

1. INTRODUCCIÓN

PARES (Portal de Archivos Españoles) es una plataforma informática avanzada y abierta que integra la gestión interna de los archivos estatales con la difusión en Internet de los fondos y colecciones documentales.

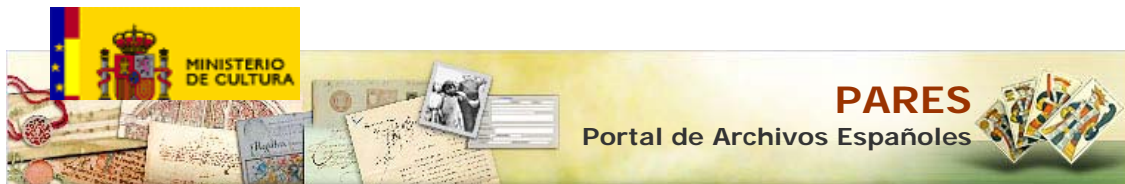
Tecnológicamente es un proyecto de colaboración entre los servicios archivísticos e informáticos del Ministerio de Cultura, que propicia el trabajo en red y la difusión centralizada, facilitando el acceso a los usuarios, investigadores y ciudadanos a un alto volumen de datos e imágenes digitalizadas en constante aumento.

PARES es un proyecto abierto que quiere facilitar la comunicación y la cooperación archivística, no solo en el nivel interno de los archivos estatales, sino también con instituciones públicas o privadas de ámbito nacional e internacional.

El PORTAL está desarrollado en su primera fase, que se corresponde con los módulos de búsqueda y descripción de los datos.

El Portal de Archivos Españoles es una herramienta destinada a la difusión de los archivos y a facilitar el trabajo interno de gestión y descripción. Proporciona a los investigadores y usuarios en general una serie de funcionalidades que facilitan la localización de los documentos:

- **BUSCADOR**, con dos modalidades, tanto de búsqueda sencilla, como de búsqueda avanzada, con diferentes filtros.
- **INVENTARIO DINÁMICO DE CONTENIDOS**, que ofrece al usuario información detallada e instantánea sobre qué fondos y colecciones de cada archivo están difundidas en Internet a través del Portal. Además, se ofrece información sobre las novedades que han tenido lugar en los últimos treinta días.
- **AGENDA DEL INVESTIGADOR**, que puede exportarse vía correo electrónico a la dirección que desee el usuario.
- **ENVÍO TELEMÁTICO DE DOCUMENTOS**, que permite al usuario la exportación de imágenes digitalizadas de un documento cualquiera vía correo electrónico.



Además, como Portal Web sirve de plataforma para la difusión de proyectos archivísticos de interés que son integrados en la base de datos general, facilitando su consulta a los usuarios. Los primeros proyectos de esta naturaleza son:

- **CATASTRO DEL MARQUÉS DE LA ENSENADA**, mediante un buscador en base de datos que localiza los documentos originales digitalizados para su consulta.
- **CARTELES DE LA GUERRA CIVIL**, una de las principales colecciones de carteles de esta etapa bélica producidos en ambos bandos contendientes.
- **DEPORTADOS ESPAÑOLES A CAMPOS DE CONCENTRACIÓN NAZIS**. Aporta información sobre más de 8.000 personas que estuvieron presas o murieron en los campos de Mauthausen o Gusen, fundamentalmente.

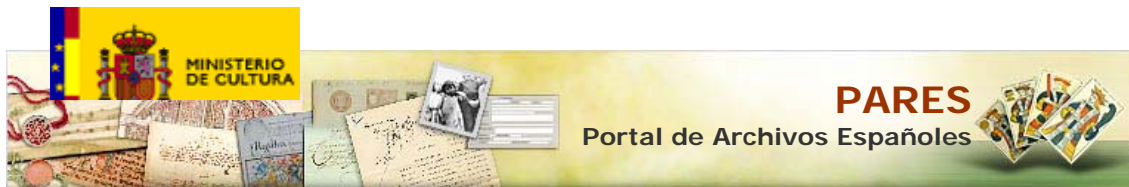
Como herramienta de trabajo archivístico, PARES proporciona un módulo de descripción que permite la elaboración e introducción de la información documental e imágenes digitalizadas dentro del sistema. La plataforma es nutrida diariamente con el trabajo de descripción en red de los profesionales de los archivos, así como con la carga de sus correspondientes bancos de imágenes digitalizadas, conforme se van materializando los distintos proyectos en curso.

