

Propuesta de una metodología para el proceso de Migración de Datos en entornos empresariales

Washington Oswaldo Pérez Argudo
Departamento de Ciencias de la Computación. Universidad de las Fuerzas Armadas ESPE.
Correo: woperez@espe.edu.ec

RESUMEN

El momento en que las instituciones deciden cambiar sus sistemas informáticos, aparece una disyuntiva sobre la accesibilidad a la data histórica, optando en ocasiones por mantener una mínima plataforma tecnológica a fin de que ciertos usuarios accedan a esa información, o migrar esos datos fusionando su almacenamiento con los nuevos repositorios. Esta última opción no ha sido acogida con frecuencia debido a que el proceso es tedioso y complejo. Se plantea en este estudio una metodología que ordene el proceso de migrar esos datos históricos hacia los nuevos repositorios a través de Pasos, Tareas, Actividades y Herramientas que al ser agrupadas en Etapas y Fases. A fin de validar la metodología propuesta, se realiza un caso de estudio en la empresa Automotriz Quito Motors, y se determina que el paso de información histórica fue óptimo en cada una de las iteraciones en las que se aplicaron diferentes *reglas de validación* y *reglas del negocio*, concluyendo con un 64% de registros migrados contra un 36% de registros rechazados, garantizando la integridad de la información migrada.

ABSTRACT

The time when institutions decide to change their computer systems, a dilemma appears on accessibility to historical data, sometimes opting to maintain a minimum technological platform so that certain users access to that information, or migrate the data merging their storage in the new repositories. The last option has not been met with frequency because the process is tedious and complex, with no guarantee of optimal results. The purpose of this paper is to raise a methodology to order the process of moving towards such historical data to new repositories through Steps, Tasks, Activities and Tools which when grouped into Stages and. In order to validate the proposed methodology, it was applied in a case study in the Quito Motors Company, concluding that the passage of historical information was optimal for each of the iterations that different validation rules and business rules were applied, concluding with a 64 % migrated records against 36% of rejected records, ensuring the integrity of the migrated information.

I. INTRODUCCIÓN

El cuidado de la información siempre será considerado el principal activo de las empresas, si

las empresas pierden datos, pierden dinero, afectando su competitividad y participación en el mercado, por lo que es de vital importancia que al momento que los sistemas informáticos son reemplazados por nuevas tecnologías, también se formalice el cuidado de datos procesados en el sistemas anteriores. La propuesta de esta nueva metodología pretende aportar al cuidado de los datos históricos, ordenando su procesamiento y entregando herramientas que contribuyan a la posibilidad de éxito en el paso de datos de anteriores repositorios a nuevos contenedores.

La literatura formal no ha identificado un proceso consecuente para la migración de datos. Reconocidos académicos y notables implementadores de sistemas, no han establecido un acercamiento formal sobre este tema, mucho más cuando la cantidad de información, las exigencias del negocio y la tecnología, avanzan aceleradamente [1]. En [Sommerville 2002] o [Kendall & Kendall 2002], no se hace mención a la necesidad que tienen las empresas de cuidar sus datos al pasarlos a nuevos repositorios. En [Pressman 2008] hace una pequeña referencia y lo plantea como un tema “que pueden tener un profundo impacto sobre la ingeniería del software en los años por venir” [2].

Las mayores referencias hacia el proceso de migración de datos se las encuentra en documentos publicados en el internet y de propiedad de empresas desarrolladoras de software y casas fabricantes de bases de datos que pretenden normar sus procedimientos y entregar una imagen de tranquilidad y confianza a sus clientes.

Es necesario anotar que se han documentado varios aportes al proceso de migración de datos, uno de ellos recopilado del internet a manera de un curso formal que establece una muy buena rigurosidad de pasos a seguir [3]. Este curso establece una serie de factores críticos de éxito, el primero de los cuales indica que debería “establecerse una metodología de trabajo”. El orden que establece es apropiado, aportando claridad en el proceso a seguir, convirtiéndose en un primer acercamiento válido al proceso de migración de datos.

El presente trabajo consta de una sección que detalla los Trabajos Relacionados con el proceso de Migración de Datos. La segunda sección habla sobre la Migración de Datos como tal, para en una tercera sección detallar la Propuesta Metodológica. La cuarta sección valida esta propuesta a través de la aplicación en un caso de estudio y se adjuntan los Resultados Obtenidos. La última sección presenta

las Conclusiones y el Trabajo Futuro que debe realizarse.

II. TRABAJOS RELACIONADOS

La producción de trabajos científicos se ha centrado en las aportaciones de estudiantes de pre grado y post grado que han trabajado sobre este tema. Se ha identificado un trabajo de graduación que toma los datos de un repositorio que ha quedado obsoleto (FOX, 1996) y se lo ha pasado a tecnologías modernas (bases de datos relacionales). Es el caso del trabajo de graduación presentado por los señores Danny Fernando Puente Carrera y Sylvia Estefanía Ramírez Tapia, estudiantes de la facultad de Ingeniería de Sistemas de la Escuela Politécnica Nacional (Quito-Ecuador) [4] que en el año 2008 desarrollaron una plataforma intermedia que permitía la migración de datos de una manera "ágil y confiable".

