



Modelo de conciliación de atributos de identificación Bancolombia valores.

Daniel Arango Obando

Informe de práctica presentado para optar al título de Ingeniero de Sistemas

Asesor

Jaime Fonseca Espinal

Universidad de Antioquia
Facultad de Ingeniería
Ingeniería de Sistemas
Medellín, Antioquia, Colombia
2023

Cita

D. Arango Obando [1]

Referencia [1] D. Arango Obando, “Modelo de conciliación de atributos de identificación Bancolombia Valores, 2023”, Trabajo de grado profesional, Ingeniería de Sistemas, Universidad de Antioquia, Medellín, Antioquia, Colombia, 2023.
Estilo IEEE (2020)



Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

TABLA DE CONTENIDO

1. RESUMEN	5
2. INTRODUCCIÓN	6
3. OBJETIVOS	7
2.1. Objetivo General	7
2.2. Objetivos Específicos	7
IV. Marco Teórico	8
V. METODOLOGIA	9
5.1. Diseño y propuesta del proyecto:	9
5.2. Construcción	9
5.3. Implementación	10
5.4. Pruebas	10
5.5. Documentación	10
6. RESULTADOS Y ANALISIS	11
6.1. Diseño y propuesta del proyecto.	11
6.2. Construcción	11
6.3 Implementación	14
6.4 Salidas del proceso:	15
6.5. Pruebas y documentación	16
6.6. Documentación	18
6.6.1 Valoración de criticidad	18
7.RESULTADO FINAL	18
8. DIFICULTADES	19
9. CONCLUSIONES	20

1. RESUMEN

La sección servicios calidad de datos perteneciente a la Vicepresidencia de servicios para los clientes propone realizar un modelo de conciliación entre las fuentes de clientes Valores y clientes Bancolombia. El termino clientes Valores hace referencia a aquellos clientes que tienen acciones. El modelo de conciliación tiene como objetivo principal buscar los cambios más recientes en las fuentes de información para realizar la conciliación entre ambas.

La principal fuente de información para obtener los datos actualizados fue clientes Valores, el motivo es que la base de datos tiene un tratamiento mucho más riguroso debido a que contiene la información de clientes de alto valor para Grupo Bancolombia.

Para el desarrollo de esta propuesta, fue necesario investigar sobre SAS Enterprise Guide que fue la herramienta base utilizada para lograr una buena comunicación entre las consultas SQL y los servicios cloud de Grupo Bancolombia, también se estudió sobre la conexión ODBC de Python con servicios cloud con el fin de lograr la calendarización del proceso, sin embargo, todo el despliegue del proyecto se realiza con herramientas SAS.

Se definió una metodología para llevar a cabo cada uno de los pasos del proyecto, los resultados fueron muy positivos, principalmente destacando que el proyecto fue calendarizado dentro de la malla de servicios Bancolombia.

Palabras clave — Client values, SQL, Scheduling, services

ABSTRACT

The data quality services section belonging to the Vice-Presidency of client services proposes a reconciliation model between the sources of securities clients and Bancolombia clients. The term "securities clients" refers to those clients who hold shares. The main objective of the reconciliation model is to look for the most recent changes in the information sources in order to perform the reconciliation between both.

The main source of information to obtain the updated data was Valores clients, the reason is that the database has a much more rigorous treatment because it contains the information of high value clients for Grupo Bancolombia.

For the development of this proposal, it was necessary to research on SAS Enterprise Guide which was the base tool used to achieve a good communication between SQL queries and cloud services of Grupo Bancolombia, also was studied on the ODBC connection of Python with cloud services in order to achieve the scheduling of the process, however, the entire deployment of the project is done with SAS tools.

A methodology was defined to carry out each of the steps of the project, the results were very positive, mainly highlighting that the project was scheduled within the Bancolombia services mesh.

Keywords — **Scientific article, review article, research, citation styles.**

2. INTRODUCCIÓN

El Grupo Bancolombia actualmente cuenta con una gran cantidad de datos que por motivos históricos se han almacenado en diferentes fuentes de información, el Grupo Bancolombia utiliza un sistema de bases de datos relacionales con distintos modelos de información y tablas.

