

III. OTRAS DISPOSICIONES

MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE

- 11739** *Resolución de 23 de octubre de 2014, de la Dirección General de Política e Industrias Culturales y del Libro, por la que se publica el Convenio de colaboración con la Xunta de Galicia para la generación de un recurso digital procedente del escaneo y tratamiento bibliográfico de publicaciones en papel custodiadas en la Biblioteca de Galicia.*

La Dirección General de Bellas Artes y Bienes Culturales y de Archivos y Bibliotecas del Ministerio de Educación, Cultura y Deporte y la Consellería de Cultura, Educación y Ordenación Universitaria de la Xunta de Galicia han suscrito un Convenio de colaboración para la generación de un recurso digital procedente del escaneo y tratamiento bibliográfico de publicaciones en papel custodiadas en la Biblioteca de Galicia, por lo que, en cumplimiento de lo dispuesto en el artículo 8.2 de la Ley 30/1992, de 26 de noviembre, de Régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común, y en el Acuerdo del Consejo de Ministros de 2 de marzo de 1990, procede la publicación en el «Boletín Oficial del Estado» de dicho Convenio, que figura como anexo a esta resolución.

Madrid, 23 de octubre de 2014.—La Directora General de Política e Industrias Culturales y del Libro, Teresa Lizaranzu Perinat.

ANEXO

Convenio de colaboración entre la Dirección General de Bellas Artes y Bienes Culturales y de Archivos y Bibliotecas del Ministerio de Educación, Cultura y Deporte y la Consellería de Cultura, Educación y Ordenación Universitaria de la Xunta de Galicia para la generación de un recurso digital procedente del escaneo y tratamiento bibliográfico de publicaciones en papel custodiadas en la Biblioteca de Galicia

En Madrid, a 30 de septiembre de 2014.

REUNIDOS

De una parte, el Sr. D. José Ignacio Wert Ortega, Ministro de Educación, Cultura y Deporte, nombrado por Real Decreto 1826/2011, de 21 de diciembre, en ejercicio de las competencias que le confiere el Real Decreto 257/2012, de 28 de enero, en relación con el artículo 6 y la disposición adicional decimotercera de la Ley 30/1992, de 26 de noviembre, de Régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común y el artículo 13.3 de la Ley 6/1997, de 14 de abril, de Organización y Funcionamiento de la Administración General del Estado.

De otra parte, el Sr. D. Jesús Vázquez Abad, nombrado por el Decreto 9/2012, de 3 de enero de 2012, que actúa en nombre y representación de la Consellería de Cultura, Educación y Ordenación Universitaria, en ejercicio de las facultades que le confiere el artículo 34 de la Ley 1/1983, del 22 de febrero, reguladora de la Xunta y de su Presidente, así como el artículo 2 de la Resolución del 8 de abril de 1991 por la que se publica el acuerdo del Consello de la Xunta de Galicia del 27 de marzo de 1991.

EXPONEN

Primero.

Que según las Directrices para proyectos de digitalización elaboradas por un grupo de expertos de IFLA (The International Federation of Library Associations and Institutions) e ICA (International Council on Archives) a petición de la UNESCO, las principales razones para la digitalización son mejorar y facilitar nuevas formas de acceso y uso, preservar las colecciones bibliográficas y documentales, ofrecer a las instituciones oportunidades para el desarrollo de sus infraestructuras técnicas y para la formación técnica de su personal e impulsar el desarrollo de recursos cooperativos.

Segundo.

Que el artículo 44.1 de la Constitución Española establece que «los poderes públicos promoverán y tutelarán el acceso a la cultura, a la que todos tienen derecho».

Tercero.

Que el artículo 46 de la Constitución Española establece que «los poderes públicos garantizarán la conservación y promoverán el enriquecimiento del patrimonio histórico, cultural y artístico de los pueblos de España y de los bienes que lo integran, cualquiera que sea su régimen jurídico y su titularidad. La ley penal sancionará los atentados contra este patrimonio».

