

Plan de migración de datos

Proyecto: Sistema para gestión de actividades en el gimnasio del CEET - GymSenApp

Versión: 1.0

SENA

2024

Historia de revisiones

|  |  |  |  |
| --- | --- | --- | --- |
| Versión | Autor(es) | Descripción | Fecha |
| 1.0 | Mariana Lizeth Marín Rojas, Santiago Salamanca Díaz, Marco Antonio Ortíz Ramírez | Creación del documento | Feb 2024 |

Índice

[1. Introducción](#_Toc159791761)

[*1.1. Alcance del documento.*](#_Toc159791762)

[*1.2. Generalidades del proyecto.*](#_Toc159791763)

[*1.3. Glosario.*](#_Toc159791764)

[2. Justificación del plan de migración de datos](#_Toc159791765)

[3. Objetivos del plan de migración de datos](#_Toc159791766)

[4. Planeación](#_Toc159791767)

[*4.1. Generalidades.*](#_Toc159791768)

[*4.2. Descripción del proceso de migración de datos.*](#_Toc159791769)

[*4.3. Estrategia de migración.*](#_Toc159791770)

[*4.4. Criterios de aceptación.*](#_Toc159791771)

[*4.5. Metodología.*](#_Toc159791772)

[4.5.1. Descripción de las fases.](#_Toc159791773)

[*4.5.1.1. Análisis de las fuentes de datos.*](#_Toc159791774)

[*4.5.1.2. Extracción y transformación de los datos en el formato destino.*](#_Toc159791775)

[*4.5.1.3. Validación de los datos.*](#_Toc159791776)

[*4.5.1.4. Prueba y cargue de datos en el sistema destino.*](#_Toc159791777)

[*4.6. Recursos.*](#_Toc159791778)

[4.6.1. Personal.](#_Toc159791779)

[4.6.2. Hardware y software.](#_Toc159791780)

[*4.7. Cronograma de actividades.*](#_Toc159791781)

[5. Referencias](#_Toc159791782)

# **1. Introducción**

## ***1.1. Alcance del documento.***

El propósito del presente documento es proporcionar la información correspondiente para planificar y desarrollar el plan de migración para el software GymSenApp.

En este sentido, se deben tener claras unas nociones generales sobre el proyecto, las cuales permitirán obtener una contextualización idónea sobre el proyecto de software. Veamos pues las consideraciones:

## ***1.2. Generalidades del proyecto.***

El sistema para gestión de actividades en el gimnasio del CEET, llamado “GymSenApp”, es un proyecto formativo el cual busca optimizar los procesos que engloba el uso del gimnasio del centro, permitiendo así una mejor planificación de las rutinas de ejercicios de los aprendices, así como también permite apoyar el acompañamiento y comunicación por parte del personal del gimnasio hacia los aprendices que asistan al lugar.

El sistema está compuesto de tres (3) grupos de actividades generales agrupadas según el tipo de usuario que puede acceder a ellas:

* + Administradores
  + Aprendices
  + Instructores

Cada uno de estos grupos o roles, está compuesto por actividades específicas:

* + Administradores
    - Gestión de anuncios
    - Gestión de usuarios
    - Generación de reportes
  + Aprendices
    - Visualizar anuncios
    - Visualizar recomendaciones saludables
    - Planificar rutinas
  + Instructores
    - Gestionar asistencias

## ***1.3. Glosario.***

En este documento se utilizan los siguientes términos abreviados:

* GymSenApp: Nombre del software

# **2. Justificación** **del plan de migración de datos**

Este plan se está creando debido a su importancia a la hora de desplegar el sistema y permitirle su uso a todo tipo de usuarios en la gran red actual llamada internet, ya que es menester mirar hacia el futuro y tener en cuenta que en algún momento por x o y razón, se puede requerir manipular los datos del sistema de forma masiva, ya sea para migrar de entorno, realizar algún tipo de prueba global, insertar, editar o mover grandes volúmenes de datos para algún proceso específico, lanzar una actualización al sistema, entre otros (determinados según la necesidad de la empresa o del contexto del proyecto, en el caso, del CEET y del personal relacionado al SENA, el centro de formación y a los instructores del gimnasio).

Por consiguiente, se desarrolla este plan para ejecutar a corto plazo una pequeña migración de información, que nos permita evaluar en esta etapa del proyecto, si se puede ejecutar un proceso tan amplio como el de migrar información, al igual que nos permita aclarar el panorama sobre cómo se ejecutaría un proceso de estos aplicado a este software, de tal modo que si en un futuro se requiere realizar un proceso parecido o más grande, se tenga una guía sobre más o menos cómo se podría aplicar este proceso.

