

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	1 de 27

ANEXO TECNICO

BIBLIOTECA DEL CONGRESO

**Digitalización de libros y
Gestión Documental
de la Colección patrimonial de la
Biblioteca del Congreso**

FEBRERO 2020

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	2 de 27
SENADO DE LA REPÚBLICA		

PRESENTACIÓN

El presente documento denominado **“Digitalización de libros y gestión documental de la colección patrimonial de la Biblioteca del Congreso”**, constituye una guía para implantar un proceso que permita digitalizar el total de los libros históricos con los que cuenta la Biblioteca del Congreso de la República, posteriormente almacenar sus datos de catalogación en una base de datos, indexar el contenido completo del texto de todos y cada uno de los libros, y así brindar la disponibilidad de dicha información en una plataforma digital, de modo que se puedan realizar búsquedas profundas del contenido de los mismos.

Biblioteca del Congreso de la República, cuenta con cerca de 90 mil volúmenes, dentro de los cuales se destacan los que compusieron las bibliotecas personales de los doctores Juan Lozano y Lozano, Luis Eduardo Nieto Caballero y Eduardo Nieto Caiderón. La de Juan Lozano y Lozano es, posiblemente, una de las bibliotecas particulares más importantes de América Latina.

La Colección Patrimonial de la Biblioteca del Congreso de la República, es un activo importante y valioso, porque constituye parte del patrimonio de la nación, reposan joyas bibliográficas entre las cuales vale la pena mencionar: el Acta del Congreso y la Constitución de Cúcuta de 1821; los Censos Generales de Población de 1835 y 1843 y la Constitución Política de Estados Unidos de Colombia, publicada en 1863.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	3 de 27
SENADO DE LA REPÚBLICA		

1. NECESIDAD Y JUSTIFICACIÓN

La Biblioteca posee una colección básica de libros, muchos de ellos provenientes de la Ley de Depósito Legal, que enriquecen el acervo documental de más de 90 mil volúmenes. Como parte de sus colecciones, se dispone de las publicaciones editadas por el Congreso que constituyen parte de su memoria institucional.

Por otra parte, cuenta con tres colecciones de gran valor como son las de Juan Lozano y Lozano, Luís Eduardo Nieto Caballero y Eduardo Nieto Calderón. Las dos últimas que suman cerca de 10.000 títulos sobre diversas áreas del conocimiento.

En el caso de la colección de Juan Lozano y Lozano, aunque desde el punto de vista temático contiene un alto porcentaje de documentos no pertinentes, representa una gran riqueza cultural. Esta biblioteca se considera como una de las más importantes de América Latina y contiene cerca de 34.000 volúmenes sobre historia universal, filosofía, ciencias naturales, biografías, derecho, sociología, política, literatura, artes y viajes, etc., ediciones originales escritas en español, inglés, francés, alemán, italiano, sueco, noruego y portugués, muchas de ellas primeras ediciones

La documentación existente en la Biblioteca constituye un activo importante, valioso, no sólo desde el punto de vista económico, sino también, porque constituye parte del patrimonio de la nación. En la colección patrimonial, reposan joyas bibliográficas que ameritan una cuidadosa evaluación y la formulación de políticas de preservación y difusión, entre las cuales vale la pena mencionar: el Acta del Congreso y la Constitución de Cúcuta de 1821; los Censos Generales de Población de 1835 y 1843 y la Constitución Política de Estados Unidos de Colombia, publicada en 1863. Así mismo, documentos que por sus características o por la importancia que tienen para la historia de Colombia y la labor legislativa, constituyen, una colección que requiere tratamientos especiales, digitalizar los libros y permitir el libre acceso y consulta de la misma por medio de herramientas tecnológicas que aseguren su conservación y lograr que la Biblioteca del Congreso pueda alcanzar una categorización de biblioteca digital e implementación de nuevos servicios en un portal web.

Digitalizar la colección patrimonial nos permitirá incrementar las visitas virtuales, esta es una razón principal y la más obvia, cuando se sabe que hay una alta demanda por parte de los usuarios y mejorar los servicios para un grupo creciente de usuarios proporcionando un acceso de mayor calidad a los recursos de la institución, reducir la manipulación y el uso de materiales originales frágiles o utilizados intensivamente y crear una "copia de seguridad" para el material antiguo, ya que en el momento la consulta de estos documentos es restringida para préstamo domiciliario por su valor histórico y los riesgos de pérdida, por lo que se requiere un proceso de digitalización que tenga como prioridad la conservación y el uso de nuevas tecnologías para su correcta preservación, consulta y recuperación en caso de ser necesario.

El crecimiento acelerado de la información, la variedad de formatos en que aparece y las facilidades para tener acceso a ella, han creado mayores presiones sobre las bibliotecas y otras unidades de información, ya que los usuarios son más autónomos, demandan servicios más especializados y tienen mayores expectativas. Lo anterior, exige nuevas formas de organización y gestión de la información, aunque la Biblioteca dispone de algunas bases de datos, sus recursos documentales aún siguen centrados en el documento impreso y la información disponible se limita, en gran parte, a la que existe en sus acervos, afectando el nivel y la calidad de los servicios, el reto es modernizar los servicios partiendo de la incorporación de tecnología y digitalización de colecciones.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	4 de 27
SENADO DE LA REPÚBLICA		

La transformación digital de la Biblioteca es una necesidad, realizar la digitalización de libros de la colección patrimonial y su posterior proceso de gestión documental utilizando modernas técnicas informáticas.

2. OBJETIVO GENERAL

Digitalizar la colección de libros y recursos que hacen parte de la Colección Patrimonial del Senado de la República, que contiene la memoria histórica del legislativo y los informes de los ministerios al Congreso, además de libros de gran valor histórico.

2.1 OBJETIVOS ESPECIFICOS

Los objetivos específicos del proyecto de digitalización de la Colección Patrimonial del Senado de la República se presentan a continuación:

- Generación de la estructura de datos descriptivos (metadatos) correspondientes a los documentos a ser digitalizados. En esta primera tarea se recogerán y/o extraerán, los metadatos que caracterizan de manera única a cada documento.
- Digitalización de los documentos utilizando herramientas cuya característica es la utilización de técnicas no invasivas. Esto significa que los documentos serán tratados según su característica física, de modo que se preserven de acuerdo con su antigüedad, tipo de papel o de medio físico, desgaste de los materiales, importancia del documento, entre otros.
- Realizar el poblamiento de la Plataforma digital inteligente con los documentos convertidos en archivos digitales, garantizando que el contenido profundo de los documentos esté disponible para el sistema
- Utilizar para el modelo de metadatos una codificación en línea con los estándares MARC 21, garantizando con ello uniformidad en el proceso de formalización documental en la institución.
- Entregar una herramienta de gestión documental que aproveche de manera integrada los “meta metadatos” y la información de contenido profundo, cuya característica es permitir la navegabilidad a lo largo de todos los repositorios digitales, utilizando heurísticas y la arquitectura funcional que facilita el acceso eficiente a la información.
- El acceso a la plataforma digital estará restringido mediante el uso de usuarios, contraseñas y niveles de acceso que garanticen el acceso seguro a los objetos digitales.
- La plataforma será utilizada de acuerdo con los roles y esquemas de acceso definidos por la institución, de modo que los documentos tendrán niveles de visibilidad público y privado, dependiendo del grado de importancia o confidencialidad requerida. Se debe establecer la categoría de visibilidad en la institución, la cual determinará dicho nivel cuando alguno de los documentos tuviera carácter reservado, aspecto que compete determinar a la

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	5 de 27
SENADO DE LA REPÚBLICA		

institución.

