

ANÁLISIS E IMPLEMENTACIÓN DE UN SISTEMA AUTOMATIZADO DE DIGITALIZACIÓN DE DOCUMENTOS (SADO) PARA SOLUCIONES INTELIGENTES.

Gabriela Aguilar Baquero, Cristóbal Arboleda Terán, Mauricio Loachamin Valencia, Germán Ñacato Caiza

1 Escuela Politécnica del Ejército, Ecuador,
{ *graguilar, ocarboleda, mloachamin, gnacato* }@espe.edu.ec

RESUMEN

Soluciones Inteligentes es una empresa de desarrollo de software que tiene la necesidad de prestar a sus clientes un sistema de digitalización masiva, para evitar el trabajo manual que se ha venido realizando.

Por lo cual se ha desarrollado el sistema SADO (Sistema Automático de Digitalización de Documentos) basado en la metodología de OOHDMM orientada a objetos y ampliamente aceptada para el desarrollo de aplicaciones hipermedia.

Se ha usado la plataforma .NET 3.0, que es un framework de Microsoft, que provee un entorno de desarrollo con componentes propios, que facilitan la construcción de aplicaciones de escritorio.

El objetivo del sistema es mejorar la administración de documentos digitales de los diferentes clientes de Soluciones Inteligentes, permitiendo que la pérdida de documentos desaparezca como es en los archivos de papel.

Palabras Clave: Digitalización, Documentos, Masiva Adobe Capture, OOHDMM, OCR

ABSTRACT

Soluciones Inteligentes is a enterprise of the software developed, that have the necessity of give its clients a mass digitization system for avoid handwork that have came doing.

Whereby it has developed SADO system (Sistema Automático de Digitalización de Documentos based in OOHDM object oriented methodology and widely accepted for develop hypermedia application.

Also It has used .NET 3.0 platform that is a Microsoft framework, This framework provide a developed environment with own components that facilitate to build desktop application.

The system goal is improve the digital documents administration of the different clients of Soluciones Inteligentes, allowing that the lost documentation will not be a problem as with the paper documents.

KeyWords: Mass Digitization Documents, Adobe Capture , OOHDM, OCR

1. INTRODUCCION

El Sistema Automatizado de Digitalización de Documentos (SADO) ayudará a simplificar el almacenamiento de varios documentos que se encuentran en forma física, y que pueden llegar a extraviarse, o ser alterado, SADO propone garantizar la fiabilidad en cuanto a la existencia y administración de documentos. Se utilizará el producto Adobe Capture [5], para mejorar la productividad a través de la optimización del tiempo de administración de los documentos, logrando así garantizar el almacenamiento y organización de la información en formato digital utilizando OCR (Optical character recognition) que Adobe Capture[5] facilita.

El valor que la información representa actualmente para las diferentes empresas, negocios, escuelas, etc. se convierte en uno de los factores más importantes que se debe resguardar por lo cual surge la necesidad de gestionar, organizar y asegurar la información por esta razón se busca obtener una herramienta que resuelva este tipo de necesidades y sobre todo que el usuario mantenga la iteración y el control a la información que el maneja y permitirle que esta información se guarde en diferentes formatos y subsiguiente dale un formato más general independiente de plataformas y aplicaciones. El archivo puede ser enviado utilizando correo electrónico, también se utilizará la indexación para obtener una búsqueda más adecuada y el versionamiento para apoyo a la empresa en su proceso de auditoría, y brindará un mejor control histórico de dicho documento.

SADO además de brindar la administración de documentos, también garantizará algunas principales características como:

- Implementación Escalable
- Seguro
- Inicio de sesión único
- Tolerancia a Fallas
- Control de Usuarios

Por lo cual se pretende que el producto puede satisfacer las necesidades de los usuarios de la empresa Soluciones Inteligentes.