III. MIGRACIÓN DE DATOS

El proceso de Migración de Datos es cada vez más utilizado y demandado, debido a la necesidad de tomar los datos de unos repositorios y trasladarlos a otros más eficientes, con mayores capacidades de almacenamiento, con mejores mecanismos de seguridad y provistos de mejores posibilidades de explotación de la *data* [5]. Debido a su magnitud e importancia empresarial estos procesos son tratados como proyectos paralelos a los propios de desarrollo. Algunas de las razones por las cuales se aborda este proceso son:

- Cambio de plataforma tecnológica.
- Cambio y actualización de aplicativos informáticos.
- Mejoramiento en tiempos de respuesta.
- Mejores Políticas de seguridad.
- Compatibilidad con otros aplicativos.
- Facilitar el intercambio de información.
- Optimización de ambientes de TI.
- Aplicación de nuevas reglas del negocio.
- Adaptabilidad a exigencias del mercado.

Las tareas que deben llevarse a cabo, no necesariamente son planificadas, habiéndose evidenciado varios problemas que han traído consigo ineficiencia en la utilización de recursos ya que se descuidan aspectos como: criticidad de los datos; rentabilidad y seguridad del proceso; minimización de costos y riesgos; automatización de tareas y agilidad en los procesos; integridad y consistencia de datos [6]. Por otro lado, el problema se agrava debido a que se cuenta con:

- Entornos heterogéneos y no estandarizados.
- Imposibilidad de interrumpir la continuidad del negocio.
- Repositorios de datos físicamente distantes.
- Hardware obsoleto o caduco.
- Incompatibilidad de formatos de datos.

Con el objetivo de minimizar estos riesgos, se sugiere tomar en cuenta las siguientes recomendaciones:

- Establecer una metodología de trabajo.
- Supervisar la calidad de los datos.
- Generar un entorno de prueba.
- Establecer la fuente y responsables de los datos.
- Obtener respaldos completos.
- Considerar el ingreso manual de datos, en donde la migración no pueda realizarse automáticamente.
- Seleccionar/desarrollar herramientas de migración.
- Considerar tiempos de desconexión.
- Costos de personal (horas extras, personal temporal, OUTSOURCING, etc.)

Es necesario establecer una metodología de trabajo que incremente la probabilidad de éxito en el proceso de migración de datos en el cual se adopta el establecimiento de pasos, tareas y actividades que aportarán a la consecución de los objetivos [7], cuyo orden se recomienda sea:

- Visión, Misión y Valores de la empresa.
- Establecimiento del equipo de trabajo.
- Levantamiento de la información.
- Análisis de criticidad y mapeo de datos.
- Generación de Ambiente de Pruebas.
- Establecimiento de planes de Prueba.
- Respaldo completo de los datos originales.
- Construcción ETL (Extract, Transform, Load).
- Ejecución de pruebas y análisis de resultados
- Cronograma definitivo de ejecución.
- Carga y procesamiento definitivo (iteraciones).
- Revisión de resultados.
- Entrega de documentación técnica.

IV. PROPUESTA METODOLÓGICA

El objetivo de la metodología propuesta es ordenar el proceso de migración de datos a través de fases, tareas e hitos que incrementen la probabilidad de éxito en momentos en que los sistemas serán reemplazados por nuevas alternativas. Se establece una Fase Inicial que comprende el análisis de la empresa y la funcionalidad/arquitectura del aplicativo actual [8]. Luego de pasar el primer Hito de control, se analiza la Data a Migrar, la que cuenta con el análisis del dominio, alcances y objetivos, los que quedan documentados en Historias de Usuario [9], para luego realizar el Modelado de la base de datos que servirá de repositorio temporal de datos. La propuesta termina con la aceptación, documentación y la aplicación del instrumento ERICd (Especificación de Reglas de Integridad sobre la Calidad de los Datos), completando el proceso de migración con la ejecución del segundo Hito de Control.

4.1. Visión general.

Un proceso migratorio requiere:

- Gestionar la configuración inicial de todo el ambiente de trabajo.

- Conocer el sistema actual, desde el punto de vista de: Expectativas futuras, Funcionalidad y Arquitectura.
- Justificación de migración de datos.
- Migración de la data.
- Evaluación de continuidad.

El eje transversal lo constituye un conjunto de acciones sistémicas que propician la calidad observando políticas institucionales y el cumplimiento de normativas gubernamentales, sin que ninguna de ellas transgreda el bien común. La figura 1 muestra el proceso en su conjunto.

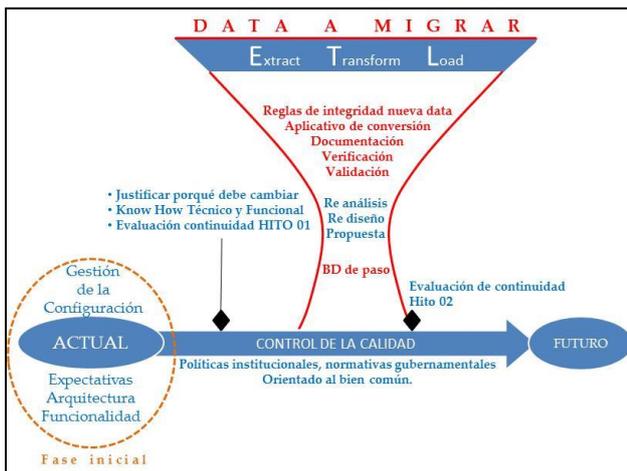


Figura 1: Metodología para la Migración de Datos

4.2. Fase Inicial.

Se analizan aspectos generales de la institución, su plataforma tecnológica y su realidad informática para iniciar el proceso gestionando la configuración del ambiente en el cual se procederá a la migración de datos. Es necesario conocer su razón de ser, su objetivo principal, su estructura financiera, sus accionistas/dueños y los valores en los cuales basa su funcionamiento. Se conocerán detalles como si posee o no un Plan Estratégico, la apertura hacia cambios tecnológicos e inversiones económicas necesarias. Se pretende determinar la realidad actual de la empresa desde el punto de vista legal, financiero y su actitud hacia el cumplimiento de controles y obligaciones gubernamentales. Debe llegarse a determinar la actitud de los involucrados con respecto a la utilización de la tecnología y sobre todo, llegar a determinar los riesgos que podrían entorpecer el desarrollo normal del proyecto informático.