El objetivo principal fue identificar los cambios más recientes en las fuentes maestras de información sobre los atributos de identificación de clientes comunes y generar un modelo analítico o solución en lenguaje SQL que permita generar salidas de información, para que de forma transitoria y a nivel de proceso, se homologue la información en los sistemas. Esta propuesta surge debido a que actualmente los sistemas maestros de información no cuentan con un servicio de sincronización entre ellos.

Un ejemplo para ilustrar la situación es tener un cliente común en ambas fuentes, su nombre y documento coinciden, sin embargo, el cliente cambio de tarjeta de identidad a cédula de ciudadanía en la base de datos Bancolombia, sin embargo, dicha información no sincroniza con la base datos de Valores Bancolombia que hace referencia a los clientes que tienen acciones en Grupo Bancolombia.

En la fase final el servicio debe estar en la capacidad de realizar extracciones de las fuentes de información, comparar los datos en búsqueda de clientes comunes entre las fuentes, luego identificar cual registros cuenta con la fecha de actualización más reciente, el registro seleccionado es tomado como base para proceder a conciliar la información en las Fuentes.

3. OBJETIVOS

2.1. Objetivo General

- Identificar los cambios más recientes entre las fuentes de información de Clientes Bancolombia y Valores Bancolombia, para generar acciones de conciliación.

2.2. Objetivos Específicos

- La actualización de la información debe suceder de manera bidireccional, por ende, si un registro es actualizado en Clientes Bancolombia de manera recíproca debe ocurrir en Valores Bancolombia.
- El servicio debe ser eficiente en términos de consumo de memoria de las herramientas analíticas, debido a que el volumen de datos a manejar es de un tamaño considerable y el servicio se ejecutará de manera diaria.
- El producto final debe ser certificado por las políticas de herramientas de usuario final Bancolombia, identificadas como EUC por sus siglas en inglés (End User Computing).
- El servicio debe generar estadísticas donde se vea reflejado los tiempos de ejecución, fallos en conexiones, inactividad de la fuente maestra e indicar registros que deben ser tratados de forma manual debido a inconsistencias.
- Establecer y aplicar reglas de calidad con el fin de asegurar que el servicio entrega a las Fuentes de información el dato más actualizado y sin inconsistencias.

4. MARCO TEORICO

La minería de datos es el descubrimiento de estructuras y patrones en conjuntos de datos grandes y complejos. Hay dos aspectos de la minería de datos: creación de modelos y detección de patrones [5].

Entre los trabajos relacionados que utilizan la minería de datos se encuentran, “Data mining with an ant colony optimization algorithm”. [6] en el documento se propone un algoritmo para la extracción de datos llamado Ant-Miner (minería de datos basado en colonias de hormigas). El objetivo de este método es extraer las reglas de clasificación de los datos.

SQL: En el funcionamiento día a día de una empresa es común escuchar sobre este concepto como base para el manejo de fuentes de información, el cual durante el desarrollo del proyecto es evidente debido a las numerosas consultas que se deben realizar para extraer información.

La eficiencia es un objetivo clave para el desarrollo del proyecto, el artículo “Performance Comparison of Hive, Impala and Spark SQL” [1], muestra una comparativa de rendimiento entre tres servicios SQL, además de exponer formas de trabajo para grandes volúmenes de información como es el caso de este proyecto.

Continuando con el problema a abordar encontramos en el artículo “Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies” [2], una detallada comparación entre distintas fuentes de información y la forma de abordar cada una con diferentes tipos de formatos, lenguajes y validación de resultados.

Algoritmos de comparación: La funcionalidad principal del proyecto es comparar información entre las fuentes de información mencionadas, el escenario ideal es utilizar algoritmos de comparación, se manejan varias opciones una de ellas es basada en el artículo “CANDID: comparison algorithm for navigating digital image databases” [3], la propuesta de los autores del artículo es la creación de un algoritmo de comparación de imágenes para calcular su similitud.

