

Газеты маленьких городков: все ещё приносят новости

Джон Герберт

[John Herbert]

Библиотека имени Дж. Уилларда Мэриотта

Университет штата Юты, США

E-mail: [john.herbert\[at\]utah.edu](mailto:john.herbert[at]utah.edu)

Ренди Олсен

[Randy Olsen]

Библиотека Церковной истории, Церковь Иисуса Христа

Святых последних дней, Солт Лейк Сити, США

E-mail: [randy.olsen\[at\]ldschurch.org](mailto:randy.olsen[at]ldschurch.org)

Session:

**119 – Users and portals: digital newspapers, usability, and genealogy –
Newspapers Section with Genealogy and Local History Section**

Краткий обзор:

Местные газеты в период с XIX по XX век сегодня, пожалуй, самый пренебрегаемый ресурс в исторических исследованиях. Эти газеты зачастую упускаются исследователями, не потому что они недооценивают историческую ценность газет, но потому, что их ценность очень трудно извлечь. Некоторые местные газеты не сохранились в своей первоначальной печатной форме до наших дней, а если и сохранились, то их трудно найти, и они очень повреждены. Перенесенные на микроноситель, они слишком сложны в использовании.

В течение последних десяти лет общественные и частные сектора с усердием взялись за проблему доступа к этим газетам путем массовой оцифровки. Национальные библиотеки, государственные службы и фонды объединились в преобразовании миллионов страниц газетных бумаг в онлайн текст, который стал бы доступен для ученых и других научных сотрудников. Часто с помощью исследовательских библиотек, коммерческие организации переиздавали газеты, к которым часто обращалась широкая публика, и которые могли пользоваться большим спросом.

Эти общественные и частные усилия по переформатированию были сосредоточены в основном полностью на газетах, представлявших историческую значимость. Значимыми часто считались газеты, содержащие мировые и национальные новости с публикациями, посвященными столичным рынкам, или также включающими группы особых интересов. Забытыми, по большей части, оставались газеты небольших городков, которые некогда выпускали очень разные новости -- однозначно местные и рассчитанные на узкий, но избранный круг читателей.

Цель этого исследования состоит в том, чтобы определить, какие из местных газет заслуживают большего внимания в продолжении усилий по реформатированию до того, как большинство газет небольших городов буквально превратится в пыль. Авторы будут подойдут к их изучению с учетом двух гипотез. Во-первых, специалисты по генеалогии являются наиболее частой группой пользователей оцифрованных газет, а, во-вторых, газеты этих маленьких городков также важны для пользователей данного общества, как и газеты крупных городов.

Чтобы проверить эти гипотезы, авторы проанализируют пользовательскую статистику и данные исследований Программы цифровых газет штата Юты (UDN). Эта программа проводилась Библиотекой им. Дж. Уилларда Мэриотта Университета штата Юты в течение более чем десяти лет и была широко признанной как одно из наиболее преуспевающих начинаний по реформатированию газет в Соединенных Штатах. UDN была запущена при поддержке субсидирования служб и техники Библиотеки. После успешной оцифровки 30000 страниц, Библиотека им. Мэриотта получила грант Национального лидерства в размере 1,02 миллиона долларов от Института музейных и библиотечных служб и преобразовала в дополнение 240000 страниц. В период с 2005 по 2008 годы было получено еще три гранта от Национального гуманитарного фонда общей суммой в 1,1 миллион долларов, чтобы реформатировать 365000 страниц газет штата Юты, как часть Национальной программы цифровых газет. Сегодня UDN гордится 60 печатными изданиями, 1,3 миллионами страниц и полным распознаванием сегментов текста.

Собрание UDN идеально подходит для исследований исторической ценности местных газет. Две трети 60 печатных изданий UDN получены из сельских районов с содержанием, идущим из почти каждого округа Юты -- в большей степени сельского штата. Авторы надеются, что заключения их исследования стимулируют больший интерес к оцифровке газет небольших городков, чтобы обеспечить их долговечность и удовлетворить требования местных историков, специалистов по генеалогии и других научных сотрудников.

Темы, которые будут обсуждаться в этой работе, будут включать следующее:

- 1. краткое описание международного движения к цифровому переформатированию газет;*
 - 2. тщательная подготовка к программе UDN;*
 - 3. краткая характеристика газет, включенных в программу UDN;*
 - 4. отчет и анализ использования статистики газетами для UDN;*
 - 5. отчет и анализ использования данных исследования для UDN;*
 - 6. рекомендации для будущих мероприятий по оцифровке, основанные на полученных данных.*
-

Введение:

Цель данной статьи заключается в стимулировании ответственных лиц в области цифровой библиотеки к тому, чтобы отдать более высокий приоритет реформатированию газет. Данные будут представлены так, чтобы изучить использование цифровых газет в пределах большого онлайн хранилища «Цифровые газеты штата Юты». Полученные данные также будут отражены по опросу пользователей, проведенного для того, чтобы проверить две гипотезы: во-первых, что специалисты по генеалогии используют цифровые газеты больше, чем любые другие исследователи, а, во-вторых, что газеты маленьких городков также или даже более важны для исследований, как и большие столичные газеты. При представлении этих данных мы надеемся, что управляющие библиотеки и заинтересованные лица повысят приоритет проектов по оцифровке газет, в частности относящихся к газетам маленьких городков, и направят больше ресурсов на их осуществление.