Definir y controlar características técnicas, legales y humanas, propias de su organización, previo a iniciar el trabajo de análisis de la situación actual y previa a realizar la propuesta de solución. Requiere una atención aparte de toda aquella norma, procedimiento o política anterior, actual o futura que pueda afectar al proceso de repotenciación del nuevo aplicativo y su migración de datos. Se pretende controlar el ambiente de desarrollo a fin de iniciar adecuadamente los procesos de documentación, conocimiento, experiencias y

procesos que intervienen en la migración de datos. Existen varias actividades de orden documental que deben desarrollarse como: Definir el dominio del problema; determinar los Acrónimos y Abreviaturas; conformar el Equipo de trabajo y determinar el Jefe de proyecto; especificar recursos; carga del software inicial; contar con acceso a la Intranet; Socialización del proyecto; contar con las Normativas internas y externas y Planes Estratégicos; Generar un ambiente de pruebas.

Es necesario documentar las expectativas que tienen los usuarios finales, en base a las restricciones evidenciadas en el actual sistema, conjugando estas expectativas con las estrategias empresariales y tecnológicas de la Institución. Para ello se consideran las siguientes áreas:

- Reconocimiento de restricciones, limitaciones y problemas del aplicativo.
- Establecimiento de Planes Estratégicos Departamentales y Empresariales.
- Establecimiento de objetivos departamentales y empresariales.
- Documentación de expectativas.
- Acoplamiento de expectativas a Visión, Misión y Objetivos departamentales y empresariales.

Las expectativas con respecto a las áreas involucradas dependerán de cada una de ellas, enmarcadas en la consecución de los objetivos institucionales. Algunos ejemplos son:

- Control de todos los momentos del proceso de desarrollo del software.
- Verificar la continuación de un paso a otro, solo si el primero se ha completado adecuadamente.
- Garantizar la integridad de la información.
- Permitir la generación de reportes ágiles, rápidos y oportunos.
- Disponibilidad del aplicativo las 24 horas del día, los 7 días de la semana.
- Asegurar la manipulación de la información solamente de aquellas personas autorizadas para tal efecto.
- Permitir la interoperabilidad del sistema con otras herramientas de colaboración.
- Envío de reportes a medios en diferentes formatos de manera automática.
- Permita la consulta de información por varios campos de acceso.
- Controlar diferentes monedas, locales o dominantes del mercado.
- No haya pérdida de información histórica y que ésta esté disponible, en su totalidad y con integridad, en el nuevo aplicativo.

Al final de esta actividad, se generará un documento formal en el cual se describirán los trabajos realizados, recomendando la continuación o no del proyecto, siendo necesario un detalle de argumentos, en caso de que la decisión sea negativa.

4.2.1. Funcionalidad.

Se inicia con el estudio del sistema que actualmente está siendo utilizado, independientemente de que su estado y funcionalidad sea el óptimo y de que se adapte a los nuevos desafíos. Se trata de determinar, de una manera objetiva las características técnicas y funcionales del software actual. Es necesario conocer el ambiente institucional identificando las áreas que intervienen y determinando quienes son los beneficiarios directos e indirectos de su explotación. Sus partes constitutivas deberían ser:

- Entorno Empresarial.
- Estructura Organizacional
- Áreas involucradas.
- Procesos e Involucrados.
- Compromiso del Dpto. de Sistemas.
- Dominio (Lo que no es ni incluye).
- Glosario de términos y Reglas del Negocio
- Requerimientos funcionales y no funcionales.
- Restricciones, limitaciones, problemas.
- Informe.

4.2.2. Arquitectura.

Se estudia el sistema desde el punto de vista tecnológico, especificando los Sistemas Operativos, Lenguajes de Programación, Bases de Datos, Comunicaciones (redes de datos), Software (herramientas CASE), Metodologías de trabajo (CVDS) y Hardware (servidores, equipos clientes) del que se sirve. Se recomienda que éste sea un examen crítico y objetivo, sin que se emitan preconceptos ni se califique de buena o mala la tecnología utilizada. En caso de que no exista información documentada, ésta deberá ser generada. Es necesario se configuren los ambientes de prueba, tanto para probar el software actual como para emitir reportes e informes con el objeto de comparar los resultados con el nuevo aplicativo. Sus partes constitutivas son:

- Plataforma tecnológica.
- Servidores, clientes y sus Sistemas Operativos.
- Repositorios de almacenamiento (BD).
- Lenguajes de programación.
- Relación con otros sistemas e interfaces.
- Diagrama general del Sistema.
- Documentos asociados y Metodologías utilizadas.
- Informe.

Establecimiento y detalle de las inversiones monetarias que tendrán que ser realizadas a fin de concretar la propuesta ya que no necesariamente implica una reutilización de lo ya existente.