- La plataforma debe permitir, a través de un formulario de registro que un ciudadano pueda darse de alta como usuario para acceder a la información pública, generando estadísticas de acceso y uso de la información.
- Documentar los procesos desarrollados.
- Realizar la trazabilidad sobre todos los procesos ejecutados.

3. ALCANCE

El alcance del proyecto de digitalización de la Colección Patrimonial estará definido a partir de la clasificación de los libros a digitalizar y la entrega de la herramienta en funcionamiento, así:

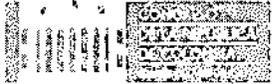
3.1. VOLUMEN DE INFORMACION A PROCESAR

El número de documentos a ser procesados en la plataforma es de hasta 2500 libros o recursos, equivalentes a 800.000 paginas aproximadamente.

3.2. PLATAFORMA DIGITAL INTELIGENTE

La plataforma digital inteligente, que representa la base del proceso de desarrollo e implantación, debe incorporar el flujo de trabajo de las siguientes funcionalidades, siendo preferente que todos los procesos y sus resultados se elaboren e integren en una única interfaz de uso. Las funcionalidades incorporadas en el flujo de trabajo de la Plataforma Digital Inteligente son las siguientes:

- Permitir la generación de taxonomías y categorías de almacenamiento
- Permitir el almacenamiento de los documentos a partir de los originales digitalizados
- Permitir la visualización del contenido de los documentos a partir de las estructuras taxonómicas definidas, es decir, de las categorías establecidas previamente.
- Permitir la búsqueda profunda de información documental, lo cual significa que el sistema encontrará todos los documentos que contengan el texto digitado por el usuario, ya sean palabras aisladas o frases, resaltándolas en el documento seleccionado. Las búsquedas serán realizadas sobre los tipos de archivos previamente descritos, tanto si poseen texto seleccionable como si fueron generados a partir de imágenes escaneadas.
- Generación automática de estadísticas e indicadores referidos a los procesos de búsqueda, a los tipos de documentos encontrados, a los términos consultados y a los temas relacionados, entre otros.
- Cumplir con los estándares establecidos por la UNESCO, con el documento "Directrices UNESCO/PERSIST sobre selección del patrimonio digital para su conservación a largo plazo", la conservación del patrimonio propio de las bibliotecas, entre los que se cuenta

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	6 de 27
SENADO DE LA REPÚBLICA		

de manera usual los documentos físicos, es tan importante como la conservación, a largo plazo, de los documentos digitales.

El primer elemento que debe cumplirse, acorde con las directrices, es la selección inteligente de los documentos de toda índole que deben conservarse. Para el caso de la Biblioteca del Congreso, este ítem ya se cumplió, habida cuenta de la preselección documental realizada con anterioridad.

Los restantes elementos considerados están ligados a la naturaleza digital de los documentos procesados, y a su conservación con base en los siguientes criterios:

Redundancia: El patrimonio digital significativo, incluyendo los ficheros maestros con metadatos asociados, debe existir en copias múltiples almacenadas en, al menos, dos distintas ubicaciones físicas. Se debe implementar un sistema de back up de incremento idealmente fuera de la institución, que sirva como respaldo tanto a las imágenes digitales del patrimonio de la biblioteca, como al código del sistema que las procesan.

Administración activa: Las instituciones culturales deben de administrar activamente su patrimonio digital para garantizar su accesibilidad e integridad a largo plazo. El patrimonio digital debe ser preservado en formatos de archivos abiertos y bien documentados, sin datos encriptados, y con una mínima compresión sin pérdida. Los documentos, tanto en su estructura descriptiva como de contenido, deben estar disponibles a partir de interfaces visuales que incorporen un buscador sobre los metadatos y sobre el contenido profundo de los documentos, según se indica a lo largo del presente documento.

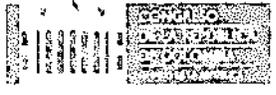
Administración de los metadatos: Generalmente los metadatos son definidos como "datos sobre datos". Los metadatos le proporcionan a la institución la información necesaria para tener acceso y conservar el patrimonio digital a futuro. Usualmente se preservan tres tipos fundamentales de metadatos asociados con el patrimonio digital:

- **Estructurales** (necesarios para tener la capacidad técnica para leer el contenido digital)
- **Descriptivos** (que contengan información contextual bibliográfica, archivística, o museográfica, la cual puede ser generada por el sistema o creada por los profesionales del patrimonio, creadores de contenidos, y/o usuarios)
- **Administrativos** (que documentan la gestión de un objeto digital que se encuentra en su colección)

Si el patrimonio digital es el "contenido", entonces los metadatos aportan el "contexto".

Identificación: Los metadatos deben permitir que cada objeto digital sea identificado de forma única y sin ambigüedades. En general, esto requerirá que se le asigne a cada objeto un identificador único a nivel global.

Ubicación: Los metadatos deben permitir que cada objeto digital pueda ser localizado y recuperado. Esta característica hace referencia a una URL única para cada objeto digital.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	7 de 27
SENADO DE LA REPÚBLICA		

Descripción: Se requiere de una descripción del objeto digital para facilitar su recuperación e interpretación. Los metadatos descriptivos caen en dos categorías: datos sobre el contenido y acerca del contexto. Estos metadatos le permiten al buscador seleccionar taxonómicamente el objeto digital, y desplegarlo en el medio preparado para ello.

Legibilidad: Hace referencia a la estructura, formato y codificación de los objetos digitales. Este elemento debe almacenarse en la estructura documental digital, de manera que el sistema de despliegue permita su fácil recuperación y presentación.

Gestión de derechos: Se debe incluir la visibilidad y acceso al documento, en donde se debe disponer de al menos tres categorías: privado, institucional y público. El carácter privado implica que el objeto digital está restringido y para su consulta se requiere de un permiso especial. El tipo denominado institucional significa que el documento podrá ser consultado al interior del Senado. Finalmente, el carácter de público implica que el objeto digital se podrá consultar fuera de la institución.

Almacenamiento de los metadatos: Los metadatos deben almacenarse de manera externa al objeto digital. Esto garantiza que se disponga de una modalidad de localización documental externa a los propios objetos digitales.

Meta metadatos: Para poder establecer la confiabilidad y autenticidad de los metadatos se requieren datos acerca de su origen y de cómo fueron compilados; ¿Cuándo fueron compilados los metadatos y quien lo hizo? ¿Los metadatos fueron cosechados automáticamente o manualmente? ¿Qué instrumentos y técnicas se utilizaron? Los metametadatos deben estar almacenados en un objeto digital descriptivo independiente de los objetos digitales almacenados.

La UNESCO establece como premisa sólida de conservación digital a largo plazo, la existencia de metadatos y “meta metadatos” que preserven la unicidad e identificación de los objetos digitales almacenados.

