

Preservar y difundir antes de que sea tarde: evolución desde papel a bits mediante repositorios de prensa canaria

Víctor Macías Alemán

Universidad de Las Palmas de Gran Canaria
Biblioteca Universitaria. Sección de automatización y repositorios digitales
victor.macias@ulpgc.es

Comunicación presentada en el *II Congreso Historia del Periodismo Canario "De la autarquía franquista a la globalización, 1936-2016"* (Universidad de La Laguna, 24 y 25 noviembre 2016)

Resumen: La hemerografía histórica canaria constituye un recurso imprescindible para la investigación. Sin embargo, no existen planes regionales para el conocimiento de su inventario, tipología, localización, estado de conservación y aplicación de pautas para su preservación. El material del que está constituido dicho fondo documental favorece la autodestrucción progresiva de nuestro pasado periodístico por lo que, a falta de medios y financiación para tratar debidamente los originales, la única posibilidad actual de preservarla es la digitalización masiva. La Biblioteca de la Universidad de Las Palmas de Gran Canaria (BULPGC), de forma colaborativa con otras entidades públicas y privadas canarias, ha conseguido gestionar y ofrecer la versión digital de la prensa canaria de forma sistemática desde 1995. También se han podido digitalizar más de 1,1 millones de páginas de la prensa anterior a la llegada de la democracia. Sin embargo, está aún pendiente de acometer la digitalización y difusión de la mayor parte de la hemerografía canaria del pasado. Los servicios que será necesario ofrecer incluirán mejoras y enriquecimientos como la conversión de texto a voz, transcripción colaborativa, reusabilizar los contenidos mediante el uso de datos abiertos y enlazados, para que los periódicos y restantes publicaciones periódicas históricas canarias puedan sobrevivir y adaptarse al futuro.

Palabras clave: Prensa, publicaciones seriadas, Canarias, Digitalización, Difusión libre, Patrimonio.

Keywords: Newspapers, Serials, Canary Islands, Digitization, Open access, Heritage.

1. Estado de la cuestión

A lo largo de nuestra trayectoria en entidades que albergan fondos documentales, primero como usuarios y luego como gestores preocupados por satisfacer las demandas de acceso a nuestro patrimonio documental por parte de la población, hemos constatado que las hemerotecas han sido tradicionalmente vagón de cola en las prioridades, en las políticas de conservación y en la asignación de medios humanos y materiales frente a archivos y bibliotecas, particularmente cuando las comparamos dentro del ámbito patrimonial.

Pese a todo, los materiales que albergan son los más populares entre la población y suelen ser en muchos casos la única fuente para la historia local, donde las personas pueden encontrar detalles sobre sus familias y sobre los lugares donde han vivido a lo largo de generaciones¹.

La hemerografía, entendida en su sentido más amplio, incluye no sólo la prensa, sino cualesquiera otras publicaciones seriadas, ya que su conjunto constituye un legado documental imprescindible para la investigación. Ateniéndonos al ámbito de las Islas Canarias no existen sin embargo planes regionales para el conocimiento de su inventario, tipología, localización, estado de conservación o aplicación de pautas sistematizadas para su preservación, a diferencia de otros ámbitos como el del patrimonio fotográfico².

Hasta la fecha no contamos con una política de preservación en nuestras entidades públicas o privadas, ni con un plan para confeccionar los instrumentos de control que nos permitan conservar los ejemplares impresos de cada una de las publicaciones periódicas de cualquier género (prensa, revistas, boletines, anuarios, publicitarias, programas de fiestas, boletines oficiales, etc.) publicadas en el Archipiélago, de autoría canaria o que, aparecidas en cualquier otro lugar, traten sobre las Islas Canarias y su cultura.

Aunque hay bastantes estudios sobre la prensa canaria publicada hasta fecha relativamente reciente, contamos con escasa información pormenorizada sobre las mismas para localizar de forma unificada³ el paradero de los ejemplares de cada cabecera histórica, cuantificar el número de ejemplares conservados y de las páginas o, al menos, el promedio de las mismas, su grado de conservación, si se encuentran completos, encuadernados o sueltos, formato y características físicas detalladas, sus posibilidades de consulta y condiciones para la misma, incluyendo las publicaciones periódicas canarias publicadas y/o ubicadas en el extranjero.

En este género documental, el soporte papel utilizado en casi todas las cabeceras de prensa y otras publicaciones seriadas históricas, ha sido

¹ Es emblemático el lema del repositorio British newspapers archive: "millions of stories : from your street to the World".