En una segunda fase, fase en la que nos encontramos actualmente, se perfeccionarán los módulos de la primera, y se desarrollarán distintas herramientas necesarias para la gestión, tanto de los archivos como de los usuarios, el control y el movimiento de fondos o la gestión del servicio de reprografía.

Los usuarios que acceden a PARES tienen la opción de comunicarse constantemente con la administración del sistema para resolver cualquier necesidad de información que se plantee.

PARES arrancó proporcionando a sus usuarios cerca de 13 millones de imágenes digitalizadas, que están asociadas a 1,7 millones de registros descriptivos de los documentos. En un futuro próximo se quieren llegar a incorporar a PARES hasta 30 millones de imágenes digitalizadas y 1 millón más de fichas descriptivas.

PARES es una herramienta de gestión e información electrónica de archivos que se ve favorecida por las ventajas en acceso en red, que en conjunción con el volumen creciente de datos, propicia y abunda



en la superación de las barreras tradicionales del acceso a los archivos:

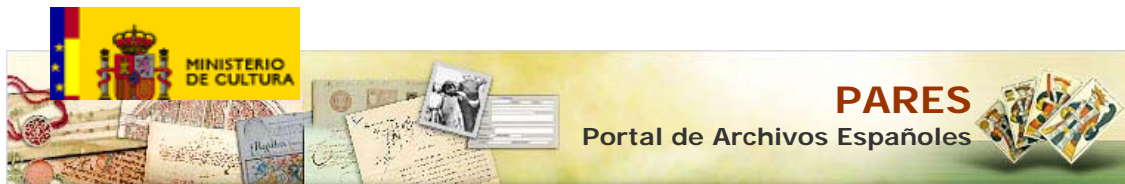
- TEMPORALES. PARES permite el acceso libre y gratuito desde cualquier lugar del mundo las veinticuatro horas del día. El usuario ya no tiene que estar supeditado a los horarios de los centros.
- GEOGRÁFICAS. El usuario no tiene que desplazarse de su lugar de residencia habitual, facilitando el ahorro de costes de viajes.
- DISPERSIÓN DE LA INFORMACIÓN. PARES permite al usuario obtener, en una sola operación, información tradicionalmente repartida por distintos archivos.

Frente a los instrumentos de descripción tradicionales, PARES actúa como un gran inventario, en el que cada orden de búsqueda por un criterio subjetivo supone la obtención de un listado de resultados equivalente a los tradicionales catálogos de una manera más ágil, rápida y efectiva.

1.1. ANTECEDENTES

Los hitos en la informatización de los archivos estatales podría resumirse de la siguiente manera:

- Informatización del Archivo General de Indias, en 1992. Informatización que no será comentada por su poca relación con el sistema de información PARES.
- AER, en 2000.
 - o AER, Archivos Españoles en Red, es el Proyecto del Ministerio de Cultura, pionero en el mundo, para difundir los contenidos culturales de los Archivos Españoles.
 - o Responde a una ciberestrategia de gran alcance para conectar los Archivos con las Tecnologías de la Información y Comunicación.
 - o Desarrolla un servicio público, accesible a todos, promoviendo políticas solidarias de promoción de conservación y difusión del Patrimonio Documental, es decir, de la Memoria Histórica.



- Se adecua a los principios y acciones de la UNESCO sobre la Sociedad de la Información.

- PARES, en 2007.

2. DEFINICIÓN ARQUITECTÓNICA DEL PROYECTO PARES

Inicialmente y, a grandes rasgos, se puede definir la arquitectura física del sistema de la siguiente forma: un entorno servidor que da soporte a una aplicación web (PARES) y que da servicio a diferentes grupos de usuarios.

La arquitectura técnica del sistema PARES, en su primera fase, está basada en las siguientes tecnologías:

- Tecnología J2EE: J2EE es un conjunto de estándares y especificaciones para el desarrollo de aplicaciones empresariales basado en la tecnología Java.
- Servidor de aplicaciones JBoss 4.0.2: Este servidor de aplicaciones emplea la tecnología J2EE, que desarrolla aplicaciones basadas en lenguaje Java, proporcionando rasgos básicos tales como escalabilidad, seguridad, rendimiento, mantenibilidad y flexibilidad. La arquitectura PARES se fundamenta en este servidor de aplicaciones, para poder implementar el concepto Model View Controller, de tres capas lógicas.
- Base de datos Oracle: El gestor de base de datos Oracle es una base de datos relacional segura y fiable, con soporte nativo de XML para la integración de aplicaciones. Este gestor proporciona módulos de gestión documental, lo que facilita el tratamiento de los documentos archivados en el sistema.
- Oracle Text: Oracle Text usa SQL standard para indexar, buscar y analizar texto y documentos almacenados en la Base de datos ORACLE, en ficheros o en la web. Oracle Text permite realizar análisis lingüísticos de documentos, así como búsquedas de texto usando múltiples estrategias (búsqueda por palabra clave, por contexto, operaciones booleanas, búsquedas por patrón, búsquedas HTML/XML, etc.).
- XML/XSL: Tecnología de publicación de páginas basada en la generación de páginas HTML a partir de los datos almacenados

en formato XML, con la presentación de las páginas recogida en plantillas XSL. Esto permite una lógica separación entre la capa de presentación y los datos.

A continuación se muestra un diagrama de la arquitectura técnica del sistema:

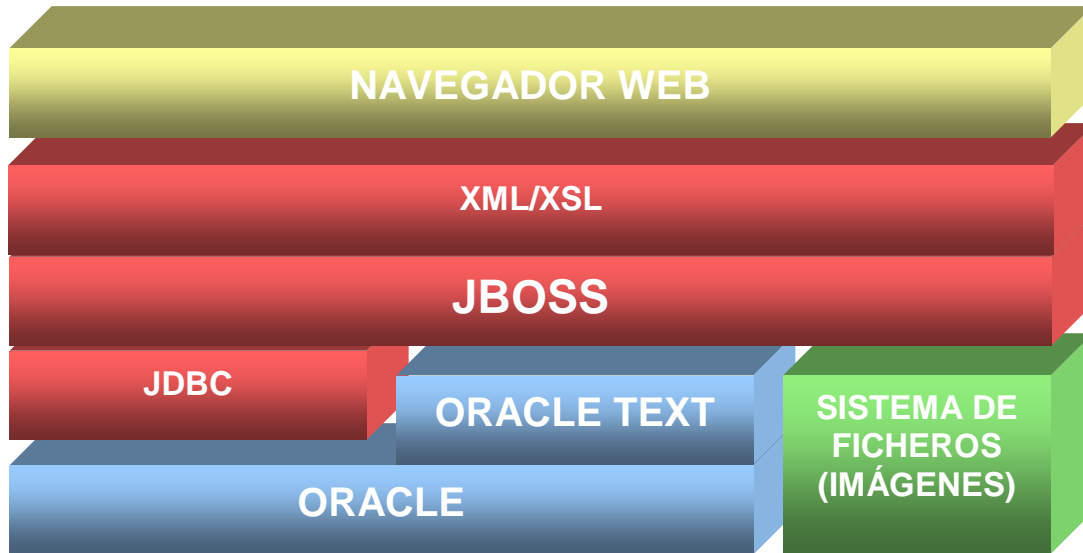
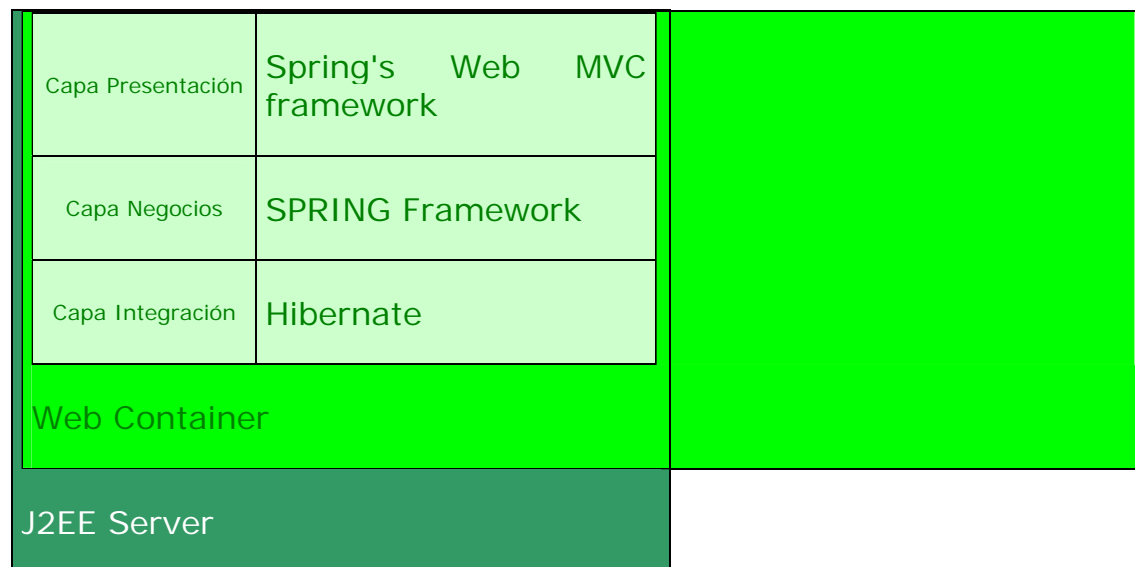
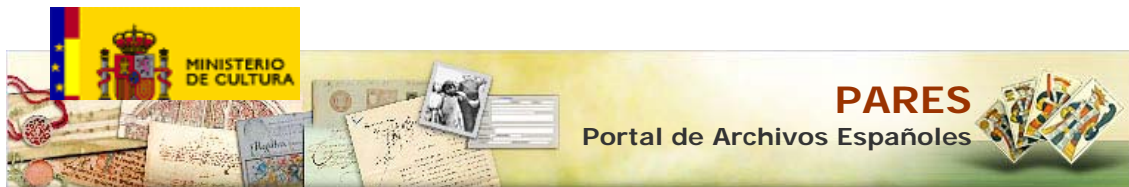