Como extensión de fondo, se debe garantizar que el contenido digital asociado a cada objeto digital se almacene de manera externa al objeto, lo cual significa que se debe extraer el contenido textual de cada objeto, y almacenarlo en un sistema digital para su posterior consulta. El sistema podrá, de manera exacta y eficiente, permitir la búsqueda de cada contenido interno, garantizando con ello que se cumpla de un modo más completo el cometido propuesto por la UNESCO. El término “eficiente” hace referencia a que los contenidos internos buscados deben estar disponibles en un rango que va entre uno y tres segundos, como máximo, pues de ello depende que se puedan consultar grandes volúmenes de información de manera sistemática y organizada.

El modelo que debe cumplir el sistema de información, en este punto, atiendo a los siguientes elementos:

- Extracción profunda de contenido

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	8 de 27
SENADO DE LA REPÚBLICA		

- Organización sistémica del contenido profundo
- Herramienta de búsqueda profunda, a lo largo de todos y cada uno de los documentos almacenados.

La extensión funcional indicada en este apartado, implica la incorporación a la plataforma digital de mecanismos basados en Inteligencia Artificial y tratamiento digital avanzado aplicado a los documentos físicos.

El sistema informático debe incorporar elementos de Inteligencia Artificial que permitan:

Reconocimiento de caracteres en documentos escaneados de diverso tipo, incluyendo aquellos que están basados exclusivamente en imágenes.

El sistema debe permitir la posibilidad de extraer de forma automática, tanto los metadatos requeridos por el protocolo MARC, como otros metadatos requeridos por la institución, que faciliten las búsquedas específicas de temas de su interés.

Organización estructural del contenido interno que facilite la búsqueda y consulta de términos y frases, generando resultados de contenido de manera eficiente y ágil, según se describe en párrafos previos.

Incorporación de heurísticas basadas en Inteligencia Artificial que faciliten la relación entre términos, contenidos, datos y metadatos, generando resultados de búsqueda útiles para los usuarios, en el sentido de responder con información referida a documentos que cumplen de una mejor manera con las necesidades expresadas por los usuarios en su consulta. Esta característica permite que a partir de una consulta realizada por un usuario, el sistema de información presente solo aquellos documentos que cumplan a cabalidad con lo buscado, y desplieguen solamente aquellas páginas que contienen (de manera resaltada) los términos de interés para el usuario.

La extensión de los estándares de la UNESCO, que deben ser cumplidas por el sistema informático, garantizan no solo una mayor eficiencia en el uso de la información y en la búsqueda de datos de interés, sino en un mecanismo que de manera integral sirve al propósito de extender en el tiempo la calidad del patrimonio documental en su versión de objeto digital, haciendo visible la historia almacenada en sus documentos, convirtiéndose por tanto en un activo vivo para la institución.

3.3. DIVULGACIÓN Y PUESTA A PUNTO

En cuanto a la divulgación y puesta a punto, se tiene lo siguiente:

- Configuración e instalación de las bases de datos y el software en el servidor entregado por el contratista.
- Sistema de back up de incremento idealmente fuera de la institución
- Los documentos digitalizados se almacenarán tanto en FILE SYSTEM como en tablas que guarden los textos extraídos de los documentos, de modo que se puedan realizar búsquedas sistemáticas en todos ellos.
- El modelo de las bases de datos debe ser de tipo relacional.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	9 de 27
SENADO DE LA REPÚBLICA		

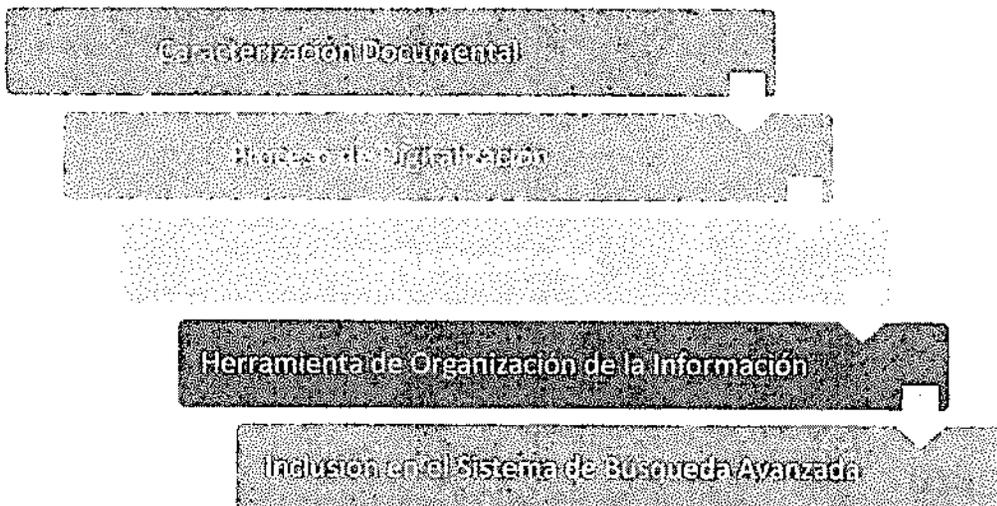
- El Front End (Interfaz Visual) debe ser responsive y multiplataforma.
- Entregar manuales de administración y operación para la plataforma requerida
- Capacitar al personal que requiera el Senado en la administración y operación de la plataforma en un mínimo de 60 horas y entregar un certificado de capacitación por el mismo tiempo.

4. MARCO TÉCNICO

El acceso a la información relevante y pertinente almacenada en la Biblioteca del Congreso, va más allá de una simple búsqueda: es el uso del conocimiento que se deriva de la rica historia del país, sintetizada en los documentos que se guardan en las instalaciones de la Biblioteca.

El proceso, visto de manera global, se descompone en los siguientes pasos:

- Caracterización de los documentos y archivos a ser procesados
- Proceso de digitalización de los documentos, libros y textos, utilizando tecnologías y herramientas que preserven la calidad de los mismos
- Definición de una taxonomía de la documentación procesada, facilitando con ello el acceso y la búsqueda de información por parte de los ciudadanos
- Poblamiento a partir de los documentos originales, de todos los sistemas informáticos de búsqueda y gestión automática de información puestos a disposición
- Utilización de modernas tecnologías que permitan la recuperación de texto de manera ágil y eficiente, con base en herramientas de Inteligencia Artificial
- Generación de funciones avanzadas de búsqueda



	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	10 de 27
SENADO DE LA REPÚBLICA		

4.1. PLATAFORMA DIGITAL INTELIGENTE

En el cuerpo del presente documento se define la Plataforma digital inteligente, como un sistema integrado que brinda soporte a todos los procesos indicados, desde la caracterización, pasando por la digitalización documental y subida de información, y la inclusión de un sistema de búsqueda profunda documental. La Plataforma referencia las mencionadas fases, y posee características integradoras que facilitan la trazabilidad de todos los procesos.

4.2. CARACTERIZACIÓN DOCUMENTAL

La caracterización documental de la información disponible en la Biblioteca, consiste en la generación previa de datos descriptivos de cada documento, conocidos también como meta datos, los cuales describen cada libro o recurso, facilitando su almacenamiento y posterior consulta.

La caracterización documental se lleva a cabo utilizando bases de datos que almacenan la información descriptiva de los documentos, libros y recursos de la Biblioteca. Esta primera fase permite realizar un inventario físico de los libros y recursos de la Biblioteca, de modo que a través un sistema de búsqueda informático se dispone de información sobre la ubicación, disponibilidad y estado de los recursos que la Biblioteca pone a disposición para su procesamiento según se describe en este documento.