² Inventario de fotografía de Canarias. <http://www.ifcanarias.com/guia-inventario.html>

³ A semejanza del Catálogo colectivo español de publicaciones periódicas

http://catalogo.bne.es/uhtbin/cgisirsi/x/0/0/57/49?user_id=CCPPWEB

confeccionado a base de celulosa de baja calidad con un alto índice de acidez. Este factor, sumado a las condiciones ambientales de Canarias, no favorece que el soporte en que se basa nuestro pasado periodístico perdure.



Fuente: British newspaper archive

Cuando hemos consultado ejemplares del pasado, no necesariamente pretérito, hemos podido observar los síntomas de degradación: cambio a sepia y ocre del color original, pérdida de la flexibilidad de las hojas con mayor riesgo de rotura en su consulta, descascarillamiento progresivo del papel sobre todo en bordes y pliegues, difuminación o borrado progresivo de las tintas, etc. Dicho de otro modo, casi todo nuestro legado hemerográfico está irremisiblemente condenado a la autodestrucción.

Es sumamente improbable, además, que en un plazo razonablemente cercano se cuente con los medios económicos necesarios para asegurar o, al menos, dilatar considerablemente en el tiempo, la vida de los ejemplares impresos más antiguos, ya que hablamos -según la fecha en la apliquemos el criterio de lo que se considera prensa histórica- de al menos, cientos de miles, cuando no de millones de páginas a tratar en nuestra región.

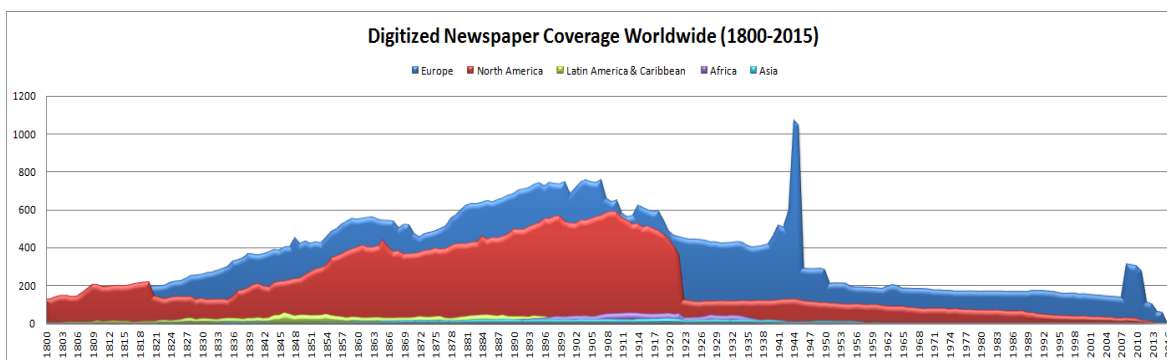
2. Digitalización, única posibilidad

A la vista de este panorama, la digitalización y puesta en línea puede conseguir lo que no sería posible por otros medios. La digitalización masiva es actualmente la única posibilidad a nuestro alcance para afrontar esta realidad, ya que permite tanto la preservación digital de la imagen facsímil del propio impreso, como de la información que contiene. El aserto tantas veces repetido de lo que no se conoce no existe, conlleva un especial valor en este caso, por cuanto, como sabemos, las colecciones hemerográficas, por la dificultad para ser completadas, almacenadas y debidamente descritas e inventariadas, no

son de fácil acceso para la población, lo que contribuye a su desconocimiento y por lo tanto a que no sean debidamente valoradas en la importancia que tienen.

Las razones para digitalizar, según la UNESCO, comportan ventajas evidentes como:

- ahorro de espacio
- ahorro económico en la gestión
- compartir sin límite, copiar sin límite, consultar sin límite, descargar sin límite.
- preservación del original evitando su uso directo
- fidelidad del original mediante la reproducción facsímil
- las herramientas de búsqueda asociadas a la consulta de la imagen digital son muy potentes e inexistentes fuera de la digitalización
- multiplica el número de las consultas
- aumenta el prestigio y visibilidad de las entidades que aportan material y lo difunden
- abre la posibilidad de formar parte de proyectos nacionales e internacionales
- favorece el acceso a fuentes de financiación
- permite reutilizar el contenido de los originales



Fuente: Center for research libraries

Sin embargo y para el caso que nos ocupa, las digitalizaciones retrospectivas son iniciativas complejas que requieren un estudio previo para establecer los procedimientos a seguir en las sucesivas fases del proyecto, como son la búsqueda de la financiación necesaria, el estudio de la tecnología a adoptar, la organización y descripción de la documentación y el cuidado de los derechos legales a aplicar. Además, hemos de preocuparnos por la preservación digital, de lo que trataremos más adelante.