К написанию данной статьи отчасти нас побудила недавняя неудачная попытка коллеги реформатировать и предохранить газету небольшого городка -- самую ценную в его программе по оцифровке библиотеки. Другие администраторы засомневались в разумности этого решения, основанного на приоритетах учреждения. Можно было бы избежать последовавших долгих, иногда острых дискуссии, если бы все вовлеченные в них были лучше информированы в ценности цифровых газет для пользователей библиотеки и важности газет маленьких городков для достаточного большого научно-исследовательского общества. Будем надеяться, что эта статья поможет другим избежать опыта, постигшего нашего коллегу.

Контекст исследования -- важность политики проведения и использования информации:

Каждая библиотека, реально погруженная в программу по оцифровке сталкивается с подобным множеством сложных вопросов. Решения должны быть приняты относительно оборудования по оцифровке, базовых программных средств, дизайну интерфейса, стандартов метаданных и очень многих последовательных процессов. Но ответить всего на три, самых основных вопроса о развитии коллекции, возможно, даже намного труднее, чем на оперативные проблемы. Эти три основных вопроса таковы: (1) кто будет пользоваться цифровой библиотекой, (2) какие материалы должны быть оцифрованы для пользователей и (3) какие материалы должны быть оцифрованы в первую очередь.

Ответы на эти вопросы крайне важны для продолжительного успеха любой программы по оцифровке. Ответы должны быть найдены до того, как программа будет запущена, а также проверены и перепроверены по мере продвижения программы. Четко сформулированные однажды, эти ответы должны быть документально оформлены в широко изученной политике в области развития коллекции, которая поддерживается управляющими библиотек, служащими, пользователями и другими заинтересованными сторонами.

Политика в области развития хорошей цифровой коллекции четко определяет коллектив пользователей их информационные потребности. Эффективная политика также распознает заинтересованные стороны и показатели, по которым эти заинтересованные лица будут оценивать успех данной программы. Политика должна быть довольно гибкой, чтобы реагировать на изменения в операционной среде, а данные операционной среды должны собираться и непрерывно сообщаться. В идеале, данные будут подтверждать полезность того, что было оцифровано в прошлом, и обосновывать планы по реформатированию на будущее. Если нет, то планы по оцифровке должны быть откорректированы и изменены, чтобы политика в области развития была принята. Другими словами, использование данных применяется для создания

обратной связи, чтобы информировать политику в области развития и отбирать материалы для оцифровки.

Эта статья послужит отчетом по использованию данных в хранилище оцифрованных газет штата Юты -- очень успешной программы, которая до недавнего времени была частью инициативы по оцифровке газет Библиотеки Конгресса. До тех пор, пока данные из проекта Юты не могут быть непосредственно применены к другим попыткам реформатирования, они могут использоваться для предположения возможной ценности оцифровки газет в других учреждениях. Мы надеемся, что данные Юты проинформируют обсуждение политики в области развития коллекции об исследованиях библиотек, включая те национальные библиотеки, которые спонсируют крупномасштабные программы по оцифровке газет.

Газеты как приоритет для оцифровки в Соединенных Штатах Америки:

Оцифровка исторических и культурных материалов в Соединенных Штатах Америки стала приоритетной, когда в 1995 году Библиотека Конгресса запустила Национальную программу цифровой библиотеки. Хотя эта программа спонсировалась Библиотекой Конгресса, она предполагала совместные усилия, которые вели бы к распространению хранилищ управляемого содержимого. Эти хранилища были созданы свободно доступными обществу через разные интерфейсы, некоторые из которых могли бы быть изготовлены на заказ для конкретных потребителей и иметь специальное содержание¹.

Когда Библиотека Конгресса оценила предстоящую деятельность, она определила десять трудностей, с которыми столкнутся библиотечные исследования 21-ого века при создании своих цифровых коллекций². Восемь из десяти связаны с техническими препятствиями и одна -- с авторским правом. Оставшаяся трудность выражает проблему развития коллекции: для успеха Национальной цифровой библиотеки она должна быть «полезной различным пользовательским коллективам и для других целей».

Это условие для успеха признало важным ответ на основные вопросы, упомянутые выше, а именно, кто будет пользоваться цифровой библиотекой, какие материалы должны быть оцифрованы для пользователей, и какие материалы должны быть оцифрованы первыми? Библиотека Конгресса ответила на вопрос, кто будет пользоваться Национальной цифровой библиотекой, максимально широкими терминами описав это как «дополнение к каждому рабочему столу, классной комнате и личной библиотеке»³. Что касается того, какое содержимое должно быть оцифровано и что должно быть оцифровано в первую очередь, Библиотека Конгресса выделила исторические материалы в качестве самых приоритетных. С помощью щедрых грантов Ameritech Corporation, Библиотека Конгресса учредила проект American Memory Project в качестве проводника для быстро реформатированных исторических документов.