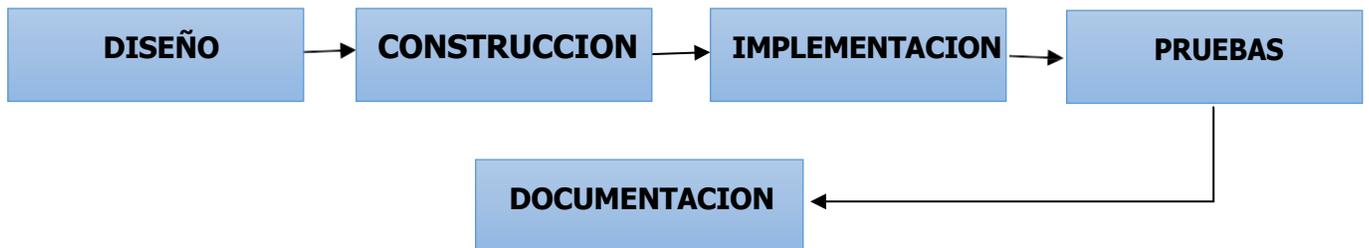
En el artículo “Design of a table-shaped comparison algorithm to improve the transformation management process in the metadata management system” [4], hacen mención a otra propuesta de algoritmo de comparación para mejorar un proceso de transformación en el sistema de gestión de metadatos.

principal herramienta utilizada a lo largo del desarrollo del proyecto fue SAS Enterprise Guide [6], que es un programa que permite a los usuarios analizar datos y publicar sus resultados. Proporciona un aprendizaje rápido para la investigación de datos, generando códigos con un índice de productividad muy alto. Lo cual acelera el despliegue de análisis de previsiones de datos.

SAS Enterprise destaca por ser una de las herramientas principales en Grupo Bancolombia y cumple con los conceptos planteados a lo largo del proyecto en el que se busca realizar minería de datos, programación y consultas SQL.

5. METODOLOGIA

El desarrollo de los objetivos anteriores se llevó a cabo de la siguiente manera:



5.1. Diseño y propuesta del proyecto:

- a) Adquirir conocimientos sobre la base de datos clientes Valores Bancolombia.
- b) Diseñar una propuesta que contenga un dashboard con los releases del proyecto el diagrama de actividades y la arquitectura de diseño.
- c) Refinar el diseño según las recomendaciones de la gerencia encargada de la sección.

5.2. Construcción:

- a) Establecer la conexión con los servicios cloud de Grupo Bancolombia utilizando SAS Enterprise Guide.
- b) Diseñar un demo en entorno SAS sin realizar despliegues hacia el cloud de Bancolombia.
- c) Crear controles para el tratamiento de las fuentes de información.
- d) Realizar las extracciones necesarias desde las fuentes de información.
- e) Homologar información entre las fuentes para que sea posible realizar las conciliaciones
- f) Tratar de forma individual cada uno los campos de identificación.
- g) Determinar diferencias entre las fuentes de información.
- h) Definir criterios de selección para determinar cual fuente contiene la mejor versión del dato.
- i) Establecer los datos de salida con los cambios propuestos.
- j) Asegurar la eficiencia y calidad de cada una de las consultas del servicio.

- k) Realizar pruebas de monitoreo a las salidas del servicio.
- l) Dar un diagnóstico riguroso a los campos considerados como críticos, por ejemplo: Tipo de documento de identificación.

5.3. Implementación:

- a) Establecer reglas de calidad para las salidas del servicio.
- b) Asegurar el cumplimiento de las reglas de calidad.
- c) Conciliar con el encargado del servicio de actualización el paso a producción del proyecto.
- d) Homologar la salida del proceso con la entrada del servicio de actualización.
- e) Crear un dashboard que controle las estadísticas diarias del proceso.
- f) Crear un log que indique los cambios puntuales a realizar en las fuentes de información.
- g) Exportar cada una de las salidas del proceso a la ruta pública de la Vicepresidencia. Las salidas son (Homologado para el proceso de actualización, Log de cambios y Dashboard de estadísticas).
- h) Calendarizar el proceso, para llevar a cabo este paso es necesario cumplir previamente con la fase de pruebas.