Pero digitalizar no es simplemente sacar una “foto” a cada una de las páginas u hojas de un documento. Es necesario llevar a cabo este proceso de forma normalizada, ajustándose a directrices y protocolos nacionales e internacionales. Y, sobre todo, con la imprescindible organización, catalogación y descripción previa del material a procesar, ya que sin organización en el

plano real no puede haberla tampoco en el digital. De forma somera, las fases a considerar son las siguientes:

- Conocimiento del estado de los originales
- Organización y descripción del fondo documental
- Selección de los originales a procesar
- Recopilación según ubicaciones y formatos
- Escaneado desde copia en microfilmación
- Escaneado desde originales
- Generación de formatos secundarios
- Asignación de metadatos
- Estructuración de metadatos y objetos digitales
- Subida al repositorio
- Preservación digital

En la realización del propio trabajo es necesario hacer previamente –dentro de los parámetros escogidos para el tipo documental a procesar- las correspondientes pruebas que permitan conocer los ajustes más adecuados para cada tipo (luminosidad, contraste, brillo, resolución, profundidad de bits, corrección gamma, etc.). Basarnos, como se suele hacer en muchos casos, únicamente en que cuanto más resolución, mejor calidad final, conduce en según que casos a disparar el tamaño de los ficheros resultantes y, sobre todo, cuando aplicamos escala de grises o color sin necesidad, perjudicar comparativamente la calidad del reconocimiento de texto. Lo mismo sucede en la aplicación de profundidades de color a documentos que no lo precisan. Sin embargo, aplicar la digitalización mediante escala de grises es imprescindible en publicaciones con ilustraciones, grabados o fotografías.

Tras el trabajo de digitalización del impreso original, es necesario también generar una serie de ficheros subsidiarios y que deben ser, entre otros, los siguientes:

- Un fichero METS XML para cada uno de los ejemplares de la publicación periódica. Este archivo describe dicho ejemplar e incluye vínculos a todos los demás archivos que conforman la cabecera completa digitalizada.
- Un fichero ALTO XML para cada página de cada número de la publicación periódica. Estos ficheros contienen el texto del reconocimiento óptico de caracteres (OCR), así como información de diseño estructurado de la página en cuestión.
- Un fichero TIFF 6.0 ISO12639:2004 sin compresión para cada página de la publicación seriada.
- Un fichero JPEG ISO10918:1994 sin compresión para cada página de la publicación seriada.
- Un fichero PDF ISO32000-1 multipágina por cada ejemplar de la publicación seriada, y que incluirá inserto el texto obtenido mediante OCR.
- Un fichero PDF/A ISO19005-1:2005 de preservación para cada página de la publicación seriada.

El almacenamiento de estos ficheros debe hacerse en una estructura de directorios y carpetas coherente a la que debe aplicársele una sintaxis normalizada que nos sirva para identificar la localización de cada uno de forma unívoca, además de que el propio nombre del fichero actúe como identificación de su contenido.

3. Digitalizar para difundir

Las entidades que albergan nuestro patrimonio documental, sean privadas o públicas, al margen de cual sea su finalidad estatutaria, deben favorecer el conocimiento generalizado y sin cortapisas de la documentación que custodian. La conservación sin difusión sólo es de utilidad para las élites intelectuales o sociales que puedan acceder de forma restringida, mientras que el resto de la población potencialmente destinataria ignora su existencia por imposibilidad práctica de poder acceder a su contenido. La documentación que no puede ser consultada no puede formar parte de una investigación y, por consiguiente, no puede ser difundida para que sus datos puedan ser conocidos por todos. Hasta hace apenas dos décadas, antes del uso generalizado de la digitalización, muchas colecciones de prensa histórica eran escasamente empleadas para el análisis social debido a las dificultades de consulta en las hemerotecas donde se encontraban. A pesar de los avances habidos, este es todavía un grave inconveniente para los investigadores, por cuanto la prensa es una fuente de primer nivel para conocer la vida cotidiana y la realidad social de cada momento. A día de hoy, diversas mejoras lo han sido sólo en el plano físico dentro del propio lugar donde se encuentra la documentación, pero no tanto para hacer desaparecer las barreras que hagan posible su rápida consulta de forma remota y eficiente. Este problema subyacente en algunas entidades custodias se basa por lo general en un sentimiento arraigado de propiedad que creen incompatible con el libre acceso remoto y, por lo tanto, teóricamente “sin control”. Sin embargo, la difusión es la reversión a la sociedad de lo que el bien representa.