Цель проекта American Memory Project заключалась в оцифровке «первостепенных исторических сокровищ» в Библиотеке Конгресса и других наиболее важных исследовательских архивов. В сущности, все форматы приветствовались в новом онлайн хранилище -- от печатных трудов до звуковых записей и движущихся изображений⁴. Газеты никогда не были упомянуты в описаниях проекта, но за эти годы были добавлены выпуски новых газет. Учитывая важность газет, как первоначального ресурса материалов, они, тем не менее, возможно не были предоставлены в значительной мере в появляющейся национальной коллекции.

Этот недостаток был более чем компенсирован, когда в 2004 году Библиотека Конгресса и Национальный гуманитарный фонд объявили о Национальной программе цифровых газет. Эта программа была разрекламирована как преемница длительной и высоко успешной Программы газет Соединенных Штатов -- усилия двух организаций по описи, каталогу и микрофильму «фонда газетных материалов с повышенным риском». Под новой инициативой Национальный гуманитарный фонд оказал материальную поддержку учреждениям культурного наследия для целей отбора, оцифровки и передачи Библиотеке Конгресса приблизительно 100000 газетных страниц согласно гранту⁵. Было прогнозировано, что более 5,6 миллионов страниц будет создано благодаря Программе к 2013 году, но к моменту написания этой статьи, эта цель уже была перевыполнена.

Определение приоритетов между газетами маленького городка и мегаполиса⁶:

Когда Библиотека Конгресса и Национальный гуманитарный фонд организовали Национальную программу цифровых газет, они ответили на третий вопрос развития коллекции -- что должно быть оцифровано в первую очередь. Ответ был сформулирован в нормативных терминах для присвоения грантов. Нормативный раздел, озаглавленный «Критерии отбора интеллектуального содержания» и благоприятствующий большим газетам столиц, гласит: «Первоочередная причина отбора печатного издания для оцифровки в том, что эта газета в значительной степени отражает политическую, экономическую и культурную историю штата, что включает печатные издания, имеющие влияние на уровне штата или региона, но не ограничивается ими»⁷.

Печатные издания маленьких городков могли быть выдвинуты согласно нормативам, дающим постановление признать ценность газет, которые предлагают «общественные» новости, такие как объявления о рождениях, смертях и браках. Газеты маленьких городков также могли бы включать «статьи о сиротах» и были определены как газеты, которые перестали выпускаться, и для которых не было именитых владельцев авторских прав или планов по оцифровке. Многие газеты маленьких городков наверняка встречались с таким определением. Все же местные нормативы гранта Национальной программы по оцифровке газет, безусловно, оказывали содействие столичным газетам, заявляя: «Предпочтение должно быть отдано тем печатным изданиям, которые охватывают штат или как минимум макро-округ большинства населения».

Акцент в программе Соединенных Штатов на газетах мегаполисов также прослеживается и в других инициативах по оцифровке национальных газет. Например, Государственная библиотека в Австрии объяснила смысл данного предпочтения так: «Газеты с широким географическим охватом публиковались еще до 1900 года, то есть являются 'государственными' печатными изданиями»⁹. На большинство попыток по реформатированию крупномасштабных газет существенно влияли модели проектов Библиотеки Конгресса и Государственной библиотеки Австралии. Тем не менее, это не является универсальной истиной. Цифровые газеты штата Юты -- это прекрасный пример программы, которая имела сознательно предпринятые попытки, чтобы подвести баланс между газетами мегаполисов и газетами маленьких городков. Оставшаяся часть данной статьи будет посвящена рассмотрению результатов попыток Юты и, что наиболее важно, реакции пользователей на возможность доступа как к изданиям маленьких городков, так и к столичным.

Цифровые газеты штата Юты:

Программой цифровых газет штата Юты руководит Библиотека Университета штата Юты им. Дж. Уилларда Мэриотта (США). С ее скромного начала в 2002 году, программа разрослась. Сейчас, на ее одиннадцатом году, UDN содержит 1,3 миллиона страниц содержимого и признана

национальным и даже международным лидером в оцифровке газет. Она остается полностью открытой для доступа, бесплатным веб-сайтом, который можно посмотреть по адресу: <http://digitalnewspapers.org>. Она остается первой строкой результатов для «цифровые газеты» в поисковых системах Google и Yahoo.

С мая 2012 года UDN владеет 80 различными газетными изданиями с самых первых выпусков the Deseret News, изданных на Территории Юта в июне 1850 года до the Vernal Express, изданного в декабре 1982 года. Она владеет изданиями 27 округов в штате Юта из 29.

Некоторая другая статистика, относящаяся к размеру коллекции UDN:

Количество изданий	80
Количество штатов	27 из 29 ¹⁰
Количество выпусков газет	128 378
Количество страниц	1 318 349
Количество статей	15 368 563
Количество частных коллекций в базе данных	231
Общее количество объектов в базе данных	16 798 920

История цифровых газет штата Юты:

Программа UDN началась в начале 2002 года с гранта в размере 93 000 долларов от Государственной библиотеки штата Юты, которая приобрела серверное оборудование и обеспечила оцифровку 30000 страниц исторических газет Юты. В декабре 2002 года после нескольких месяцев экспериментирования процесса оцифровки первоначальный веб-сайт UDN был запущен с 10000 страниц каждого из трех печатных изданий. Весть об этом новом уникальном ресурсе быстро распространилась по сообществу библиотеки Юты.