5.4. Pruebas

- a) Validar la veracidad de los cambios propuestos por el modelo de conciliación.
- b) Monitoreo diario de los resultados arrojados por el proceso de actualización.
- c) Prueba de calendarización.
- d) Verificar que el proceso funciona de forma automática.
- e) Validar tiempos de ejecución.

5.5. Documentación

- a) Realizar valoración de criticidad establecida por Grupo Bancolombia.
- b) Documentar el código con las normativas y formatos establecidos por las herramientas de usuario final (EUC).
- c) Documentar el proceso en la base de datos que contiene las EUC de Grupo Bancolombia.
- d) Realizar videos explicando el funcionamiento técnico del proceso y la gestión manual.

6. RESULTADOS Y ANALISIS

Abordaremos la fase de resultados desde cada uno de los puntos propuestos en la metodología

6.1. Diseño y propuesta del proyecto.

El diseño del proyecto se divide en dos fases:

- Diagrama de actividades: Flujo que indica las actividades a cumplir según las fases del proyecto. El diagrama comprende desde el inicio del proyecto hasta que pasa a ser parte el proceso de actualización.
- Backlog: Se definieron los releseaes a entregar según los tiempos establecidos. Fueron divididos como:

R1: Conciliación de clientes Valores a clientes Bancolombia.

R2: Conciliación de cliente Bancolombia a clientes Valores.

R3: Pruebas y documentación.

Los resultados de esta fase tuvieron muy buena repercusión por parte de la gerencia encargada de la sección. Como recomendación por parte de la gerencia se sugirió incluir dentro del proyecto la realización del Log de cambios.

El log de cambios es un proyecto realizado por otros miembros de la gerencia con el objetivo de llevar un control de todos los cambios realizados en las fuentes de información.

6.2. Construcción

Como se menciona a lo largo del informe, la herramienta SAS Enterprise se utiliza para realizar la minería de datos y despliegues hacia los servicios cloud de Grupo Bancolombia.

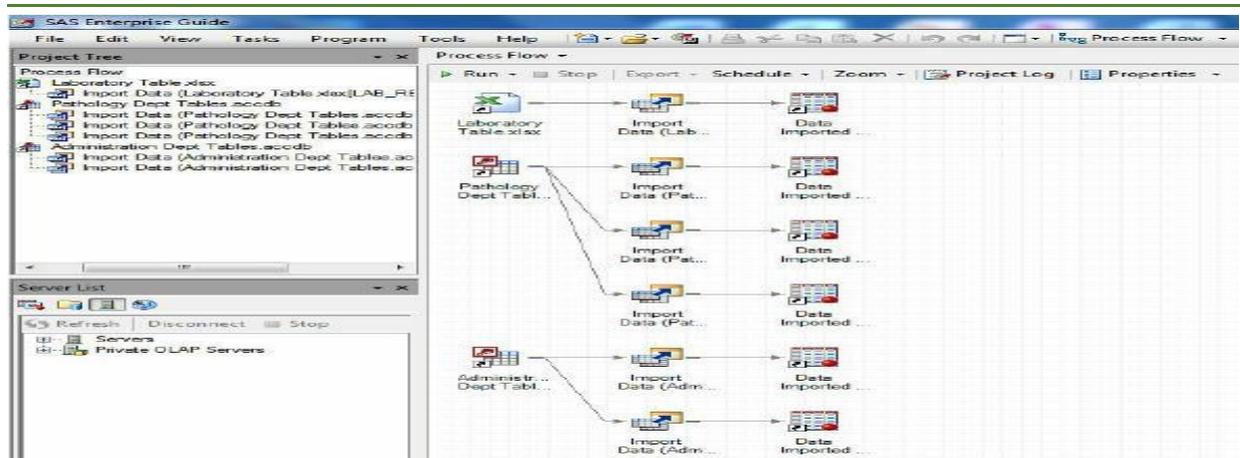


Fig. 1: Entorno SAS Enterprise Guide.

La figura 1 muestra de forma general como es un entorno SAS, las funcionalidades de la herramienta ofrecen muchas posibilidades para el desarrollo de los proyectos. Una ventaja a destacar es que el despliegue se realiza por nodos, lo cual hace posible monitorear cada consulta de una forma eficiente.