La generalización de la digitalización de la documentación de cualquier género y, en particular, de la prensa histórica, se ha convertido en un yacimiento prácticamente inagotable para conocer nuestro pasado en casi todos los campos, pudiendo emplearse casi para cualquier propósito o acción didáctica. La noticia de hoy es la historia de mañana, lo que nos permite – debidamente estructurada y difundida- ofrecer y obtener un enorme caudal de información que, además, crece en abrumadora proporción. El conocimiento de la documentación a través de la digitalización accesible sirve también para paliar su expolio, sustracción o pérdida que, de otro modo, podría permanecer desconocida por no estar a disposición general, lo que a fin de cuenta es patrimonio de todos.

Ejemplares de periódicos canarios antiguos disponibles a través de Internet

🕒 14 de enero de 2008

Más de 145.000 páginas de periódicos canarios publicados en el siglo XIX y durante las primeras décadas del XX pueden consultarse en la Red gracias al acuerdo al que han llegado la Universidad de La Laguna y la Dirección General del Libro del Ministerio de Cultura. Esta acción se enclava dentro del Plan Nacional para Preservar y Difundir el Patrimonio Bibliográfico Español que, de momento, ya ha digitalizado cuatro millones de páginas de publicaciones históricas.

Fuente: Noticias Universia. <http://noticias.universia.es>

Y lo que es más, no sólo sirve para acceder a la información, sino para divulgar también formas de pensamiento y culturas de una a otra parte del mundo. La forma de ser de los habitantes de las Islas Canarias, su historia, sociedad y cultura subyace en cada una de las páginas de nuestros periódicos y boletines. Si no los divulgamos y hacemos plenamente accesibles, ¿cómo podrá ser conocido lo nuestro y lo que somos frente a otras culturas y sociedades que cuentan con más medios y producción y que no paran de difundir a través de Internet su patrimonio, dando la equívoca idea de que no hay nada más que merezca ser considerado?.

En abril del pasado año 2015, el Center for Research Libraries convocó en la Biblioteca nacional de Suecia una cumbre de la International Coalition on Newspapers (ICON)⁴ para vindicar que los periódicos impresos del mundo constituyen una fuente incomparable e insustituible de historia crítica y de evidencia cultural. En los últimos años, bibliotecas y editores como ProQuest, Gale o Readex han invertido sumas considerables en la preservación y digitalización de esos materiales para hacer posible su acceso. Este encuentro reunió a representantes de estas organizaciones para evaluar el estado actual de la digitalización de periódicos y los principales logros hasta la fecha, como el National Digital Newspaper Program norteamericano o Europeana Newspapers, incluyendo el World Newspaper Archive del CRL. La idea del encuentro fue también auspiciar un consenso sobre las prioridades a adoptar en la digitalización futura y favorecer acuerdos sobre el intercambio de recursos, objetos digitales y metadatos necesarios para que el esfuerzo sea sostenible y abierto.

Para ello, se elaboró un informe de la cobertura actual de los diarios del mundo en las principales bibliotecas y de los esfuerzos de digitalización en los

⁴ <http://icon.crl.edu>

periódicos comerciales. También se generó un sistema de acreditación para evaluar la sostenibilidad y la integridad del contenido de los repositorios digitales de periódicos históricos.

Los resultados incluyeron compartir información entre los participantes de los problemas, costos y beneficios de los principales programas de digitalización de periódicos hasta la fecha; acuerdo sobre los términos generales para hacer posible una división del trabajo que fuera viable entre bibliotecas y editores comerciales en la digitalización masiva de periódicos mundiales, y el compromiso de los participantes para contribuir y mantener metadatos a nivel de ejemplar en periódicos digitalizados disponibles, así como otros protocolos para compartir información sobre proyectos de digitalización.

4. Plataformas y repositorios para la difusión

En Canarias se realiza desde hace mucho la digitalización de publicaciones periódicas, y en particular de prensa. Sin embargo, la mayoría de las acciones son de escaso calado y cobertura exclusivamente local, sin más propósito que servir a la entidad que los custodia y, en los más de los casos, no incluye el debido procesamiento y aplicación normalizada. Tampoco se pretende su pública difusión, al menos con la sistematización y protocolos que deben seguirse.