Cuarto.

Que, según el artículo 1 del Real Decreto 257/2012, de 27 de enero, por el que se desarrolla la estructura orgánica básica del Ministerio de Educación, Cultura y Deporte (BOE de 28 de enero de 2012), le corresponde al Ministerio de Educación, Cultura y Deporte la «promoción, protección y difusión del patrimonio histórico español (dentro del cual se encuentra el patrimonio bibliográfico), así como «de las artes, del libro, la lectura y la creación literaria, de las actividades cinematográficas y audiovisuales y de los libros y bibliotecas estatales, y la promoción y difusión de la cultura en español.

Quinto.

Como consecuencia de todo lo anterior, el Ministerio de Educación, Cultura y Deporte impulsa y ejecuta, en colaboración con las Comunidades Autónomas, con el resto de administraciones públicas y con todo tipo de entidades privadas, proyectos de digitalización de ámbito supracomunitario consistentes en la generación de recursos digitales procedentes del escaneo y tratamiento bibliográfico de publicaciones en papel con el objetivo de contribuir a su preservación y mejorar el acceso a las mismas haciéndolas accesibles a través de internet.

Sexto.

En relación con la propiedad intelectual y los derechos de autor, solo se generan recursos digitales procedentes del escaneo y tratamiento bibliográfico a partir de publicaciones en papel cuyos derechos de explotación han pasado al dominio público por haberse extinguido conforme a lo establecido en el Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual. De otro modo, esto es, en el supuesto de publicaciones en papel cuyos derechos de explotación no hayan pasado al dominio público por haberse extinguido conforme a lo establecido en dicho Real Decreto Legislativo 1/1996, pero cuya inclusión en un proyecto de digitalización se considere de interés conforme a lo establecido en el punto primero de esta exposición, el Ministerio de Educación, Cultura y Deporte recabará la autorización de

los titulares de dichos derechos de explotación para poder generar recursos digitales –mediante escaneo y tratamiento bibliográfico– procedentes de dichas publicaciones en papel, así como para difundir dichos recursos digitales a través de Internet.

Séptimo.

Dado que las publicaciones en papel o microfilmes u otro soporte digital custodiados por la Consellería de Cultura, Educación y Ordenación Universitaria en la Biblioteca de Galicia forman parte del patrimonio histórico-artístico gallego y español, deberán tenerse en cuenta las disposiciones de la Ley 16/1985, de 25 de junio, del Patrimonio Histórico Español y la Ley 5/2011, del 30 de septiembre, del patrimonio de la Comunidad Autónoma de Galicia.

En consecuencia, el Ministerio de Educación, Cultura y Deporte y la Xunta de Galicia acuerdan formalizar el presente Convenio de Colaboración, conforme a las siguientes

CLÁUSULAS

Primera. *Objeto del convenio.*

El presente convenio tiene por objeto la generación de un recurso digital procedente del escaneo y tratamiento bibliográfico de las publicaciones en papel *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926), custodiado en la Biblioteca de Galicia, con la finalidad de mejorar el acceso al contenido de dichas publicaciones y lograr su preservación.

Segunda. *Propiedad intelectual y derechos de autor.*

Los derechos de explotación sobre las publicaciones en papel reseñadas en la cláusula primera, a partir de las cuales se va a generar el recurso digital mencionado en dicha cláusula, se han extinguido y han pasado al dominio público, todo ello conforme a lo dispuesto en el Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual. En consecuencia, es lícito generar un recurso digital mediante el escaneo y tratamiento bibliográfico de dichas publicaciones en papel y difundir dicho recurso digital a través de Internet sin recabar, en ambos casos, autorización alguna.

Tercera. *Actuaciones de la Consellería de Cultura, Educación y Ordenación Universitaria (Consellería, en lo sucesivo).*

1. Siguiendo las instrucciones del Ministerio de Educación, Cultura y Deporte, la Consellería entregará a la empresa contratada por aquel o, en su caso, al propio Ministerio de Educación, Cultura y Deporte, los ejemplares de la publicación en papel de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926), al objeto de proceder a la generación de un recurso digital a partir de ella, procedente del escaneo y tratamiento bibliográfico de la misma.

