



ПЕРЕДМОВА

Випуск дайджесту присвячено досвіду установ світу щодо зберігання і використання мікрофільмів та електронної інформації в сучасному інформаційному суспільстві.

У публікації «Дэвид Розенталь: Обзор технологии хранения данных в ДНК» наведено огляд історії та сучасного стану з розробки технологій молекулярного зберігання цифрових даних з використанням ДНК.

У публікації «Электронная архивация – дело, связанное с рисками» розповідається про ризики зберігання електронного контенту, поняття цих ризиків з метою забезпечення його надійного збереження.

У публікації «Круглый стол ВНИИДАД «Актуальные вопросы хранения, комплектования, учета и использования научно-технической документации в организациях»» розповідається про роботу круглого столу в якій прийняли участь більше 50 фахівців різних організацій та установ. Під час роботи круглого столу відбувся обмін досвідом роботи з сучасною науково-технічною документацією, у тому числі і електронною.

У публікації «В России будет создана национальная система управления данными» розповідається про роботу в Росії національної системи управління даними.

У публікації «Франция: Первое практическое руководство по работе с архивными документами» розповідається, що ця книга не є керівництвом по роботі архівів, а є практичним посібником для не фахівців, мета якого - надати поради та рекомендації, які допоможуть тим людям, які хочуть проводити дослідження у фондах архівів, не володія при цьому правилами та тонкощами цього великого багатогранного миру.

У публікації «США: «Конференция Седона» опублікувала для публичного обговорення проект документа по определению юрисдикции при трансграничной передаче персональных данных» наведено стислий зміст документа, принципи визначення законодавства яке буде застосовуватись.

У публікації «Веб-архивация и языки» розповідається, що типовий сучасний веб-сайт не є фіксованим документом; він швидше сервіс або додаток. Більшість сайтів використовують JavaScript-скрипти, які часто завантажуються з зовнішніх сайтів. Сторінки розроблені так, щоб виглядати по-різному в різних умовах. Чи повинен архів вибирати зовнішній вигляд для одного розумного набору параметрів, або ж він повинен спробувати зберегти динамічний характер сторінки? Чи вплине той факт, що сторінку запитує архів, а не інтерактивний браузер, на те, що у відповідь видасть сервер? У статті також розглядаються причини недоліків у виборі мови.

У публікації «ИСО: Спецификации самодостаточного контейнерного формата SIRF для хранения информации» розповідається про публікацію нового стандарту ISO/IEC 23681:2019 «Інформаційні технології – Специфікації самодостатнього формату для зберігання інформації SIRF» .

У публікації «США: Продолжается развитие технологии объёмного сканирования документов» розповідається, що дана технологія здатна читати 9 – 20 сторінок закритої книги, використовуючи Терагерцеві хвилі. В майбутньому відкривається перспектива читання цілих коробів паперу без відкриття кришки.

У публікації «ИСО: Начата работа над стандартом менеджмента рисков, связанных с искусственным интеллектом» розповідається, що цей документ буде містити рекомендації щодо менеджменту ризиків, з якими стикаються організації при розробці та застосуванні методів і систем штучного інтелекту.

У публікації «Ответ на вопрос коллеги: С чего начать свой путь в профессии?» наведено рекомендації для починаючих архівістів та молодих фахівців.

У публікації «Моя «мечта архивиста»» наведено добірку постів, присвячених майбутньому архівів, їх місцю в суспільстві, про професію архівіста та професійне співтовариство, про роль держави.

У публікації «Великобритания: Почему Архивы служат всем» розповідається для обслуговування кого існують архіви, і чому архіви в усьому світі повинні прагнути бути якомога краще.

У публікації «США: Национальные Архивы рассматривают использование блокчейна для проверки подлинности документов в момент роста количества фейковых видео» розповідається про те, що майбутнє блокчейна виглядає багатообіцяючим і що блокчейн може зіграти певну роль в аутентифікації електронних примірників фото- і відеоматеріалів з фондів Національних архівів.



ДЭВИД РОЗЕНТАЛЬ: ОБЗОР ТЕХНОЛОГИИ ХРАНЕНИЯ ДАННЫХ В ДНК

Источник: DSHR's Blog <https://blog.dshr.org/2019/05/review-of-data-storage-in-dna.html>

Луис Сезе (Luis Ceze), Джефф Нивала (Jeff Nivala) и Карин Штраус (Karin Strauss) из Вашингтонского университета и исследовательской группы Microsoft опубликовали захватывающий обзор истории и современного состояния молекулярного хранения цифровых данных с использованием ДНК (Molecular digital data storage using DNA, <https://doi.org/10.1038/s41576-019-0125-3>). В аннотации на статью сказано:

«Молекулярное хранение данных является привлекательной альтернативой для целей плотного и долговечного хранения информации, которое крайне необходимо для преодоления растущего разрыва между производством информации и возможностями для хранения данных. ДНК является ярким примером эффективного архивного хранения данных в молекулярной форме.

В настоящем обзоре мы даём обзор самого процесса, текущего положения дел в этой области и проблем, которые необходимо решить для массового внедрения данной технологии.

Мы также даём обзор области систем молекулярной памяти *in vivo* (в живом организме), которые записывают и хранят информацию в ДНК живых клеток, и которые, вместе с хранением данных в ДНК *in vitro* (вне живого организма, в культуре клеток), принадлежат к расширяющейся области пересечения компьютерных систем и биотехнологий.»

Обзор включает в себя обширную библиографию.

На данном этапе развития технологии крайне полезно иметь авторитетный обзор этой области, особенно для борьбы с ажиотажем, который, по-видимому, всегда возникает вокруг хранения данных в ДНК. Специалисты совместной исследовательской группы Вашингтонского университета и Microsoft (UW/MSFT) по своей квалификации лучше, чем кто-либо, подходят для подготовки такого обзора.

Некоторые, возможно, считают, что я в своих постах преувеличивал сложность продвижения продукта для хранения данных в ДНК на рынок, когда писал (см. <https://blog.dshr.org/2018/02/dnas-niche-in-storage-market.html>):

«Инженеры, перед Вами стоит вызов за менее чем 10 лет увеличить скорость синтеза в четверть триллиона раз при одновременном снижении стоимости в пятьдесят триллионов раз, при этом расходуя не более 24 миллионов долларов в год.» (1 триллион = 10^{12})

Одна из вещей, относительно которых группа UW/MSFT всегда была впечатляюще реалистичной, - это масштаб технологических проблем, с

которыми они сталкиваются. Они более-менее согласны с моей оценкой, когда пишут (см. <https://doi.org/10.1038/s41576-019-0125-3>):

«Текущая результирующая пропускная способность при записи данных в систему хранения на основе ДНК, скорее всего, будет порядка нескольких килобайт в секунду.

По нашим оценкам, система, способная через 10 лет конкурировать с популярными облачными системами архивного хранения, должна будет обеспечивать пропускную способность при записи и чтении, измеряемую гигабайтами в секунду. Это разрыв на 6 порядков в части синтеза и примерно на 2–3 порядка в части расшифровки последовательности генов (sequencing).

Что касается разницы в затратах, то стоимость хранения на ленте в 2016 году составляла около 16 долларов США за терабайт, и ежегодно снижается примерно на 10% в год. Затраты на синтез ДНК, как правило, не раскрываются, но ведущий аналитик отрасли Роберт Карлсон (Robert Carlson) оценивает стоимость синтеза массива приблизительно в 0,0001 долл. США за базовую аминокислоту, что в итоге составляет 800 млн. долл. США за терабайт – т.е. на 7–8 порядков выше, чем у ленты.»

К сожалению, чтобы вытеснить ленту с рынка, недостаточно обеспечить конкурентоспособную стоимость записи по сравнению с лентой. ДНК-хранение должно быть значительно дешевле.

Обзор в академическом журнале - не место для маркетингового анализа, который я провел в своей публикации «Ниша ДНК-решений на рынке хранения» (DNA's Niche in the Storage Market, <https://blog.dshr.org/2018/02/dnas-niche-in-storage-market.html>), поэтому следующие три придирки - это именно придирки.

Во-первых, Сесе и соавторы не учитывают самый важный фактор стоимости, когда пишут:

«Плотность, долговечность и расходы на энергию в состоянии покоя являются основными факторами для архивного хранения, целью которого является хранение огромных объемов данных для использования в будущем в долговременной перспективе.»

Одной из фундаментальных экономических проблем, которую я не обсуждал в статье «Архивные носители информации – не самый процветающий бизнес» (Archival Media: Not A Good Business, <https://blog.dshr.org/2018/03/archival-media-not-good-business.html>, (перевод на русский язык см. https://rusrim.blogspot.com/2018/03/blog-post_15.html), является барьер, связанный с эпидемией ориентации общества на краткосрочную перспективу (short-termism, см. <https://blog.dshr.org/2014/07/discounting-far-future.html>).

Как показала наша работа над экономическими моделями долговременного хранения (The Economics of Long-Term Digital Storage, <http://www.lockss.org/locksswp/wp-content/uploads/2012/09/unesco2012.pdf>), долговечные носители информации, требующие высоких капитальных затрат и затрат на запись, но с низкими эксплуатационными расходами, - находятся в крайне невыгодном положении

по сравнению с недолговечными носителями с низкими капитальными затратами, но более высокими эксплуатационными расходами (включая затраты на регулярную миграцию на носители следующего поколения).

Во многом именно из-за данного барьера лента занимает столь небольшую часть общего рынка средств и систем хранения. В приведенной выше цитате авторам следовало включить «капитальные затраты на систему».

Во-вторых, стоимость роботизированного жидкостного оборудования для записи и чтения в системах ДНК-хранения, вероятно, будет оставаться довольно высокой. Как и в случае с «холодным» Blu-Ray-хранилищем Facebook (см. <https://blog.dshr.org/2014/09/more-on-facebooks-cold-storage.html>), эти расходы необходимо амортизировать посредством хранения большого объема данных. Таким образом, с экономической точки зрения ДНК-хранение, вероятно, лучше всего подойдет для систем масштаба центров обработки данных, что делает маркетинговые проблемы еще более сложными, потому таких потенциальных клиентов очень немного. Все они намного крупнее любого поставщика устройств хранения данных и, ввиду этого, способны отжимать маржу поставщика устройств, как они это делают на рынках жестких дисков и флеш-памяти (см. <https://blog.dshr.org/2019/05/demand-is-even-less-insatiable-than-it.html>).

В-третьих, Сесе и соавторы пишут (см. <https://doi.org/10.1038/s41576-019-0125-3>):

«Использование ДНК для хранения данных обеспечивает плотность до 10^{18} байт на кубический мм, что примерно на шесть порядков плотнее, чем у самых плотных носителей информации, доступных сегодня.»