Через несколько месяцев, за первым грантом в размере 93 000 долларов последовал второй, более солидный, грант от Государственной библиотеки штата Юты в размере 278 000 долларов. Это финансирование обеспечило проект руководителем, работающим весь день, и 106 000 страниц оцифрованного содержания, практически утраивающих размер базы данных.

Когда Библиотека получила грант Национального лидерства в размере 470 000 долларов от Института музейной и библиотечной службы (IMLS), федерального агентства Соединенных Штатов, в сентябре 2003 года, это стало поворотным событием.

Когда в 2005 году грант IMLS был использован, Национальный гуманитарный фонд (NEH), в сотрудничестве с Библиотекой Конгресса (LC), запустил Национальную программу цифровых газет (NDNP). Университет штата Юты был одним из шести учреждений, присудившим грант на первой испытательной стадии программы с 2005 по 2007 годы. В последствии Библиотека получала дополнительно от Национального гуманитарного фонда двухлетнюю награду -- в 2007 и 2009 годах -- гранты на общую сумму в 863 000 долларов для оцифровки 380 000 страниц содержимого.

На протяжении этих лет, Библиотека имела очень большой успех, увеличивая фонды штата от различных учреждений, на оцифровку местных газет. Она работала с педагогическими библиотеками, государственными библиотеками, самими газетами, историческими обществами и другими учреждениями культурного наследия. Самый большой из этих проектов был профинансирован на 527 000 долларов Историческим обществом штата Юты для оцифровки 247 000 страниц газеты Salt Lake Telegram. Это содержимое охватывает 50 лет (1902-1952) самой крупной ежедневной газеты Солт-Лейк-Сити и более 18% общей базы данных UDN¹¹.

Принципы действия:

На протяжении курса программы по оцифровке, UDN следовало шести простым принципам действия, все из которых предназначались для того, чтобы улучшить результативность работника. Первый: с самого начала проект цифровых газет штата Юты был сосредоточен на широких границах в масштабе штата и отображении. Особенно в первые годы UDN противостояла искушению оцифровывать крупные столичные печатные издания. Фактически, она осознанно занималась противоположной целью -- направляла свои силы исключительно на более маленькие еженедельные издания маленьких селений. Это позволило UDN сформировать спрос по всему штату, одновременно расширяя его хронологический охват еженедельных, вместо ежедневных документов.

Второй: после выбора заголовка для оцифровки, стратегия UDN состояла в том, чтобы сканировать материалы, начинающиеся с более ранних дат, а затем двигаться дальше во времени, насколько позволит доступное финансирование. Эта тактика дает программе возможность оцифровать набор материалов, которые вероятнее всего имели наибольшую потребность в сохранности, и пользовались наибольшим спросом у пользователей¹².

Третий: всегда, когда это было возможно, UDN использовала современные технологии для съемки изображений оригинальных печатных копий газет вместо сканирования потертых и устаревших кадров микрофильмов. Эти технически созданные изображения достойны 21-ого века. Изображение высокого разрешения, в свою очередь, способствовало более высокой точности для программного процесса оптического распознавания символов (OCR), что обеспечило более точный поиск результатов для пользователей.

Четвертый: протоколы процесса UDN включают обеспечение изображений и метаданных к каждой газетной статье. Весь текст OCR прикреплен к своему изображению статьи так, что полное изображение статьи может быть включено в результаты поиска. Это позволяет пользователям быстро просматривать и понимать контекст ответов, вернувшихся из поиска по базе данных. Большинство других программ по оцифровке газет в Соединенных Штатах не разделяют страницы изображений на их индивидуальные статьи из-за значительно повышающейся стоимости за это¹³. Кроме того, требуется гораздо более комплексная структура базы данных для того, чтобы управлять информацией, основанной на статьях. UDN, тем не менее, твердо знает, что метаданные уровня статей обеспечивают гораздо более полезный опыт работников и вполне достойны дополнительной платы. К тому же, до этого момента базе данных UDN удавалось удовлетворительно справляться с более комплексной структурой газетных выпусков.

Пятый: чтобы способствовать увеличению точности поиска, поставщик услуг по оцифровке UDN вручную зашифровывает заголовки статей. Фактически, они дважды зашифрованы и проверены, что означает, что два разных человека зашифровывают каждый заголовок и любые несоответствия улаживаются. Этот процесс гарантирует почти 100%-тную точность текста заголовка. И опять же, эти дополнительные процессы более дорогие, но программа UDN полагает, что улучшение соответствия в точности поиска оправдывает дополнительные затраты.

Шестой: чтобы оставаться на связи с работниками и получать их отклики, UDN на веб-сайте предлагает пользователям простой опрос об их использовании UDN. Опрос был запущен непрерывно с 2005 года и собрал почти 1500 ответов руководителей. Среди того многого, что было выявлено из этого опроса, было следующее:

- 84% пользователей оценили свой опыт на «хорошо» или «отлично»
- 79% скоро вернуться
- 74% расскажут другим о UDN
- 66% оценили точность поиска на «хорошо» или «отлично»

65% нашли новый ресурс для своих исследований
63% в результате использования UDN стали более осведомленными о собственной семейной истории
Самой частой просьбой на вопрос о том, что следует улучшить, было: «больше содержания!»