# **3. Objetivos del plan de migración de datos**

El plan de migración de datos siendo planificado en este documento, tiene como objetivo general transferir cierta cantidad de información a la base de datos del software. Adicionalmente, comprende 4 objetivos específicos:

1. Analizar y diseñar los procesos de creación, transformación y carga de los datos estipulados.
2. Desarrollar el o los instrumentos necesarios que contengan la información fuente a ser migrada a la base de datos teniendo en cuenta el modelo de datos de esta.
3. Cargar información válida (que cumpla con las reglas de validación de la base de datos destino).
4. Utilizar archivos en formato CSV y usar técnicas de importación de archivos en MySQL (motor de base de datos que contiene la base de datos del sistema).

# **4. Planeación**

## ***4.1. Generalidades.***

La migración de datos planteada contempla el análisis, diseño, implementación y pruebas a los datos inmersos dentro del proceso. La idea contempla crear un conjunto de registros en archivos CSV de tal forma que se realice un proceso de importación de datos a la base de datos del sistema, de tal forma que se pruebe la capacidad de nosotros como equipo encargado del proyecto, y del sistema, para encaminar procesos relacionados a las prácticas de migración de información a una base de datos.

De acuerdo con lo anterior, podemos asegurar que las fuentes de información para realizar este proceso de migración de datos son, por un lado, los esquemas que rodean al sistema de información y la estructura de los datos que maneja, hablamos del MER, de la base de datos, las tablas existentes, las columnas y los tipos de datos que aceptan, y demás elementos relacionados a las características de los datos del sistema; por otro lado, el software GymSenApp concebido como un todo, conteniendo tanto los procesos que ejecutan los usuarios, así como los elementos que componen al sistema en general, dando una guía sobre el contexto del proyecto y lo que se debe tener en cuenta para lograr que el proceso de migración de datos sea un éxito.

Aparte de lo anterior, la información objeto de la migración va a estar relacionada directamente a la base de datos, teniendo en cuenta que se realizará (como se explicó anteriormente) una importación de datos hacia la base de datos; por lo cual los datos a migrar tienen que ver con la información que el sistema guarda.

## ***4.2. Descripción del proceso de migración de datos.***

En primera instancia, se creará la información a migrar a la base de datos, se creará en un archivo XLSX (de Microsoft Excel) de tal forma que allí se registrarán los datos e información relacionada al sistema, y será el repositorio o fuente origen de los datos a pasar a la base de datos, representando el sistema destino. Los datos a generar serán creados por nosotros como desarrolladores y con ayuda de agentes externos, para determinar variaciones en los datos que puedan existir en la realidad buscando probar al sistema en su capacidad de almacenar y gestionar correctamente esos datos que debería guardar y manipular correctamente.

Luego se ejecutará el proceso de comprensión de los datos, para lo cual, tanto se reestudiarán los documentos y la información que tenemos sobre el sistema y cómo gestiona los datos, qué tipos de datos acepta en qué tablas, qué condiciones existen, las relaciones y dependencias que existen entre tablas, entre otros delimitantes que es necesario conocer que maneja la base datos; y adicionalmente, se analizarán los datos creados en el archivo a migrar, de tal forma que se comprenda la estructura de estos datos, el contexto en el que se enmarcan y qué representa cada cosa, con el objetivo de, posteriormente (habiendo comprendido los datos origen), pasar a transformarlos y limpiarlos de tal forma que se adapten al modelo esperado por la base de datos, sin perder de vista que estos datos inicialmente recogidos tienen su razón de ser y se les tiene que tratar con cautela, para no perder información valiosa por hacer las cosas rápidamente y tomarlas a la ligera, velando por comprender a los usuarios, lo recolectado, y buscando trasladar la información de la forma más completa y transparente posible al sistema para poder ser entendido por la base de datos.

Posteriormente pasaremos a transformar los archivos origen (del archivo XLSX) a un tipo de archivo que pueda ser entendido por la base de datos (en el caso se pasarán a CSV, cuidando elegir el formato UTF-8 del mismo), luego se pasará al proceso de carga, para lo cual se tendrá que crear el entorno propicio para que la base de datos acepte la importación del archivo y arroje los resultados del proceso (en el caso se creará la base de datos con los datos que incorpora inicialmente y se realizará la importación de los datos); y por último se realizarán las validaciones correspondientes para determinar cómo se desarrolló el proceso, qué situaciones inesperadas ocurrieron, y en general qué resultados se obtuvieron del proceso, permitiendo determinar si es necesario repetir el proceso, modificar algún elemento (bien sea del servicio origen o del servicio destino), o ejecutar más pruebas para lograr que el proceso de migración de datos sea concluido con éxito.