Sin embargo, existen portales que, desde la pasada década, compilan y divulgan este sector documental, como *Jable. Archivo de prensa digital de Canarias* (649 títulos con cobertura cronológica desde 1808), *Prensa canaria digitalizada* de la Universidad de La Laguna (103 títulos desde 1808), *Revistas en la Memoria digital de Canarias* (30 títulos desde 1762), *Publicaciones periódicas canario-americanas* del Centro de documentación de Canarias y América y también la *Memoria digital de Lanzarote*. Más allá de nuestro Archipiélago, la Biblioteca virtual de prensa histórica del Ministerio de Cultura ofrece también algunas cabeceras pertenecientes a la ULL.

Existen iniciativas de similar género en otras regiones, como *Arca. Arxiu de revistes catalanes antigues* o *Prensa digitalitzada* de la Generalitat catalana, aunque la inmensa mayoría de los fondos hemerográficos se concentran en repositorios de ámbito nacional como la ya mencionada Biblioteca virtual del Ministerio, la Hemeroteca digital de la Biblioteca Nacional de España o la Hemeroteca de la Biblioteca virtual Miguel de Cervantes. Todos estos repositorios de prensa, debidamente normalizados y que, por lo tanto, cumplen el protocolo OAI-PMH⁵ –caso de los portales gestionados por las dos universidades públicas canarias- son recolectados por Hispana, el agregador nacional de los proyectos de digitalización realizados en España.

Pasando a nivel europeo, es de obligada referencia y consulta Europeana, uno de los mayores portales patrimoniales del mundo, con más de 54 millones de ítems y donde se incorporan portales abiertos que cumplen los estándares

⁵ <https://www.openarchives.org/pmh/>

ESE/EDM⁶. Prueba de la importancia que la hemerografía y la prensa tienen dentro de este proyecto internacional europeo la atestigua *Europeana newspapers*⁷. Su objetivo es agrupar lo relativo a la prensa y sus contenidos dentro de *Europeana* para hacer que los periódicos digitales sean más accesibles, más fáciles de buscar, crear herramientas que ayuden a evaluar su calidad, realizar recomendaciones para la elaboración de mejores metadatos, y sensibilizar a través de jornadas y talleres formativos.

En el resto del mundo podemos localizar encomiables iniciativas de amplio calado como el National Digital Newspaper Program (NDNP)⁸ de la Biblioteca del Congreso de Estados Unidos, el British newspapers archive "*One place, millions of stories*"⁹, Elephind. *Search the world's historical newspapers archives*¹⁰, el Australian newspapers plan¹¹ de los cuales podemos tener consultar un amplio –aunque no exhaustivo– censo en el Newspaper digitization projects del ICON¹², en el ya citado World newspaper archive¹³ del CRL e incluso en la Wikipedia¹⁴.

5. Acceso permanente y, por lo tanto, de la preservación digital

La puesta en línea de forma debida, organizada, estructurada y protocolizada, asumiendo la normativa y estándares nacionales e internacionales no es, pese a todo, el único de los trabajos que debemos asumir a la hora de divulgar la prensa y restantes publicaciones periódicas digitales y digitalizadas en un repositorio. Hay que incorporar todo lo necesario para que dicho acceso quede garantizado a lo largo del tiempo en lo que se conoce como preservación digital.

La preservación es costosa de aplicar –tanto en trabajo como financieramente– y mantener. Es necesaria la adopción, como mínimo, del modelo OAIS (ISO 14721:2003)¹⁵, la incorporación, comúnmente de metadatos PREMIS¹⁶ para gestionar la durabilidad a largo plazo, el refresco sucesivo de formatos informáticos, de los soportes donde se contienen, de la redundancia de los mismos en aras de la seguridad y de emulación a la hora de reproducir y visualizar las digitalizaciones, sus contenidos y los metadatos estructurados, caso de ser necesario.

⁶ <http://pro.europeana.eu/page/ese-documentation>
<http://pro.europeana.eu/page/edm-documentation>

⁷ <http://www.europeana-newspapers.eu>

⁸ <https://www.loc.gov/ndnp/>

⁹ <http://www.britishnewspaperarchive.co.uk>

¹⁰ <http://elephind.com>

¹¹ <http://www.nla.gov.au/australian-newspaper-plan>

¹² <http://icon.crl.edu/digitization.php>

¹³ <https://www.crl.edu/collaborations/global-resources-partnerships/news/world-newspaper-archive>

¹⁴ https://en.wikipedia.org/wiki/Wikipedia:List_of_online_newspaper_archives

¹⁵ https://en.wikipedia.org/wiki/Open_Archival_Information_System

¹⁶ <http://www.loc.gov/standards/premis/>

Pero, yendo más allá, la *data curation* o preservación de datos es aún más amplia, puesto que asume el ciclo de vida completo del documento digital desde su generación, atendiendo a la selección, recopilación, archivo, conservación y mantenimiento de activos en repositorios digitales. Se entiende por tanto, que la “curación digital” es un proceso que engloba entre otras tareas las propias de la preservación digital.