La caracterización debe cumplir con los siguientes parámetros:

- Elaborar la indización para el material digitalizado con cumplimiento de las normas del Archivo General de la Nación y otros que le apliquen
- Catalogar el material digitalizado que incluya alfabetización informacional, ficha técnica del material con título, autor, ubicación, estado, resumen, palabras claves, entre otros. En general que posibilite hacer uso de distintos criterios de búsqueda.

4.3. PROCESO DE DIGITALIZACIÓN

El proceso de Digitalización consiste en transformar los documentos físicos en entidades digitales que serán almacenadas en los repositorios de la institución. Su propósito es permitir su transformación desde lo físico en un recurso de carácter digital, utilizando para ello herramientas basadas en OCR (Optical Character Recognition – Reconocimiento Óptico de Caracteres), como base para su posterior utilización en la plataforma de búsqueda.

Los procesos de reconocimiento de texto basado en tecnología OCR deben estar acompañados por un proceso de conversión de cada documento, en el cual se preserva el estado del libro sin alterar su forma física ni su contenido, a partir del uso de escáner de tecnología punta.

El sistema de digitalización contempla las siguientes fases:

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	11 de 27
SENADO DE LA REPÚBLICA		

- Verificar y catalogar las condiciones del inventario a digitalizar
- Realizar el embalaje, traslado y transporte del material a digitalizar en condiciones que garanticen su conservación fuera de las instalaciones del Senado de la Republica, si es que se va hacer afuera
- Limpieza y desinfección de todo el material entregado
- Reporte por escrito del proceso de limpieza para cada soporte y su estado
- Digitalización del material entregado

Estos procesos deben estar debidamente sustentados por software en el ciclo de vida del proceso.

4.4. TAXONOMÍA DOCUMENTAL

La Taxonomía Documental permite establecer una estructura a la cual se podrá recurrir para procesos de búsqueda, y así mismo para la inclusión sistemática de documentos a la plataforma digital.

La definición de la taxonomía documental obedece a dos principios organizativos. El primero de ellos se basa en la estructura lógica de la entidad, y por tanto define las categorías sobre las cuales se organizará de manera estructural el conjunto de documentos digitalizados. Su propósito es el de permitir disponer de un sistema de indexación simple que el buscador avanzado pueda reconocer para realizar búsquedas eficientes y rápidas. La definición de dichas categorías es flexible, y propende por la generación de un entorno en el que se describa el modo en que se encuentra dividida la entidad por secciones y departamentos, y los tipos de documentos que se almacenan en cada una de ellas. Esta primera definición facilita el proceso automático de subida de información a la plataforma, el cual se describe más adelante. Las categorías hacen alusión a los departamentos internos de la entidad o a los procesos que describen el uso y/o almacenamiento de los documentos. La plataforma permite la definición dinámica de estos elementos.

El segundo mecanismo de definición taxonómica se corresponde con la utilización del modelo estándar descriptivo para Bibliotecas, y cuya referencia es el protocolo MARC (Machine Readable Cataloging – Modelo de descripción bibliográfica), elaborado por la Biblioteca del Congreso de los Estados Unidos. Existen varias versiones de MARC, pero en los procesos de implantación de la propuesta se elige MARC 21 por su naturaleza universal. Los registros basados en este estándar permiten la descripción de cada recurso bajo el estándar ISO 2709. Con base en este modelo, la descripción de los campos asociados a cada documento se establece de manera universal, y el tratamiento de software admite aplicativos que cumplan con esta norma.

Los campos de los registros MARC 21 se encuentran descritos en la norma disponible para describir su naturaleza, pero se pueden citar los siguientes: tipo de registro (material textual, material cartográfico, grabación sonora musical, entre otros), nivel bibliográfico (colección, monografía, entre

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	12 de 27
SENADO DE LA REPÚBLICA		

otros), tipo de control archivístico, nivel de codificación, forma de catalogación descriptiva, nivel de un registro multiparte (conjunto, parte con título independiente, entre otros), y todos los demás contemplados en la norma.

Se dispone de software de catalogación para el uso de la norma MARC 21, y así mismo los registros y bases de datos se integran dentro de la plataforma de organización y búsqueda de información, lo cual permite que el sistema se encuentre integrado en todas sus facetas, desde el proceso de caracterización, digitalización, subida de información y procesos de acceso y búsqueda documental avanzada.

4.5. HERRAMIENTA AUTOMÁTICA DE ORGANIZACIÓN

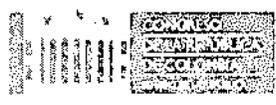
La inclusión en la Herramienta Automática de Organización de la Información está centrada en el uso sistemático de los meta datos (datos descriptivos). Ello hace referencia a los elementos básicos, como título, colección, año, categoría, entre otros, disponibles desde la norma MARC 21, lo que facilita que la subida de información a la plataforma digital responda a las necesidades propias de organización basada en estándares propios del área, como el poder responder al modelo organizativo disponible en el sistema de software de búsqueda avanzado, el cual debe poder permitir búsquedas sobre datos descriptivos y búsquedas sobre el contenido interno de los documentos. En esta segunda alternativa, la plataforma debe poder permitir que los usuarios realicen consultas referidas al contenido específico de los documentos, y la plataforma digital debe poder visualizar los documentos en la página o punto en donde se encuentre dicho contenido, resaltando o haciendo visible la palabra o frase buscada. Este ítem se refiere a lo largo del documento como búsqueda profunda documental.

En esta fase se realiza la subida de documentos a la plataforma, haciendo notar que dentro del sistema se almacenan dos tipos de datos ligados a cada documento: a) Datos descriptivos del documento, b) Contenido interno del documento.

Los datos descriptivos (meta datos) se utilizan para reconocer de manera sistemática los documentos, y desplegarlos según se lo solicite. El contenido interno de los documentos es un valor agregado que se deriva de los procesos avanzados de OCR disponibles en la plataforma de software que da soporte al sistema, a través del cual los usuarios dispondrán de la posibilidad de buscar textos o frases que formen parte del contenido interno y propio de cada documento.

4.6. SISTEMA AVANZADO DE BÚSQUEDA

La inclusión en el Sistema Avanzado de Búsqueda tiene como propósito permitir la búsqueda sistemática y profunda de información documental, facilitando que los usuarios busquen frases y palabras contenidas en los documentos, desplegando de manera sistemática, rápida y eficiente, dicha información. Esto facilitará los procesos de búsqueda para todos los usuarios.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	13 de 27
SENADO DE LA REPÚBLICA		

La plataforma base de software debe permitir que los usuarios realicen solicitudes en lenguaje natural, es decir, en sus propias palabras, o navegando entre las distintas taxonomías disponibles en el sistema, de modo que el motor interno de búsqueda encuentra los documentos que mejor se corresponden con la búsqueda solicitada, despliega la página en la cual se encuentra lo buscado, y lo resalta de modo que resulte visible para los usuarios.

Esta modalidad implica que todos los contenidos internos de los documentos estarán disponibles para búsqueda profunda de los mismos.

El tiempo de respuesta de la Plataforma digital inteligente ante una consulta de usuario que involucre contenido profundo, debe ser menor a cinco segundos bajo el supuesto de un ancho de banda de Internet considerado normal, tomando como base 10 Megabytes en tiempo actual.