Ilustración 1 Arquitectura técnica del sistema

Dada la experiencia en el entorno tecnológico J2EE, se ha realizado una evolución lógica, para esta segunda fase del proyecto, dentro de la arquitectura de las aplicaciones que se integran dentro de este gran sistema de información. La arquitectura J2EE que se está implementando es la siguiente:

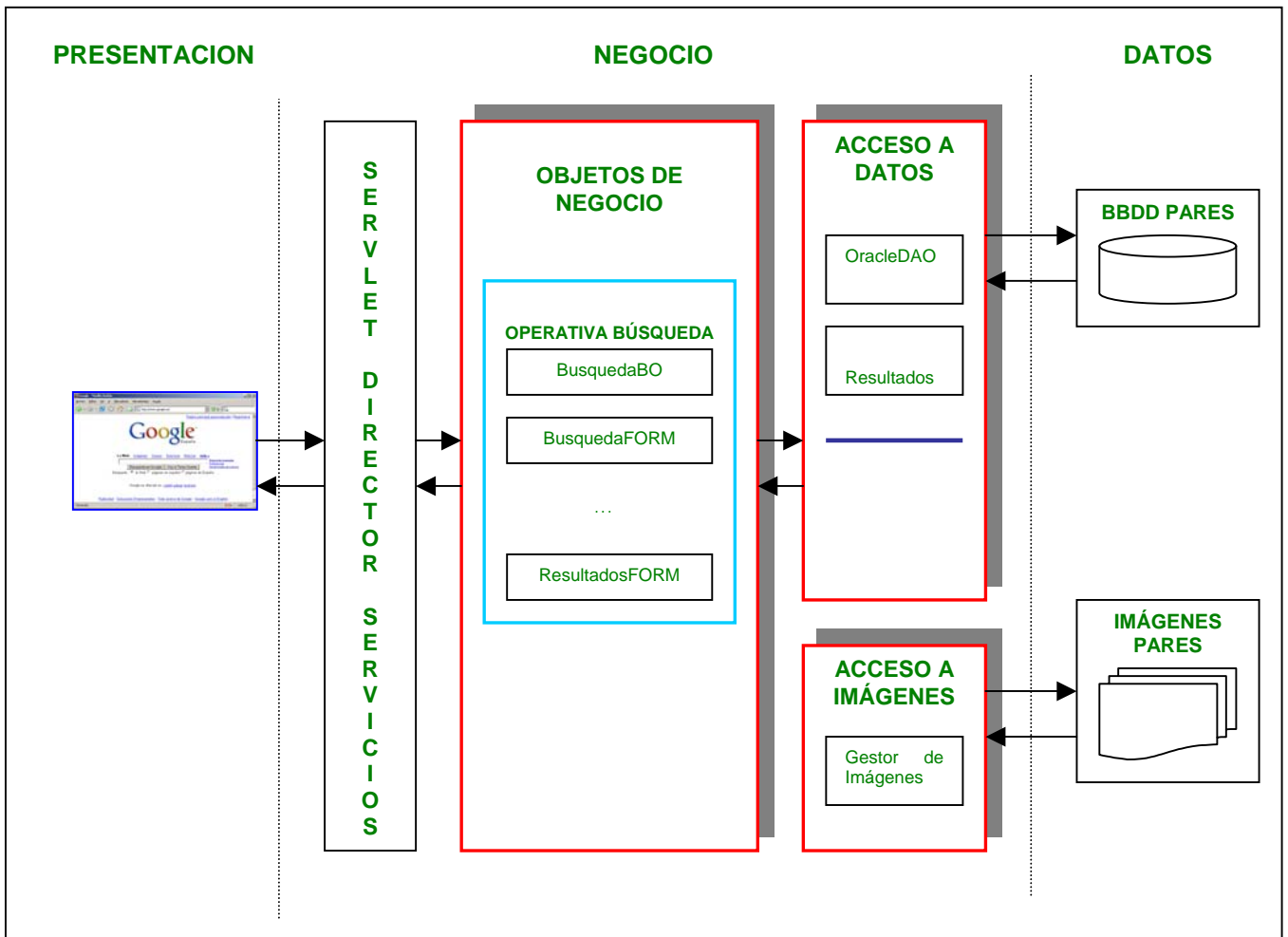




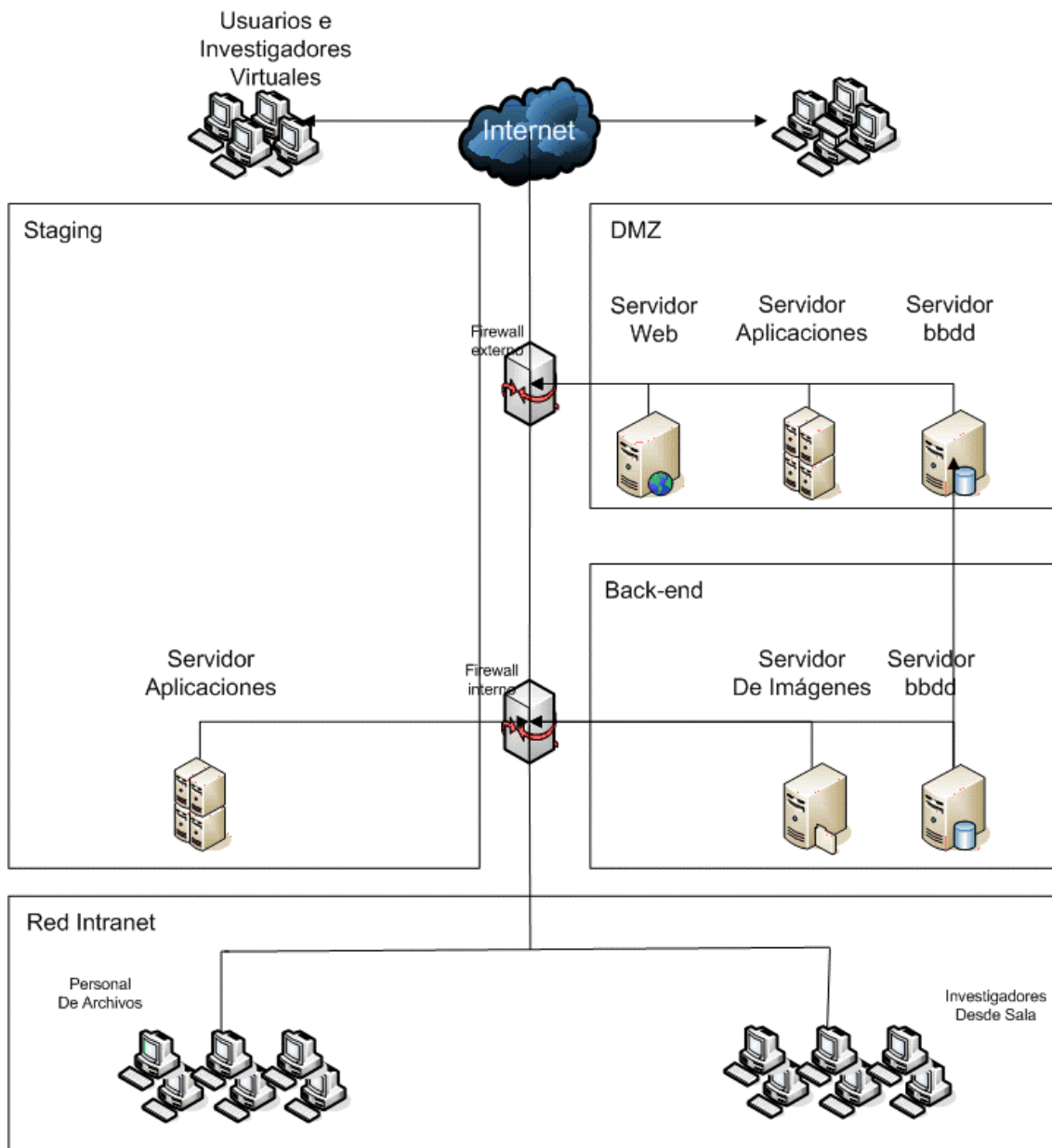
Las herramientas empleadas para el desarrollo de PARES son las siguientes:

- Eclipse 3.1.1 con Plug-in WEB Tool Project y Exadel: Este software es la herramienta de modelado empleada para realizar la codificación de las soluciones en lenguaje java, en base al análisis de las clases y casos de uso detectados.
- UML / Enterprise Architect 5.0: Es el lenguaje de modelado visual para entender, diseñar y controlar la información sobre los sistemas que se van a construir. La especificación UML proporciona apoyo a la mayoría de los procesos de desarrollo orientados a objetos detectados en PARES. Como herramienta para modelar visualmente los procesos se utiliza Enterprise Architect.
- ERWin: para realizar el Modelo Entidad/Relación de los diferentes esquemas de bases de datos que conforman el sistema.
- CVS: será la herramienta de control de versiones empleada para realizar el backup del código desarrollado, integrándose perfectamente con Eclipse.
- Macromedia DreamWeaver: Software comercial para el diseño de páginas HTML/XHTML para la interfaz de la aplicación.

A continuación se muestra la arquitectura en capas de la aplicación, independientemente de las diferentes implementaciones de código llevadas a cabo en las dos fases del proyecto y de la arquitectura del software utilizada:



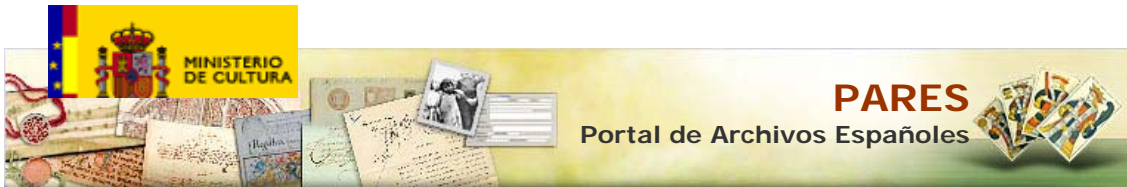
Además, se puede ver a continuación la topología de red de PARES.



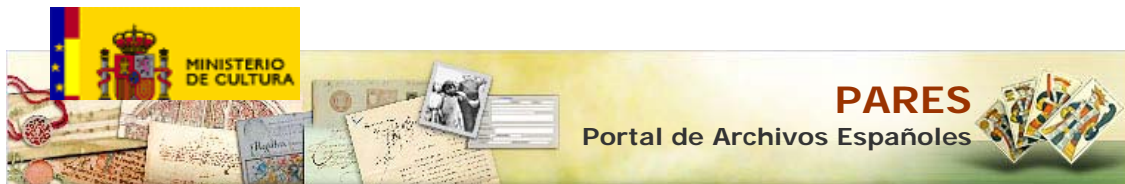
2.1. PROBLEMAS ENCONTRADOS

Los problemas que se plantean a la hora de implementar este Sistema de Información están en los diferentes parámetros que incorpora PARES, que hacen del mismo un sistema que debe estar lo suficientemente dimensionado, como para no crear cuellos de botella en ninguno de sus subsistemas.