2. El Ministerio de Educación, Cultura y Deporte autoriza a la Consellería a realizar, permanentemente y en todos los países del mundo, los actos de reproducción, distribución, comunicación pública y transformación que estime convenientes del recurso digital generado procedente del escaneo y tratamiento bibliográfico de la publicación en papel de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926), en cualquiera de las modalidades de explotación existentes y conocidas al día de la fecha.

Cuarta. *Actuaciones del Ministerio de Educación, Cultura y Deporte.*

1. El Ministerio de Educación, Cultura y Deporte, por medio de la empresa que, en su caso, contrate, generará un recurso digital procedente del escaneo y tratamiento

bibliográfico de la publicación en papel de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926), financiando el coste completo de dicha generación conforme a sus disponibilidades presupuestarias.

2. Los trabajos de generación del recurso digital a los que se refiere el número anterior se realizarán conforme a la normativa técnica y demás recomendaciones de organismos internacionales que se vienen aplicando por dicho Departamento ministerial. En concreto, se ajustará a las condiciones técnicas y controles de calidad del Ministerio de Educación, Cultura y Deporte que se especifican en el anexo I.

3. El Ministerio de Educación, Cultura y Deporte facilitará una copia del recurso digital procedente del escaneo y tratamiento bibliográfico de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - año XXX, n.º 13.500 (26 de abril de 1926) a la Consellería.

4. Corresponde al Ministerio de Educación, Cultura y Deporte la conservación del recurso digital generado procedente del escaneo y tratamiento bibliográfico de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926). El master de la reproducción en formato TIFF Grupo 6 quedará en posesión plena y total del Estado español para su conservación.

5. El Ministerio de Educación, Cultura y Deporte difundirá a través de Internet, en el marco de los proyectos de «Bibliotecas Virtuales» que tiene en marcha, el recurso digital generado procedente del escaneo y tratamiento bibliográfico de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - año XXX, n.º 13.500 (26 de abril de 1926). Además, el Ministerio de Educación, Cultura y Deporte podrá difundir dicho recurso digital por cualquier medio y en el marco de cualquier otro proyecto, actuación o iniciativa de cualquier clase que dicho Ministerio adopte para el cumplimiento de los fines que tiene encomendados.

6. El Ministerio, directamente o a través de la empresa contratada para la digitalización, restituirá a la Consellería la publicación en papel de *El Eco de Santiago* - Año I, n.º 1 (1 de marzo de 1896) - Año XXX, n.º 13.500 (26 de abril de 1926) en las mismas condiciones en que le fue entregada, en el plazo máximo de dos meses desde la finalización de los trabajos objeto de este convenio.

Quinta. *Financiación.*

Las actuaciones que llevará a cabo el Ministerio de Educación, Cultura y Deporte en base al presente convenio implican la digitalización, asignación de los metadatos y generación de los formatos de 50.000 páginas de *El Eco de Santiago*, desde el 1 de marzo de 1896 al 31 de diciembre de 1938. A tal fin, la Dirección General de Bellas Artes y Bienes Culturales y de Archivos y Bibliotecas del Ministerio de Educación, Cultura y Deporte ha abierto el correspondiente proceso de contratación conforme a lo establecido en el Real Decreto Legislativo 3/2011, de 14 de noviembre, por el que se aprueba el Texto Refundido de la Ley de Contratos del Sector Público, mediante el cual aportará las cantidades necesarias (35.000 €) con cargo a la aplicación presupuestaria 18.11.640.06 del Programa 332-B, código de proyecto 2005.24.04.0001, del Presupuesto de Gastos del año 2014.