2. DESARROLLO DEL SISTEMA

Se empleó OOHDM [6] (Método de Diseño Hipermedia Objeto Orientado) como metodología de desarrollo y UML (Unified Modeling Language) como metodología de diseño, adicionalmente los requerimientos del sistema están definidos por parte de Soluciones Inteligentes, a estos requerimientos se realizaron las respectivas validaciones e incrementar algunos requerimientos que se vieron necesarios para el desarrollo del sistema.

Al utilizar UML se garantiza que en el ciclo de vida del producto se defina claramente los requisitos y cumplir con las expectativas del cliente.

La modularidad del esquema de desarrollo y las interfaces utilizadas, permiten garantizar que soluciones inteligentes en un futuro pueda crecer ilimitadamente así mismo como deja un estándar de desarrollo para sus futuras aplicaciones.

Las Fases de la metodología OOHDM[6] son:

a. Requisitos: Se identifica los Roles de Administrador, Editor y Usuario a continuación se describe cada una de las tareas asignas por los roles.

➤ Usuario

1. Utiliza el Servicio de Históricos
2. Utiliza Servicio de Consulta de Documentos
3. Solicita Prestamos e Impresiones

➤ Editor

1. Utilizar Servicio de Administración de Documentos
 - Aprobación de Documentos Nuevos
 - Actualizaciones
 - Aprobación de Impresiones y Prestamos
2. Utilizar Servicio de Históricos
3. Utilizar Servicio de Consulta
4. Utilizar Servicio de Reportes

➤ Administrador

1. Gestionar Usuarios
2. Gestionar Seguridades
3. Gestionar Reportes
4. Gestionar Servicio de Administración de Documentos
 - a. Aprobación de Documentos Nuevos
 - b. Actualizaciones
 - c. Aprobación de Impresiones y Prestamos

b. Diseño Conceptual: En esta actividad de se elaboró un Diagrama de Base de Datos que nos permitió identificar la estructura de la Base de datos y organizar las tablas del Sistema

e identificar las clases elaborado en el Diagramas de Clases. En la Figura 2.1 se presenta el Diagrama de Base de Datos SADOB.

