

Такую свободную альтернативу «Архивному фонду» создает сейчас специалист по информационным технологиям из Государственного архива социально-политической истории Кировской области. Его программа называется Архивный фонд - WEB. Она основана на языке программирования «Python», языке преобразования XML-документов «XSLT» (eXtensible Stylesheet Language Transformations) и системе управления базой данных «MySQL». Интерфейс представляет собой интернет-страницу, с которой можно взаимодействовать практически из любой операционной системы с любого устройства, хоть с мобильного телефона. Есть возможность переноса данных из 3-ей и 4-ой версии «Архивного фонда». Все сторонние компоненты, использованные автором в программе, бесплатны и распространяются по свободным лицензиям, поэтому архив Кировской области не несет от этой разработки никаких материальных затрат. Будем надеяться, что энтузиасту из Кировской области удастся довести свой проект до завершения.

Теперь вновь посмотрим на 4-ую версию «Архивного фонда». Эта программа умеет формировать учетные документы для вывода на печать, но только посредством проприетарного текстового редактора «Microsoft Word», с которым связаны дополнительные проблемы.

Формат хранения данных «Word» - файлы с расширением DOC (Document). Этот формат - интеллектуальная собственность «Microsoft», поэтому любой такой файл можно рассматривать как арендованный у иностранной коммерческой структуры контейнер для хранения документа, что вызывает вопрос о владении информацией. И до сих пор ни один независимый разработчик не сумел создать программу, которая полностью поддерживала бы этот формат. «Microsoft Word» остается единственной возможностью наиболее комфортной работы с ним. И когда, например, Управление архивами Свердловской области публикует документы на своем сайте в формате DOC, оно тем самым принуждает граждан, желающих ознакомиться с работой архивной службы, к покупке вышеназванной офисной программы, т.е. бесплатно служит интересам иностранной коммерческой организации.

Альтернатива офисным форматам «Microsoft» - открытый, свободный от патентов формат «OpenDocument Format» (ODF). Он принят в качестве международного стандарта 1 мая 2006 года, сейчас стандартизирован и активно используется в странах НАТО, Евросоюза, Южной Африки, Южной Америки и Азии. В России формат ODF станет стандартом 1 июня 2011 года, согласно ГОСТу (Р ИСО/МЭК 26300-2010 «Информационная технология. Формат Open

Document для офисных приложений (OpenDocument) v1.0»), принятому Росстандартом (Российское федеральное агентство по техническому регулированию и метрологии) в декабре 2010 года.

Поддержка ODF реализована в двух десятках офисных программ, часть из которых распространяется бесплатно под свободными лицензиями. Наиболее популярные из них - «OpenOffice», разработанный компанией «Oracle», и «LibreOffice», созданный некоммерческой организацией «The Document Foundation» на основе «OpenOffice».

Одна из самых свободных систем для обмена информацией - сеть Интернет. Но закрытое программное обеспечение там тоже используется и это имеет последствия.

Эти последствия можно продемонстрировать на примере электронного банка документов «Подвиг Народа в Великой Отечественной войне 1941-1945 г.г.» (podvignaroda.mil.ru) основой для которого стала технология «Flash». «Flash» - средство для создания интернет-страниц с мультимедийным содержимым. Оно проприетарное, поэтому разработчики некоторых полностью свободных Linux-систем не включают его поддержку в свои продукты. Также «Flash» принципиально не поддерживается в довольно известной операционной системе «Mac OS», работающей на компьютерах корпорации «Apple». Таким образом, сайт «Подвиг Народа», предназначенный для популяризации архивных документов среди всех граждан страны, тем не менее допускает их дискриминацию по программному обеспечению.

Похожая история произошла и с сайтом «Хроника Победы. 1941-1945 гг.» (www.pobeda-info.ru), где отображение мультимедиа-контента происходит при помощи технологии «Silverlight» от «Microsoft». По понятным причинам поддержка «Silverlight» в ОС «Linux» первоначально отсутствовала и появилась лишь благодаря усилиям независимых энтузиастов и программе «Moonlight», использование которой нарушает патенты «Microsoft».

Наиболее правильно создавать интернет-ресурсы полностью на основе формата HTML (HyperText Markup Language - «язык разметки гипертекста»). Он был разработан в 1991 году как свободный формат для обмена научной информацией, независимый от программного и аппаратного обеспечения. Сегодня страницы большинства интернет-сайтов - это документы в формате HTML, а интернет-браузеры, программы для их просмотра, бесплатны и встроены в большинство операционных систем на самых различных устройствах.

Первые версии HTML умели выводить на экран компьютера лишь форматированный текст и изображения, поддержка же сложной графики, анимации, видео и аудио отсутствовала. Этот недостаток и ликвидировали такие проприетарные технологии, как «Flash» и «Silverlight». Но недавно была опубликована 5-ая версия HTML, которая почти полностью лишена всех прошлых минусов и в силу своей лицензии обречена на популярность, поэтому будущее «Flash» и «Silverlight» довольно туманно, как и будущее интернет-ресурсов на их основе.

Впрочем, программы и форматы хранения данных, предназначенные для работы с мультимедийными, чаще всего проприетарны, но свободные альтернативы имеются.

Полностью свободные форматы для хранения изображений - JPEG и PNG. JPEG был разработан как формат для хранения фотоизображений Объединенной группой экспертов по фотографии (Joint Photographic Experts Group). PNG предназначен для изображений с текстовой информацией, чертежей, схем, планов и карт. Он использует сжатие без потерь, принят в качестве стандарта для обмена графической информацией в сети Интернет и иногда

позиционируется как альтернатива формату TIFF.

Свободная альтернатива популярному проприетарному аудиоформату MP3 - «Vorbis». Это формат хранения аудиоданных и алгоритм их сжатия, созданный некоммерческой организацией «Xiph.Org», которая занимается созданием и продвижением свободного программного обеспечения для работы с мультимедиа-информацией. Чаще всего, аудиофайлы, сжатые методом «Vorbis», имеют расширение OGG.

Для хранения видеоданных «Xiph.Org» предлагает алгоритм сжатия «Theora» с расширением OGV. Именно в формате OGV распространяется, например, свободный документальный фильм «Революционная ОС» («Revolution OS»), дающий хорошее представление о том, что такое свободное программное обеспечение и каковы преимущества его использования.

Именно этот фильм в какой-то мере и пытался заменить данный доклад, преследующий схожие цели. Свободные программы и форматы файлов - вещь, на которую стоит обратить внимание, особенно архивам как государственным организациям.

Федеральные органы исполнительной власти уже обязаны до 2015 года перейти на свободное программное обеспечение согласно плану, утвержденному в декабре 2010 года распоряжением Правительства РФ № 2299-р. Вполне возможно, что органы государственной власти субъектов в ближайшем будущем ждет то же самое.