Опрос 2012 года:

Для постоянно проводимого веб-опроса пользователей авторы организовали новый опрос в течение марта-апреля 2012 года, чтобы задать более специфические вопросы об опыте работы с UDN. Вопросы были в основном направлены на две гипотезы данных документов:

специалисты по генеалогии составляют самую большую выявленную группу пользователей UDN; и
печатные издания маленьких городков более важны пользователям UDN, чем столичные газеты.

Опрос¹⁴ был отправлен по электронной почте 500 пользователям цифровой газеты штата Юты, чью контактную информацию программа собирала на протяжении нескольких лет. На эти письма было получено 137 ответов. Итоговые результаты были таковыми:

72% посещают UDN для генеалогических исследований
20% посещают для разного рода исторических исследований
87% находят полезными некрологи
Более 60% находят полезными другие виды генеалогических статей (объявления о рождениях и свадьбах)
Только 7% не находят генеалогические статьи полезными
Многие записывают историю семьи и поэтому также ищут основные биографические сведения
Старое содержание ценится гораздо выше, чем недавнее (смотри более детальное объяснение этому ниже)
44% находят меньшие, сельские газеты более полезными, в то время, как только 15% таковыми находят большие столичные газеты.

В опросе респондентов попросили расположить особые временные периоды в порядке от 1 (наиболее используемого) до 5 (наименее используемого). Результаты этого вопроса четко демонстрируют, что ранее содержание больше ценится пользователями:

<u>ОЦЕНКА</u>	<u>ПЕРИОД ВРЕМЕНИ</u>
1	1850-1874
2	1875-1899
3	1900-1939
4	1940-1979
5	1980-present
	1980-настоящее время

Интересно, что каждый последующий временной период оценен на один балл ниже (т.е. один шаг менее используем), чем предыдущий.

В опросе также спросили о пользе маленьких сельских газет по сравнению с пользой больших столичных. Результаты, отраженные по шкале из пяти пунктов, были такими:

- 19% находят маленькие сельские газеты гораздо более полезными
- 24% находят маленькие сельские газеты значительно более полезными
- 42% были нейтральны
- 6% находят большие столичные газеты значительно более полезными
- 9% находят большие столичные газеты гораздо более полезными

Две границы этого спектра показывают, что вдвое больше респондентов (19% против 9%) заявили, что маленькие сельские газеты гораздо более полезны, чем большие столичные периодические издания, и в четыре раза больше респондентов (24% против 6%) находят маленькие публикации как минимум значительно более полезными, чем издания больших населенных пунктов.

Эти данные полностью поддерживают авторские разногласия в том, что газеты маленьких городков на столько же, если не больше, важны пользователям, как и большие столичные печатные издания, а 1,3 миллиона страниц оцифрованного содержания UDN намеренно отражают информационные предпочтения своих пользователей.

Для иллюстрации, пять самых лучших печатных изданий с наибольшим количеством страниц в базе UDN -- это все большие столичные газеты штата Юты, четыре из которых принадлежат Солт-Лейк-Сити и одна -- Огдену:

1. Salt Lake Telegram	247 785 страниц
2. Salt Lake Tribune	132 264 страницы
3. Salt Lake Herald	130 820 страниц
4. Odgen Standard	120 548 страниц
5. Deseret News	<u>107 761</u> страница

Итого пять крупнейших изданий 739 178 страниц

Совокупно эти пять газет составляют 56% от общего количества страниц в базе данных. Уравновешивание содержание больших столичных ежедневных изданий все из маленьких, часто реже публикуемых печатных изданий. В более поздних категориях есть 75 изданий, включающих 587 959 страниц содержания, которые компенсируют выразительные 44% от всех страниц UDN. Достигнутое отношение больших газет к маленьким удивительно близко к 50 на 50, в частности давая общий спад маленьких периодических изданий в последние годы. Равновесие в содержании базы данных UDN -- это доказательство ее миссии в настоящий момент, заключающейся в управлении инициативой в масштабе штата, балансировании информации и исследовании потребностей всех регионов штата Юты.

Заключение:

Результаты опроса программы UDN подтвердили гипотезы сделанные в работе о том, что 1) около 72% респондентов-пользователей UDN являются специалистами по генеалогии, и 2) газеты маленьких городков более важны для пользователей UDN, чем газеты столичные. Авторам встречались случаи, доказывающие, что специалисты генеалогии составляют большинство или по крайней мере, множество пользователей и других заметных газетных веб-сайтов. Хотя специалистов по генеалогии, в других программах может быть меньше 70% от всех пользователей, авторы глубоко уверены, что специалисты по генеалогии будут составлять большинство пользователей. Полагаясь на подтвержденные предположения, авторы предлагают следующие рекомендации:

Масштабные программы по оцифровке должны проводить опросы для того, чтобы определить, кто пользуется их информацией, с какой целью ее используют, и какого рода информацию они хотели бы видеть следующей в оцифрованном виде. Было бы интересно узнать, являются ли специалисты по генеалогии крупнейшей группой пользователей архивов цифровых газет в мире. Мы советуем всем программам использовать такие же вопросы и, если необходимо, то дословно звучащие, как из проведенного нами в этом году опроса. В таком случае легко будет сопоставлять варианты ответов.