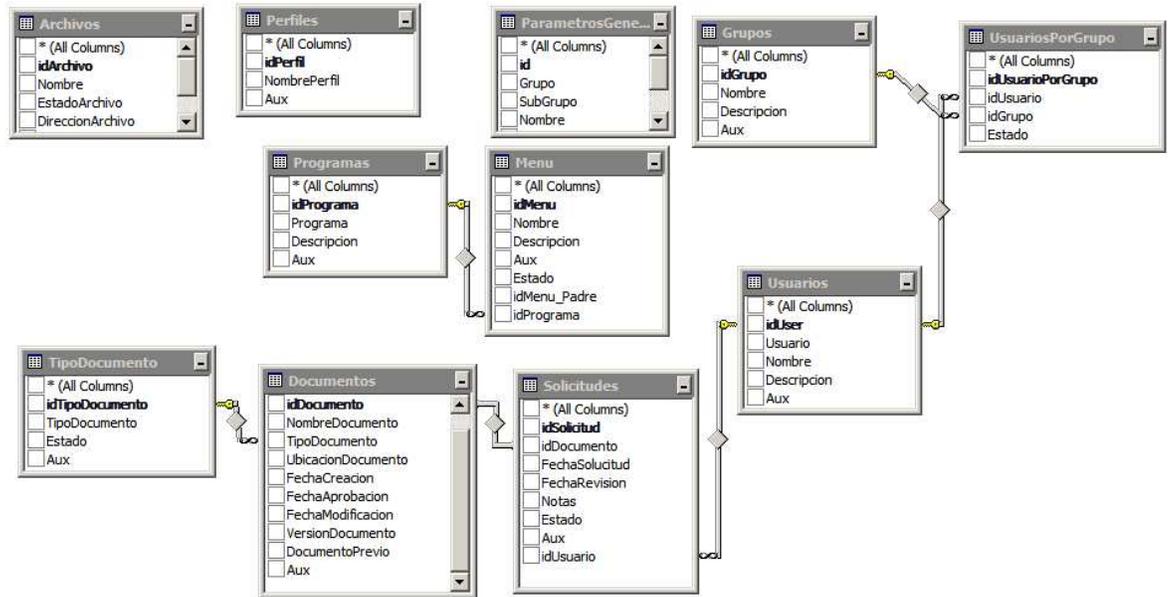


Figura 2.1 Diagrama de Base de Datos

c. **Diseño Navegacional:** A continuación el diseño de navegación se expresó en esquemas de contexto de clases navegacionales. Que para el Sistema SADO son modelos diferentes de acuerdo con los perfiles definidos (Administrador, Editor, Usuario) siendo el primero el Esquema para el Menú Principal, que se puede observar en la Figura 2.2.

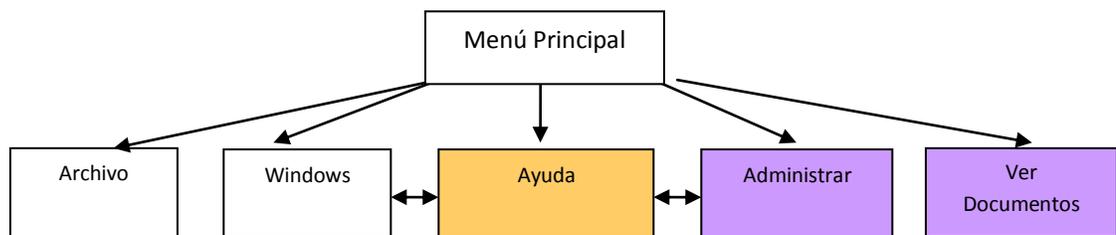


Figura 2.2 Esquema Navegacional del Menú Principal

d. **Diseño de Interfaz Abstracta:** Una vez que las estructuras Navegacionales son definidas se procedió a definir los aspectos de la interfaz por lo cual se crearon dos

diagramas: Diseño estético y Diagrama de Estado este último que se presenta en la Figura 2.3.