Sexta. *Vigencia del convenio.*

La vigencia del presente Convenio está supeditada a la realización de su objeto, sin perjuicio de los compromisos acordados. El plazo máximo de ejecución será el 31 de diciembre de 2014.

Séptima. *Modificación y resolución.*

1. El presente Convenio de Colaboración podrá ser modificado de mutuo acuerdo entre las partes, siempre que el oportuno acuerdo de modificación sea formalizado y se produzca antes de la expiración de su plazo de duración.

2. Serán causas de resolución de este Convenio:
- El mutuo acuerdo.
 - El incumplimiento de las cláusulas del mismo por alguna de las partes.
 - La denuncia motivada de una de las partes, que deberá comunicarse por escrito a la otra parte, al menos con un mes de antelación a la fecha prevista de extinción del convenio que figura en la cláusula sexta.
 - Por las demás causas de extinción previstas en la legislación vigente.

Octava. Difusión de la colaboración de las partes.

En las actuaciones informativas, divulgativas y publicaciones que se realicen relacionadas con las actividades objeto de este convenio se hará constar, en un lugar destacado, la colaboración de las dos partes y se usará, cuando proceda, los anagramas de cada una de las entidades aprobados por la normativa de imagen corporativa.

Sea cual fuere el medio de difusión, cada una de las partes se compromete a hacer mención de la otra parte y de la existencia de este convenio.

Novena. Comisión de vigilancia y seguimiento. Memoria final.

1. Una comisión de vigilancia y seguimiento del convenio resolverá los problemas de interpretación y ejecución que pudieran plantearse y controlará su ejecución. La comisión constará de cuatro miembros nombrados a partes iguales por las personas que suscriben el convenio. La presidirá uno de los miembros nombrados en representación del Ministerio de Educación, Cultura y Deporte. Se reunirá a iniciativa de cualquiera de sus miembros cuando considere que existen problemas de interpretación o ejecución del convenio. La comisión ajustará su funcionamiento a lo establecido sobre órganos colegiados en el título II, capítulo II, de la Ley 30/1992, de 26 de noviembre, de Régimen jurídico de las administraciones públicas y del procedimiento administrativo común.

2. A la finalización del Convenio, el Ministerio de Educación, Cultura y Deporte elaborará una memoria en la que se acredite el cumplimiento de los compromisos asumidos por las partes a través del Convenio.

Décima. Interpretación y jurisdicción.

El presente convenio tiene naturaleza administrativa. Las controversias que puedan surgir sobre la interpretación, modificación, resolución y efectos que puedan derivarse del presente convenio se resolverán entre las partes, agotando todas las formas posibles de conciliación para llegar a un acuerdo amistoso extrajudicial.

En su defecto será competente para conocer de las cuestiones litigiosas los órganos jurisdiccionales del orden contencioso-administrativo.

Y en prueba de conformidad, se suscribe el presente Convenio, en lugar y fecha arriba indicados, por cuadruplicado ejemplar y a un solo efecto, quedando uno de ellos en poder de cada una de las partes.

El Ministro de Educación, Cultura y Deporte, Jose Ignacio Wert Ortega.–El Conselleiro de Cultura, Educación y Ordenación Universitaria, Jesús Vázquez Abad.

ANEXO I

Pliego de prescripciones técnicas

1) Requisitos técnicos

La digitalización de los originales de la prensa histórica y de las revistas se realizará con una resolución mínima de 400 ppp. y en una escala de 256 gamas de grises, excepto

las páginas correspondientes a las revistas culturales que se digitalizarán en una escala de 16,7 millones de colores.

Se entregarán los ficheros resultantes en formato TIFF 6.0/ISO 12639:2004 y en formato JPEG/ISO/IEC 10918:1994, como formato de máxima calidad y formato comprimido y optimizado para su lectura en Internet, respectivamente. Para ambos formatos, se entregarán tres copias, una en discos Blu-Ray DL 2-4x 50GB Doble-Capa Impresión Térmica y otras dos en disco duro externo USB 2.0. Las imágenes correspondientes a cada una de las bibliotecas o editores se entregarán en discos independientes y la estructura de directorios deberá agrupar en un único directorio todas las imágenes correspondientes a un mismo título en un único formato. Los ficheros se estructurarán en un directorio formado por un primer nivel perteneciente al año, un segundo nivel perteneciente al mes y un tercer nivel perteneciente al día.