В своей заметке «Ниша технологий на основе ДНК на рынке хранения информации» (DNA's Niche in the Storage Market, <https://blog.dshr.org/2018/02/dnas-niche-in-storage-market.html>) я сравнил плотность ДНК-носителя с плотностью жёсткого диска:

«Современные жёсткие диски хранят 1,75 ТБ пользовательских данных на пластину в магнитном слое толщиной 20 нм с каждой стороны. Пластина диаметром 95 мм имеет в центре отверстие диаметром 25 мм, поэтому объем слоя, который фактически содержит данные, равен $\pi * (95^{2} - 25^{2}) * 40 * 10^{-6} \cong 1$ кубический мм. Этот том содержит $1,4 * 10^{13}$ используемых битов, соответственно, каждый бит занимает около $7 * 10^{-14}$ кубических мм».

Таким образом, реально нужно сопоставлять $1,25 * 10^{-19}$ мм³ / бит против $7 * 10^{-14}$ мм³ / бит, таким образом, разница составляет около $5,6 * 10^{5}$. Различие в шесть порядков возможно, но оно вводит в заблуждение по двум причинам. Во-первых, сравнивается максимальная теоретически возможная плотность ДНК-носителя с реальной плотностью жёстких дисков 2018 года в массовом производстве. Во-вторых, как я писал в том же посте (<https://blog.dshr.org/2018/02/dnas-niche-in-storage-market.html>), плотность устройств хранения намного ниже, чем плотность «первичной» среды хранения. Для жесткого диска:

«Накладные расходы на упаковку, окружающую первичные биты, примерно в полмиллиона раз больше, чем на сами биты. Если бы эти издержки можно было устранить, мы могли бы хранить 7 экзабайт в накопителе 3,5-дюймового формата.

В системе ДНК-хранения будут аналогичные накладные расходы».

Насколько большими окажутся накладные расходы на упаковку, зависит от ряда вопросов проектирования системы, которые еще предстоит решить, но в конечном итоге системы ДНК-хранения вряд ли будут в миллион раз плотнее, чем их конкуренты.



ЭЛЕКТРОННАЯ АРХИВАЦИЯ – ДЕЛО, СВЯЗАННОЕ С РИСКАМИ

Источник: блог Национальных Архивов Великобритании
<https://blog.nationalarchives.gov.uk/digital-archiving-is-a-risky-business/>

Дэвид Андедаун - старший электронный архивист в департаменте электронной архивации Национальных Архивов Великобритании. Получил образование в области ИТ-технологий, в Национальных Архивах работает с 2005 года. Имеет степень бакалавра по математике от Имперского колледжа в Лондоне.

Данная заметка публикуется в рамках Международной недели архивов (<https://www.ica.org/en/international-archives-week-2019>), основной темой которой является «Проектирование Ваших архивов в 21-м веке».

Обеспечение долговременной сохранности электронного контента целиком и полностью связано с пониманием рисков и управлением ими. По сути дела это ничем не отличается от того, как мы управляем нашими более старыми физическими коллекциями.

Например, то, что хранилища Национальных Архивов располагаются в местечке Кью (Kew), связано со своим собственным риском: возможен разлив Темзы и наводнение! Однако здание Национальных Архивов имеет различные конструктивные особенности, предназначенные для того, чтобы уменьшить вероятность попадания паводковой воды внутрь и причинения ущерба коллекциям. Даже наши замечательные пруды отчасти созданы для того, чтобы помочь контролировать сток воды.

Цифровые риски

В то время, как для электронно-цифровых коллекций существует ряд физических рисков (таких, например, как повреждение жестких дисков и магнитных лент), большинство цифровых рисков труднее наглядно себе представить.



Для данных, хранящихся на магнитных носителях, существует крохотный риск того, что отдельные магнитные частицы, с помощью которых записываются данные, «перевернутся». В результате считываемое с носителя значение может измениться с единицы на ноль (или наоборот). Для некоторых файлов это может не иметь большого значения:

- В простом текстовом файле в этом случае изменился бы один символ, внося тем самым небольшую ошибку;
- В несжатых графических образах (tiff-файлы) мог измениться бы цвет одного пикселя.

Однако для других типов файлов, таких как zip или jpeg, подобное изменение может оказаться более существенным. Может даже так случиться, что файл вообще невозможно будет открыть.

Одной лишь уверенности в том, что файл удастся открыть, может быть недостаточно. Если этот файл не является в точности идентичным тому файлу, который мы первоначально приняли на хранение, можно ли считать его аутентичным документом? Будут ли доверять нам исследователи, если мы не сможем доказать аутентичность наших документов?

Мы можем защищаться от таких рисков, сохраняя несколько экземпляров файла. Мы также создаем для каждого файла контрольные суммы (*сейчас это чаще всего хеши, которые имеют ряд дополнительных полезных свойств, и в первую очередь то, что хеш каждого документа уникален*), выполняющие роль цифровых «отпечатков пальцев», которые позволяют показать, что файл не изменился.

Понимание рисков

Мы очень хотим найти способ наглядно увидеть, как взаимодействуют все эти различные риски. В идеале такая картина также показала бы нам, какие риски наиболее велики - и это могут оказаться совсем не те риски, которые мы традиционно считаем важными. Или же мы могли бы увидеть, что защиту от ряда рисков можно обеспечить легко и дешево, и суммарная отдача от этого могла бы оказаться больше, чем смягчение одного большого риска.

Как нам кажется, такую возможность может нам дать байесовская динамическая сеть (Bayesian Dynamic network, https://en.wikipedia.org/wiki/Dynamic_Bayesian_network). Для изучения возможностей этого подхода мы провели ряд внутренних работ по статистическому моделированию. В интересах дальнейшего продвижения в данном вопросе мы начали сотрудничать с экспертами из группы прикладной статистики и рисков Университета Уорика (University of Warwick, см. <https://warwick.ac.uk/fac/sci/statistics/asru/>). Мы стремимся к тому, чтобы в нашей работе был представлен широкий спектр точек зрения.

Если Вас интересует дополнительная информация об этой работе, посетите веб-сайт британской Коалиция по электронной сохранности (Digital Preservation Coalition, DPC) по адресу <https://www.dpconline.org/blog/creating-an-evidence-base>.

Дэвид Андедаун



КРУГЛЫЙ СТОЛ ВНИИДАД «АКТУАЛЬНЫЕ ВОПРОСЫ ХРАНЕНИЯ, КОМПЛЕКТОВАНИЯ, УЧЕТА И ИСПОЛЬЗОВАНИЯ НАУЧНО-ТЕХНИЧЕСКОЙ ДОКУМЕНТАЦИИ В ОРГАНИЗАЦИЯХ»

Источник: сайт ВНИИДАД / YouTube
http://www.vniidad.ru/index.php?option=com_content&view=article&id=2610:-19-20-2019-l-r-&catid=96:sample-news&Itemid=453
https://www.youtube.com/watch?v=_z2qXPJtm6Y

Автор: Наташа Храпцовская

30 мая 2019 года во ВНИИДАД прошёл круглый стол на тему «Актуальные вопросы хранения, комплектования, учета и использования научно-технической документации в организациях». Принять участие в нём могли все желающие, и таких нашлось много – зал был заполнен полностью.

Для справки: «План работы Федерального архивного агентства на 2019 год» утвержденный руководителем Федерального архивного агентства 26 декабря 2018 года предусматривает, что ведомство будет принимать участие в нормативном обеспечении работы с научно-технической документацией и электронными документами, в том числе в рамках реализации плана мероприятий («дорожной карты») по совершенствованию законодательства и устранению административных барьеров в целях

обеспечения реализации Национальной технологической инициативы по направлению «Технет» (передовые производственные технологии) (см.: http://rusrim.blogspot.com/2019/02/blog-post_41.html)

До декабря 2019 года Росархив должен выпустить приказ «О правилах организации хранения, комплектования, учета и использования документов Архивного фонда Российской Федерации и других архивных документов в органах государственной власти, органах местного самоуправления и организациях (включая хранение, учет и использование научно-технической документации, в том числе в электронном виде)».

Новые нормы правил должны быть направлены на оптимизацию «нагрузки на хозяйствующие субъекты по долгосрочному хранению электронной документации, обеспечение доступа хозяйствующих субъектов к необходимым электронным техническим архивам, обеспечение технологических возможностей последующего использования документации, хранящейся в электронном техническом архиве».

По данному пункту Федеральное архивное агентство является основным исполнителем. Приказом Росархива от 28 декабря 2018 года № 217 утвержден «План научно-исследовательских и опытно-конструкторских работ, выполняемых на основе государственного задания» в котором запланирована разработка проекта «Правил организации хранения, комплектования, учета и использования научно-технической документации в организациях» (п.1.1.). Исполнять эту работу будет ВНИИДАД.

В работе круглого стола приняли участие более 50 специалистов из организаций различных ведомств, в.т.ч. госкорпорации Росатом и Роскосмос, ОАО РЖД, Российская Академия Наук и др. Они поделились опытом своей работы с современной научно-технической документацией, в том числе и электронной.

Была организована онлайн-трансляция круглого стола для всех желающих. Видеозапись круглого стола выложена на YouTube, СМ.https://www.youtube.com/watch?v=_z2qXPJtm6Y

Мой комментарий: Данное мероприятие было очень интересным, поскольку на нём в основном выступали практики, которые делились собственным практическим опытом.

ВНИИДАД планирует наладить сотрудничество со специалистами по научно-технической документации для подготовки проекта правил, и я желаю успеха всем участникам этой нужной работы!



В РОССИИ БУДЕТ СОЗДАНА НАЦИОНАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ДАННЫМИ

Источник: Консультант Плюс <http://government.ru/news/36857/>
<http://government.ru/news/36838/#noskov>

Автор: Наташа Храмцовская

На заседании Правительства 29 мая 2019 года был рассмотрен вопрос о национальной системе управления данными. Она будет создана в рамках федерального проекта «Цифровое государственное управление» национальной программы «Цифровая экономика Российской Федерации». Эта система будет обеспечивать управление, интеграцию, совместимость и мониторинг качества данных, содержащихся в уже существующих государственных информационных ресурсах, на основе единой информационной платформы.

С докладом на заседании выступил Министр цифрового развития, связи и массовых коммуникаций Российской Федерации К.Ю.Носков.

В ближайшее время Правительство утвердит:

- Концепцию создания и функционирования национальной системы управления данными;
- План мероприятий («дорожная карта») по созданию национальной системы управления данными.

Для справки: На заседании были приведены следующие данные: На сегодняшний день в России функционирует более 800 информационных государственных ресурсов. По оценкам Минкомсвязи, в них хранится более 80 петабайт информации, и в течение следующих пяти лет ожидается удвоение этого объёма. В одном только Едином государственном реестре недвижимости содержится более 300 млн записей; система межведомственного электронного взаимодействия функционирует уже девять лет, и в прошлом году через эту систему прошло более 32 млрд. транзакций.