Es importante hacer mención a instrumentos como el elaborado por los ya citados CRL y la ICON, consistente en un esquema de evaluación de repositorios de periódicos digitalizados¹⁷ para chequear repositorios de periódicos digitalizados con el objetivo de medir objetivamente la viabilidad de estos portales para proporcionar acceso persistente a las noticias digitalizadas y mantener la integridad documental de su contenido a lo largo del tiempo.

6. La imprescindible cooperación

Canarias fue pionera en España en la preservación de la prensa corriente publicada a partir del año 1995. Este hecho fue posible gracias a la cooperación espontánea de una serie de entidades públicas y privadas –las bibliotecas de la Universidad de Las Palmas de Gran Canaria y de La Laguna, las municipales de Santa Cruz de Tenerife y de Las Palmas de Gran Canaria, la de El Museo Canario y otras que se fueron sumando al cabo de los siguientes años– que colaboraron financieramente. Su finalidad fue abaratar costes y hacer posible que se digitalizaran los periódicos que, por entonces, sólo estaban disponibles en papel y suponían un grave problema de almacenamiento, de organización física para hacer posible su consulta y aumento del deterioro de los ejemplares debido a su uso directo. Gracias a dicha iniciativa, se ha preservado digitalmente la totalidad de la prensa corriente del Archipiélago de los últimos veinte años, labor que continúa en la actualidad con propósito de ser mantenida de forma indefinida.

¹⁷https://www.crl.edu/sites/default/files/d6/attachments/events/PART_1_Digitized_Newspaper_Repository_Assessment_Scheme.pdf



Fuente: Memoria digital de Lanzarote

Esta cooperación se vió posteriormente ampliada cuando la Biblioteca de Las Palmas de Gran Canaria puso en marcha en 2006 un primer sistema de consulta en línea de forma remota, abriendo la posibilidad de comenzar la digitalización retrospectiva de prensa histórica. Este hecho fue posible mediante acuerdos establecidos al efecto con algunas de las más importantes hemerotecas de Canarias, como la de la Biblioteca de la Universidad de La Laguna o la Biblioteca pública municipal de Santa Cruz de Tenerife. A ellas se sumaron posteriormente las de la Biblioteca pública municipal de Puerto del Rosario o del Centro de datos del Cabildo Insular de Lanzarote, que ha gestionado ejemplarmente el acceso a la casi totalidad de la prensa y publicaciones seriadas de aquella Isla con propósito de digitalizarlas y difundirlas. También se ha contado con la aportación de numerosos particulares, así como editores que han mostrado su interés para que las publicaciones seriadas que generan puedan tener más amplia divulgación y preservación para el futuro por este medio. Por fin, en 2008, se contó con un repositorio específico plenamente capaz para este propósito: *Jable. Archivo de prensa digital de Canarias*¹⁸ que, gracias a esta amplia participación colectiva, compila ya más de 1,1 millones de páginas de prensa y publicaciones seriadas canarias aparecidas hasta la llegada de la democracia. La cooperación no sólo ha servido para digitalizar y difundir de forma directa este patrimonio, sino también para revisar las colecciones con las que ya se cuenta e intentar completarlas a la búsqueda de la mayor exhaustividad posible.

¹⁸ <http://jable.ulpgc.es>

7. Renovarse o quedarse atrás

Da igual lo interesante que sea el patrimonio que se ofrezca. Ha de competir en Internet, en la red de redes, contra el mundo entero consultable a través de una pantalla. Por ello, no basta con ofrecer, de forma más o menos normalizada, más o menos eficiente, nuestros contenidos. Además, ha de hacerse con las tendencias y tecnologías que marca cada momento para que los usuarios, bombardeados por mil y un reclamos digitales, sepan que podemos ofrecerles no sólo lo mejor, si no además de la forma más atractiva y acompañada de servicios complementarios.

Entre las numerosas posibilidades a desarrollar, la más perentoria es, actualmente, la adaptación *responsive* –redimensionamiento de los contenidos y funcionalidades para su uso con dispositivos móviles- por parte de cualquier usuario y desde cualquier lugar, convirtiendo en realidad lo que hasta ahora era impensable: que los ordenadores no sean imprescindibles para acceder a los repositorios.