Cada título se entregará junto con un fichero XML estructurado según el esquema Metadata Encoding & Transmission Standard (METS), versión 2.0 y se ajustará al perfil METS definido por el Ministerio de Cultura para la prensa histórica y contendrá los siguientes etiquetas: título, año, mes, día, número, número de página, incidencias detectadas en la digitalización, ficheros de imágenes TIFF 6.0 y ficheros de imágenes JPEG, y la descripción bibliográfica detallada de cada título junto con los datos completos de fondos y localizaciones en formato MARC21, actualización 17, incluyendo lo prescrito en las RDA tal y como lo codifica dicha versión del formato MARC21, así como de los aspectos que hagan referencia al formato, al algoritmo de compresión, densidad de grabación, modo de acceso, etc.

Los ficheros: TIFF 6.0, JPEG y XML deberán validarse automáticamente a través de un software de validación de objetos digitales como JHOVE (JSTOR/Harvard Object Validation Environment) o similar. Los METS incluirán la información relativa al proceso de validación mediante el elemento <digiprovMD>, dentro de la sección de metadatos administrativos (amdSec), indicando esta información como una entidad Event, de acuerdo con lo especificado en el PREMIS Data Dictionary 2.1.

El adjudicatario deberá describir las publicaciones periódicas, tanto la prensa histórica como las revistas culturales, de acuerdo con MARC21/RDA v. 17 así como los datos completos de sus fondos y localizaciones según ISO 10324:1997/Z39.71:2006 y generar los correspondientes ficheros MARC21/ISO 2709 a los que asociará las correspondientes imágenes digitalizadas.

En el caso de que alguna de las cabeceras se encuentre parcialmente digitalizada en la Biblioteca Virtual de Prensa Histórica, la Dirección General de Bellas Artes y Bienes Culturales y de Archivos y Bibliotecas entregará al adjudicatario las descripciones de los registros bibliográficos y sus fondos. El adjudicatario deberá completar esas descripciones incluyendo los campos y subcampos necesarios que completen los registros de fondos ya existentes y modificando, si fuera necesario, el registro bibliográfico.

El adjudicatario deberá describir, asimismo, las revistas culturales de acuerdo con MARC21/RDA v. 17 así como los datos completos de sus fondos y localizaciones según ISO 10324:1997/Z39.71:2006 y generar los correspondientes ficheros MARC21/ISO 2709 a los que asociará las correspondientes imágenes digitalizadas. La Dirección General de Bellas Artes y Bienes Culturales y de Archivos y Bibliotecas proporcionará al adjudicatario los registros de analíticas correspondientes a las revistas culturales que se digitalicen con cargo a este contrato a partir de la base de datos REBECA <http://rebeca.mcu.es>. Estos registros de analíticas se vincularán al registro bibliográfico correspondiente mediante el campo 773 subcampo \$g.

Las materias de los registros de analíticas se vincularán automáticamente con el correspondiente término codificado en SKOS que se encuentra en la base de datos <https://id.sgcb.mcu.es>

En la sección Descriptive Metadata (dmdSec): se incluirá un elemento External Descriptive Metadata (mdRef) con MDTYPE="MARC" cuyo valor debe coincidir con el número de control del registro bibliográfico del documento y quedará embebido en el registro XML MARC21.