Las características de PARES se podrían resumir de la siguiente manera:



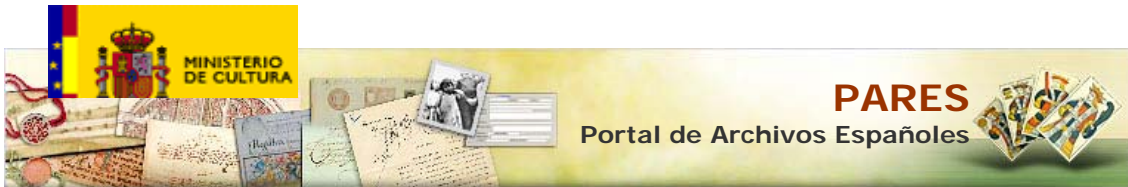
- **Gran cantidad de registros:** Como se indicaba, ahora mismo se tienen 1,7 millones de registros descriptivos de los documentos, pero en breve se pretende que estos sean 1 millón más, y esperando aún más; dado que el volumen de información que hay en los Archivos es inagotable. Se estima que la información contenida en los 11 archivos estatales, si se pusiera en línea tendría más de 170 kilómetros de longitud en la actualidad.
- **Gran cantidad de usuarios internos y externos:** El número de usuarios de PARES es elevado, encontrando usuarios internos y externos:
 - o Usuarios Internos: Como se indicaba en el anterior punto se esperan describir 1 millón de fichas documentales en breve. En los archivos hay cerca de 200 personas describiendo y/o revisando.
 - o Usuarios Externos: Los usuarios que hasta ahora se acercaban a los archivos a investigar o consultar sobre un fondo, se está intentando desviarlos a consultas e investigaciones vía Web. Después de 6 meses desde que se publicó PARES, el número medio de usuarios distintos a diario se cifra en 1.558 en el mes de septiembre, con una media de 203.324 páginas visitadas y más de 250.000 solicitudes.
- **Registros documentales:** Un registro documental es la recopilación de las informaciones que describen cada una de las unidades mínimas de información de un archivo, el documento. Es decir, estamos hablando de registros con gran cantidad de campos y grandes cantidades de información. Nos debemos imaginar que un registro está compuesto por cientos de campos, pudiendo ser éstos de muy diversos tipos:
 - o Enteros,
 - o caracteres,
 - o caracteres largos,
 - o documentos binarios, e
 - o imágenes.



- **Información estructurada de manera jerárquica:** Las relaciones entre los diferentes elementos de bases de datos incorporados en PARES son de dos tipos:
 - o Relacionales: Como base de datos se ha utilizado ORACLE.
 - o Jerárquicas: Dentro de este esquema relacional se han establecido relaciones jerárquicas entre los elementos de una misma tabla.
- **Otros trabajos realizados en el sistema:** La entrada masiva de datos a PARES (procesos de migración y carga de datos) y de salida (informes y estadísticas) son, a día de hoy, caóticos y de un alto coste.
- Poca experiencia con el sistema documental de Oracle: Oracle Text.
- La interfaz con el cliente está construida de tal manera que resulta muy difícil de mantener, modificar y/o corregir. Además, es difícilmente adaptable a exigencias como la Accesibilidad, Portabilidad, ...
- Convivencia con Censo-Guía. El Censo-Guía de Archivos de España e Iberoamérica es una guía electrónica de los archivos, fondos y colecciones públicos y privados de España e Iberoamérica. En la misma máquina conviven dos sistemas de información, con los inconvenientes que este detalle conlleva.

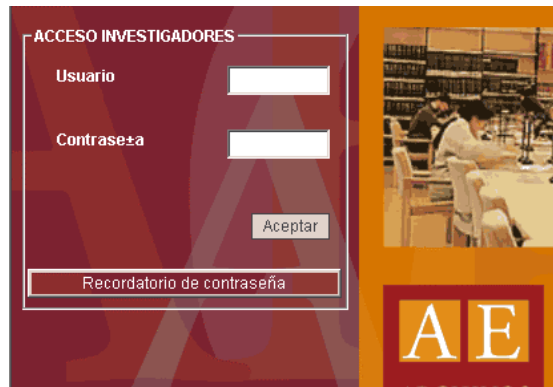
2.2. Conclusión

Dados los parámetros vistos en el punto anterior, en esta segunda fase del proyecto se está modificando todo el sistema de información para conseguir un sistema flexible a la vez que garantice la escalabilidad del mismo.



Frente al acceso tradicional a los archivos estatales

Los investigadores acceden a las Salas de Consulta de los Archivos para investigar



Se ponen los contenidos culturales de los Archivos Españoles a disposición de los Investigadores

Plataforma abierta a todos los usuarios destinada a la difusión en Internet del Patrimonio Histórico Documental Español conservado en su red de Archivos