Министр подчеркнул, что теоретически всё это позволяет нам принимать решения, основанные на данных, но на практике, к сожалению, не всё всегда так радужно. В своем докладе Носков отметил, что:

- Сведения, содержащиеся в важнейших государственных ресурсах, зачастую недостоверны, устарели или просто некачественны;
- Права доступа к информации регулируются отдельными соглашениями и не объединены общими правилами;
- Нет достоверной информации о составе и структуре данных, имеющихся в различных ведомствах.

Подходы к решению этой проблемы были изложены в представленной к рассмотрению концепции создания и функционирования национальной системы управления данными (НСУД).

Органам государственной власти предстоит определить:

- Ключевые, базовые информационные ресурсы, содержащие данные о важнейших сущностях: о человеке, о недвижимом имуществе, о транспортных средствах, о юридических лицах;
- В каких информационных системах или ключевых ресурсах будут находиться эталонные данные, и обеспечить «привязку» всех других систем с ними

Минкомсвязь совместно с Аналитическим центром при Правительстве, Минэкономразвития и Минфином проведёт аудит информационных систем госорганов, сформирует план их развития, обеспечит их модернизацию.

НСУД должна определить:

- Формат хранения информации и данных;
- Структуру и состав данных;
- Кто хранит эталонные данные;
- Кто имеет доступ к тем или иным данным.

Всё это должно быть реализовано в рамках единой информационной платформы НСУД.

Одновременно на заседании был одобрен проект постановления Правительства о проведении эксперимента по повышению качества связанности данных, содержащихся в государственных информационных ресурсах.

Эксперимент будет проведен в 2019 году по пяти направлениям, чтобы подтвердить правильность выбора методологии управления различными данными и протестировать компоненты этой создаваемой единой платформы:

- Формирование федерального регистра избирателей;

Для справки: В настоящее время существует 85 избирательных комиссий субъектов, почти 2,8 тыс. территориальных избирательных комиссий, более 92 тыс. участковых избиркомов. Зарегистрировано 108 миллионов избирателей на территории России, ещё 2 миллиона за её пределами.

Регистр избирателей сегодня ведётся путём физической передачи сведений на машиночитаемых или даже бумажных носителях от нескольких ведомств, после чего они сводятся органами местного самоуправления. На передачу и обобщение сведений отводится до 90 дней. Это обусловлено низким уровнем автоматизации всего процесса, что, конечно, значительно увеличивает сроки обновления данных и административную нагрузку. Результатом эксперимента станет автоматический сбор сведений для федерального регистра избирателей из имеющихся информационных ресурсов.

- Формирование цифрового профиля гражданина. Он позволит человеку передавать юридически значимую информацию о себе в государственные органы и бизнес структуры;

- Эксперимент с Федеральной информационной адресной системой ФИАС

Для справки: Сегодня адресная информация содержится в Росреестре, ФНС, частично в Росстате, Росавтодоре, Минюсте, муниципалитетах. В рамках этого эксперимента будет обеспечена актуальность и однозначность адресных сведений.

- Формирование паспорта муниципальных образований на основе имеющихся информационных ресурсов Росстата, Федерального казначейства, ЦИК и Минэкономразвития;

- Создание цифровой аналитической платформы для снижения отчётной нагрузки и упрощения процесса сбора статистических данных бизнесом и населением. В рамках эксперимента планируется сформировать единое информационное пространство сбора статистической отчётности Росстата и Минсельхоза и исключить дублирование собираемой информации.

Мой комментарий: Вообще-то распоряжением Правительства Российской Федерации от 15 апреля 2011 г. № 654-р был утвержден перечень базовых государственных информационных ресурсов (см.: http://rusrim.blogspot.com/2011/04/blog-post_26.html).

Документ утратил силу с 1 января 2013 года в связи с изданием постановления Правительства РФ от 14 сентября 2012 года № 928 «О базовых государственных информационных ресурсах» (см.: http://rusrim.blogspot.com/2012/10/blog-post_3.html), которым были утверждены:

- Требования к порядку формирования, актуализации и использования базовых государственных информационных ресурсов;

- Правила формирования, актуализации и использования реестра базовых государственных информационных ресурсов.

Эти документы утратили силу с 1 января 2019 года в связи с изданием постановления Правительства № 478 от 1 июня 2016 года «О первоочередных мерах, направленных на создание государственной информационной системы «Единая информационная среда в сфере систематизации и кодирования информации»», которым были утверждены «Правила создания, изменения, ведения и применения отдельных информационных ресурсов» (см.: http://rusrim.blogspot.com/2016/06/blog-post_96.html).

ФРАНЦИЯ: ПЕРВОЕ ПРАКТИЧЕСКОЕ РУКОВОДСТВО ПО РАБОТЕ С АРХИВНЫМИ ДОКУМЕНТАМИ

Источник: сайт «Архивоведческие хроники» (Chroniques chartistes)
<https://chartes.hypotheses.org/5273>

Автор: Наташа Храмцовская

В Франции было опубликовано практическое и методическое руководство по работе с архивными документами. Несмотря на важную роль исторической и архивной литературы в данной области, - это первая публикация такого рода.



Джонатан Барбье (Jonathan Barbier) и Антуан Мандрет-Дегейл (Antoine Mandret-Degeilh) «Работа с архивными документами. Практическое руководство» (Le travail sur archives. Guide pratique), Париж: из-во Armand Colin, 2018 год.

Авторы книги Джонатан Барбье и Антуан-Мандрет-Дегейл - не архивисты, а научные сотрудники и преподаватели Института политических исследований в Тулузе (Science-Po Toulouse), причём специализацией первого из них является история, а второго - политология. Они отмечают, что взяли за основу ряд методологических руководств, существующих в других социальных науках: руководство Говарда Беккера (Howard S. Becker) в социологии [1] и Стефана Бода (Stéphane Beaud) и Флоранс Вебер (Florence Weber) в этнологии [2].

Данная книга - не руководство по работе архивов, а практическое руководство для неспециалистов, цель которого - дать советы и рекомендации, которые помогут тем людям, которые хотят проводить

исследования в фондах архивов, не зная при этом правил и тонкостей этого большого многогранного мира, охватывающего сотни архивных служб и сотни погонных километров документов.

Существуют руководства для пользователей, задача которых - ориентировать читателей в лабиринте фондов, такие, например, как справочник по фондам Национальных Архивов Франции, первое издание которого вышло в 1964 году и который затем регулярно обновлялся до 1993 года [3]. Другие руководства по архивным исследованиям для трех исторических периодов (история средних веков, эпохи возрождения и новая история) были опубликованы под руководством Пола Делсалля (Paul Delsalle) [4].

Но до сих пор не было «сквозного» справочника по всем французским архивам. Рекомендации Джонатана Барбье и Антуана-Мандре-Дегейля вращаются вокруг трёх видов использования архивных документов: для научных целей и в журналистике; для генеалогических исследований и для использования в государственном управлении. Книга делится на четыре части и восемь глав.

Первая часть книги является вводной и даёт общее представление о ней. Первая глава является терминологической, она разъясняет юридические термины французского права и технические термины архивной терминологии, определяющей понятие «архивов, архивных документов».

Затем во второй главе рассказывается о том, где сохраняются архивные документы, с учетом их типологического, организационного и географического разнообразия: это государственные службы на разных уровнях государственного управления (национальном, региональном, муниципальном и окружном). Описана их история, текущая организация и системы классификации. Не забыты также библиотечные службы (Национальная библиотека Франции, муниципальные библиотеки, библиотеки музеев и университетов) и частные центры хранения (компаний, профсоюзов, политических партий, ассоциаций), что справедливо, поскольку они тоже хранят архивные документы.

Вторая часть книги посвящена подготовке к проведению исследования: от прочтения основополагающих руководств до просмотра онлайн-каталогов, заранее определяется, какой тип архивных документов соответствует целям исследования, а затем – где эти документы могут храниться.

Для трёх видов использования архивных документов - научного/журналистского, генеалогического и административного – даются ссылки на соответствующие печатные ресурсы, такие, как генеалогические руководства по исследованиям, и электронные, такие как доступные в Интернете путеводители. Ссылки на эти инструменты приводятся в приложении, и подробно описана навигация по этим веб-сайтам.

Идея дать рамочную тематическую классификацию, охватывающую различные архивные службы, особенно полезна для понимания взаимодополняемости подлежащих исследованию источников, что иллюстрируется на примере судебных архивов.

Третья часть книги посвящена работе с архивными документами и взаимодействию с соответствующими архивными службами: работе с аналитическим и синтетическим научно-справочным аппаратом в читальном зале или с помощью виртуального зала с описями, определению архивных шифров.

Даются советы о том, как эффективно вести поиск: это использование ключевых слов, ранжирование по приоритету и важности нужных архивных шифров, формирование репрезентативной выборки в случае, если весь массив документов слишком велик, ведение исследовательского журнала с указанием уже просмотренных архивных шифров, чтобы к ним не обращаться снова и не тратить зря время, и, наконец, оценка возможности раскрытия интересующих исследователя документов.

Обсуждаются даже практические аспекты ознакомления с документами: преимущества и недостатки конспектирования вручную или с помощью компьютера, а также преимущества и недостатки фотокопирования в сравнении с архивным описанием дел (*dépouillement des cartons*).

В руководстве также рассматриваются научные аспекты работы с документами, которая не ограничивается его прочтением: изложены принципы и методы внутренней и внешней критики, а также основные правила транскрибирования. Одной из основных проблем, с которыми сталкиваются пользователи-новички, является знакомство со сроками раскрытия для общественности современных архивных документов. Интересующие исследователя документы могут оказаться государственными архивными документами, доступ к которым ещё не открыт, или же которые всё ещё хранятся в организациях-создателях или в труднодоступных частных архивах. Руководство предлагает ряд советов относительно того, как общаться с хранителями документов; как формулировать свои запросы, подаваемые компетентным лицам; как строить стратегию переговоров с государственными администрациями и с хранителями архивов.



В том же 2018 году было издано еще одно пособие, облегчающее доступ к архивам французских департаментов (*это единица административного деления Франции – Н.Х.*): это рекомендации, написанные Яном Гильермом (Yann Guillerme) и Мари-Одиль Мерняк (Marie-Odile Mergnac), являющимися членами генеалогических ассоциаций - в новой обновленной версии по сравнению с первым изданием 2017 года [5]. Это руководство состоит из трех частей. Первая содержит краткое описание миссии архивов департаментов, сохраняемых ими фондов, с указанием основных серий, наиболее интересных с точки зрения генеалогических исследований. Также даются общие рекомендации по использованию веб-сайтов и организации работы исследователя в читальном зале.

Вторая, самая ценная часть представляет собой перечень, включающий 101 архив департамента и зарубежных территорий в алфавитном порядке кода Insee (*код, используемый французским Национальным институтом статистики и экономических исследований для идентификации территориальных образований – Н.Х.*). Для каждого архива указаны его координаты, часы работы, функциональные возможности и ресурсы веб-сайта, особенно наличие на сайте оцифрованных документов.