4.2.3. Primer hito de control

Resumir en un solo documento la necesidad de continuar con el proyecto o cancelarlo. Se pretende conocer la importancia que tienen los datos almacenados en el aplicativo actual, a fin de considerar o no la necesidad de conservarlos y

determinar hasta qué nivel de detalle se desearía hacerlo. De aquí en adelante todo el proceso, tanto de desarrollo de software como de migración de datos, tendrá un fuerte contenido de control de calidad a fin de garantizar que los productos de software entregados observen políticas institucionales, cumplan normativas gubernamentales y estén orientadas al bien común. Se establece, al término de esta fase, el primer documento formal sobre evaluación y conveniencia o no de continuar con el proyecto. Este se convertirá en el primer punto de control o Hito 01. La metodología planteada, sugiere preocuparse de cuatro aspectos:

- **Justificativo.** Importancia que tienen los datos a fin de conservarlos y llevarlos hacia la nueva plataforma.
- **Conocimiento.** Ratificación (usuarios y desarrolladores) sobre el conocimiento que tienen sobre el giro del negocio, sus particularidades, reglas, normas y expectativas.
- **Evaluación de continuidad.** Documento formal y producto de software que debe generarse como punto de control inicial (Hito 01).
- **Instrumento a aplicar.** El instrumento chequea que el conocimiento del grupo de desarrolladores sea el adecuado.

4.3. Data a Migrar a través de un sub aplicativo.

La información resultante como producto de la migración de datos puede ser continuamente revisada y refinada, permitiéndole al usuario validar los resultados esperados, comparándolos con las expectativas iniciales. Con el objeto de facilitar este proceso iterativo, esta propuesta sugiere se desarrolle una sub aplicación informática, la cual, recomienda se lo haga bajo la metodología XP (Extreme Programming) la que mantiene una gran flexibilidad y apertura sobre la aceptación de requerimientos recurrentes y poco estables que por lo general se manifiestan en estos procesos [10]. El proceso definido para la migración de datos se muestra en la figura 2.

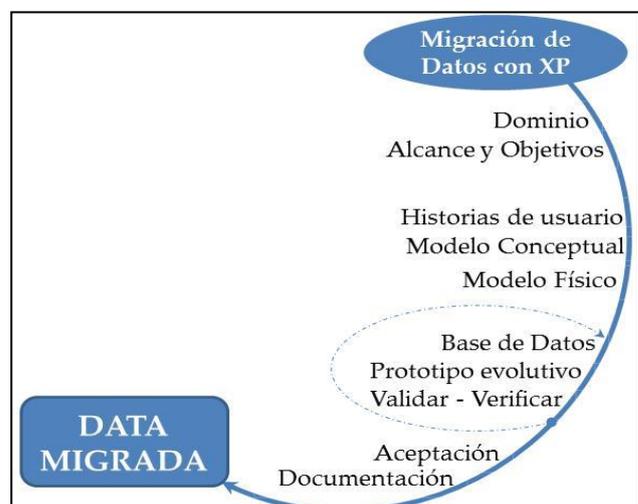


Figura 2: Migración de Datos con XP

4.3.1. Dominio

Acorde con los objetivos primordiales del proyecto y encaminado a aportar en su consecución, el dominio de la migración de datos, se refiere a entender la necesidad que tiene la institución de conservar los datos históricos. Aquí se justificará la realización de la migración de datos, en base a los requerimientos del usuario.

4.3.2. Alcance y objetivos

Especificar hasta dónde se va a proceder con la migración de datos y detallar cuáles son los objetivos a perseguir. Como ejemplo se podría mencionar los siguientes alcances:

- Se requiere conservar los datos desde un año determinado en adelante.
- Se requiere conservar los datos manejados por un departamento, respetando las regulaciones gubernamentales actuales.
- Se conservan los datos, solo de aquellos clientes que aún tengan registradas relaciones comerciales.
- Los datos a conservar se refieren exclusivamente a aquellos procesados en las sucursales que aún mantienen movimiento, no se requiere la información de sucursales cerradas.
- Los datos que importan a la institución son solo aquellos que cuentan con toda la información completa, cualquier registro que no tenga información completa no será trasladada al nuevo repositorio.

Como ejemplos de objetivos se podría mencionar:

- Rescatar toda aquella información que es vital para la institución.
- Garantizar que ni un solo dato procesado en el sistema actual se pierda.
- Mantener la oportunidad, disponibilidad y veracidad de los datos procesados desde que la empresa inició sus operaciones.
- Asegurar que los datos manejados por la institución serán cuidados con el mismo esmero que hasta la fecha se lo ha realizado.

4.3.3. Historia de usuarios

La necesidad de que los datos manejados por las instituciones sean cuidados, no nace como un requerimiento de los desarrolladores, sino como una iniciativa propia de quienes trabajan diariamente con ellos. Estos requerimientos quedarán plasmados en las *Historias de Usuarios*, mecanismo que utiliza la metodología XP. Los aspectos a considerar son:

- Auto numeradas.
- Fecha en la que se realiza la entrevista.
- Tipo de requerimiento (Nuevo, Corrección, Ampliación).
- Prioridad asignada por el usuario.
- Referencia, de existir, a otra Historia de Usuario.
- Indicador de atención del requerimiento.

- Descripción de las actividades y tareas realizadas.
- Notas relevantes que podrían impactar al proyecto.