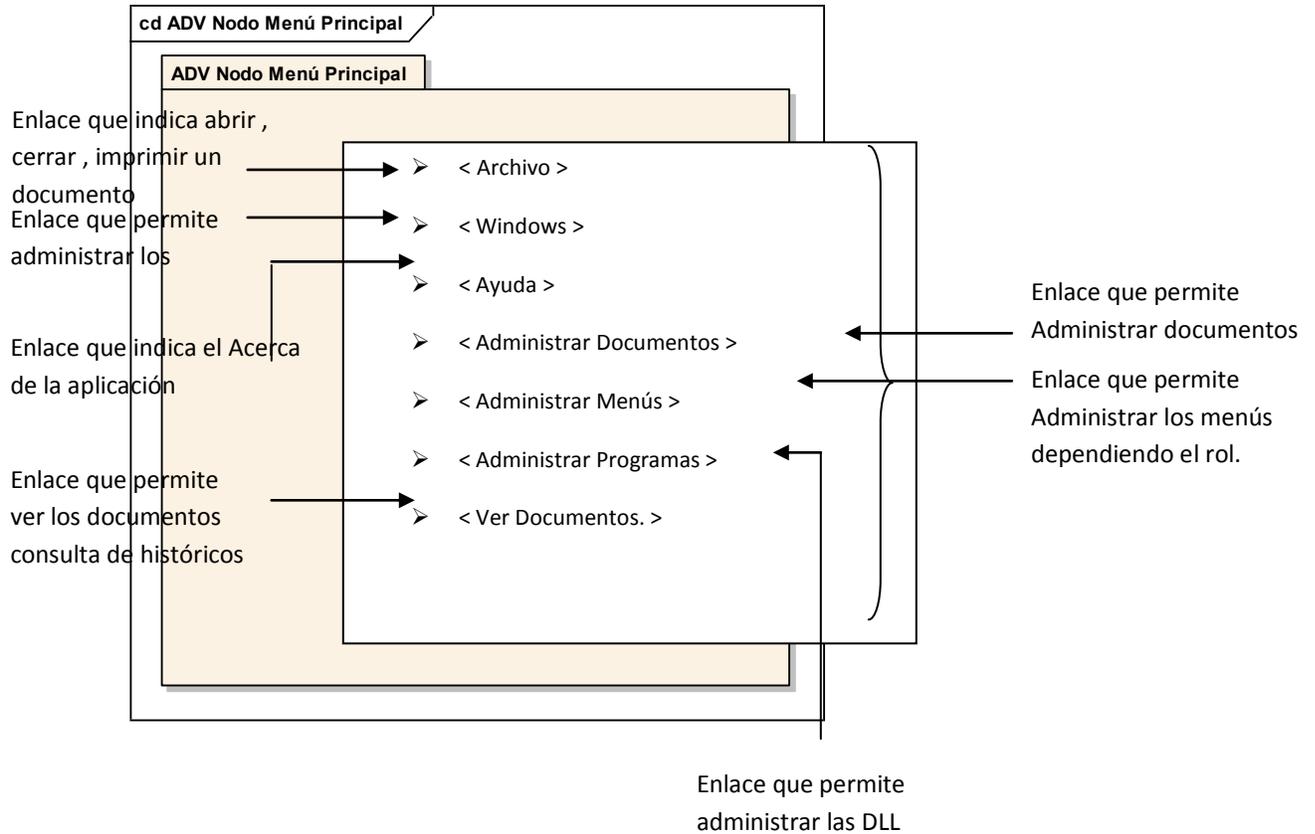


Figura 2.3 Diagrama de Estado – SADO

2.1 DEFINICIÓN DE ROLES

Según las necesidades de Soluciones se definieron los siguientes roles:

- *Product Owner:* Soluciones Inteligentes
- *Project Master:* Ing Edwin Jaramillo
- *Project Team:*

1.1.1 Cristóbal Arboleda,

1.1.2 Gabriela Aguilar.

3. ARQUITECTURA

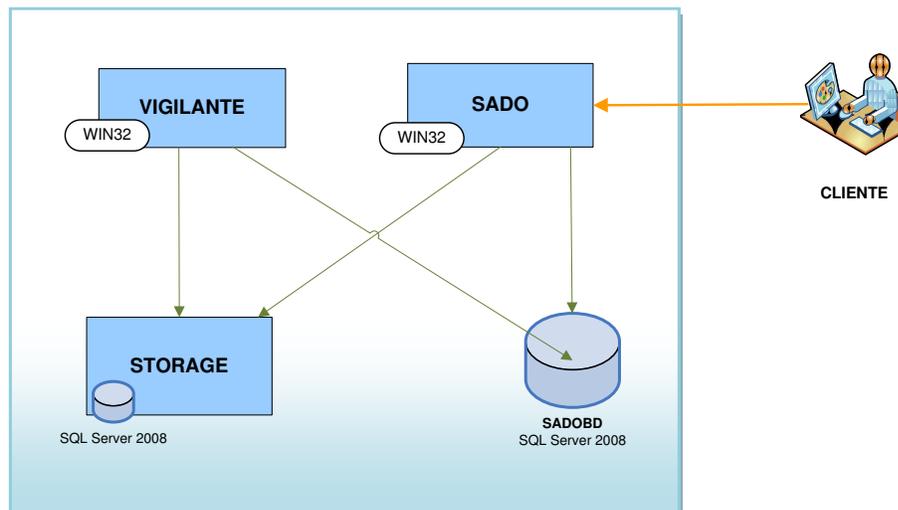


Figura 3.1 Diagrama de Arquitectura

La arquitectura que presenta SADO es de Cliente/Servidor de 3 capas. En donde se integra todas las herramientas de forma sólida y eficiente.

