



## PLIEGO DE PRESCRIPCIONES TÉCNICAS DEL SERVICIO DE DIGITALIZACIÓN DE FONDO ANTIGUO DE LA BIBLIOTECA UNIVERSITARIA: CREACIÓN Y TRANSFORMACIÓN DE RECURSOS DIGITALES Y SU DIFUSIÓN Y PRESERVACIÓN MEDIANTE REPOSITARIOS

1. La digitalización de los documentos impresos y manuscritos, así como la del resto de materiales especiales (fotografías, postales, carteles, planos, grabados, mapas, etc.) se realizará en color. La resolución de imagen mínima será de 400 ppp.

2. Los formatos de entrega serán TIFF, con opción CCITT Fax 4, como copia de seguridad para la preservación de los originales, otra copia en formato JPEG más reducida, comprimida y optimizada para su visualización en Internet, y una tercera copia en formato PDF multipágina para cada documento completo digitalizado con texto oculto. De igual forma imágenes thumbnail para miniatura en formato JPG.

3. Para cada título se realizará un control de paginación, la indexación del documento, su índice de contenidos, así como la asignación de metadatos Dublin Core Cualificado en XML con elementos ESE (Europeana Semantic Elements) versión 3.2.1 o posterior si existe. Se deberá tener en cuenta, en cualquier caso, las directrices de la iniciativa europea DRIVER. Igualmente se contemplará la integración, por la empresa adjudicataria del contrato, de los registros digitalizados en el repositorio institucional de la Universidad de Granada, con dirección <http://digibug.ugr.es>

4. Se entregará una copia de los ficheros resultantes, tanto TIFF 6.0, JPEG, como PDF multimedia con texto oculto. De los ficheros JPG se entregará una copia en formato de máxima calidad y otra en formato comprimido y optimizado para su lectura en Internet.

5. De los ficheros resultantes, tanto TIFF como JPEG Y PDF multipágina con texto oculto, se entregarán copias en disco duro externo USB 2.0. La estructura de directorios deberá agrupar en un mismo directorio todas las imágenes correspondientes a un mismo título en todos sus formatos.

6. Para el almacenamiento más detallado de las imágenes y el nombrado de directorios se seguirán las siguientes instrucciones:

Para cada título se creará una carpeta cuya denominación se corresponderá con la signatura de la obra pero solo con los elementos que siguen a la barra todos seguidos y sin guiones, ejemplos:

- Para la obra con signatura BHR/A-003-890, la carpeta se denominará **A003890**
- Para la obra con signatura BHR/A-004-908 (1), la carpeta se denominará **A004908-1**
- Para las obras cuya signatura se corresponde con la CAJA FUERTE, esta denominación se abreviará en **CJ**, manteniendo el orden en la estructura de la signatura:

La obra con signatura BHR/Caja 2-004, tendrá una carpeta denominada **CJ2004**.

La obra con signatura BHR/Caja 3-134 (1), tendrá una carpeta con denominación **CJ3134-1**.

La obra con signatura BHR/A-032-009 Caja, tendrá una carpeta denominada **A032009CJ**.



La obra con signatura BHR/A-003-009 (1) Caja, tendrá denominación de carpeta **A003009-1CJ**.

7. Cada título se entregará junto con un fichero XML estructurado según el esquema Metadata Encoding & Transmission Standard (METS) y un fichero con los metadatos Dublin Core Cualificado en XML con elementos ESE que contendrá los datos descriptivos, administrativos y técnicos de cada copia digital (tipo de material, título, autor, número de páginas, materias, incidencias detectadas en la digitalización, tipos de ficheros de imágenes TIFF, JPEG, PDF, etc., algoritmos de compresión, densidad de grabación, modo de acceso, etc).

8. Las entregas resultantes del proceso OCR serán:

- Fichero PDF con el texto por debajo para cada obra, volumen o número. Para denominar los ficheros se utilizará el nombre de la carpeta (signatura) con la extensión 'pdf'.
- Ficheros ALTO (uno por página) con el resultado del OCR y las coordenadas de cada uno de los elementos de la página (bloques de texto o imagen, líneas, palabras, etc.) marcando los elementos obligatorios HEIGHT (alto), WIDTH (ancho), VPOS (vertical position) y HPOS (horizontal position), según especifica el esquema *ALTO: Analyzed Layout and Text Object*. Estos ficheros se denominarán con el mismo nombre de las imágenes en formato TIFF con la extensión 'xml'.
- Un fichero METS/ALTO, que relacionará por cada obra los ficheros de imágenes (TIFF), de reconocimiento óptico de caracteres (ALTO) y PDF con la extensión METSRights para la declaración de derechos de propiedad intelectual y con la extensión PREMIS para los metadatos de preservación. La correspondencia de páginas entre los ficheros TIFF y ALTO se realizará en el elemento structMap (puede verse ejemplos en el anexo de esta oferta). Estos ficheros se denominarán con el número de identificación de la carpeta complementado con la especificación del número, el número de página y la extensión 'xml'.
- Un fichero MS Access con la identificación del título, número, directorio y nombre de las páginas TIFF, número de caracteres reconocidos, número de caracteres dudosos y relación entre ambos en forma de porcentaje.