Las descripciones bibliográficas de las cabeceras y de los artículos y su vinculación con las imágenes correspondientes deberán integrarse en la aplicación de gestión bibliográfica de la Subdirección General de Tecnologías y Sistemas de Información. Las imágenes se vincularán a través de los METS con los registros MARC21 de fondos y localizaciones correspondientes. Para ello, en los registros METS deberán especificarse al menos los siguientes elementos:

– En la sección File Section (fileSec): incluir un único elemento de agrupación fileGrp con un elemento file por cada imagen digitalizada de ese título. En el elemento file se indicarán los atributos correspondientes al formato (MIMETYPE), su tamaño (SIZE) y la fecha de creación (CREATED). Además, en el elemento Flocat de cada elemento file se especificará el atributo LOCTYPE="URL" y su valor debe indicar, siguiendo el protocolo file, el camino (estructura de directorios) y el nombre del fichero imagen.

– En la sección Structural Map (structMap): con TYPE="logical", en la que se incluya la información de paginación, es decir, el número de página asociado a cada imagen.

También se incluirán los metadatos necesarios que describan los términos que se precisan para la preservación de la copia digital. Estos metadatos se obtendrán del Diccionario de Datos de PREMIS, versión 2.1, denominado PREMIS Data Dictionary, y elaborado según el Modelo de Referencia del Open Archival Information System (OAIS), ISO14721.

La oferta describirá con el mayor grado de precisión los metadatos que se encapsularán dentro del fichero METS XML.

Asimismo, el adjudicatario entregará un fichero XML estructurado según el esquema METSRights en el que se detallarán las características de derecho de uso de las publicaciones según determine el Ministerio de Cultura a través de la Subdirección General de Coordinación Bibliotecaria. Este fichero se integrará en el esquema general METS que se menciona en el párrafo anterior.

De igual modo, el adjudicatario entregará un fichero XML estructurado según el esquema MADS en el que se detallará el título uniforme de las publicaciones periódicas así como todas las variantes y transformaciones que ese título haya podido experimentar y que pueda desprenderse de la catalogación completa de la totalidad del ejemplar disponible. Con estos ficheros XML MADS se constituirá el fichero de autores y títulos uniformes y variantes de título.

En ambos casos, tanto en lo que se refiere al formato METSRights como MADS se utilizará la última versión publicada por la Biblioteca del Congreso de los Estados Unidos dentro del conjunto de normativa definido por la IFLA y la CDNL y conocido por sus siglas en inglés ICADS, así como de las recomendaciones del informe final del W3C Library Linked Data Incubator Group.

Las imágenes deberán incluir por el procedimiento de marcas de agua la siguiente leyenda, escrita al pie de cada página: Ministerio de Educación, Cultura y Deporte.

En el caso de los artículos de las revistas culturales las ofertas deben especificar el proceso de creación de los registros de autoridad (tanto de nombres personales y de nombres de entidades como de títulos y materias) que deben generarse a partir de los registros bibliográficos de acuerdo con el MARC format for authority data, act. 17., publicado por la Biblioteca del Congreso de los Estados Unidos. El mismo proceso se llevará a cabo con los registros de fondos y localizaciones.

En lo que se refiere a las materias, se utilizará el fichero de la Lista de Encabezamientos de Materia para Bibliotecas Públicas, previamente convertido a SKOS (Simple Knowledge Organization System), que la Subdirección General de Coordinación Bibliotecaria pondrá a disposición del adjudicatario, que deberá vincular a los correspondientes campos MARC21 los registros SKOS. Idéntica labor se llevará a cabo con los nombres geográficos que pudieran aparecer, tanto en el campo 260 como en el 651 o 752 para lo que se utilizarán los Geonames producidos por el Instituto Geográfico Nacional y que, asimismo, la Subdirección General de Coordinación Bibliotecaria pondrá a disposición del adjudicatario.

Por cada unidad bibliográfica se entregarán tres ficheros XML conformando un objeto digital ePub en formato texto en la versión que se haya publicado en la fecha de inicio de los trabajos.

6) Reconocimiento óptico de caracteres

Deberán realizarse las siguientes actuaciones:

1) Sobre las públicas periódicas que se digitalicen o que ya formen parte de la Biblioteca Virtual de Prensa Histórica, se realizará un proceso de reconocimiento óptico de caracteres sobre 1.000.000 de páginas. El licitador identificará el software utilizado para este proceso, incluyendo la versión del mismo.