Es importante resaltar que realizaremos este proceso de migración de datos de forma local y en entornos de prueba, de tal forma que podamos corroborar la generación correcta de los archivos CSV, los resultados de las importaciones, al igual que podamos como controlar y/o reaccionar a las situaciones que puedan ocurrir (por ejemplo, si en algún caso ocurre un problema, se puede tranquilamente eliminar información, o hasta eliminar la base de datos); así como se realizarán las verificaciones correspondientes antes y después de realizar las importaciones, de tal forma que se consiga al final migrar la información completa y correctamente.

## ***4.3. Estrategia de migración.***

De acuerdo con lo anteriormente expuesto, la estrategia de migración a implementar será la Big-Bang, debido a que el sistema dentro del entorno de pruebas va a ser probado con la migración, se va a implementar por completo la misma buscando lograr que toda la información esperada sea transferida (o importada) a la base de datos. Posterior a la satisfacción de aquel objetivo, se pasará a realizar la importación en el entorno real o de producción, en donde se espera que todo se desarrolle satisfactoriamente, y por lo tanto solo bastará con importar los archivos correspondientes, dado a que con las pruebas ejecutadas se asegura que en el entorno real los datos se transferirán correctamente de tal forma que el sistema en producción, con una sola operación, se actualice e integre la nueva información correctamente.

## ***4.4. Criterios de aceptación.***

El criterio de aceptación de la migración de datos se determina a través del indicador de carga de información, el cual (dado a las condiciones actuales del proyecto y la cantidad de información planeada a migrar) se espera que sea del 100%; es decir, que todos los datos que se espera migrar a la base de datos logren efectivamente ser importados en ella y con todos sus atributos tal cual estén ideados en el sistema origen de los datos (archivos XLSX).

## ***4.5. Metodología.***

La metodología a implementar para desarrollar el proceso de migración de datos va a ser una basada en una secuencia cíclica que permite el cargue ordenado y esperado de los datos en el sistema origen, al sistema destino.

Las cuatro fases en las que consistirá este proceso de migración de datos son las siguientes:

1. Análisis de las fuentes de datos.
2. Extracción y transformación de los datos en el formato destino.
3. Validación de los datos.
4. Prueba y cargue de datos en el sistema destino.

Este proceso se desarrollará de forma cíclica, de tal modo que si llegan a presentarse problemas al mover los datos al sistema destino (en el caso la base de datos), se pueda repetir el proceso hasta que se logre consumar la migración exitosamente.

### **4.5.1. Descripción de las fases.**

#### ***4.5.1.1. Análisis de las fuentes de datos.***

En esta fase se crearán los archivos Excel con la información que se desea migrar, de tal forma que se registrarán los datos en el formato requerido.

Las actividades a realizar en esta fase, basándonos en el sistema iterativo a implementar, se dividen en las siguientes:

* Descubrimiento inicial: Este es el subproceso que permite entender los datos a migrar, teniendo en cuenta comprender lo que contienen y cómo se relaciona la información entre ella.
* Mapeo: Este subproceso se ejecuta para crear una comparación entre los datos con los que actualmente se está trabajando (fuente) y el formato en el que se espera que sean convertidos (destino), de tal forma que en esta fase se debe entender los datos con los que se cuenta, las relaciones que estos pueden establecer con la base de datos y con los datos esperados por esta, así como determinar los cambios requeridos para que sean acordes con lo esperado por el sistema destino.

El resultado de esta fase es directamente proporcional a la calidad del entendimiento y especificación de los datos tanto fuente como destino, de tal modo que, si se quieren obtener buenos resultados, hay que ejecutar esta etapa con mucho detenimiento.

#### ***4.5.1.2. Extracción y transformación de los datos en el formato destino.***

En esta fase se desarrollarán los procedimientos correspondientes para transformar los datos fuente y validar su probabilidad de éxito en la importación a la base de datos.

Las actividades a realizar en esta fase, basándonos en el sistema iterativo a implementar, se dividen en las siguientes:

* Manejo de irregularidades de datos y excepciones: En este subproceso se realizarán los análisis y correcciones pertinentes para enfrentar las situaciones que se puedan presentar relacionadas al movimiento de la información fuente al formato destino, determinadas por eventos como la incompatibilidad de tipos de datos, relaciones imposibles planteadas, entre otros elementos que deberán ser corregidos para lograr que la información sea migrada correctamente.
* Maximización de la eficiencia de la migración: Aquí tendremos en cuenta efectuar las modificaciones necesarias para lograr que la migración sea un éxito, de tal forma que, si se detectan elementos, registros o campos que puedan ser modificados y que su cambio represente una mayor probabilidad de eficiencia y ejecución correcta de la migración de los datos, se efectúe para lograr cumplir con el cometido.

#### ***4.5.1.3. Validación de los datos.***

En esta fase se crearán los archivos CSV con la información que se desea migrar, de tal forma que se desarrollarán los procedimientos adecuados para validar que los datos puedan ser importados a la base de datos correctamente.