Правда, не упоминаются некоторые особенности определённых веб-сайтов, как, например, наличие руководства по работе в архиве - за исключением упоминания многочисленных ресурсов архивов департамента Верхняя Вьенна (Haute-Vienne) [6]. Например, в архивах департамента Морбиан (Morbihan) есть онлайн-конвертер перевода дат из революционного календаря в григорианский [7]. Архивы департамента Приморские Альпы (Alpes-Maritimes) предлагают исследователям многочисленные справочники и путеводители (карты по исторической географии, инструменты пересчета мер и весов, биографический словарь, глоссарий итальянских слов) [8]. Наконец, архивы департамента Атлантические Пиренеи (Pyrénées-Atlantiques) предлагают онлайн-глоссарий учреждений, налогов, религиозных терминов и архивной лексики [9]. Эти инструменты могут быть полезны для всех исследователей, а не только для тех, кто регулярно посещает архивную службу своего департамента. Кроме того, наиболее важные элементы системы классификации архивных документов, глоссарий и таблица сроков раскрытия архивных документов для общественности помещены в конце книги.

Несомненно, что публикация Джонатана Барбье и Антуана Мандрет-Дегейля является и останется в будущем полезным справочником для любого посетителя архивов: ясность изложения и актуальность многочисленных советов и упомянутых в ней инструментов позволяют ей эффективно помогать пользователям архивных документов в их исследованиях. Важные моменты хода исследований выделены в тексте, и каждую главу завершает сводка основных положений с тем, чтобы подчеркнуть основные идеи, которые необходимо взять на вооружение. Такой подход авторов также нашёл отражение и в приложениях, содержащих информацию первостепенной важности: сроки раскрытия документов для общественности,

типовой бланк запроса на раскрытие документов в виде исключения, руководство по использованию устройства чтения микроформ, фотографии ряда видов архивных документов и выписка из описи. Эту информацию дополняют библиография и список интернет-ресурсов. Мы очень надеемся, что данное руководство приведет в архивы новых читателей.

Литература:

[1] Howard S. Becker « Les ficelles du métier : comment conduire sa recherche en sciences sociales » (пер. с английского издания Sociological Work: Method and Substance – «Социологическая работа: Метод и сущность», 1977 г.), 2002 г.

[2] Stéphane Beaud, Florence Weber «Guide de l'enquête de terrain : produire et analyser des données ethnographiques», Paris : Éditions de la Découverte, 1997.

[3] Gérard Ermisse, Philippe Le Tréguilly, «Guide du lecteur», Paris, Archives nationales, 1993, 6-е издание.

[4] Paul Delsalle (ред.) «La recherche historique en archives du Moyen-Âge», Gap, Orphys, 1995; «La recherche historique en archives: XVIe, XVIIe, XVIIIe siècles», Gap, Orphys, 2007; «La recherche historique en archives: XIXe et XXe siècle: de 1789 à nos jours», Gap, Orphys, 1996.

[5] Yann Guillerm, Marie-Odile Mergnac «Archives départementales mode d'emploi», Paris, Archives and Culture, 2018.

[6] Сайт архивов департамента Верхняя Вьенна, см. <http://archives.haute-vienne.fr/r/14/boite-a-outils/>

[7] Сайт архивов департамента Морбиан, см. <https://archives.morbihan.fr/rechercher/conseils-et-outils-de-recherche/calendrier-republicain/>

[8] Сайт архивов департамента Приморские Альпы, см. <https://www.departement06.fr/outils-de-recherche-et-archives-numerisees/aide-a-la-recherche-2896.html>

[9] Сайт архивов департамента Атлантические Пиренеи, см. <http://archives.le64.fr/faire-une-recherche/preparer-vos-recherches/outils-pratiques/glossaires.html>

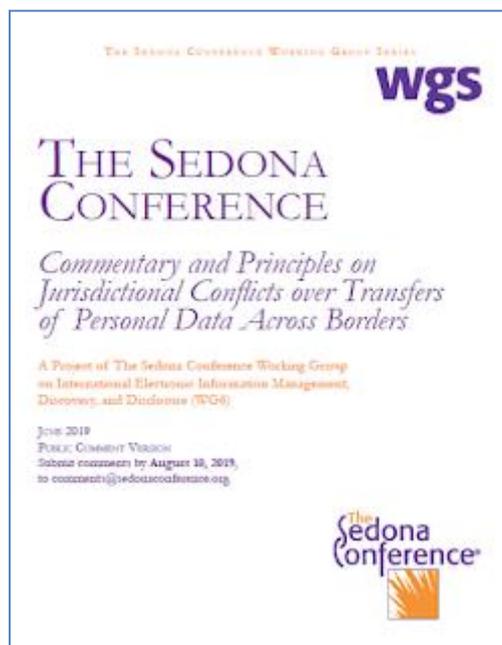


США: «КОНФЕРЕНЦИЯ СЕДОНА» ОПУБЛИКОВАЛА ДЛЯ ПУБЛИЧНОГО ОБСУЖДЕНИЯ ПРОЕКТ ДОКУМЕНТА ПО ОПРЕДЕЛЕНИЮ ЮРИСДИКЦИИ ПРИ ТРАНСГРАНИЧНОЙ ПЕРЕДАЧЕ ПЕРСОНАЛЬНЫХ ДАННЫХ

Источник: сайт Конференции Седона

https://thesedonaconference.org/publication/Commentary_and_Principles_on_Jurisdictional_Conflicts_over_Transfers_of_Personal_Data_Across_Borders

Как сообщил 11 июня 2019 года сайт «Конференции Седона» (Sedona Conference®) – очень авторитетного американского некоммерческого правового идейного центра, в основном занимающегося вопросами раскрытия в ходе судебных разбирательств и расследований сохраняемой в электронном виде информации (э-раскрытия), - для публичного обсуждения выложен проект «Комментария и принципов Конференции Седона в отношении юрисдикционных конфликтов в ходе трансграничной передачи персональных данных».



«Конференция Седона» и ее рабочая группа WG6 по управлению, выявлению и раскрытию электронной информации в международном контексте (Working Group 6 on International Electronic Information Management, Discovery, and Disclosure) рады объявить о том, что для общественного обсуждения опубликован **«Комментарий и принципы Конференции Седона в отношении юрисдикционных конфликтов в ходе трансграничной передачи персональных данных»** (The Sedona Conference

Commentary and Principles on Jurisdictional Conflicts over Transfers of Personal Data Across Borders).

Цель данного «Комментария» заключается в том, чтобы:

- Предоставить практическое руководство для корпораций и других заинтересованных лиц, которым приходится ежедневно принимать оперативные решения относительно трансграничной передачи данных;
- Обеспечить концептуальную основу для анализа вопросов относительно того, какие законы применимы при трансграничной передаче персональных данных.

Документ объёмом 48 страниц можно бесплатно скачать с веб-страницы сайта Седона по адресу <https://thesedonaconference.org/download-publication?fid=3959> (как правильно подсказал один из читателей, свои персональные данные можно не давать, достаточно согласиться с условиями использования, поставив «галочку»).

Публичное обсуждение продлится до 10 августа 2019 года. Вопросы, замечания и предложения в отношении «Комментария» следует направлять по адресу comments@sedonaconference.org. Редакционная группа внимательно рассмотрит все поступившие замечания и предложения и определит, какие изменения будет уместно внести в окончательную версию документа.

Вебинар по «Комментариям» будет запланирован на ближайшие недели. О нём будет объявлено по электронной почте и на веб-сайте Конференции Седона. Вебинар даст Вам возможность задать вопросы и получить дополнительную информацию по этой важной теме.

Комментарий: Содержание документа следующее:

Сводка принципов определения применимого законодательства

I. Введение

II. Принципы определения применимого законодательства

Приложение: Текущая ситуация и проблемы обеспечения защиты персональных данных

Предлагается 6 следующих основных принципов:

Принцип 1. Страна обладает исключительной юрисдикцией и может применять свои законы о защите неприкосновенности частной жизни и персональных данных в отношении физических лиц и организаций, которые находятся или ведут деловую деятельность на её территории, независимо от того, происходит ли обработка соответствующих персональных данных на её территории.

Принцип 2: Страна обычно имеет исключительную юрисдикцию и может применять свои законы о защите неприкосновенности частной жизни и персональных данных в отношении обработки персональных данных, неразрывно связанной с её территорией.

Принцип 3: В коммерческих транзакциях, в которых договаривающиеся стороны имеют сопоставимые позиции на переговорах, юрисдикция или применимое законодательство в отношении обработки персональных данных в связи с соответствующей коммерческой сделкой

должны определяться осознанным выбором, сделанным сторонами договора, и такой выбор должен уважаться, пока он находится в разумной связи между сторонами и сделкой.

Принцип 4. За пределами коммерческих сделок, когда физическое лицо свободно делает выбор юрисдикции или применимого законодательства, это лицо не должно лишаться его/её защиты, которая в противном случае была бы применима в отношении его/её данных.

Принцип 5: Данные в процессе передачи (Data in Transit) из одной суверенной страны в другую должны подчиняться юрисдикции и законам суверенной страны, из которой эти данные происходят, таким образом, что при отсутствии чрезвычайных обстоятельств обращение с данными было таким же, как если бы они всё ещё находились в месте своего происхождения.

Принцип 6: Если персональные данные, находящиеся или иным образом подпадающие под юрисдикцию или законы суверенной страны являются существенными для судебного иска, расследования или иного судебного разбирательства в другой суверенной стране, такие данные в обязательном порядке должны предоставляться, если обеспечиваются адекватные меры защиты, регламентирующие использование, распространение и уничтожение данных.

Ключевым является последний шестой принцип. Говоря попросту, если американское правосудие или правоохранительные органы хотят добраться до персональных данных граждан другой страны и/или хранимых на территории другой страны, то они имеют на это полное право – достаточно лишь «поклониться мамой», что будут приняты надлежащие меры защиты этих данных. Фактически американцы отказываются от каких-либо уступок в их давнем конфликте с Евросоюзом.



ВЕБ-АРХИВАЦИЯ И ЯЗЫКИ

Источник: блог «Mad File Format Science»
<https://madfileformatscience.garymcgath.com/2019/06/10/web-archiving-languages/>

Заметка профессионального разработчика программного обеспечения Гэри Мак Гэфа была опубликована 10 июня 2019 года на его блоге «Сумасшедшая наука о файловых форматах» (Mad File Format Science), который посвящён «файловым форматам, проверке структуры файлов, программному обеспечению для архивации и прочим странным вещам».

Веб-архивация – дело непростое. Немногие сайты состоят исключительно из статического самодостаточного контента. Большинство

используют JavaScript-скрипты, часто загружаемые с внешних сайтов. Адаптирующиеся страницы разработаны так, чтобы выглядеть по-разному в разных условиях. Архив должен сделать снимок веб-страницы, который отражает её внешний вид в определённый момент времени, но что именно это означает? Должен ли архив выбирать внешний вид для одного разумного набора параметров, или же он должен попытаться сохранить динамический характер страницы? Повлияет ли тот факт, что страницу запрашивает архив, а не интерактивный браузер, на то, что в ответ выдаст сервер?