El tiempo indicado para la respuesta no toma en consideración lo que se tarda en desplegar la información propiamente dicha. Esta restricción se refiere a que una vez recibida la consulta en la interfaz del buscador, el sistema de software debe ser capaz de sugerir los documentos más relevantes que contienen en su interior la cadena de texto buscada, y una vez que se elija uno de ellos, la plataforma lo debe obtener desde la ubicación en la web en donde se encuentre almacenado, para desplegarlo a continuación en la interfaz visual del aplicativo.

El tiempo total de búsqueda, en el cual se incluye la obtención de la lista de documentos y el tiempo de obtención del documento seleccionado, así como el texto a resaltar, no debe superar los cinco segundos. El tiempo adicional que tarde la plataforma en desplegar en la interfaz visual la información propiamente dicha, va a depender de la complejidad y el tamaño del documento, y por tanto no se tomará en cuenta dentro del tiempo señalado.

Se valorará positivamente cualquier disminución de tiempo que se obtenga en la Plataforma digital inteligente en relación con la obtención de la página y texto asociado en búsqueda profunda.

4.7. HERRAMIENTAS DE DIGITALIZACIÓN

Para el Proceso de Digitalización se utilizan equipos de tecnología de punta, descritos más adelante, a través de las cuales se hace el despliegue digital de los textos almacenados en el repositorio institucional, preservando los elementos físicos de manera eficiente, segura y no invasiva con los documentos.

El proceso de digitalización incorpora fases descritas previamente: caracterización de los documentos a digitalizar, determinación del tipo y herramienta más adecuada para el proceso de digitalización, registro en bases de datos, cronograma de trabajo, aplicación del modelo de digitalización (registrando el proceso realizado), generación del documento digitalizado, inclusión del archivo digital en la plataforma de gestión de software, y gestión de control del proceso realizado, en el cual se determina el cumplimiento de todas las fases descritas (almacenado también en el sistema de información)

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	14 de 27
SENADO DE LA REPÚBLICA		

Las condiciones de digitalización implican la disponibilidad de las herramientas y plataformas de software requeridas, colocadas a disposición en la Biblioteca para la realización de las tareas indicadas.

Las herramientas y dispositivos físicos de digitalización de documentos, no requieren ningún tipo de manipulación física que altere o estropee en ninguna forma los documentos digitalizados.

Para la Digitalización de los libros no se deberá hacer con descuadernación. Cada libro será preservado en su totalidad, dado que las tecnologías utilizadas de última generación no requieren alterar ni maltratar en lo más mínimo la estructura física del documento. Para ello se emplean sensores electrónicos y dispositivos laser que realizan la labor de digitalización de modo automático y controlado

Resolución óptica (300 dpi), opcional 400 ppp.

Tamaño del libro máximo: 35.5 cm x 34 cm.

Grosor máximo del libro: 15 cm.

Profundidad de color y Tipos de imágenes

36-bit, color, escala de grises, b & w

Para aquellos documentos que tienen parámetros diferentes a lo anterior, se deberá contar con los equipos requeridos y el personal altamente calificado que asegure la conservación y protección de los mismos, la especificación anterior es referencial por lo cual el proveedor deberá tener el cubrimiento técnico necesario para este tipo de eventualidades.

En la medida en que se va digitalizando cada página, se debe permitir la visualización en pantalla de la hoja y facilitar un proceso posterior de edición, mediante el cual se asegura una alta calidad en el resultado final.

Se debe disponer de soporte para digitalización de planos y grandes formatos.

Los formatos de salida de los documentos deben obedecer a los formatos más comunes del mercado, PDF, JPG, JPEG, jpeg2000, tiff, tiff G4, png, gif, bmp, pdf (capa incluido OCR), XML, DjVu.

4.8. PROCESO DE DIGITALIZACIÓN LOCAL

Para el Proceso de Digitalización se debe disponer de los equipos en las instalaciones de la institución, garantizando con ello la seguridad en el proceso y la conservación de los bienes patrimoniales de la Biblioteca del Congreso.

Si por el contrario la institución lo dispone así, la digitalización se podrá realizar en el espacio del proveedor, garantizando todas las medidas de seguridad que sean requeridas para el transporte, cuidado y conservación de los bienes patrimoniales.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	15 de 27
SENADO DE LA REPÚBLICA		

5. ESPECIFICACIONES TECNICAS

- SERVIDOR
- PROCESADOR:
- # DE NÚCLEOS DE CPU 12
- # DE HILOS 24
- RELOJ BASE 3.5GHZ
- RELOJ DE AUMENTO MÁX. HASTA 4.0GHZ
- CACHÉ L1 TOTAL 1.125MB
- CACHÉ L2 TOTAL 6MB
- CACHÉ L3 TOTAL 32MB
- DESBLOQUEADOS SÍ
- CMOS 14NM
- PACKAGE STR4
- VERSIÓN DE PCI EXPRESS PCIE 3.0
- TDP/TDP PREDETERMINADO 180W
- TEMP. MÁX. 68°C
- DOS DISCOS DUROS PCIE DE 1 TERABYTE CON LAS SIGUIENTES ESPECIFICACIONES: VELOCIDADES DE ESCRITURA DE LECTURA INFERIOR: NIVELES DE RENDIMIENTO DE LECTURA Y ESCRITURA SECUENCIALES DE HASTA 3.500 MB / SY 2.500 MB / S, RESPECTIVAMENTE
- LECTURA ALEATORIA (4KB, QD32): HASTA 500,000 IOPS DE LECTURA ALEATORIA
- TARJETA DE VIDEO: 4 GB 256 BITS GDDR5
- 1 X DL-DVI-D 1 X HDMI 3 X DISPLAYPORT 2048
- PROCESADORES DE TRANSMISIÓN; CONECTORES DE PANTALLA ESTÁNDAR: DL DVI-D/HDMI/DISPLAYPORT X3.
- PCI EXPRESS 3.0
- FACTOR DE FORMA: ATX.
- MEMORIA RAM:2 X 16 GB
- VELOCIDAD DE RELOJ DE MEMORIA2,400 MHZ
- VELOCIDAD DE MEMORIA 2,400 MHZ
- TECNOLOGÍA RAM DRAM
- Sistema Operativo: Linux
- Servidor: Apache
- Motor de base de datos: MySQL
- Lenguajes de desarrollo: PHP, Node.JS, JavaScript, JQuery, entre otros.

Los productos desarrollados deben ser servicios web cuya característica es la de permitir que se puedan integrar de manera transparente a los sistemas de información del Senado de la República. Los servicios web se basan, inicialmente, en componentes RESTfull. El objetivo del desarrollo es permitir que se genere interacción transparente con los aplicativos de la institución, lo cual permitirá que el uso de las respuestas obtenidas (usualmente JSON y XML) puedan ser procesadas y utilizadas en el modelo XHTML, garantizando con ello la transparencia de los procesos, y así mismo permitiendo que la presentación de los resultados sea independiente de los desarrollos de las herramientas tecnológicas.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	16 de 27
SENADO DE LA REPÚBLICA		

5.1. CARACTERÍSTICAS DE LA PLATAFORMA

El FileSystem requerido debe poseer 500 GigaBytes para el almacenamiento del aplicativo, los archivos de soportes y la información requerida por el gestor de búsqueda profunda.