Национальные программы и другие инициативы масштабных программ по оцифровке должны сделать большим приоритетом газеты маленьких городков. Гранты должны быть направлены тем организациям, которые имеют проекты по спасению меньших, подверженных риску печатных изданий. Если результаты этого исследования могут быть продублированы для других архивов, тогда перенос приоритета по финансовой поддержке со столичных газет на газеты маленьких городов станет императивом.

Разного рода национальные программы и другие инициативы масштабных программ по оцифровке должны смотреть и думать, не только об историках и исследователях, как об основных пользователях газетных хранилищ. С повсеместностью Интернета мы должны создать «граждан-историков», которые численно будут намного превосходить традиционных историков и других формальных исследователей. Исторические газетные исследования претерпевают фундаментальные изменения, и эти программы должны принять новую парадигму, как им строить и вводить в действие свои хранилища.

Инициативы по оцифровке газет должны рассматривать сбор метаданных и адаптировать интерфейсы для удовлетворения нужд специалистов по генеалогическим исследованиям. Если эти исследователи в действительности являются крупнейшей категорией пользователей цифровых газет, то инвестирование в метаданные и интерфейсы будет правильным расходом средств.

Далее следуют приложения 1-5

ПРИЛОЖЕНИЕ 1

БИБЛИОТЕКА КОНГРЕССА / НАЦИОНАЛЬНЫЙ ВКЛАД В КРИТЕРИИ ГРАНТА ГУМАНИТАРНЫХ НАУК

1. Ценность исследования

Главным критерием в отборе статьи для оцифровки является то, существенно ли она отражает политическую, экономическую и культурную историю государства. Это включает (но не ограничивается) статьи, имеющие влияние на государственном или региональном уровне, и статьи, являющиеся источником информации о разнообразии этнических, расовых, политических, экономических, религиозных и иных особых читателей или групп по интересам.

Предпочтение должно быть отдано статьям, известным как «авторитетные издания» на уровне штата или округа и содержащим опубликованное надлежащее уведомление, новости штата или местные государственные дела и объявления о новостях сообществ, включая рождения, смерти и браки.

2. Географическое представление

Одной из целей NDNF является обеспечение широкого географического охвата американских газет в цифровом формате. Предпочтение должно быть отдано статьям, которые охватывают штат, или, по крайней мере, несколько округов с большинством населения.

3. Временной охват

Статьи, охватывающие географический регион или группу в течение длительных временных периодов, более предпочтительны, чем статьи, охватывающие краткий промежуток времени, либо имеющие пробелы в хронологии. Изначально сканирование будет ограничено периодом с 1836 по 1922 годы, но, в конечном счете, подходящие статьи могут находиться в диапазоне за много десятилетий до или после указанного периода. Непрерывность изложения предпочтительней коротких периодов и разбросанной хронологии¹⁵.

В категории «Уточненные критерии отбора»:

Статьи о сиротах

Дан огромный объем газетных материалов, опубликованных в течение обозначенного временного периода (1836-1922), и, скорее всего, понадобятся как частные, так и общественные усилия, чтобы сделать эти материалы в цифровом формате. Статьи с высокой исследовательской ценностью, переставшие публиковаться и не имеющие активных владельцев, возможно, войдут в будущие цифровые коллекции и должны особо рассматриваться NDNF¹⁶.

ПРИЛОЖЕНИЕ 2

ПРАВИЛА ДЛЯ АВСТРАЛИИ

Австралийская программа по оцифровке газет

Отбор статей -- Обзор

Как отбираются газетные печатные издания и выпуски?

Газетные печатные издания и выпуски были выбраны в большом количестве библиотеками штатов и территорий как часть национальной совместной программы. Отбор стал переходящим процессом, потому что:

Большинство австралийских газет, находящихся на микрофильмах, не являются собственностью Национальной библиотеки Австралии.

Важно иметь многостороннее суждение о первоначальном содержании службы.

В 2005 году Национальная библиотека Австралии начала планирование программы по оцифровке газет и очертила некоторые правила библиотекам штатов и территорий для первоначального отбора газет. Основные критерии отбора, очерченные для них, приводятся ниже:

Использование газет не должны нарушать авторские права, т.е. любой материал, опубликованный до 1955 года, приемлем.

Газеты должны существовать и в формате микрофильмов.

Микрофильмы, которые должны быть оцифрованы, должны иметь подходящее качество и полноту.

Каждый штат и территория должны быть представлены в равном объеме.

Газеты с широким географическим покрытием, опубликованные до 1900 года, имеют самую высокую значимость, т.е. являются «государственными» печатными изданиями.

Общее количество страниц, выбранное для стадии проекта 2007-2010 годов, для занесения в Национальную библиотеку Австралии, будет равно 3,4 миллионам.

На основе этих правил каждая из библиотек штатов и территорий выберет по одному из основных «государственных» печатных изданий для первоначального содержания. Некоторые библиотеки уже подготовились (либо делают это сейчас) к переоцифровке печатных изданий перед оцифровкой, поскольку качество и полнота существующих микрофильмов не удовлетворяют правилам.