A continuación se describe cada uno de los componentes de la arquitectura SADO.

- a. **VIGILANTE:** Este modulo se encarga de Monitorear una carpeta definida, en donde se depositan los archivos digitalizados. Cuando se encuentra un archivo nuevo, Vigilante lo mueve y lo ingresa al Sistema SADO para que posteriormente, el Editor o Administrador ejecuten tareas específicas sobre el archivo.
- b. **SADO :** Se encarga de la Administración de Documentos como es la aprobación, consulta y solicitudes de documentos, dependiendo de los accesos que posee cada uno de los perfiles definidos para la utilización del sistema.
- c. **STORAGE:** Es un lugar donde se almacena la información físicamente. Windows Server con el rol de File Server monitorea el Storage, para administrar tanto el ingreso y el tipo de archivos, a su vez controla quien ingresa a dicha información denegando el acceso a la misma de ser necesario.
- d. **SADODB :** Es la base de datos del Sistema SADO , en donde se organiza , almacena la información, versionamientos y propiedades del archivo digitalizado. Y a su vez permite recuperar datos de archivos digitalizados para el proceso de históricos que el Sistema SADO posee.

3.1 HERRAMIENTAS UTILIZADAS

Para la elaboración de SADO se utilizó herramientas actuales open y closed source, las mismas que se menciona a continuación:

- a. **.Net Framework 3.0 [4]** : Es parte integral de las muchas aplicaciones que se ejecutan en Windows y proporciona funcionalidad común para ejecutar las aplicaciones.
- b. **SQL Client:** Es un utilitario que permite acceder a base de datos.
- c. **SQL Server 2005:** Es un motor de la base de datos que ofrece almacenamiento más seguro y confiable tanto para datos relacionales como estructurados.

4. REQUERIMIENTOS

Soluciones Inteligentes generó un listado de requerimientos los cuales fueron validados para verificar su factibilidad de implementación, adicionalmente durante el desarrollo de SADO se identificaron varios requerimientos más sin los cuales el proyecto no hubiera sido factible.

Entre los Requerimientos se puede resaltar los siguientes:

- Sistema Modular, basado en el estándar de Soluciones Inteligentes, tablas no recursivas, implementación de componentes por esquema de plug ins, basado en .Net 2005[4] o mejor, instalación de componentes automática sin necesidad de registro, sincronización con Active Directory [2], perfiles basados en instalación de componentes.
- Administración de información mediante store procedures, dejar abierta la posibilidad de inclusión de triggers para manejar auditoría.
- Módulo de vigilancia (SADO Vigilante) (servicio de windows) monitoreo de archivos ingresados, organización por categorías, fechas, tipos, parámetros configurables desde parámetros generales.
- Manejo de histórico de documentos
- No existe eliminaciones

Adicionalmente Incluímos requerimientos como:

- Instalación de Active Directory[2]
- Instalación de Adobe Capture [5]
- Administración de Document Services
- Módulo de Migración de LDAP (LDAP Migrator) [1]

5. INTERFAZ DEL SISTEMA

La interfaz del sistema fue concebida de manera que sea de fácil comprensión, intuitiva y practica enfocándonos en que los usuarios finales conocen poco o nada acerca de la administración de Documentos.

Los Módulos implementados además de manejar una GUI (Graphical User Interface) similar, son simples y ofrecen búsquedas específicas basadas en criterios preestablecidos de información.