El gestor de bases de datos debe estar disponible para el contratista, con las credenciales que garanticen el acceso seguro al sistema

Se debe disponer de un acceso FTP a través de una VPN, para acceder externamente al sistema de gestión del FileSystem, con las credenciales y permisos necesarios para el acceso al sistema

Se requiere disponer de una capacidad de un TeraByte en disco, para el almacenamiento de la información en las bases de datos.

Se requiere disponer de un sistema de bases de datos de tipo relacional, tipo MySQL, para el almacenamiento y la gestión de los datos.

El ancho de banda debe permitir una funcionalidad adecuada para el acceso eficiente a los datos e información disponible en el sistema. Valor sugerido: 100Megas.

El servidor web puesto a disposición para la gestión de la aplicación debe poseer las últimas versiones de: Apache, MySQL, PHP, Node.JS.

5.2. CARACTERÍSTICAS DE LA PAGINA WEB

- ✓ La página debe ser responsive
- ✓ Debe poseer acceso a través de usuarios, cada uno con su respectivo nombre de usuario y contraseña
- ✓ Se debe establecer un esquema de perfiles para usuario del sistema: administrador, usuario biblioteca, visitante
- ✓ La página debe poseer manejo basado en sesiones, para garantizar la seguridad en el acceso
- ✓ La página debe ser diseñada para operar en modo https
- ✓ La página debe cumplir con los estándares de la Internet, como la W3CSS, Y W3XHTML
- ✓ La página debe cumplir con los lineamientos establecidos por el Ministerio de TIC
- ✓ La página debe poseer un diseño institucional, bajo el marco formal establecido por la institución
- ✓ La página debe responder al objetivo de permitir acceder de manera simple y eficaz a la información almacenada en las bases de datos, permitiendo la búsqueda sistemática de los contenidos brindados por la Plataforma Digital Inteligente

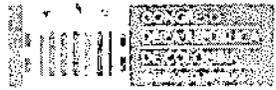
	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	17 de 27
SENADO DE LA REPÚBLICA		

- ✓ La página debe permitir el cambio de contraseña para cada usuario
- ✓ La página debe admitir hasta al menos cuatro idiomas para el despliegue de información, conteniendo al menos el Español y el Inglés
- ✓ La página debe enlazarse con el sitio web oficial de la institución
- ✓ La página debe permitir que el usuario y la contraseña sean manejados, en caso de requerirlo la institución, a través de directorios LDAP
- ✓ La página web en términos de contenido debe permitir extraer todos los documentos, mostrar la taxonomía o agrupación de los títulos e información, una característica fundamental es facilitar las búsquedas inteligentes de títulos, contenidos individuales, frases o búsquedas compuestas y facilitar descarga de información

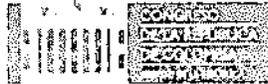
5.3. CONSIDERACIONES TÉCNICAS

En términos específicos, las consideraciones técnicas a cumplir son las siguientes:

SOFTWARE PARA LA GESTIÓN INTELIGENTE DE INFORMACIÓN DISTRIBUÍDA
Requerimientos mínimos y especificaciones:
RESTRICCIONES ESPECÍFICAS
Las taxonomías deben permitir establecer las rutas específicas de búsqueda, cada una de ellas definida a partir de una URL en concreto.
El servicio debe organizar los resultados derivados de la búsqueda, utilizando diferentes criterios: taxonomía, fecha, relevancia, entre otros.
El sistema debe brindar ayudas de autocompletar al momento de digitar la búsqueda y/o brindar alternativas que le faciliten al usuario la elección de la respuesta buscada
El sistema debe permitir la generación automática de búsquedas relacionadas con base en la taxonomía definida.
El sistema debe realizar sugerencias en el campo de búsqueda con base en las actividades realizadas previamente y como consecuencia de la estructura interna del sistema, como diccionarios de datos y componentes taxonómicos de la plataforma.
El sistema debe permitir agrupar resultados a partir de filtros que permitan acotar la información buscada
El sistema de información debe permitir la inclusión de contenido de archivos ya existentes, al menos de los siguientes formatos:
<ul style="list-style-type: none"> • Acrobat (PDF) • Documentos de Office

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	18 de 27
SENADO DE LA REPÚBLICA		

<ul style="list-style-type: none"> • RTF • HTML • XML • Dbf, dbt
El sistema no debe indizar contenido duplicado.
El sistema debe poder indizar el texto al interior de los documentos, generando objetos HTML relacionados.
El sistema debe poder realizar búsquedas af interior de las bases de datos, en todos los campos relacionados con la taxonomía definida.
El sistema deberá poder soportar actualizaciones y crecimiento en el número de documentos a indizar (sin la necesidad de cambios de hardware) hasta para al menos 20.000.000 de documentos.
El sistema deberá proveer mecanismos de autenticación única.
El sistema debe proveer un mecanismo de administración (back-end) de usuarios, artículos y taxonomías.
El sistema debe permitir la generación de reportes ligados a los siguientes aspectos: <ul style="list-style-type: none"> • Consultas por hora • Consultas por día • Ranking de las palabras y/o artículos más buscados (encontrados y no encontrados) • Frases más buscadas (encontradas y no encontradas)
El sistema presentará el documento simplificado como objeto de selección, de modo que el usuario pueda decidir si continúa o no con el proceso de despliegue de la información.
La plataforma deberá soportar mecanismos de autenticación tales como LDAP y formularios en HTTP.
El sistema debe estar en capacidad de indizar los documentos en múltiples idiomas, acorde con los contenidos profundos de los documentos.
RESULTADOS GENERADOS COMO RESPUESTA A CONSULTAS DE USUARIO
Una vez definida la estructura taxonómica del sistema de información asociada a cada dependencia, la plataforma de búsqueda elegirá las mejores respuestas ante las solicitudes de los usuarios, navegando por el árbol de conocimiento generado a través de la plataforma.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	19 de 27
SENADO DE LA REPÚBLICA		

El sistema clasificará los documentos y recursos utilizando criterios como: fecha de creación de los artículos, repositorio específico de almacenamiento y ubicación en el árbol taxonómico en la plataforma.

El sistema brindará resultados alternativos con base en el árbol taxonómico de la plataforma, o contenidos relacionados que coincidan con palabras similares, sinónimos y términos contenidos en el diccionario base de la plataforma.

El sistema posicionará en los primeros lugares de despliegue de información, los contenidos previamente definidos en la taxonomía, o los que tengan un mayor número de selecciones cuando la búsqueda no sea exacta.

El sistema ajustará la búsqueda al resultado más cercano, a pesar de que el usuario cometa errores ortográficos, con base en las taxonomías definidas y los target documentales.

El sistema permitirá la utilización de filtros utilizando lenguaje natural, de modo que se restrinja el espacio de búsqueda de información

El sistema debe brindar sugerencias de manera dinámica en el proceso de búsqueda, permitiendo autocompletar el texto a medida que introduce la consulta.

Permitir filtrar los resultados generados acorde con las taxonomías establecidas y los contenidos de los documentos almacenados en los repositorios de datos.