4.3.4. Modelado.

Una de las actividades más importantes a ejecutar es modelar los requerimientos de usuario, obligando a los desarrolladores a conjugar las necesidades de paso de datos históricos (migración) con las necesidades del nuevo aplicativo. Uno de los factores críticos de éxito es compaginar la estructura de bases de datos para que soporte, de manera íntegra, oportuna y verás, los datos procesados en el anterior aplicativo y las nuevas estructuras. Se requiere de mucha flexibilidad e imaginación, para crear un modelo que logre armonizar, la existencia de unos datos (importantes para la empresa) y de otros formatos que satisfagan los nuevos retos [11]. Un fuerte control en la generación de documentos garantizará un proceso de digitalización eficiente, al conservar las diferentes versiones por las cuales ha pasado el modelo definitivo. Los modelos a desarrollarse son: Modelo Lógico o Conceptual y Modelo Físico. A partir del modelo Físico se generará el SCRIPT respectivo.

4.3.5. Base de Datos.

Generación automática de la estructura (tablas, índices, campos y restricciones) que servirán como repositorio de datos. Primer Producto de Software palpable, resultado del análisis, tanto sobre los requerimientos del nuevo sistema como de la necesidad de la información histórica. De acuerdo a la metodología propuesta, se inicia un proceso cíclico de varias iteraciones que persigue perfeccionar la estructura de base de datos [12]. Junto con las actividades de construcción (prototipo) y verificación, se irá estructurando de mejor manera el repositorio de datos. El usuario final, al constatar y evidenciar el resultado del proceso de migración de datos y al verificar la idoneidad de los datos históricos, tiene la posibilidad de rectificar, ampliar o cambiar sus criterios de selección de información, haciendo de este proceso, un proceso de perfeccionamiento continuo.

4.3.6. Prototipos.

A través de la creación de un prototipo evolutivo, se pretende mostrar el resultado sobre los datos a migrar. Es en este momento en el cual los procesos de Base de Datos, Prototipo y Validar-Verificar, se justifican plenamente al catalogarlos como iterativos, debido a que los cambios que el usuario final realice deben ser trasladados a la base de datos y a las nuevas reglas del negocio.

4.3.7. Validar y verificar.

El usuario final ratifica, cambia o niega, la idoneidad de los datos migrados, ratifica la integridad

de los mismos, comparando los resultados obtenidos con los objetivos planteados, basándose en evidencia que permita tomar decisiones objetivas. Cuantas veces sea necesario el criterio del usuario final, éste deberá ser tomado en cuenta, analizado, validado y consensuado. Las observaciones entregadas serán plasmadas en nuevas Historias de Usuario las que servirán de base para una nueva definición de requerimientos y su posterior implementación en una nueva iteración.

4.3.8. Aceptación.

Exponer formalmente que los datos pasados del sistema anterior al nuevo repositorio, se lo ha realizado de manera exitosa habiéndose respetado todas las reglas de validación y del negocio que los involucrados definieron en su momento. Este documento es generado por los desarrolladores del proyecto, ratificado por los usuarios finales y elevado a los principales directivos de la institución. Este instrumento de validación, formará parte de toda la documentación que tendrá que ser recopilada y guardada. Si los resultados son satisfactorios se continuará con el proceso, caso contrario, en plenaria se revisarán las observaciones y si es el caso, se procederá a corregir los errores encontrados.

4.3.9. Documentación.

Almacenamiento de todo aquel documento, sea generado por este proceso o recopilado, que se considera como importante a fin de:

- Respaldo de las decisiones tomadas.
- Servir como evidencia de los requerimientos establecidos.
- Control de trazabilidad en la atención de necesidades.
- Aportar al establecimiento de nuevos procesos o mejoramiento de los actuales.
- Aportar al mejoramiento de las reglas del negocio.

El objetivo es:

- Acceder de fácil manera a información crítica.
- Asegurar que la información confidencial está bien resguardada y cuidada su integridad.
- Garantizar su disponibilidad.

4.3.10. Instrumento a aplicar (ERICd)

Instrumento de evaluación completo conocido como ERICd (Especificación de Reglas de Integridad sobre la Calidad de los Datos) que trata de evidenciar cada uno de los datos que componen la *DATA* a migrar. El personal que ejecuta esta revisión debe contar con poder suficiente en la toma de decisiones ya que se podría suspender, dar por terminado o continuar el proceso de migración de datos. Se trata de responder a preguntas como:

- Qué datos deben ser migrados y qué datos no son indispensables.

- Qué reglas del negocio deben ser tomadas en cuenta.
- Cuáles son las reglas de validación de datos que deben aplicarse.
- Cuáles son las acciones que deberán tomarse en caso de que los datos a migrar presenten ambigüedades o contradicciones.
- Qué sucederá con los campos de las nuevas estructuras que antes no eran necesarios.

Existen tres parámetros de integridad que regirán la suspensión, continuidad o terminación del proceso de migración de datos, estos son:

- Pedir una nueva revisión de datos si el 5% de las preguntas realizadas fueron respondidas bajo la opción de "EN ESPERA".
- Cancelar el proceso de migración de datos, de manera definitiva, si acaso, una o más preguntas fueron contestadas con la opción de "NO SE SABE".
- Continuar con el proceso de migración de datos si más del 95% de preguntas fueron contestadas bajo la opción "COMPLETO".