La interfaz del sistema se encuentra distribuida de la siguiente manera como se puede observar en la Figura 5.1:



Figura 5.1 Interfaz del Sistema

6. TRABAJOS RELACIONADOS

Entre los trabajos relacionados, podemos mencionar los siguientes:

- **Empresa ComputingManagement** con su producto Digitalización de Documentos con FULL TEXT (<http://www.cmsrl.com.ar/Servicios.html>)
- **FAI Solución de Digitalización de Documentos**
<http://www.todomercado.com/Productividad-Negocios/-SOFTWARE-SISTEMA-DE-DIGITALIZACION-DE-DOCUMENTOS.view?id=1226592270952>
- **Sistema de Gestión Documental Quipux**
<http://www.gestiondocumental.gob.ec/>

- **Lead Solutions** con su producto Inforouter
www.lead solutions.ec/Inforouter/Inforouter_Brochure.pdf

7. IMPACTO DEL SISTEMA

SADO al ser un producto que no va enfocado a un cliente final si no hacia una empresa de desarrollo no podemos contar con las estimaciones de los usuarios finales ni como afecto dentro del entorno profesional la implementación del mismo.

Nuestras Expectativas sobre el impacto son las siguientes:

- Reducir las pérdidas de información por mal manejo de documentos. Al mantener históricos de los documentos con un espacio de almacenamiento adecuado e implementando un respaldo correcto de la información la administración de documentos prevalecerá de manera indefinida.
- Simplificar la Administración de Información Documental. Con un sistema de gestión único, dicha administración y aprovechamiento no tienen límites.
- Reducir los tiempos de búsquedas de información documental. Al contar con un repositorio centralizado y con nuestro módulo de búsqueda la ubicación de documentos y las versiones que estos puedan presentarse podemos reducir dichos tiempos hasta en un 90% de lo que tomaría ubicar manualmente un documento.
- Administración de Acceso a la Información. De manera implícita en SADO podemos controlar quien o que accede a la información así llevamos un control total sobre información que podemos definir como Privada o Secreta.

8. CONCLUSIONES

En base a los requerimientos presentados por parte de soluciones inteligentes se diseñó el plan de desarrollo de la aplicación SADO con el objetivo de cumplir las expectativas del cliente y sus futuras implementaciones.

Con respecto a la parte investigativa de los autores se concluye que para alcanzar el objetivo propuesto, fue necesario adquirir nuevos conocimientos de los ya recibidos en la Universidad debido a que las aplicaciones en la vida profesional cubren una gama mucho más amplia de componentes, interfaces, lenguajes, artificios de desarrollo, que los aprendidos en la universidad.

Para el desarrollo de este sistema se elaboró un análisis FODA en donde se evaluó la empresa "Soluciones Inteligentes" en donde como producto estrella para alcanzar una de sus oportunidades es el aplicativo SADO.

Al final del presente trabajo se puede concluir que el diseño inicial de interfaces y del estándar modular de la aplicación fue la forma más correcta del desarrollo aplicado, su escalabilidad es única e ilimitada y el presente diseño base de la aplicación así mismo como sus librerías particulares serán aplicadas dentro de Soluciones Inteligentes como el nuevo Estándar de Desarrollo para sus futuros proyectos.

Las pruebas de contenido, función, estructura, facilidad de uso, navegabilidad, y desempeño; ayudaron a detectar y corregir los errores antes de la puesta en producción del producto SADO para Soluciones Inteligentes.

9. REFERENCIAS

- [1] Wikipedia, LDAP , Actualización 27 julio 2011; <http://es.wikipedia.org/wiki/LDAP>

- [2] Wikipedia ,Active Directory, Actualización 22 agosto 2011;
http://es.wikipedia.org/wiki/Active_directory

- [3] Fabian Tang, 21 Diciembre del 2005, Referencias de Interfaces de Sistemas MDI;
<http://www.codeproject.com/KB/bugs/MgdMDIClientActivation.aspx>
<http://www.codeproject.com/KB/cs/MDITabBrowsing.aspx>

- [4] Microsoft, 2009-2010, Librería de Referencia de .Net,
<http://msdn.microsoft.com/es-ec/library/>

- [5] Adobe Systems Incorporated, 14 Julio 2009 , Adobe Capture,
<http://www.adobe.com/products/acrcapture/>

- [6] PRESSMAN Roger, Ingeniería de Software Un Enfoque Práctico, Mc. Graw Hill, Madrid – España, 2002.