Grupo Bancolombia utiliza esta herramienta para establecer conexión con la Nube donde se encuentran todas las fuentes de información. Para que fuese posible el despliegue de información, la creación de tablas, extracción de información entre otras funcionalidades, fue necesario crear dos nodos encargados de establecer la conexión con las fuentes de información y métodos que permitieran crear tablas, obtener tablas existentes y modificar tablas. El objetivo de los métodos es mejorar la eficiencia del proceso ahorrando pasos en las consultas realizadas a lo largo del desarrollo del proyecto.

Luego de establecer las conexiones, hay que establecer controles que garanticen que la ejecución del flujo tiene altas probabilidades de ser exitosas, algunos controles incluidos fueron:

- Validar que las fuentes de información Clientes Valores y Clientes Bancolombia realizaron ingestión en la fecha actual.
- Verificar las últimas fechas de ingestión de cada fuente de información.
- Control para determinar que las fechas de ingestión fueran iguales en ambas fuentes de información.
- Validar que las fechas de ingestión fueran coherentes, por ejemplo: **20201208! =0**.

Si algún control falla entonces se envía la orden de abortar la ejecución.

Luego de que la ejecución apruebe los controles establecidos, el proceso continuo con la extracción de información de cada una de las fuentes, para realizar la extracción se utilizó la fecha de ingestión más actual en cada una de las fuentes. Los resultados obtenidos por la extracción también es necesario aplicarle controles para determinar:

- Controlar las fechas de modificación en las fuentes de información. Es importante establecer este control porque es posible que algunos días las modificaciones más recientes no fueran una cantidad suficiente para justificar la ejecución del proceso.
- Homologación de campos con el objetivo de establecer igual nomenclatura en cada una las fuentes. Un ejemplo de esto es que en clientes Valores el tipo de documento de identificación se maneja de forma descriptiva y en clientes Bancolombia se maneja de forma numérica por tanto fue necesario homologar la información para que no se presentaran inconvenientes al momento de realizar comparaciones.

El siguiente paso fue buscar las desconciliaciones entre las fuentes de información, para comparar las fuentes de información se realizan varios cruces:

- **Por tipo y número de documento:** Cruce utilizado para determinar las desconciliaciones en los campos nombre completo, primer nombre, segundo nombre, primer apellido, segundo apellido, fecha de nacimiento y fecha de expedición.
- **Por número de documentos:** Utilizado para determinar las desconciliaciones en el tipo documento.
- **Por tipo de documento, número de documento y nombre:** Comparación solamente con el objetivo de obtener datos estadísticos sobre la conciliación de los nombres entre las fuentes de información.

Criterios utilizados para determinar si el mejor dato se encuentra Clientes Valores o Clientes Bancolombia:

- Fecha de modificación más actualizada.
- En caso de la fecha de modificación coincida el dato a tomar será el de Clientes Valores, esta decisión se toma porque esta base de datos tiene un tratamiento mucho más riguroso.

Criterios para tratar las desconciliaciones por tipo documento:

- Validar la similitud de los nombres utilizando una función de comparaciones que indica de forma numérica si un nombre es similar a otro.
- Validar que la fecha de nacimiento sea igual en ambas fuentes de información.

Si estos criterios se cumplen entonces se toma en cuenta el cambio propuesta por el proceso en el campo tipo de documento.

Luego de determinar los cambios en cada uno los campos, el proceso genera un primer consolidado con la nueva información de cada uno de los clientes. Con este paso queda finalizado el proceso de construcción del proyecto.

Los resultados de esta fase fueron muy interesantes, los tiempos de ejecución eran alrededor de 7 minutos lo cual es una excelente cifra teniendo en cuenta que las fuentes de información son muy grandes.

6.3 Implementación

Luego de la fase de construcción, continuamos con la fase de implementación que consiste dar tratamiento a la salida de la fase anterior. Para llevar esto a cabo iniciamos con la creación de nodos que se encarguen de asegurar la calidad de la información.