Какой у меня язык?

Одна из менее очевидных проблем связана с тем, на каком языке страница заархивирована. Здесь есть два аспекта. Один из них - язык, объявляемый атрибутом **lang**. Тег `<html>` должен иметь атрибут lang; и его также могут иметь другие элементы - возможно, указывая другой язык. Второй аспект - это язык, используемый для контента. Архивы не должны быть полностью ориентированными на английский язык, и указанные в HTTP-запросе допускаемые языки (accepted languages) будут влиять на возвращаемый контент.

Дополнительным осложнением, которое я наблюдал во время моей недавней поездки в Германию, является то, что некоторые сайты игнорируют заголовок «Допускаемые языки» (Accept-language), предпочитая ориентироваться на IP-адрес клиента. Большинство сайтов делают это корректно, однако некоторые крупные сайты, такие как Google, используют IP-адрес. Это означает, что очевидное местоположение архивного клиента будет влиять на то, что он получает. Такое же влияние оказывают и другие факторы.

Ах, каннада!

Мой комментарий: Каннада – язык, распространённый в юго-западной Индии, преимущественно в штате Карнатака. На нём говорит около 40 миллионов человек. – см. [Википедию, https://ru.wikipedia.org/wiki/Каннада](https://ru.wikipedia.org/wiki/Каннада)

Системы архивации веб-сайтов сталкиваются с еще одной сложностью. Они проходят по ссылкам, и переход по некоторым из них может давать побочные эффекты. Именно этим объяснялась загадка, с которой столкнулись архивисты в университете Олд Доминион (Old Dominion University, см. <https://ws-dl.blogspot.com/2018/03/2018-03-21-cookies-are-why-your.html>). Когда они просмотрели пять различных архивов твитов Барака Обамы, выяснилось, что только 53% из них были на английском языке. То есть сами-то его твиты были на английском, как он их написал, однако расставленные Твиттером метки, ссылки, заголовки и т.д. были на другом языке. Самым распространенным языком после английского оказался язык каннада. На каннаде говорят миллионы людей в Индии, но в других местах он малоизвестен, и даже самый фанатичный сторонник теории рождения бывшего президента вне США не станет утверждать, что это был родной язык Обамы.

В упомянутой статье рассматривается множество причин неправильного выбора языка. Авторы Савуд Аллам (Sawood Alam) и Плинио Варгас (Plinio Vargas) считают, что объяснение нужно искать в списке альтернативных ссылок для разных языков (когда используется конструкция `<link rel="alternate">`), которые не предназначены для отображения в браузере, но помогают программному обеспечению подыскать подходящую версию страницы.

При переходе по этим ссылкам устанавливается куки-файл (cookie), сообщающий Твиттеру предпочитаемый язык. При переходе по ссылке для французского языка устанавливается куки-файл, запрашивающий последующий контент на французском языке. Существует длинный список этих альтернативных ссылок, и последняя из них - для языка каннада! Если программа-клиент сканирует каждую из этих ссылок, она каждый раз получает разные языковые куки-файлы, заменяющие предыдущие. В конце концов будет создан куки-файл, запрашивающий доставку контента на языке каннада.

В обычных обстоятельствах такое поведение кучки-файлов является вполне разумным. Тот, кто просматривает Твиттер на польском языке, вероятно, захочет и впредь просматривать его на польском, независимо от геолокации или настроек браузера. И лишь тогда, когда проводится сканирование сайта, всё становится очень странным.

В статье предлагается несколько возможных решений, таких, как использование куки-файлов короткого срока действия или помещение каждого запроса в отдельную «песочницу». Отключение куки-файлов вряд ли будет хорошим решением; в статье отмечается, что некоторые веб-сайты ожидают, что куки будут сохраняться, по крайней мере, в ходе операций перенаправления. Лучшим выбором может быть использование метода «песочницы», когда каждый из запросов эквивалентен отдельной операции curl.

Как прибить желе к стенке

Для архивистов проблема заключается в том, что типичный современный веб-сайт не является фиксированным документом; он скорее сервис или приложение. То, что вы видите, зависит от модели вашего компьютера, операционной системы, браузера, параметров экрана, географического местоположения, времени и сервера, к которому вы подключились. Создание непротиворечивого снимка сайта - это искусство, особенно если сайт вроде Твиттера, который спроектирован так, чтобы никогда не выглядеть одинаково дважды. Снимки, которые люди находят в «Интернет-архиве» (Internet Archive) и других архивах, порой сбивают с толку или разочаровывают.

Гераклит как-то сказал, что нельзя дважды войти в одну и ту же реку. Эти слова также очень хорошо описывают и Интернет.

Гэри МакГэф (Gary McGath)

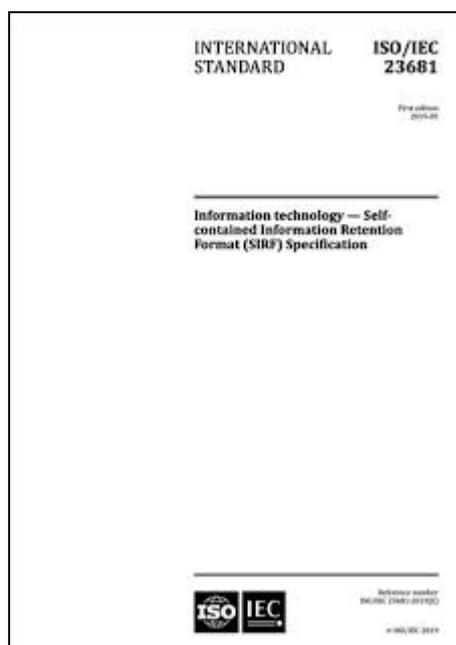


ИСО: СПЕЦИФИКАЦИИ САМОДОСТАТОЧНОГО КОНТЕЙНЕРНОГО ФОРМАТА SIRF ДЛЯ ХРАНЕНИЯ ИНФОРМАЦИИ

Источник: сайт ИСО <https://www.iso.org/standard/76648.html>
<https://www.iso.org/obp/ui/#!iso:std:76648:en>
<http://www.doc88.com/p-2149912837399.html>

В конце мая 2019 года сайт Международной организации по стандартизации (ИСО) сообщил о публикации нового стандарта **ISO/IEC 23681:2019 «Информационные технологии – Спецификации самодостаточного формата для хранения информации SIRF»** (Information technology - Self-contained Information Retention Format (SIRF) Specification) объёмом 44 страницы, см. <https://www.iso.org/standard/76648.html> и <https://www.iso.org/obp/ui/#!iso:std:76648:en>.

Стандарт разработан Объединённым техническим комитетом ИСО/МЭК JTC1 «Информационные технологии» (Information technology).



Во вводной части документа, в частности, сказано следующее:

«Во многих организациях в настоящее время требуется в неопределённо длительной перспективе сохранять и поддерживать доступ к большим объёмам цифрового контента. Для исполнения законодательно-нормативных требований и в связи с правовыми проблемами требуется обеспечивать долговременную сохранность архивов электронной почты, медицинских документов и сведений о правах интеллектуальной собственности.

Веб-сервисы и приложения конкурируют друг с другом как средства обеспечения хранения, организации и коллективного использования фотографий, видеоматериалов и других продуктов творческой деятельности. Да и на многие другие хранилища фиксированного контента возлагается задача сбора и предоставления доступа к научным данным, знаниям, библиотекам, фильмам и музыке. При решении этих задач ключевой проблемой является создание нейтральных по отношению к поставщикам контейнеров хранения, которые возможно будет интерпретировать по прошествии времени.

Архивисты и специалисты по управлению документами, имеющие дело с физическими объектами, такими, как материалы разного рода, документы и т.д., избегают обработки каждого объекта по отдельности. Вместо этого они формируют группу объектов, которые каким-то образом взаимосвязаны - по характеру использования, по связи с каким-то конкретным событием, по времени и т.д. - и затем обрабатывают всю группу как единое целое. Сама группа может называться серией, коллекцией, фондом или, в некоторых случаях, документацией или группой документов.

После того, как группа сформирована, архивист помещает её в физический контейнер (например, в папку с делами или в короб стандартных размеров), присваивает контейнеру имя и регистрационный номер и помещает контейнер в известное место. Информация о группе включается в ярлык, который физически прикрепляется к контейнеру, а также в научно-справочный аппарат, такой, как онлайн-каталог, который организован по определённой схеме и содержит названия групп, их местоположение, объём и краткое описание содержания.

В настоящем стандарте предлагается подход к обеспечению долговременной сохранности цифрового контента, в рамках которого используются процессы архивного дела, тем самым помогая архивистам по-прежнему чувствовать себя комфортно уже в цифровой среде.

Чтобы сделать эту стратегию возможной, одной из первоочередных потребностей является создание электронного эквивалента физического контейнера – архивного короба или папки с делами - который определяет группу и который может быть снабжён стандартным описанием в определенном формате, способствующему поиску и извлечению данных, когда в этом возникнет необходимость.

Предполагается, что роль такого аналога должен играть самодостаточный формат для хранения информации (Self-contained Information Retention Format, SIRF) – это формат контейнера хранения для набора подлежащих долговременному сохранению объектов, который также содержит каталог с метаданными, относящимися как ко всему содержимому контейнера, так и к отдельным объектам и их взаимосвязям.

Данный логический контейнер упрощает и повышает эффективность выполнения многих процессов, которые нужны для предотвращения угроз цифровому контенту. В свою очередь, более простые и эффективные процессы обеспечения долговременной сохранности делают сохранение

цифрового контента в целом более масштабируемым и менее затратным делом.

Компоненты формата SIRF, варианты использования и функциональные требования к нему были перечислены в публикации Международной ассоциации индустрии сетевого хранения данных (Storage Networking Industry Association, SNIA, <http://www.snia.org/>) под названием «Варианты применения и функциональные требования к формату SIRF» (Self-contained Information Retention Format (SIRF) Use Cases and Functional Requirements, версия 0.5a, сентябрь 2010 года), см. https://www.snia.org/sites/default/files/SIRF_Use_Cases_V05a_DRAFT.pdf.

Позднее они были более детально описаны в публикации группы авторов «На пути к самодостаточному формату для хранения информации SIRF» (Towards SIRF: Self-contained Information Retention Format, см. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.397.8547&rep=rep1&type=pdf>).

Настоящий стандарт делает ещё один шаг вперёд и детализирует состав метаданных, категорий и элементов в каталоге контейнера. В нём также описывается, как логический формат SIRF реализуется («сериализуется») в виде контейнеров хранения в облаке и контейнеров на основе лент. С облачной реализацией SIRF экспериментирует хранилище объектов OpenStack Swift, и соответствующее представление предлагается в качестве открытого источника в рамках инициативы OpenSIRF (см. <http://github.com/opensirf>).