4.3.11. Segundo hito de control

El último Punto de Control determina si, luego de que se ha evidenciado el paso de información, éstos cuentan con la integridad, veracidad y completitud requerida por los usuarios finales. Se determinará, bajo un criterio netamente objetivo, si el proyecto debe continuar o debe suspenderse o hasta darse por terminado, dependiendo del nivel de cumplimiento de los resultados, comparados éstos con las necesidades documentadas en las Historias de Usuario. El documento final es elevado a las máximas autoridades del proyecto y comunicado, por primera vez, a todos los involucrados, indicando la decisión final y las razones que justifican tal decisión. Luego de la presentación de este informe, y de la aceptación del mismo por parte de los Directivos del proyecto, el mismo podrá continuar con las siguientes fases. El último de los instrumentos a aplicar se refiere a una lista de chequeo que verifica que el proceso de migración de datos fue llevado a cabo satisfactoriamente y coincide con los trabajos realizados en el Hito No. 01. Las preguntas se refieren específicamente a aspectos técnicos como:

- Documentación de Historias de Usuario.
- Congruencia en la estructura de la Base de Datos.
- Evidencia de validación de datos.

Los resultados de este chequeo podrían sugerir la suspensión, temporal o definitiva del proceso de migración de datos, o confirmar su continuidad.

V. APLICACIÓN DE LA PROPUESTA

El Sistema de Importación de Vehículos (SIV) que maneja el Grupo Automotriz Quito Motors, fue el seleccionado para aplicar la metodología propuesta ya que cumplía con las características y exigencias del mercado que lo hacían un candidato óptimo debido a la gran cantidad de información que maneja

y a la importancia que tiene la misma para la operatividad de la empresa.

Todas las fases, actividades, tareas e Hitos, descritos en la propuesta, han sido aplicados en el proceso de desarrollo del nuevo aplicativo, generando diferentes productos de software los que van completando la documentación exigida. Los productos de software se explican en Tabla No. 1:

Fase	Contiene	Entregables
Gestión de la configuración	Primeras acciones, antes de iniciar la construcción del software	Especificación de estándares, almacenamientos y roles.
Estudio actual	Expectativas, Funcionalidad y Arquitectura	Tres documentos que describen la realidad actual del sistema a migrar.
Hito 01.	Justificativo, Know How y Evaluación de continuidad	Documento que sugiere la continuidad o no de los datos a migrar.
Data a migrar.	Dominio, Alcance y Objetivos, Historias de usuario, Modelado, Base de Datos, Prototipos, Validar y verificar, Aceptación y Documentación.	Modelado, BD, script y Programas fuente que viabilizan la migración de los datos.
Hito 02.	Evaluación de continuidad.	Documento que sugiere la continuidad o no de los datos a migrar.

Tabla 1: Productos de Software generados.

El arranque de la metodología establece la observancia de todo aquello que sea necesario considerar a fin de preparar el ambiente que se requiere, previa la iniciación de cualquier otra actividad. Los documentos de esta primera fase se describen en la Tabla No. 2:

Documento	Propósito
Cronograma de actividades.mpp	Representación gráfica de las diferentes actividades a realizar, identificando fechas, responsables y recursos.
MSIV_GC_Informe Final_Ver001.doc	Memorando interno sobre la confirmación de cambio del sistema a una nueva plataforma.
MSIV_GC_Sistema Actual_Componentes_Ver001.vsd	Diagrama de componentes del sistema actual, utilizando UML para su diagramación.
MSIV_GC_Sistema Actual_Despliegue_Ver001.vsd	Diagrama de despliegue del sistema actual, utilizando UML para su diagramación.
MSIV_GC_Ver001.doc	Gestionar el ambiente futuro de trabajo, así como conocer varias particularidades y singularidades propias del negocio, incluyendo normas, políticas y regulaciones gubernamentales, que los rigen.

Tabla 2: Documentos que intervienen.

Se requiere tener un conocimiento total del aplicativo a reemplazar por lo que los siguientes documentos, muestran su aplicación:

- Caso de Uso general del sistema actual.
- Descripción gráfica de todos los componentes.
- Memorando interno que explica la funcionalidad del sistema actual, tanto operativo como técnico.
- MER del sistema actual.

- Descripción detallada (IP, configuraciones, software, etc.) de los equipos servidores y clientes que requiere el sistema actual.
- Descripción técnica, identificando el hardware, lenguajes de programación, sistemas operativos, bases de datos, medios de almacenamiento, metodología de desarrollo, etc.
- Memorando interno final, de esta etapa, dirigido a los Directivos de la empresa en el que se resumen aspectos de funcionalidad y expectativas.
- Especificación y detalle de expectativas de los usuarios tanto de a nivel gerencial como operativo.
- Memorando interno que describe: entorno empresarial; dominio del problema, requerimientos funcionales; restricciones, limitaciones y problemas del sistema actual; Reglas del negocio.
- Documentación de pantallas.
- Determinar objetiva, clara, precisa y documentada, las características técnicas y funcionales del software actual, independientemente de que sea flexible y se adapte o no a los nuevos desafíos institucionales

Caso especial de este objetivo se refiere a los programas fuente con los cuales estuvo trabajando el sistema actual, los que representan 178 programas escritos en lenguaje FOX.

El primer control que sugiere la metodología se presenta en los siguientes formularios y documentos, que terminan por definir la continuidad o no del proceso de migración de datos y de la construcción del nuevo aplicativo.