2) Se realizarán los procesos automáticos de corrección necesarios para que el porcentaje de errores en cabeceras, titulillos, titulares, secciones, epígrafes, y otros elementos de primer y segundo nivel tienda a 0. Tanto las cabeceras como los diferentes elementos tipográficos mencionados deberán estar segmentados correctamente por el procedimiento que estime apropiado el licitante, siempre que se ajuste a la normativa ALTO (Analyzed Layout and Text Object), de forma análoga a como se lleva a cabo en proyectos similares como ChroniCling America o el proyecto de digitalización de prensa de Australia. Esta segmentación se reflejará en la composición digital de los ePub.

3) El porcentaje de errores en el resto del contenido no podrá ser superior al 20%. En cualquier caso, se detallarán en la memoria las tasas de errores estimadas a partir de una muestra que se proporcionará a los licitantes. Para ello se suministrarán las correspondientes tablas Excel de un porcentaje aleatorio de páginas que no será inferior a 5.000 páginas.

4) La corrección no supondrá modernización alguna de grafías o actualización ortográfica de las palabras. No se corregirán los errores que pudieran detectarse en el original aunque parezcan evidentes.

5) El resultado del OCR se proporcionará en objetos normalizados XML según el esquema METS/ALTO 2.0 o la versión que se encuentre disponible a la fecha de la firma del contrato ((Metadata Encoding Transmission Standard / Analyzed Layout and Text Object) mantenido por la Network Development and MARC Standards Office de la Biblioteca del Congreso de los Estados Unidos incluyendo dentro de la estructura METS las direcciones de las páginas en formato PDF de imagen con el texto oculto. Los ficheros PDF estarán provistos a su vez de los marcadores necesarios para permitir el acceso por los diferentes niveles de contenido (tomos, números, secciones, y epígrafes). Los ficheros XML deberán incluir una estructura jerárquica con los niveles de contenido apropiados (tomos, números, páginas, columnas, secciones, epígrafes e ilustraciones) e incluirán además por cada división de página un vínculo con el fichero imagen correspondiente.

6) Es imprescindible la utilización del perfil METS para prensa histórica elaborado por el Ministerio de Cultura. En cualquier caso, el licitante presentará en la memoria tres ejemplos totalmente desarrollados de la aplicación del referido esquema conteniendo tanto la descripción bibliográfica detallada como el etiquetado de las páginas PDF y de las páginas OCR/ALTO.

7) El adjudicatario proporcionará toda la información siguiendo el esquema METS 2.0 con el perfil citado en el que se incluirán los metadatos descriptivos, la relación de ficheros, su estructura lógica y los metadatos de preservación conforme a PREMIS 2.1, referidos a los datos en OCR/ALTO. Estos esquemas estarán vinculados a la descripción bibliográfica de cada título y estructurados en un directorio por cada unidad básica de numeración. El nombre de los ficheros será el número almacenado en el campo 001 del registro MARC que describe la publicación.

8) Por cada ejemplar se proporcionará un fichero METS organizado según los directorios y subdirectorios que establezca la Subdirección General de Coordinación Bibliotecaria. La identificación de los metadatos descriptivos se realizará a través de los números de control de los registros bibliográficos y de los de fondos y localizaciones

utilizados en la Biblioteca Virtual de Prensa Histórica que se almacenarán obligatoriamente en el campo 035 del formato MARC21.

9) Los ficheros XML deben validarse automáticamente a través de un software de validación de objetos digitales como JHOVE (JSTOR/Harvard Object Validation Environment) o similar. Los METS incluirán la información relativa al proceso de validación mediante un elemento <digiProvMD>, dentro de la sección de metadatos administrativos (amdSec), indicando esta información como una entidad Event, de acuerdo con lo especificado en el referido PREMIS Data Dictionary 2.1.