VISUALIZACIÓN DE CONTENIDOS OBTENIDOS A PARTIR DE LAS BÚSQUEDAS

El sistema desplegará como resultado de la búsqueda, un objeto HTML como hiper enlace al recurso solicitado. Al seleccionar el objeto HTML, se visualizará un documento estándar HTML sobre la interfaz del navegador. El documento tendrá asociado un enlace a la categoría taxonómica correspondiente, de modo que se pueda navegar a través del árbol de búsqueda asociado. Se resaltarán las palabras buscadas, facilitando así mismo la navegación entre los diferentes contenidos relacionados.

CARACTERÍSTICAS TÉCNICAS DE LA PLATAFORMA

El sistema debe poseer un mecanismo que permita convertir documentos PDF o imágenes a objetos HTML, de manera directa y automática, sin intervención manual. Bastará con introducir la dirección URL del documento, para que se obtenga de manera automática el objeto HTML a ser desplegado en la pantalla de resultados.

El sistema debe proveer un mecanismo de carga masiva de información, de modo que se pueda cargar de manera automática un paquete de documentos a partir de disponer de ellos en un almacenamiento estándar en FILE SYSTEM. Las categorías y taxonomías disponibles en el software se deben corresponder con la estructura del FILE SYSTEM, la cual contendrá los documentos a subir a la plataforma.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	20 de 27
SENADO DE LA REPÚBLICA		

El sistema debe ser escalable en cuanto al modelo de búsquedas a partir de un esquema basado en configuración por software, bajo el supuesto de que la plataforma hardware esté configurada en cuanto al gestor de base de datos y capacidad de los repositorios de datos.

El sistema debe permitir la administración de usuarios, estableciendo sus perfiles y niveles de seguridad, en una base de datos corporativa o bajo el protocolo que la Institución defina para tal propósito.

El sistema debe permitir tres niveles de acceso a los documentos: 1) Documentos privados, cuya consulta requiera de un permiso especial, 2) Documentos públicos, cuya consulta se pueda extender a usuarios externos, si la institución lo ve factible y, 3) Documentos locales, los cuales pueden ser consultados por todos los usuarios de la institución. Para los tres casos, el sistema de búsqueda desplegará únicamente la información que se corresponda con estos criterios. En ningún caso se desplegará información a la que no se tenga derecho. La plataforma contará con un sistema de estadísticas de búsquedas, tanto de manera grupal como individual.

El administrador del sistema podrá monitorear la plataforma a través de la interfaz web.

El sistema debe permitir acceso remoto mediante el protocolo SSH, de modo que desde consola (o herramientas asociadas) se puede brindar soporte técnico por parte del fabricante del sistema.

El sistema debe ser compatible con las plataformas más utilizadas, entre las que se incluye como mínimo: sistemas operativos Windows y Linux, AIX, Solaris, MAC, entre otros.

SOLUCIÓN REQUERIDA EN LA INSTITUCIÓN

El software debe implantarse sobre un sistema que brinde y soporte el hardware, software y procesos requeridos para su correcto funcionamiento y escalabilidad, lo cual incluye (pero no se limita) a:

- Servidor (es)
- Sistemas operativos
- Bases de datos
- Portales de administración
- Actualizaciones liberadas por el fabricante
- Instalación
- Configuración
- Operación

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	21 de 27
SENADO DE LA REPÚBLICA		

Todos los elementos mencionados deben disponer de las licencias permanentes y el soporte necesario durante un año (12 meses), para garantizar la continuidad del servicio. Incluir el valor del soporte y actualizaciones, para los siguientes años.

La garantía de funcionamiento de las herramientas tecnológicas involucradas en la prestación del servicio, debe incluir cambios en el hardware en caso de daños o actualización de versiones liberadas por el fabricante sin costo adicional para la entidad.

La vigencia de la garantía de los productos que lo requieran, debe ser de al menos un año (12 meses).

El servicio incluye la entrega en el sitio designado por la Institución, y así mismo la provisión de los equipos necesarios para la prestación del servicio.

Las actualizaciones de la plataforma, cuando estas se presenten, no tendrán ningún costo extra para la entidad en el marco del tiempo de vigencia del contrato

La única conexión a los repositorios de información será a través de puertos de red Ethernet, facilitando con ello la integración con la infraestructura de la entidad

La solución informática debe poder escalar a la capacidad para tolerancia a fallos y la continuidad de la operación en caso de requerirse.

Las labores de mantenimiento durante el periodo de licenciamiento son las siguientes:

- Mantenimiento correctivo
- Mantenimiento preventivo (por lo menos una vez al año).
- Actualizaciones de software.
- Soporte técnico sobre la solución informática por un año en sitio (cuando sea requerido) y telefónico, con una disponibilidad 5x8 en el mismo día, o como máximo al día hábil siguiente

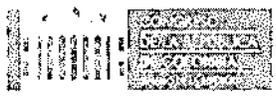
ENTREGABLES ADICIONALES

Plan detallado de implantación y cronograma de actividades.

Evidencia de la incorporación de las fuentes de información disponibles en la Institución a los repositorios institucionales y su inclusión en la taxonomía de la institución.

Reportes de resultados, al finalizar cada mes de actividades.

Entregar el documento en el que se especifiquen las licencias de los productos entregados, el servicio de alojamiento en plataforma CLOUD y los derechos del servicio de soporte por un (1) año, mantenimiento y actualizaciones por (3) años nombre de la Institución, para todos los componentes de hardware y software.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	22 de 27
SENADO DE LA REPÚBLICA		

TIEMPO DE ENTREGA: 2 meses

La Plataforma digital inteligente integradora de todos los procesos debe ser instalada y configurada en los servidores que la Institución determine, en el momento de la firma del Contrato entre las partes.

Los reportes mensuales de los procesos desarrollados en la plataforma se generarán automáticamente en el sistema de información una vez finalizado el mes, y estarán disponibles para su consulta en línea.

El plan detallado de trabajo debe entregarse en un plazo no mayor a cinco (5) días calendario a partir del perfeccionamiento del contrato.

6. LICENCIAMIENTO Y SOPORTE FORMAL

Se brinda licencia de uso del producto sin restricción de tiempo, lo cual implica que la licencia no caducará en ningún sentido en el marco especificado en este documento. Por tanto, la Biblioteca dispondrá de todas las funcionalidades del producto debidamente licenciadas por el proveedor, sin que caduque la licencia en ningún momento del tiempo.

Los datos utilizados por el Contratista serán únicamente aquellos que estén disponibles en el repositorio de datos, o los que formen parte de los portales institucionales o de acceso público. Por lo ya expresado, es responsabilidad exclusiva de la institución el garantizar la calidad, pertinencia y validez de la información desplegada.

Para validar que la información almacenada en el sistema sea fidedigna, se dispone de un sistema de usuarios y roles que garantizan que el acceso se realice a través del modelo de seguridad que el Senado de la República determine con antelación de manera concertada con el contratista.

El soporte y actualizaciones al sistema, estarán cubiertos inicialmente por un año; para los años posteriores se debe presentar el costo de estos servicios.

Por último, se debe indicar que la confidencialidad de los datos se debe manejar directamente desde la Institución, pues el contratista no realiza ningún tipo de manipulación sobre ellos.

6.1. PROCESO DE CAPACITACIÓN

Se brindará capacitación a todos los funcionarios de la institución en el manejo de la Plataforma digital inteligente, de software, cubriendo todas las dependencias internas de la Biblioteca. Se complementará la capacitación con un manual de usuario y un vídeo que ilustre el manejo del sistema.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	23 de 27
SENADO DE LA REPÚBLICA		

La capacitación toma en cuenta el perfil de los usuarios, adecuándolos al rol a ejecutar: administrador, usuario final.