- Documento: MSIV_Hito_01_Ver001.doc.
- Actualización: martes, 09 de noviembre de 2010.
- Propósito: Memorando interno. Control de Metodología No. 01. Descripción de actividades realizadas a la fecha y sugerencia de continuidad o no

A fin de ejecutar los procesos de migración de datos, se ha trabajado con los datos del sistema anterior y se han generado nuevos programas que permiten tomar esos datos, procesarlos (validarlos y/o completarlos) y dejarlos en el formato adecuado (Script de carga de datos), estos archivos son detallados en la Tabla No. 3:

Documento	Propósito
AEADI_MA, CLIENTTC, COMPANTC, FACVEHTP, IVMAESTR, IVTABLAS, MSIV_MDA, STKVEHTP	8 archivos tipo índices, creados bajo el formato del programa FOX
AEADI_MA, ARCPARTA, CLIENTTC, COMPANTC, FACVEHTP, IVMAESTR, IVTABLAS, MSIV_MDA, STKVEHTP	9 archivos tipo base de datos, creados bajo el formato del programa FOX
M.BAT	Archivo tipo ejecutable (lotes)
MSIV_TAB	Script de inserción registros
MSIV_MD0, MSIV_MDA, MSIV_MDB, MSIV_MDC, MSIV_MDD, MSIV_MDE, MSIV_MDM, MSIV_MDP, MSIV_TIP	9 archivos de código fuente a fin de procesar los datos viejos del Sistema de Importación de Vehículos en el formato adecuado bajo las exigencias de la nueva arquitectura

Tabla 3: Archivos sistema anterior.

La metodología plantea un segundo punto de control, luego de que se ha analizado la información a migrar y luego de que se ha procedido a generar, en tablas temporales, la nueva información, la que, a su vez, poblará los nuevos repositorios. En vista de que la nueva estructura de almacenamiento de datos deberá soportar, además de los datos a migrar, los nuevos requerimientos, ha sido necesario documentar la nueva solución en los cuales se muestra la recomendación o no de continuar con la construcción del software. En vista de que el proceso de migración de datos es repetitivo, es decir la carga de datos puede hacerse varias veces, hasta llegar a satisfacer adecuadamente al usuario final, ha sido necesario generar varios documentos que apoyen a este proceso, la Tabla No 4 muestra su detalle:

Documento	Propósito
00_UsuariosOpciones.xls	Generación instrucciones SQL para insertar registros de usuarios/opciones, así como sus respectivos permisos, que luego será importados al nuevo sistema para establecer el control de ingresos y roles de cada uno de ellos.
01_CargaInicialTablas.sql	Script que carga los primeros valores a las tablas del sistema, esto es Parámetros, usuarios y opciones.
02_BajarTamanoLog.sql	Instrucción SQL que permite bajar el tamaño de la nueva base de datos.
Backup.bat	Archivo de ejecución en lotes que permite obtener un respaldo de la BD.

Tabla 4: Documentos que automatizan procesos.

VI. RESULTADOS ALCANZADOS

La aplicación de la metodología ha llevado a la depuración y paso de datos históricos de una manera óptima, segura y confiable. Los resultados satisfacen las regulaciones empresariales, observando que la cantidad de data migrada es inversamente proporcional al número, complejidad y combinación de las reglas del negocio establecidas, es decir a más complejidad de reglas aplicadas, menos data histórica íntegra obtenida.

El caso de estudio da la posibilidad de analizar data histórica de varios años atrás, manejados en una moneda que ya no es la oficial (Sucre). Se ha depurado una gran cantidad de información y llenado otra gran porción de data anterior, en vista de que las reglas del negocio actuales, establecen nuevas regulaciones y características que antes no existía, observándose la aplicabilidad y pertinencia de la metodología al considerar estas variables. Si bien la data a migrar en términos de número de registros no es extremadamente grande, la aplicación de la metodología refleja que las regulaciones, hitos y validaciones establecidas, incrementan la probabilidad de éxito al desear contar con data íntegra.

La plataforma tecnológica en la cual reposó la data, siendo única, coadyuvó a que se disminuya la incertidumbre de la heterogeneidad tecnológica (lenguajes de programación, bases de datos, sistemas operativos) factores que hubiesen incrementado la complejidad del trabajo a realizar.

Entre los datos analizados se puede citar las siguientes fuentes: Clientes; Facturas de vehículos; Maestro de vehículos; Stock de vehículos. La data más antigua a analizarse proviene del 31 de marzo de 1991.

El establecimiento de las *Reglas del Negocio* cae en el campo de Directivos, Gerentes y dueños del negocio, implicando largos períodos de tiempo de trabajo en equipo, con resultados que justificaron el tiempo invertido. Al momento de procesar la información se consideraron las siguientes reglas del negocio:

- Fuente de datos (Suces o Dólares).
- Errores graves: Duplicación en número de Chasis y VIN, o inexistencia de *Tipo de vehículo*.
- Errores leves: Inexistencia de Numero de Ingreso; Fecha de Ingreso; Fecha de Recepción, etc.
- Datos completados (SUCRES): Origen del vehículo; tipo de orden; Mes de producción; Estado de fabricación; tipo de carrocería; clase del vehículo; Cantón de Matriculación; Tipo de teléfono del cliente; etc.

Una vez que este proceso se cumplió, la data nueva quedó lista para ser migrada, en este momento es en donde intervienen por segunda vez, Directivos, Gerentes o dueños del negocio con el objeto de especificar, que data será pasada al nuevo repositorio, las cuatro reglas aplicadas (solo como ejemplo) fueron:

- Pasar aquella data que no tenga ni un solo error.
- Pasar datos con información completa de clientes.
- Datos cuyo año de facturación sea superior 1999.
- Datos con precio de venta, mayor a \$50.000.

Las diferentes variantes elegidas y validadas por los desarrolladores fueron corridos secuencialmente, de tal forma que en cada uno de ellos, los principales de la empresa, pudieron analizar los resultados obtenidos y orientar los mismos a los intereses institucionales. Como la metodología lo establece, la corrida de estas pruebas puede ser realizada tantas veces como sea necesario, hasta llegar a un punto base, en el cual se sienten satisfechos con los resultados obtenidos.