La actividad a realizar en esta fase, basándonos en el sistema iterativo a implementar, se enfocará en la revisión de los datos generados al convertir los datos fuente en datos que puedan ser comprendidos por el servicio destino, de tal forma que se asegure que estos datos a importar realmente vayan a ser importados con éxito a la base de datos, validando elementos como las restricciones que tenga la base de datos con respecto a la inserción de datos dentro de ciertos campos o columnas, y corroborando que no se violen las restricciones identificadas.

#### ***4.5.1.4. Prueba y cargue de datos en el sistema destino.***

En esta fase se desarrollarán los procedimientos correspondientes para el cargue de la información esperada a la base de datos y la retroalimentación que se pueda obtener al ejecutar el proceso.

Las actividades a realizar en esta fase, basándonos en el sistema iterativo a implementar, se dividen en las siguientes:

* Ejecución de la carga de información: En esta actividad, se realiza el cargue de la información al sistema destino, es decir, se realiza la importación de la información de los archivos CSV a la base de datos.
* Análisis resultados obtenidos: En esta actividad se ejecuta un análisis de la información que arroja la base datos sobre la consecución del proceso de importación del o de los archivos (indicando si la importación se completó de manera satisfactoria, cuántos elementos fueron saltados, cuantos errores se generaron, etc.).
* Análisis datos cargados en la base de datos: En este subproceso se analizan los datos cargados en la base de datos, corroborando cómo llegaron, si hicieron falta, si el formato es el que corresponde, si hubo datos insertados de forma incorrecta, etc.

## ***4.6. Recursos.***

Para ejecutar este proceso de migración, requerimos de los siguientes recursos:

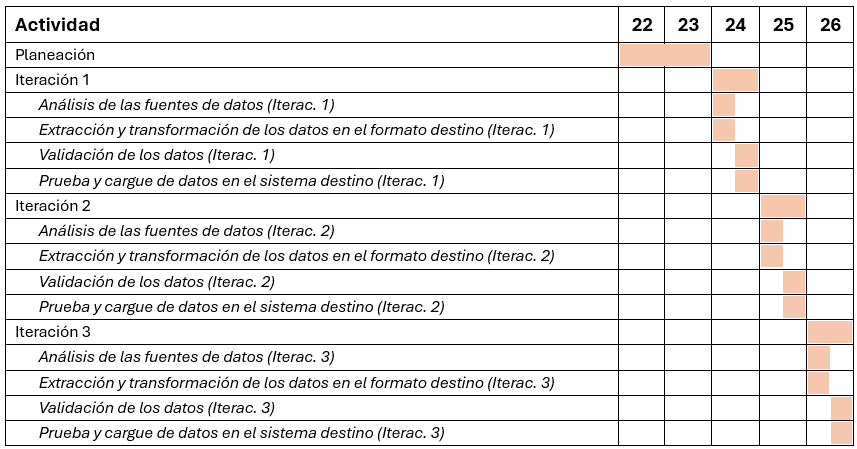
### **4.6.1. Personal.**

En cuanto a los requerimientos de personal, vamos a estar involucrados nosotros los desarrolladores del proyecto (Mariana Lizeth Marín Rojas, Santiago Salamanca Díaz y Marco Antonio Ortíz Ramírez), así como solicitaremos información de aprendices de nuestro centro de formación, con el objetivo de mezclar información para probar el sistema tanto que sea ideada por nosotros según nuestros conocimientos y lo que concebimos que puede verse en la vida real, así como obtener información de aprendices que brinden datos que puedan ser reales y que puedan llegar a ser insertados en el sistema cuando sea usado por cualquier usuario de internet.

### **4.6.2. Hardware y software.**

En este apartado, requeriremos de mínimo un computador (puede ser portátil o de mesa) con un sistema operativo (en lo posible Microsoft Windows), un buscador web, un motor de bases de datos MySQL (en lo posible xampp o MySQL Workbench) y, finalmente, requeriremos de mínimo el programa Microsoft Excel junto con sus opciones de guardado de archivos (para lograr crear archivos XLSX y CSV).

## ***4.7. Cronograma de actividades.***



Se incluyen tres iteraciones para la migración de la información, ya que consideramos que sería una cantidad realista de iteraciones que se desarrollarán debido a los problemas que pueden ocurrir al migrar la información. Sin embargo, es imperante tener en cuenta que puede disminuirse o extenderse la cantidad de iteraciones y, por lo tanto, el tiempo que demore en realizarse correctamente la migración de datos.

# **5. Referencias**

Procuraduría general de la nación. (2006, diciembre 14). Diseño, Desarrollo e Implantación del Sistema de Información Misional (SIM) de la Procuraduría General de la Nación. Plan de Migración. V.0.2.