Создание и ведение SIRF-каталога требует выполнения вычислений с интенсивным использованием данных для различных сохраняемых объектов, включая проверки неизменности и преобразования данных. Эта задача может эффективно решаться при помощи вычислительных модулей-«сторлетов» (storlets), рядом с местом хранения данных. Преимущества использования сторлетов включают уменьшенную потребность в пропускной способности (уменьшается количество байтов, передаваемых по глобальной сети), повышенную безопасность (снижаются риски утечки чувствительных данных), сокращение затрат (уменьшаются потребности в инфраструктуре на стороне клиента) и поддержку соответствия законодательно-нормативным и иным требованиям (за счёт улучшения отслеживания происхождения данных).

Сторлет-механизм (см. «Механизм Storlet для выполнения биомедицинских процессов в системе хранения», Storlet Engine for Executing Biomedical Processes within the Storage System, <https://pdfs.semanticscholar.org/78a6/b5ffd573819623ab735b8de01246bdfdf342.pdf>) - это механизм поддержки таких сторлет-вычислений в безопасных «песочницах» (sandboxes) в системе хранения, и он может использоваться для создания и обслуживания SIRF-контейнеров».

Содержание документа следующее:

Предисловие

Введение

1. Область применения
2. Нормативные ссылки
3. Термины и определения
4. Деловое обоснование
5. Обзор спецификаций
6. Метаданные контейнера (Container Information Metadata)
7. Метаданные сохраняемых объектов (Object Information Metadata)
8. Сериализация для интерфейса CDMI (Cloud Data Management Interface – «Интерфейс управления данными в облаках», описан в стандарте ISO/IEC 17826:2016) ассоциации SNIA
9. Сериализация для формата LTFS (от Linear Tape File System, описан в стандарте ISO/IEC 20919:2016 – Н.Х.) ассоциации SNIA
10. Сериализация для OpenStack Swift (*масштабируемая система хранения с избыточностью, разработанная в рамках инициативы OpenStack;* *см. также* [https://en.wikipedia.org/wiki/OpenStack#Object_storage_\(Swift\)](https://en.wikipedia.org/wiki/OpenStack#Object_storage_(Swift)))
11. Пример варианта применения
 - Приложение А: XML-схема для SIRF-каталога
 - Приложение В: Образец XML-каталога
 - Приложение С: Образец JSON-каталога
 - Библиография

Комментарий: Это ещё один документ, посвященный вопросу конструирования контейнеров для передачи информации и документов на архивное хранение. Рекомендую коллегам из Росархива и ВНИИДАД обратить на него внимание – в нём есть немало полезных идей, которыми они могут воспользоваться для превращения своей собственной «теории контейнеров» на полутора страничках во что-то приличное)



США: ПРОДОЛЖАЕТСЯ РАЗВИТИЕ ТЕХНОЛОГИИ ОБЪЁМНОГО СКАНИРОВАНИЯ ДОКУМЕНТОВ

Источник: сайт DocumentImagingReport.com
<https://www.documentimagingreport.com/?p=6267>

Автор: Наташа Храмовская

Заметка Ральфа Гэммона была опубликована 12 июня 2019 года на сайте DocumentImagingReport.com под названием «Компания Ripcord расширяет своё видение, назначив Бармака Хешмата в Консультативный совет» (Ripcord Enhances its Vision with Appointment of Barmak Heshmat to Advisory Board). Данный сайт позиционирует себя как основной новостной

источник, освещающий тенденции и анализ рынка средств управления графическими образами документов.

Исследователь из Массачусетского технологического института (MIT) привносит глубокие знания передовой технологии Terahertz создания графических образов. Он пополнит ряды отраслевых «тяжеловесов», которыми располагает компания Ripcord.

Ripcord, первая в мире компания, занимающаяся роботизированной оцифровкой и стремящаяся способствовать созданию по-настоящему безбумажного мира, сегодня объявила о назначении бывшего научного сотрудника Массачусетского технологического института (MIT) Бармака Хешмата (Barmak Heshmat) в свой Консультативный совет. Он присоединяется к звёздному составу, в который входят соучредитель Apple Стив Возняк (Steve Wozniak), бывший директор по технологиям НАСА Крис Кемп (Chris Kemp) и соучредитель Oracle Брюс Скотт (Bruce Scott). Данное назначение обещает продвинуть программное обеспечение Canopy и роботизированные системы компании за пределы традиционных ограничений оптики благодаря новаторской работе Хешмата в области технологии создания графических образов Terahertz.

Технология создания графических образов Terahertz, впервые разработанная Хешматом в MIT, позволяет получать спектроскопические изображения со стробированием по времени, что даёт возможность извлекать контент из слоистых структур.

Сегодня это означает, что данная технология способна читать 9-20 страниц закрытой книги, используя терагерцевые волны. В будущем открывается перспектива чтения целых коробов бумаги без открытия крышки. Эта технология интегрирована с существующими средствами распознавания символов Ripcord и алгоритмами выделения объектов на основе искусственного интеллекта, и сочетание всех этих технологий послужит основой для следующего рывка Ripcord, который позволит компании, уже являющейся лидером в сфере управления контентом, стать лидером в трансформации отрасли.

В настоящее время Хешмат Бармак является основателем и генеральным директором избегающего публичного внимания стартапа BRELYON. Бармак является пионером технологии Terahertz, позволяющей сканировать и читать страницы в закрытых книгах с использованием невидимого излучения. Ранее проделанные работы Хешмата изменили мир управления графическими образами, начиная от основ дизайна оптических систем и до внедрения новых приложений, таких, как пакетное сканирование.

Мой комментарий: По всем этим направлениям существуют разработки других исследователей и организаций, в том числе успешно опробованные для чтения обгоревших и впоследствии окаменевших свитков в Помпеях. Главное в том, что технологии объёмного сканирования на подходе, и они несут с собой как поразительные возможности (в том числе для делопроизводства и архивной отрасли, но также и для разведки), так и

риски – например, станет куда более рискованно передавать бумажные документы на внеофисное хранение.

«Для меня большая честь присоединиться к Консультативному совету Ripcord, с тем, чтобы помочь использовать уроки, извлеченные из опыта терагерцевого сканирования и сканирования «на лету» (time-of-flight imaging), которые могут помочь компании Ripcord радикально ускорить захват данных с бумажных документов», - отметил в своём интервью Бармак Хешмат. «Ripcord - это инновационная и дальновидная компания, которая объединяет данные, которые мы можем получить с помощью новейших неоптических методов сканирования, с машинным обучением и искусственным интеллектом, с тем, чтобы получить знания и результаты, которые раньше невозможно было себе представить. Восхитительно думать о возможностях, которые эти технологии создают при переходе от физической к цифровой среде».

«Помощь Бармака в коммерциализации этого подхода сделает реальностью миссию Ripcord по превращению всего мира в безбумажный», - сказал Алекс Филдинг, основатель и генеральный директор компании Ripcord. «Не каждый день вы встречаете столь блестящего и любопытного человека, как Бармак, не говоря уже о возможности учиться у него и работать вместе с ним для коммерциализации и масштабирования технологического подхода, который может помочь нам создать безбумажное будущее. Этот прогресс в области управления графическими образами является технологическим эквивалентом перехода от рентгеновских снимков к МРТ-томографии высокого разрешения».

Компания Ripcord, основанная бывшими сотрудниками Apple и НАСА, даёт компаниям возможность оцифровывать бумажные документы и связывать их с существующими корпоративными системами. Компания предлагает альтернативу устаревшим системам управления документами, которые являются слишком сложными, требуют дорогой настройки и на эффективное внедрение которых уходят годы.

Робототехническая технология Ripcord сокращает расходы на оцифровку до небольшой доли тех, что характерны для имеющихся на рынке традиционных вариантов, благодаря тому, что более качественные результаты достигаются в 10 раз быстрее. Приход Хешмата позволит и дальше сотрясать рынок, одновременно повышая отраслевые стандарты оцифровки данных и управления документами.

В Совет директоров компании Ripcord входят генеральный партнер GV Энди Уиллер (Andy Wheeler); генеральный менеджер Icon Ventures Джеб Миллер (Jeb Miller); партнер в КРСВ Вэнь Се (Wen Hsieh); наблюдатель в Совете директоров и соучредитель и управляющий партнер Lux Capital Питер Хеберт (Peter Hebert); и председатель и президент Ripcord, а также исполнительный директор Aspect Software Патрик Денис (Patrick Dennis).

Для получения дополнительной информации, пожалуйста, пишите по адресу sales@ripcord.com или посетите <https://www.ripcord.com/> .

Ральф Гэммон (Ralph Gammon)

Дополнительная информация: На сайт YouTube можно посмотреть 4-минутный видеосюжет, в котором Хешмат Бармак рассказывает о своей работе в MIT, под названием «Читая сквозь книги» (Reading through closed books), см. <https://www.youtube.com/watch?v=6i25SuJzb0A>



ИСО: НАЧАТА РАБОТА НАД СТАНДАРТОМ МЕНЕДЖМЕНТА РИСКОВ, СВЯЗАННЫХ С ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

Источник: сайт ИСО / сайт BSI <https://www.iso.org/standard/77304.html>
<https://standardsdevelopment.bsigroup.com/projects/9018-02222>

Автор: Наташа Храмовская

Как сообщил недавно сайт Международной организации по стандартизации, начата работа над проектом стандарта **ISO/IEC 23894 «Информационные технологии – Искусственный интеллект – Менеджмент риска»** (Information Technology - Artificial Intelligence - Risk Management), см. <https://www.iso.org/standard/77304.html>



Эта работа проводится в рамках подкомитета SC42 «Искусственный интеллект» Объединенного технического комитета ИСО/МЭК JTC1.

В свою очередь сайт Британского института стандартов, говоря об этом проекте, отмечает следующее:

«Настоящий документ содержит рекомендации по менеджменту рисков, с которыми сталкиваются организации при разработке и применении методов и систем искусственного интеллекта (ИИ), с тем, чтобы помочь организациям интегрировать менеджмент связанного с ИИ риска в существенные направления и виды деятельности. В стандарте также

описывается процесс эффективного внедрения и интеграции менеджмента связанного с ИИ риска.

Настоящие рекомендации могут быть адаптированы с учётом особенностей организации и условий её деятельности.

В настоящем документе используются требования и рекомендации, содержащиеся в международном стандарте ISO 31000 «Менеджмент риска - Руководство» (Risk management – Guidelines, в России адаптирован как ГОСТ Р ИСО 31000-2010 «Менеджмент риска. Принципы и руководство», см. <http://protect.gost.ru/v.aspx?control=8&baseC=6&id=171333> – Н.Х.). Кроме того, в нём содержатся дополнительные рекомендации по вопросам, которые возникают при применении ИИ в существующих процессах любой организации, или же тогда, когда организация предоставляет ИИ-систему для использования другим сторонам.