Complementando esta adaptación, está la promoción de nuestros contenidos y novedades mediante redes sociales a través de las ya más que utilizadas Facebook, Twitter, Flickr, los tableros de Pinterest, la elaboración de pildoras informativas, etc.



Fuente: <https://es.pinterest.com/bulpgc/prensa-y-revistas-canarias-digitalizadas-en-jable>

También es necesaria la elaboración de SiteMaps para los buscadores con el fin de que sean localizables a través de los mismos cada una de las partes que componen un título.



Need to analyse the quality of your digitised newspapers?

Fuente: Europeana newspapers

La evolución natural, desde la página a cada una de las partes que la integran, mediante segmentación de los diversos contenidos, de cada noticia, anuncio, apartado, artículo de cada página de periódico digitalizado cobra una importancia trascendental. Por ello, mediante la aplicación de los estándares METS/ALTO se puede pasar, del componente mínimo actual –la página- a cada una de sus partes coherentes. De esta forma, se podrán ofrecer resultados más pertinentes en los que se le ofrecerán los artículos o apartados coincidentes con su búsqueda, en lugar de únicamente una secuencia de páginas. Además, cada parte se puede también categorizar. Por ejemplo, se puede etiquetar como anuncios, ilustraciones o esquelas. Estas categorías permiten opciones de búsqueda mucho más eficientes, ya que los usuarios podrán restringir sus búsquedas a unos o excluir otros en sus pesquisas. Esto también permitirá que cualquiera de dichas partes pueda ser recortada digitalmente dentro de la página, con la finalidad de guardarla, enviarla o imprimirla.

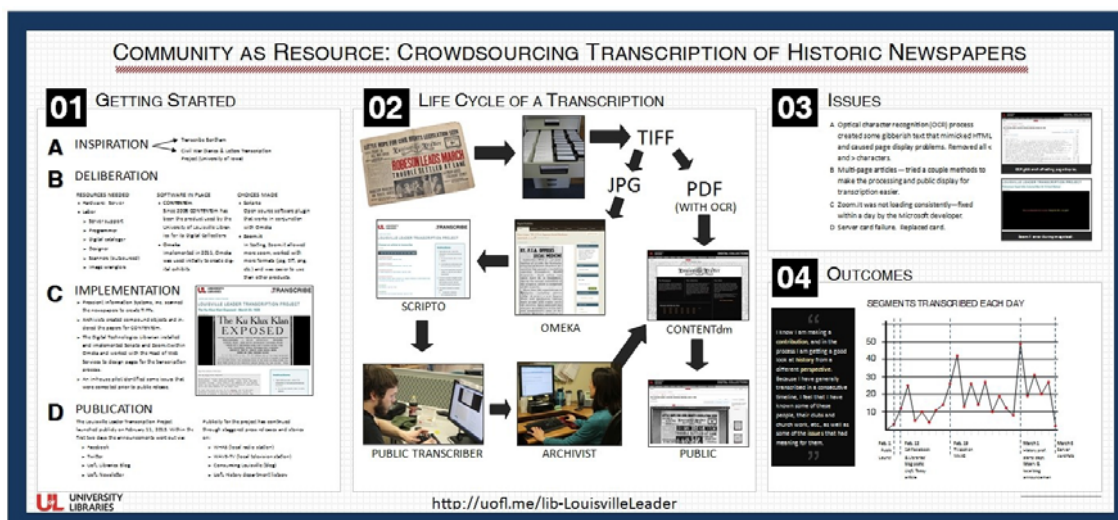
Las utilidades informáticas libres ofrecidas, entre otros por Europeana Newspapers tools¹⁹ ayudan a obtener este tipo de productos y a facilitar que nuestros repositorios evolucionen.

Pero los repositorios que, como Jable, no recogen únicamente prensa o publicaciones históricas, sino también vivas, la automatización validada de nuevos contenidos, cada vez más masivos, es una cuestión de amplio calado. En estos casos es necesario adoptar la carga mediante protocolos como SWORD²⁰ que faciliten dicho cometido.