В ноябре 2007 года Vincent Fairfax Family Foundation проспонсировал оцифровку Sydney Morning Herald до 1955 года и создание микрофильмов стоимостью 1 миллион долларов.

Это позволит достичь в сумме 4,4 миллиона страниц за период с 2007 по 2010 год.

После того, как 1,5 миллиона страниц было отсканировано, в мае 2008 года стало ясно, что количество страниц на микрофильмах было переоценено из-за расхождений между микрофильмами. Поэтому в консультации с библиотеками штатов и территорий имел место обзор отбора печатных изданий с июня по октябрь 2008 года для того, чтобы повысить количество выбранных изданий. Было предложено несколько новых изданий и дана новая оценка количеству изображений на взятых микрофильмах. Пересмотренный и расширенный список отобранных печатных изданий (все еще насчитывающих 4,4 миллиона страниц, но на данный момент содержащих 91 печатное издание) был опубликован 8 декабря 2008 года.

Как газетные печатные издания будут отбираться в будущем?

Оцифровка 4,4 миллионов страниц -- это кратковременная цель. Национальная библиотека собирается расширить содержание службы в будущем. Долгосрочной целью программы является создать свободный доступ ко всем австралийским газетам, изданным до 1955 года.

www.nla.gov.au/ndp

Последнее обновление: 8 декабря 2008 года

ПРИЛОЖЕНИЕ 3

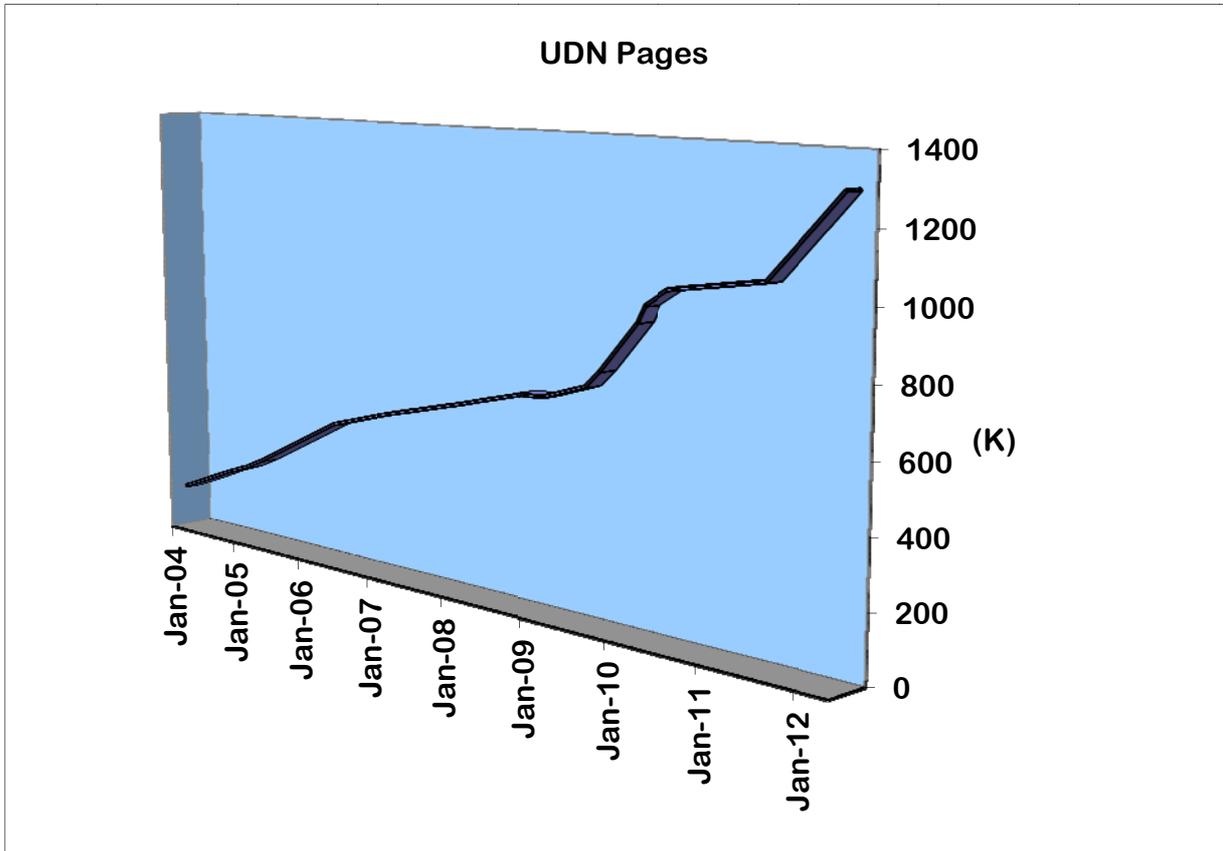
ВЕДОМСТВЕННЫЕ УЧРЕЖДЕНИЯ, ФИНАНСИРУЮЩИЕ ЦИФРОВЫЕ ГАЗЕТЫ ШТАТА ЮТЫ

National Endowment for the Humanities Национальный гуманитарный фонд	\$ 863,035
University of Utah Университет штата Юты	\$ 716,182
Utah Department of Community and Culture Департамент населения и культуры штата Юты	\$ 576,000
Institute of Museum and Library Services Институт музейных и библиотечных служб	\$ 469,514
Utah State Library Библиотека штата Юты	\$ 386,300
Utah Academic Library Consortium Консорциум академических библиотек штата Юты	\$ 161,400
Brigham Young University Университет им. Бригама Янга	\$ 103,033
Weber County Library Библиотека округа Вебер	\$ 70,000
Davis County Library Библиотека округа Дэвис	\$ 52,800
Uintah County Library Библиотека округа Уинта	\$ 43,748
National Park Service Служба национального парка	\$ 20,000
Park City Library Библиотека Парк-Сити	\$ 15,000
Southern Utah University Университет Южной Юты	\$ 15,000
Salt Lake Community College Колледж Солт Лейк Коммьюнити	\$ 10,463
Dee Foundation Dee Foundation	\$ 10,000
Murray City Library Библиотека города Мюррей	\$ 10,000
Grand County Library Большая окружная библиотека	\$ 8,100
Clipper Publishing	\$ 8,000
Utah State Historical Records Advisory Board Комиссия экспертов по историческим записям штата Юты	\$ 7,500
Bryce Canyon Natural History Association Национальная историческая ассоциация Брайс-Каньон	\$ 4,000
Delta City Library Библиотека города Дельта	\$ 2,800
Millard County Округ Миллард	\$ 2,500
Fillmore City Город Филлмор	\$ 2,500
Times Independent Times Independent	\$ 2,500