Одной из главных проблем внедрения систем искусственного интеллекта является отсутствие доверия к таким системам. Возражения связаны с последствиями предвзятости данных и/или алгоритмов, с новыми угрозами безопасности, такими, как сброс данных враждебными сторонами, угрозы для неприкосновенности частной жизни ввиду утечки персональных данных, а также недостаточные прозрачность и подотчётность.

Менеджмент риска - это концепция, которая широко рассматривается в публикациях ИСО, начиная от специфики конкретной области или приложения, переходя к сектору/отрасли и заканчивая общими рекомендациями. Структурированный, воспроизводимый процесс менеджмента риска для ИИ-систем напрямую решает следующие проблемные вопросы:

- Наличие ошибок проектирования, вопрос оценки типичных соответствующих угроз и рисков для систем на основе искусственного интеллекта, а также способы смягчения этих рисков и методики, позволяющие выявлять, классифицировать и обрабатывать риски для (и связанные с использованием) ИИ-систем;
- Создание и укрепление доверия к ИИ-системам посредством прозрачности, проверяемости, объяснимости, управляемости и т.д., с использованием хорошо понятых и документированных процессов менеджмента риска;
- Устойчивость, живучесть, надёжность, точность, безопасность, защищённость ИИ-систем, защита неприкосновенности частной жизни и т.д., посредством обеспечения прозрачности обработки рисков для выявленных заинтересованных сторон.»



ОТВЕТ НА ВОПРОС КОЛЛЕГИ: С ЧЕГО НАЧАТЬ СВОЙ ПУТЬ В ПРОФЕССИИ?

Источник: http://rusrim.blogspot.com/2019/06/blog-post_24.html

Автор: Наташа Храмцовская

***Вопрос:** Закончив институт, я не понимаю, в каком направлении двигаться. Очень хочется работать по специальности, я уверена, что молодые специалисты помогут сделать большой шаг вперёд. Только готовы ли организации к этому?*

Ответ: Моему поколению довелось жить в «эпоху перемен», и очень мало кому не пришлось переучиваться, а то и вообще менять профессию. Как мне кажется, для Вашего поколения это будет ещё более характерно. Скорее всего, каждому из Вас придётся за время своей карьеры освоить несколько профессий и постоянно повышать свою квалификацию. В этом плане не так даже важно, с чего конкретно начать – скорее, важно настроиться на постоянный рост, на то, чтобы идти в ногу со временем, всё время узнавать что-то новое.

Выбирая первую работу, желательно в первую очередь найти хорошего – строгого, но человечного – начальника и сильный сплочённый коллектив, горящий своим делом. Ещё лучше при этом иметь ясную перспективу роста, чтобы не «застрясть» на технической работе за мизерную зарплату. Если при этом придётся пойти в бизнес, или в какую-то не очень знакомую отрасль – не страшно, любые навыки пригодятся. Я многое прошла в своей жизни; были и должности, которые я не упоминаю в своём резюме – тогда нужно было кормить семью, и выбора особого не было; но даже их мне удалось использовать для дальнейшего роста).

Имейте в виду следующее: наше время перемен – это время максимальных возможностей для молодых. Право и технологии, знания и приёмы работы меняются настолько быстро, что если Вы сможете лет 5 «продержаться на волне», Вы обойдёте подавляющее большинство тех, кто сегодня находится впереди). И хотя таких молодых коллег пока не очень много, они есть.

***Вопрос:** Интересно Ваше мнение, с чего начать свой путь, если действительно интересна эта профессия, не только в рутинной деятельности документоведа, но и в способности анализировать ситуацию, предлагать и внедрять новые технологии.*

Ответ: Сегодня – если говорить об управлении документами и архивном деле, но не только – востребованы синтетические специалисты, способные разговаривать на одном языке с владельцами бизнеса, юристами, ИТ и «безопасниками», специалистами основных «зарабатывающих» подразделений. Нужны люди с хорошими навыками менеджеров и организаторов, умеющие, помимо прочего, «продавать» свои знания и

умения руководству и быть полезными своей организации. Чтобы научиться анализировать ситуацию, нужны не только узкопрофессиональные знания, но и достаточно широкий кругозор, позволяющий видеть реальные тенденции и не обращать внимание на регулярно поднимаемую шумиху; нужна практика такого анализа.

Очень важно хорошо освоить английский язык и уверенно овладеть компьютером и прочими современными гаджетами.

Время от времени, с учетом складывающихся обстоятельств, следует оценивать потенциальную отдачу от аспирантуры и защиты диссертации, а то и получения второго высшего образования.

Вообще я считаю, что в жизни для каждого из нас есть много интересных и очень разных путей, и не так важно, по какой тропе пойти – важно не тормозить и не останавливаться!

В заключение хочу пожелать Вам счастья и удачи – не только в работе, но и во всём остальном!

С самыми лучшими пожеланиями, Наташа Храмцовская.



МОЯ «МЕЧТА АРХИВИСТА»

Источник: блог МСА <https://blog-ica.org/2019/06/09/my-archivists-dream/>

Автор: Наташа Храмцовская

Заметка архивиста-консультанта Эммануэля-Фабриса Омгбва-Яссе была опубликована на блоге Международного совета архивов (<https://blog-ica.org/>) в рамках подборки постов, посвященной Международной неделе архивов 2019 года, основной темой которой является «Проектирование Ваших архивов в 21-м веке».

Я мечтаю об обществе, в котором к архивам не относятся как к вещи, которой можно попользоваться, а затем о ней забыть; об обществе, в котором архивисты пользуются доверием создателей контента и документов, а информационное взаимодействие является каждодневным делом. Это общество, в котором люди искренне беспокоятся всякий раз, когда их культурно-историческому наследию угрожает исчезновение; в котором идея общего блага имеет глубокие корни, и которое привержено делу сохранения памяти о прошлом. Это общество, где архивы переживают своих создателей, где препятствия становятся движущими силами, где коллективный интерес преобладает и является той ценностью, которая чаще всего является для людей общей.



Я мечтаю о профессии, представители которой обсуждают обмен знаниями, а не сами знания; о профессии «пожарных», готовых погасить пламя, зажженное пробелами в истории прошлого; о профессии, где обсуждаются реальные этические проблемы, а не только их поверхностные последствия. Это проактивная профессия, где вопросы обсуждаются «в настоящем времени» и где ищутся решения проблем как вчерашнего, так и завтрашнего дня. Это профессия, представители которой полны энтузиазма и привержены своему делу, а не относятся к нему равнодушно и «спустя рукава». Это профессия, которая собой гордится, представители которой шагают в ногу под бой одного и того же барабана, в рамках координированных усилий защитить культурно-историческое наследие.

Я мечтаю о профессиональном сообществе, сильном и динамичном, которое черпает силы в своих победах и которое никогда не «выбрасывает полотенце»; где образ оставляют место для воображения. Это открытое сообщество, готовое обмениваться, учиться и делиться; которое не замыкается на внутренних вопросах, но открыто для широкой публики, и деятельность которого порождает энтузиазм и интерес, а не разочарование и жалобы. Это сообщество, которое стремится убеждать, а не побеждать, в котором его представители гордятся не принадлежностью к «клубу», а тем, что чувствуют себя его составной частью. Это слушающее, заботливое, поддерживающее и понимающее сообщество.

Я мечтаю о государстве, достаточно уверенном в себе, чтобы выходить за рамки лозунгов и предпринимать реальные позитивные действия; о государстве, которое не просто предлагает поддержку, но также и партнерство; которое, потерпев неудачу, способно подняться ради общего блага. Это государство, в котором мы не живём одним лишь сегодняшним днём, потому что он уже наступил, но где каждый день имеет своё прошлое и является продолжением того, что было раньше. Это сильное государство, которое не бежит от своего прошлого, а находится в единстве с ним.

Таковы мои «мечты архивиста».



ВЕЛИКОБРИТАНИЯ: ПОЧЕМУ АРХИВЫ СЛУЖАТ ВСЕМ

Источник: блог Национальных Архивов Великобритании

<https://blog.nationalarchives.gov.uk/why-archives-are-for-everyone/>

Данная заметка Исполнительного директора и Хранителя Национальных Архивов Великобритании Джеффа Джеймса была опубликована на блоге Национальных Архивов Великобритании. Джефф Джеймс также является президентом Форума руководителей национальных архивных служб Международного совета архивов.

Тема Международной недели архивов этого года (см. <https://www.ica.org/en/international-archives-week-3-9-june-2019> - это тема «Проектирование архивов в 21-м веке») напрямую связана с усилиями, предпринятыми Международным советом архивов (МСА) и Форумом руководителей национальных архивных служб МСА по пропаганде уникальной ценности и воздействия архивов различных видов. Она также находится в центре стратегических устремлений Национальных Архивов Великобритании.

Когда в 2014 году вернулся в Кью, мне посчастливилось оказаться там в тот момент, когда наша организация начинала разработку новой стратегии. В нашем предыдущем бизнес-плане «В интересах документальной истории, ради общего блага» (For the Record, For Good – это был деловой план на 2011-2015 годы, см. <http://www.nationalarchives.gov.uk/documents/the-national-archives-business-plan-2011-2015.pdf>) говорилось о том, что именно делают архивы.

Его преемник, документ под названием «Архивы вдохновляют» (Archives Inspire, <http://www.nationalarchives.gov.uk/documents/archives-inspire-2015-19.pdf>, о нём см. пост <http://rusrim.blogspot.com/2015/12/4.html>) - наша амбициозная, ориентированная на интересы аудитории стратегия, реализация которой завершилась в марте этого года, - чётко определила, для чего нужны архивы.

Наша новая стратегия, «Архивы для всех» (Archives for Everyone, см. <http://www.nationalarchives.gov.uk/about/our-role/plans-policies-performance-and-projects/our-plans/archives-for-everyone/>), говорит о том, для обслуживания кого существуют архивы, и почему архивы во всём мире должны стремиться быть как можно лучше.



Чтобы быть национальным архивом 21-го века, мы должны быть инклюзивными, предприимчивыми (entrepreneurial), трансформационными (disruptive). Инклюзивный архив укрепляет доверие и разрушает барьеры для доступа, участия и понимания. Предприимчивый архив действует гибко и бесстрашно, создавая и реализуя полезную отдачу у себя дома и по всему миру. Трансформационный архив меняет всё, постоянно адаптируется, переосмысливает и переформирует свою практику работы для решения стоящих перед ним текущих и будущих задач.

Как и многие национальные архивы по всему миру, мы полагаем, что являемся важным ресурсом, поддерживающим нашу демократию. Мы являемся общественным достоянием, и наши коллекции являются активом для будущих поколений. Мы также знаем, что нам предстоит проделать большую работу, чтобы сделать людей фокусом нашей деятельности, и для этого потребуются коллективные усилия всего мирового архивного сообщества. Чтобы стать учреждениями, которые действительно представляют интересы всех тех, кто сталкивается с нами, нам также необходимо развивать новые навыки и осваивать нарождающиеся технологии.