¹⁹ <http://www.europeana-newspapers.eu/public-materials/tools/>

²⁰ [https://en.wikipedia.org/wiki/SWORD_\(protocol\)](https://en.wikipedia.org/wiki/SWORD_(protocol))

Otra evolución y mejora considerable para la legibilidad del texto contenido en las digitalizaciones, viene dada por la necesidad de solucionar la escasa calidad de su reconocimiento automático OCR y que, en algunos casos, es debida a ejemplares en muy mal estado de conservación, cuando la digitalización se ha realizado desde microfilm, incorpora fuentes tipográficas poco corrientes, los tipos están gastados o sucios, las letras están superpuestas o muy juntas, las columnas están fragmentadas, etc. Esta cuestión puede resolverse implementando un sistema de transcripción humana colaborativa abierta a la participación de usuarios con plataformas como Scripto²¹.



Fuente: Louisville University Library

Los resultados de dicho reconocimiento pueden servir posteriormente para implementar lectura para discapacitados mediante síntesis de voz, utilizando técnicas Text-to-speech (TTS), o bien como medio alternativo a la lectura textual tradicional y que ya ha dado lugar a incontables posibilidades como la generación de audiolibros y entre los que figuran contenidos de la Memoria digital de Canarias (mdC)²².

Para terminar, señalemos que desde 1993, la Unión Europea viene demostrando en sus documentos oficiales el impacto de las tecnologías en la sociedad, en el mercado de trabajo y en la educación, con especial incidencia en la sociedad del conocimiento. Por ello, los enfoques tradicionales han de ser sustituidos por otros que favorezcan la interactividad, conectividad y la dinámica comunicativa. Por ello, un repositorio no debe ser sólo una virtualidad que ofrezca digitalizaciones, sino que debe ofrecer servicios basados en la reutilización de sus contenidos. Esto hace imprescindible incorporar la gestión de datos abiertos y enlazados (Linked open data)²³. Cuando hablamos de LOD, nos referimos a datos estructurados que están interconectados con datos de otras fuentes de datos basados en tecnologías web estándar como HTTP, RDF y URLs unívocas o persistentes.

²¹ <http://scripto.org>

²² <http://mdc.ulpgc.es/u/?mdcm,667>

²³ <https://www.youtube.com/watch?v=arC3VEW1wWQ>

Bibliografía

- Abadal, Ernest y Guallar, Javier. *Prensa digital y bibliotecas*. Gijón: Trea, 2010. ISBN 978-84-9704-446-2.
- Adaryukov, Andrew. *Cooperative dimensions of a digitization project*. Resource Sharing & Information Networks, 2004. V. 17, n. 1-2, p.175-185. http://www.tandfonline.com/doi/pdf/10.1300/J121v17n01_14
- Chern Li Liew. *Collaborative construction of digital cultural heritage: a synthesis of research on online sociability determinants*. D-Lib Magazine. November/December 2015. Volume 21, Number 11-12
<http://dlib.org/dlib/november15/liew/11liew.html>
- Daniels, Caroline, Terri L. Holtze, Rachel I. Howard, and Randy Kuehn. *Community as resource: crowdsourcing transcription of an historic newspaper*. Journal of electronic resources librarianship 26.1 (2014): 36-48.
<http://www.crowdconsortium.org/wp-content/uploads/COMMUNITY-AS-RESOURCE-CROWDSOURCING-TRANSCRIPTION-OF-HISTORIC-NEWSPAPERS.pdf>
- *Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos*. Madrid: Ministerio de Cultura, 2005. 174 p. ISBN 84-8181-271-4.
- Ríos Hilario, Ana; Martín Campo, Diego y Ferreras Fernández, Tránsito. *Linked data y linked open data : su implantación en una biblioteca digital. El caso de Europeana*. El profesional de la información, 2012, mayo-junio, v. 21, n. 3. <http://hdl.handle.net/10366/115842>
- Rubin, Jeffrey A. *Collaboration as a means to add value in digital libraries*. En: Congreso Nacional de Bibliotecas Públicas (6º. 2012. Burgos). <http://hdl.handle.net/10421/6681>
- Summerlin, Donnie. *Selecting newspaper titles for digitization at the Digital Library of Georgia*. D-Lib Magazine, sep.-oct. 2014. V. 20, n. 9/10. <http://dlib.org/dlib/september14/summerlin/09summerlin.html>
- Vives, Josep (coord.). *Digitalización del patrimonio: archivos, bibliotecas y museos en la red*. Barcelona: UOC, 2009. 220 p. ISBN 978-84-9788-833-2.
- Widmer, Lois; Taylor, Laurie N. y Sullivan, Mark V. *Born digital newspaper preservation workflows for the Florida Digital Newspaper Library and the Caribbean Newspaper Digital Library*. George A. Smathers Libraries, University of Florida, 2012