9. Los ficheros: TIFF, JPEG, PDF y 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* citado.

10. En la sección *Descriptive Metadata (dmdSec)* se incluirá un elemento *External Descriptive Metadata (mdRef)* con MDTYPE-"MARC" cuyo valor debe coincidir con la signatura del documento.



11. Tanto las descripciones bibliográficas de los documentos como de las colecciones o fondos de las mismas y su vinculación con las imágenes correspondientes, deberán integrarse en el Recolector y Directorio de Recursos Digitales del Ministerio de Cultura y vía METS en la Biblioteca Virtual de Patrimonio Bibliográfico.

12. También se incluirán los metadatos necesarios referentes a los distintos tipos de materiales que describan los términos que se precisan para la preservación de la copia digital. Estos metadatos se obtendrán del Diccionario de Datos elaborado por el grupo de trabajo PREMIS de mayo de 2005, versión 1.0. 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 XML METS.

### **Estructura de la información de los recursos electrónicos que se incorporarán al repositorio**

La estructura de la información de los registros que se están incorporando al Repositorio es la siguiente:

Descripción bibliográfica: expuesta en el apartado 1) de este Anexo II

Digitalización: realización en color o escala de 256 gamas de grises según las características del documento original. La resolución de la imagen mínima será de 400 ppp.

Ficheros: los formatos de entrega serán TIFF, compresión LZW, como copia de seguridad para la preservación de los originales y otra copia en formato JPEG más reducida, comprimida y optimizada para su visualización en Internet. De estos ficheros se entregarán copias en discos duros de 500 Gb con formato ISO 9660, incluyendo títulos completos.

Metadatos: el esquema de XML que permite aglutinar los diferentes tipos de metadatos es la norma *Metadata Encoding & Transmission Standard (METS)*, que mantiene la Library of Congress. Consideramos como requisito indispensable que junto con los objetos digitales se realicen y entreguen los correspondientes ficheros METS. De esta forma el proyecto se asegura su continuidad técnica y su independencia de proveedores al estar basado en la norma para el intercambio y transmisión de información digital. Los ficheros METS incluirán, además de los apartados propios de la norma para la relación de ficheros, mapa estructural de los mismos, las siguientes extensiones: Metadatos descriptivos (con las extensiones IBERMARC ISO 2709, MARCXML Schema y Dublin Core Cualificado; Metadatos administrativos técnicos conforme a la extensión MIX (NISO Metadata for Images in XML); Metadatos administrativos. Declaración de propiedad intelectual conforme al esquema METSRights o The RightsDeclarationMD Extension Schema; Metadatos de preservación. Estos metadatos se obtendrán del Diccionario de Datos de mayo de 2005, versión 1.0, elaborado por el grupo de trabajo PREMIS y denominado *PREMIS Data Dictionary*, basado a su vez en el Modelo de Referencia de *Open Archival Information System (OAIS)*, ISO 14721; Metadatos de autoridad. Se crearán los correspondientes registros de autoridad de autores, títulos y materias según Metadata Authority Description Schema (MADS) que mantiene la Library of Congress.



**j) Equivalencia entre la estructura de la información descrita en el apartado anterior y la normativa ISO citada en la convocatoria de ayudas**

La norma ISO15836:2003 define el conjunto de elementos Dublin Core que se conoce habitualmente como “DC simple”. Consta de 15 elementos básicos para describir cualquier objeto de información. Estos se presentan habitualmente divididos en tres grupos que indican la clase o alcance de la información incluida en ellos, y que responden a las expectativas que tiene el usuario cuando se enfrenta a la información de la Red. La clasificación de elementos de estructura de información que establece la ISO 15836-2003 y su equivalencia con el Formato MARC 21 es la estructura de la información que actualmente se utiliza en la BUG para la descripción de los objetos digitales que forman parte del patrimonio bibliográfico de la Universidad de Granada.

13. Se entregará un fichero XML estructurado según el esquema METSRight en el que se detallarán las características de derecho de uso de las publicaciones según determine la Dirección General del Libro y del Patrimonio Bibliográfico y Documental a través del Servicio del Libro, Bibliotecas y Centros de Documentación. Este fichero podrá presentarse o bien de forma independiente o bien integrado en el esquema general METS que se menciona en párrafos anteriores.

14. En lo que se refiere al formato METSRight 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 ICABS.

15. 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 *Formato MARC21 para registros de autoridad* publicado 5 por la Biblioteca Nacional en 1999 o con la última versión del *MARC21 format for authority data*. El mismo proceso se llevará a cabo con los registros de fondos y localizaciones

16. Al finalizar el trabajo se entregará una Memoria descriptiva de lo que se ha realizado y sus incidencias; explicación clara de los ficheros que se entregan, Cd, y discos duros externos resultantes; así como factura en dónde se especifique no solo el coste general del trabajo sino el coste por imagen

**PRECIOS UNITARIOS MÁXIMOS**

1,18 sin IVA  
1,36 con IVA