En total se procesaron 143.429 registros, entre los datos de Clientes, Facturas y Vehículos, de los cuales 36.772 (26%) fueron procesados (aun cuando tenían errores graves o leves). 13.536 (37%) del total procesados, fueron tomados de los datos en Suces y 23.236 (63%) del total procesados, fueron tomados de los datos en dólares, como lo muestra la tabla No. 6:

Archivo	# registros
Vehículos repositorio AEADE	19,192
Clientes	71,777
Facturas	15,688
Maestro de vehículos importados	23,236
Stock de vehículos	13,536
Total Registros procesados	143,429

Registros validados: 36,772	
Sucres	Dólares
13,536	23,236
37%	63%

Tabla 6: Estadística datos procesados.

La validación de la data histórica depende de los criterios que apliquen los principales de la empresa. De los ejercicios ejecutados, en ambientes controlados de prueba, los resultados se muestran en la tabla No. 7:

Registros con errores graves	1,327	13,111
Registros con errores leves	11,784	
Registros limpios (sin errores)	23,661	
Registros limpios de clientes		
Dirección	18,077	
Dirección + teléfono	14,537	
Dirección + teléfono + número	5,191	
Año facturación > al 1999	9,993	
Precio de venta > a \$50.000	257	

Tabla 7: Datos validados.

El cuidado de la información histórica se ha evidenciado a cabalidad al aplicar la metodología propuesta. Del universo de registros validados (36.772), el 36% de ellos (13.111 registros) contienen algún tipo de error, mientras que el 64% restante (23.611 registros) cumplieron con las reglas de validación iniciales propuestas, como lo muestra el siguiente gráfico:

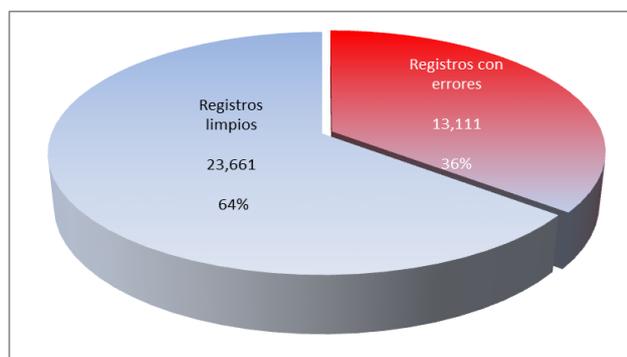


Figura 3: Datos limpios y con errores

De aquí en adelante, la decisión de qué data tomar y bajo que nuevos parámetros de validación, queda a criterio único de los directivos, gerentes o dueños de la institución, y dueños a la vez, de su información, notando además que mientras más rigurosas son esas reglas de validación, menos datos son seleccionados como idóneos.

VII. CONCLUSIONES Y TRABAJO FUTURO

Los Métodos, Técnicas y Herramientas utilizadas a fin de cuidar la información histórica, han sido escasas convirtiendo a su tratamiento en una

complicación informática que nadie quiere asumir debido a lo complicado de su tratamiento, sugiriendo sustentarse en la utilización de metodologías ágiles que alivianen el trabajo a realizar. Las sugerencias descritas en esta metodología han sido consideradas sobre una base cierta, producto de un diagnóstico completo de los sistemas informáticos analizados en el Grupo Automotriz Quito Motors, pretendiendo contribuir en algo al dilema de ¿qué hacer con la información histórica?

Queda por cerrar el ciclo de vida del desarrollo del software, incluso hasta la entrega del producto final, plasmado en nuevo sistema informático que hago uso de las mejores tecnologías y brinde las mejores soluciones, sin olvidar los datos históricos, fundamentales para la toma de decisiones.

VIII. BIBLIOGRAFÍA

- [1] Coleman K., Larsen P., Shaw P., "Software Process Improvement in Small Organizations: A Case Study", IEEE Software, 2005.
- [2] Pressman, Roger. (2008). Ingeniería del software, un enfoque práctico (4ta ed.). Editorial McGrawGill. España.
- [3] Herrera, A., "Migración de Datos", recuperado de <http://www.emagister.com/curso-migracion-datos> en Agosto 2013.
- [4] Puente, D. Ramírez, S. "Tesis de Grado ingeniería en Informática", recuperado de <http://bibdigital.epn.edu.ec/handle/15000/894> en Septiembre 2013.
- [5] Elmasri, R.; Navathe, S.B. Fundamentos de sistemas de bases de datos. 3ª ed. Madrid, Pearson Educación, 2002.
- [6] IEEE Std. 1044-1993, IEEE Standard Classification for Software Anomalies, 1993.
- [7] Ferré, Xavier: Principios Básicos de Usabilidad para Ingenieros Software, Universidad Politécnica de Madrid, 2005.
- [8] Kenneth, Kendall & Kendall, Julie. (2002). Análisis y Diseño de Sistemas (3ra ed.). Editorial Prentice-Hall. USA.
- [9] Sommerville, Ian. (2001). Ingeniería del software (6ta ed.) Editorial Pearson. España.
- [10] Jacobson, Ivar; Booch, Grady; Rumbaugh, James., El Proceso Unificado de Desarrollo de Software.
- [11] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy y W. Lorensen. Object-oriented modeling and design. Prentice- Hall.
- [12] González C., Sistemas de Bases de Datos, Tercera reimpresión, Editorial Tecnológica, Cartago, Costa Rica, 2002.