ИТОГО \$ 3,560,375

ПРИЛОЖЕНИЕ 4

РОСТ СТРАНИЦ UDN СО ВРЕМЕНЕМ



ПРИЛОЖЕНИЕ 5

ОПРОС РАБОТНИКОВ UDN В 2012 ГОДУ

May 07, 2012 8:54 AM PST

1 мая 2012 года 8:54

Опрос UDN -- 2012 год

Статус ответов: Выполнен

Фильтр: нет

1. Какова главная цель использования вами UDN?		
Генеалогия / Семейная история	96	72%
История Юты	7	5%
История местного населения	10	7%
История Запада США / региона	2	1%
Общие и/или исторические исследования	10	7%
Другое (поясните)	9	7%
Итого	134	100%

2. Для СПЕЦИАЛИСТОВ ПО ГЕНЕАЛОГИИ: какого рода информацию вы ищите? (Укажите все, что подходит)		
Объявления о рождении	74	63%
Объявления о свадьбе	79	68%
Некрологи	102	87%
Биографические сведения	79	68%
Общество в целом / исторические сведения	72	62%
Правовые/судейские записи	58	50%
Другое (поясните)	22	19%

3. Насколько для вас важна возможность просмотра и поиска UDN по особым, связанным с генеалогией, видам статей? Например, объявлений о рождении и свадьбе.

Верхний номер -- это число респондентов, выбравших данный вариант. Нижний -- это процент респондентов, выбравших данный вариант, от их общего числа.

Совсем не важна	Минимально важна	Достаточно важна	Очень важна
1	2	3	4
10	33	39	54
7%	24%	29%	40%

4. Пожалуйста, укажите тот период времени, который наиболее интересен для вас. Используйте стрелки, отмечая каждый пункт в правой ячейке от наиболее важного наверху к наименее важному внизу.

Верхний номер -- это число респондентов, выбравших данный вариант. Нижний -- это процент респондентов, выбравших данный вариант, от их общего числа

	Высокая важность				Низкая важность
1850-1874	57 49%	22 19%	16 14%	14 12%	8 7%
1875-1899	45 37%	48 39%	18 15%	12 10%	0 0%
1900-1939	25 20%	39 31%	52 42%	6 5%	2 2%
1940-1979	7 6%	16 15%	21 19%	62 57%	2 2%
1980-настоящее время	0 0%	2 2%	9 10%	7 8%	74 80%

5. Укажите ниже, что принесло бы наибольшую пользу вашим исследованиям.

Добавление документов за другие годы к тем, что уже есть в UDN	76	60%
Добавление газет, которые сейчас не являются частью UDN	51	40%
Итого	127	100%
49 ответов		

6. Находите ли вы маленькие сельские газеты более/менее полезными, чем большие столичные ежедневные газеты?					
Верхний номер -- это число респондентов, выбравших данный вариант. Нижний -- это процент респондентов, выбравших данный вариант, от их общего числа	Большие газеты гораздо более полезны, чем маленькие	Большие газеты значительно полезнее маленьких	Одинаково	Маленькие газеты значительно полезнее больших	Маленькие газеты гораздо более полезны, чем большие
	1	2	3	4	5
	12	8	56	32	25
	9%	6%	42%	24%	19%

7. Пожалуйста, перечислите ниже до 3 газет, которые вы используете наиболее часто.
116 ответов

8. Вы хотели бы посетить Университет штата Юты (Солт-Лейк-Сити) для участия в исследовании фокус-группы, связанном с UDN?		
Нет	90	68%
Да	7	5%
Если ваш ответ «Да», пожалуйста, введите контактную информацию	36	27%
Итого	133	100%