Национальные Архивы Великобритании, созданные в 2003 году в результате слияния Управления государственных документов (Public Record Office) и Комиссии по историческим рукописям (Historical Manuscripts Commission, [НМС, http://www.nationalarchives.gov.uk/archives-sector/our-archives-sector-role/historical-manuscripts-commission/](http://www.nationalarchives.gov.uk/archives-sector/our-archives-sector-role/historical-manuscripts-commission/)), изменялись и эволюционировали на протяжении всей своей истории. С момента зарождения в 1838 году и до наших дней, мы наблюдаем прогресс в направлении открытости и уход от эксклюзивности и привилегированности. У нас ушло почти 200 лет на то, чтобы понять - архивы предназначены для всех. В год, когда Комиссии по историческим рукописям отмечает свое 150-летие, пришло время в полной мере принять это понимание и сделать учреждение соответствующим его миссии! По мере того, как мы строим планы на следующее четырёхлетие, этот юбилей НМС является достойным

подтверждением того, как архивы и архивная практика способны в предстоящие годы изменить то, что мы делаем, и как мы это делаем.

Мы можем многому научиться на опыте друг друга, и я очень рад видеть, что пропагандистская кампания этого года длится целую неделю. Если Вы ещё этого не сделали, посетите веб-сайт МСА (<https://www.ica.org/en/international-archives-week-3-9-june-2019>) и примите участие в Международной неделе архивов #IAW2019.



США: НАЦИОНАЛЬНЫЕ АРХИВЫ РАССМАТРИВАЮТ ИСПОЛЬЗОВАНИЕ БЛОКЧЕЙНА ДЛЯ ПРОВЕРКИ ПОДЛИННОСТИ ДОКУМЕНТОВ В МОМЕНТ РОСТА КОЛИЧЕСТВА ФЕЙКОВЫХ ВИДЕО

Источник: сайт FederalNewsNetwork.com
<https://federalnewsnetwork.com/author/jory-heckman/>
Автор: Наташа Храмцовская

Данная заметка репортёра «Федеральной сети новостей» (Federal News Network) Джори Хекман была опубликована на сайте FederalNewsNetwork.com.

После успешного внедрения новой технологии в других органах федерального правительства, Национальные Архивы США (National Archives and Records Administration, NARA) изучают вопрос о том, может ли технология блокчейна помочь специалистам, занимающимся управлением документами, контролировать свои обширные хранилища информации.

Эрик Дуглас, специалист по политике управления документами и по поддержке программ в офисе директора по управлению документами в федеральном правительстве США (Chief Records Officer for the U.S. Government – *должность в составе высшего руководства Национальных Архивов*), сообщил, что будущее блокчейна выглядит многообещающим и что блокчейн может сыграть определенную роль в аутентификации электронных экземпляров фото- и видеоматериалов из фондов Национальных Архивов.

Возможность проверить документы на отсутствие несанкционированных изменений может оказаться полезной для защиты от высококачественных фальсифицированных видеороликов, которые создаются с помощью алгоритмов машинного обучения, способных манипулировать изображениями и звуком таким образом, чтобы всё

выглядело так, как будто государственные должностные лица сказали или сделали что-то, чего они на само деле не делали.

«Альтернативная информация стала вызовом сложившимся способам потребления информации общественностью», - отметил Дуглас в ходе своего доклада, сделанного 10 июня 2019 года на проводимом Национальными Архивами два раза в месяц заседании дискуссионной группы по вопросам управления документами и информацией (Bimonthly Records and Information Discussion Group).

В прошлом году Национальные Архивы уже протестировали версию этой технологии, когда они раскрыли около 20 тысяч государственных документов об убийстве в 1963 году бывшего президента Джона Кеннеди (John F. Kennedy). Дуглас объяснил, что каждая из массовых выгрузок содержит в своих метаданных хеш, который позволяет пользователям убедиться, что электронный материал не был изменён.

Мой комментарий: *Путать блокчейн с использованием хешей стыдно :) Скажем, те же компьютерные пираты уже лет 20 как используют хеши для контроля неизменности своего «продукта».*

«Распространение этой практики на все хранимые Национальными Архивами электронные материалы позволит общественности независимо проверять, что взятый из каталога Национальных Архивов электронный контент не был изменён», - сказал он.

Дуглас в то же время предупредил, что любые разговоры о передаче федеральными органами исполнительной власти своих блокчейн-документов в Национальные Архивы остаются «абсолютно теоретическими на данный момент», поскольку по-прежнему неясно, будут ли у архивного ведомства ресурсы для архивации и обеспечения долговременного доступа к информации, сохраняемой в блокчейне.

Национальные Архивы также хотят убедиться в том, что официальные лица смогут открывать различные типы файлов, сохранённых в блокчейне, учитывая разнообразие типов документов, которые Национальные Архивы получают от федеральных органов исполнительной власти.

Мой комментарий: *Трудно сказать, кто ничего не понимает в блокчейне – специалист Национальных Архивов или журналист (однако последняя гипотеза представляется мне более правдоподобной). В блокчейне сами документы не содержатся, в него записываются лишь их хеши – так что названная проблема явно надуманная.*

«Блокчейн переносит ответственность за обеспечение целостности электронных документов и, соответственно, доверие, с изолированных систем организации на распределенную сеть. Это отход от роли централизованного управления документами, к которой мы привыкли», - отметил Дуглас. «Переход на блокчейн может обеспечить целостность и доверие, которые традиционно обеспечивали системами управления электронными документами, и этот сдвиг может повлиять на то, как документы организуются, упорядочиваются и поддерживаются во времени»

Мой комментарий: Блокчейн (в том виде, как его пробуют использовать в архивном деле сегодня) – не более чем инструмент для независимого контроля неизменности документов (но и не менее!). Самих документов в блокчейне нет, и если традиционные структуры не справятся со своими задачами, то мы будем знать, например, что документы были изменены – но не будем иметь верных экземпляров. Задача по обеспечению сохранности документов, имеющих длительную или постоянную ценность, в этом случае всё равно будет провалена. Соответственно, не стоит преувеличивать роль блокчейна в данной конкретной области применения.

Более широкое внедрение блокчейна может также потребовать от федеральных органов исполнительной власти проведения переобучения или переподготовки их персонала, занимающегося управлением документами. И если многие федеральные специалисты по управлению документами имеют компьютерные и ИТ-навыки, необходимые для работы с электронными документами, то, как отметил Дуглас, будущим архивам понадобится «междисциплинарная» команда, в которую войдут компьютерные специалисты и инженеры-программисты.

Однако несмотря на ажиотаж по поводу того, что блокчейн в перспективе может дать федеральному правительству, соответствующие проекты большинства федеральных органов еще не вышли из стадии тестирования пилотных решений.

Тем не менее, Национальные Архивы указывают на успех решения Accelerate («Ускорение») Министерства здравоохранения и социальных услуг (Department of Health and Human Services, HHS), использующего с декабря 2018 года оперативные данные ведомства для «прочёсывания» более 100 тысяч контрактов HHS, с целью договориться о более низких ценах на товары, которые структурные подразделения ведомства закупают оптом.

Успешное развертывание решения Accelerate было отмечено в отчете, подготовленном исследовательской организацией «Фонд данных» (Data Foundation) и компанией Booz Allen - Hamilton о внедрении блокчейна в федеральном правительстве.

Мой комментарий: речь идёт об отчёте, опубликованном 10 июня 2019 года, который называется «Внедрение блокчейна в деятельность государственных органов: Путь к организации эффективных федеральных инициатив, связанных с блокчейном» (*Bringing Blockchain Into Government: A Path Forward for Creating Effective Federal Blockchain Initiatives*) объёмом 24 страницы, [см. https://www.datafoundation.org/bringing-blockchain-into-government](https://www.datafoundation.org/bringing-blockchain-into-government), полный текст которого доступен по адресу https://static1.squarespace.com/static/56534df0e4b0c2babdb6644d/t/5cfaa34d5a82320001639f17/1559929681007/Blockchain_Paper_Final.pdf

В отчете также отмечается, что Управление по контролю за качеством пищевых продуктов и медикаментов (Food and Drug Administration, FDA) экспериментировало с использованием блокчейна при обмене с исследователями чувствительной информацией о больных раком, и искало

способы использовать данную технологию для отслеживания пациентов во время эпидемии.

Отчёт призывает администрацию Трампа «поощрять дальнейшее изучение и внедрение» блокчейна посредством утверждения федеральной стратегии в области данных (Federal Data Strategy). Административно-бюджетное управление президентской администрации (Office of Management and Budget, OMB) недавно опубликовало свою стратегию внедрения, рассчитанную на один год.

«Пока ещё предстоит увидеть, окажется ли блокчейн в конечном итоге успешным в рамках федерального правительства. Однако на данный момент перспективы выглядят достаточно радужными», - говорится в отчете.

Джори Хекман (Jory Heckman)

Мой комментарий: Честно говоря, для сильных и стабильных стран я не вижу никакой необходимости в использовании блокчейна (в виде широкомасштабной распределенной системы с недоверенными участниками, вроде Биткойна) в архивном деле.

Контроль хешей можно прекрасно организовать и без него; также без него можно внедрить криптографическим образом связанные цепочки блоков – уже есть немало централизованных систем, где такой подход используется.

В то же время я с пониманием отношусь к попыткам Национальных Архивов США «отскрести копеечку» на волне ажиотажа вокруг блокчейна – да и нашим архивистам рекомендую не упускать такой возможности, если таковая представится :)

ЗМІСТ

Передмова.....	1
Дэвид Розенталь: Обзор технологии хранения данных в ДНК.....	3
Электронная архивация – дело, связанное с рисками.....	6
Круглый стол ВНИИДАД «Актуальные вопросы хранения, комплектования, учета и использования научно-технической документации в организациях».....	8
В России будет создана национальная система управления данными.....	10
Франция: Первое практическое руководство по работе с архивными документами.....	13
США: «Конференция Седона» опубликовала для публичного обсуждения проект документа по определению юрисдикции при трансграничной передаче персональных данных.....	18
Веб-архивация и языки.....	19
ИСО: Спецификации самодостаточного контейнерного формата SIRF для хранения информации.....	23
США: Продолжается развитие технологии объёмного сканирования документов.....	26
ИСО: Начата работа над стандартом менеджмента рисков, связанных с искусственным интеллектом.....	29
Ответ на вопрос коллеги: С чего начать свой путь в профессии?.....	31
Моя «мечта архивиста».....	32
Великобритания: Почему Архивы служат всем.....	34
США: Национальные Архивы рассматривают использование блокчейна для проверки подлинности документов в момент роста количества фейковых видео.....	36