6.2. MANTENIMIENTO Y SOPORTE

El soporte y mantenimiento de la aplicación, una vez realizada la instalación, configuración y poblamiento del sistema sobre el cual opera la plataforma será realizada por el Contratista. El soporte se brinda por un período durante todo el período de contratación, y su periodicidad es permanente, para lo cual la empresa dispondrá de un funcionario con la disponibilidad de tiempo pertinente para dichas actividades.

6.3. CARACTERÍSTICAS TÉCNICAS

En atención al proceso de implantación y seguimiento, el proveedor del producto se compromete a:

- Disponer de las licencias y soporte que garantice el correcto funcionamiento de la plataforma durante el primer año, sin costo adicional
- La continuidad del servicio incluye actualización de la plataforma en caso de cambios tecnológicos que así lo requieran, en el caso de ajuste de las herramientas que constituyen el sistema
- El proceso de implantación se hará directamente sobre las plataformas de la institución
- El proveedor administrará de manera remota o presencial la plataforma durante el período de garantía de la misma, la cual es de un año contado a partir de la instalación del aplicativo
- Se brindarán reportes mensuales del proceso de implantación
- Se brindará un reporte final del proceso de implantación

7. PERSONAL REQUERIDO

Para una adecuada ejecución del contrato se requiere un equipo de trabajo que cumpla como mínimo con los siguientes perfiles:

SERVICIOS PROFESIONALES BRINDADOS POR EL PROVEEDOR
El proveedor del sistema deberá brindar apoyo a la Institución para que la solución informática sea compatible con la plataforma de base tecnológica de la entidad
El proveedor brindará asesoría en cuanto a las dudas de carácter técnico que se presenten durante el proceso de ejecución e implantación del sistema.

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	24 de 27
SENADO DE LA REPÚBLICA		

Durante la ejecución del presente contrato, el proveedor administrará de manera remota o presencial, de ser necesario, la plataforma de búsqueda realizando cambios en la configuración, reportes del sistema, mantenimiento de software, y tareas correlacionadas con el normal funcionamiento y operatividad de la plataforma.

Cargo	Nro.	Responsabilidad	Dedicación	Formación Profesional	Experiencia profesional
Administrador del proyecto	1	Liderar el proceso de implantación completa del proyecto	Tiempo completo	Ingeniero afín con el área de Sistemas y Computación, preferiblemente con Maestría en el área de gestión de proyectos	Experiencia certificada de al menos 3 años durante los cuales participó en dirección y/o implantación, de proyectos de sistemas de software
Ingeniero Líder de implantación del Proyecto	1	Dirección del proceso de implantación de la Plataforma digital inteligente	Tiempo completo	Ingeniero afín con el área de los sistema y la computación, y preferiblemente con especialización o maestría en arquitectura de software, o desarrollo de software	Experiencia certificada de al menos 2 años durante los cuales participo en la implantación de productos de software
Ingeniero de soporte	1	Soporte de la plataforma integrada	Tiempo completo	Ingeniero afín con las áreas de sistemas y computación	Experiencia certificada de al menos 2 años durante los cuales participó en soporte de software
Programador	2	Desarrollo de módulos	Tiempo completo	Ingeniero afín con las áreas de	Experiencia certificada de 2

	ANEXO TÉCNICO		Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO		25 de 27
SENADO DE LA REPÚBLICA			

		requeridos por los usuarios y la plataforma		sistemas y computación, preferiblemente con formación certificada en desarrollo web	años durante los cuales participó en desarrollo web, bases de datos y manejo de frameworks
Ingeniero Web Master	1	Desarrollo sitio web y control de servidores	Tiempo completo	Ingeniero afín a las áreas de sistemas y computación, preferiblemente con formación en redes y formación en desarrollo web	Experiencia certificada de 2 años durante los cuales participó en administración de redes y construcción de portales web
Experto en sistemas de digitalización	1	Programación Proceso de digitalización	Tiempo completo	Ingeniero o Tecnólogo afín con el área de la electrónica o los sistemas de cómputo, preferiblemente con especialización y/o formación en sistemas digitales	Experiencia certificada de 2 años durante los cuales participó en actividades en el campo de los sistemas digitales, la electrónica o los sistemas aplicados al área del hardware,
Operario del sistema software	2	Tareas manuales o semiautomáticas de subida de información	Tiempo completo	Tecnólogo o Técnico afín al área de los sistemas y la computación	Experiencia certificada de 2 años durante los cuales participó en actividades en el uso técnico de sistemas de software

	ANEXO TÉCNICO		Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO		26 de 27
SENADO DE LA REPÚBLICA			

Operarios del sistema de digitalización	2	Proceso manual de digitalización	Tiempo completo	Técnico afín al área de sistemas y/o instrumentación y/o electrónica	Experiencia certificada de 2 años durante los cuales participó en las áreas de sistemas digitales y/o electrónica y/o, sistemas y computación
Tecnólogo (a) en Bibliotecología	1	Clasificar, indizar, dirigir los procesos de registro y supervisar la digitalización, registrar y realizar seguimiento a los procesos de integración con los sistemas de información	Tiempo completo	Tecnólogo (a) Bibliotecología	Experiencia certificada de al menos dos años en el área de la Bibliotecología
Restaurador	2	Caracterizar el estado de los documentos a digitalizar, y determinar los procesos y procedimientos a seguir para velar por la conservación del patrimonio de la Biblioteca del Congreso	Tiempo completo	Título profesional en conservación y restauración del patrimonio cultural	Experiencia certificada de al menos tres años en el proceso de restauración de bienes patrimoniales

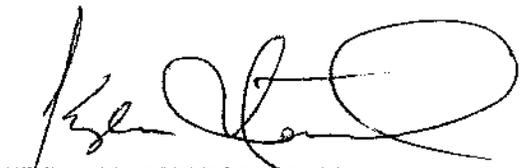
8. VALOR DE LA PROPUESTA

Se debe incluir el valor de la propuesta, acorde con el siguiente modelo:

	ANEXO TÉCNICO	Página
	DIGITALIZACIÓN DE LIBROS Y GESTIÓN DOCUMENTAL DE LA COLECCIÓN PATRIMONIAL DE LA BIBLIOTECA DEL CONGRESO	27 de 27
SENADO DE LA REPÚBLICA		

	Cantidad	Unidad de medida	Valor unitario sin IVA	IVA	Valor unitario con IVA	SUBTOTAL VALOR CON IVA
DIGITALIZACION	800.000	Paginas				
DESARROLLO DE PLATAFORMA DIGITAL INTELIGENTE	1	Unidad				
SERVICIOS, PUESTA EN MARCHA	1	Unidad				
SOPORTE 1 AÑO	1	Unidad				
MANTENIMIENTO Y ACTUALIZACIONES 3 AÑOS	1	Unidad				
PERSONAL PROFESIONAL DE APOYO	14	Unidad				
VALOR DE LA PROPUESTA						

(*) Debe presentarse adicionalmente, el costo de soporte y mantenimiento a partir del segundo año.


 KEYLA MENESES TORREGLOSA
 Biblioteca del Congreso


 PABLO ALZATE PEREZ
 Asesor II División Planeación y Sistemas