Algunas reglas de calidad definidas:

- Para los campos nombres se valida que la información contenida sea válida, por ejemplo, verificar que la actualización no sea un campo vacío, no contenga palabras reservadas por Grupo Bancolombia y que la actualización sea diferente al valor anterior.
- En Grupo Bancolombia por ahora solamente se manejan los tipos de género como Masculino y Femenino. Por tanto, se valida que la actualización a realizar corresponda exclusivamente a alguno de los géneros mencionados.
- El caso de las fechas, Grupo Bancolombia tiene definidas fechas por defecto para validar la calidad de la información. Se verifica que la fecha a actualizar no pertenezca a los Valores por defecto, validar coherencia de las fechas por ejemplo que la fecha de nacimiento no sea mayor a la fecha de expedición.
- El tipo de documento ya cuenta con una validación previa en la fase de construcción que asegura la calidad de la información a actualizar.

Una vez aseguradas el cumplimiento de las reglas de calidad previamente definidas por la sección servicios calidad de datos, se procede a crear el archivo de homologación, fue necesario conciliar con la persona encargada del proceso de actualización el paso a producción de las salidas del proceso.

Proceso de actualización: Proceso ejecutado diariamente por la sección servicios calidad de datos. Dicho proceso se encarga de llevar las actualizaciones realizadas por todos los procesos correspondientes a actualización manejados por cada uno de los miembros de la sección. Su funcionamiento consiste en una serie de Macros diseñadas en Visual Basic Application para Microsoft Excel que se conectan a los aplicativos encargados de manejar las fuentes de información de Grupo Bancolombia.

Luego de establecer los acuerdos con el proceso de actualización, se homologa la salida del modelo de conciliación con la entrada requerida por el proceso mencionado, esto se debe a que el funcionamiento de las macros funciona con unos insumos previamente definidos por ende es necesario acomodar la información según la estructura de los insumos. Una vez cumplida esta tarea tendremos la primera salida del proceso que es el Homologado de actualización.

Luego de que el homologado de actualización este funcionando correctamente, se crea el dashboard con las estadísticas del proceso. Esta funcionalidad también será una de las salidas del proceso el objetivo es controlar las cifras diarias del proceso para identificar las oportunidades de mejora propuestas por el proceso es importante destacar que el proceso se encarga de diligenciar el proceso de forma automática.

Fecha de ingestión Clientes Valores	20201208
Fecha de ingestión Clientes Bancolombia	20201208
Numero de registros Clientes Valores	500000
Numero de registros Clientes Bancolombia	2400000
Numero de coincidencias entre fuentes	183
Diferencias encontradas por cliente	54
Diferencias encontradas	76
Numero de registros cargados por el proceso de actualización	74

Fig. 2: Estructura del tablero de control

Los Valores de la figura 2 son ficticios y no corresponden a las cifras reales del tablero de control. Sin embargo, es una estimación cercana a los resultados reales de cada una de las ejecuciones.

Otra salida del proceso es el Log de cambios, su funcionalidad es identificar las diferencias encontradas de forma individual. Otra sección de la gerencia trabajaba en un proyecto que tenía como finalidad realizar modificaciones al proceso de actualización para identificar todos los cambios realizados por todos los procesos que se ejecutan diariamente en la gerencia interacción con clientes, sin embargo, el proceso aún no se encontraba terminado, de igual forma se crea el log de cambios pensando en que a futuro será necesario como insumo del proceso mencionado.

6.4 Salidas del proceso:

- Homologado final: insumo para el proceso de actualización.
- Tablero de control: Estadísticas de ejecución.
- Log de cambios: Insumo para la modificación del proceso de actualización.

Cada una de las salidas debe exportarse a la ruta pública de la vicepresidencia servicios para los clientes, es necesario realizar este proceso porque luego de terminar con la práctica académica otro miembro del equipo será el responsable de la gestión del proceso.

La fase de implementación concluye con la calendarización del proceso, el objetivo es que el modelo de conciliación se ejecute de forma automática a una hora establecida con una periodicidad diaria. La idea inicial fue ejecutar la calendarización con Python sin embargo el banco utiliza la herramienta **SAS Management Console** para calendarizar los procesos dentro de la malla de servicios Bancolombia.

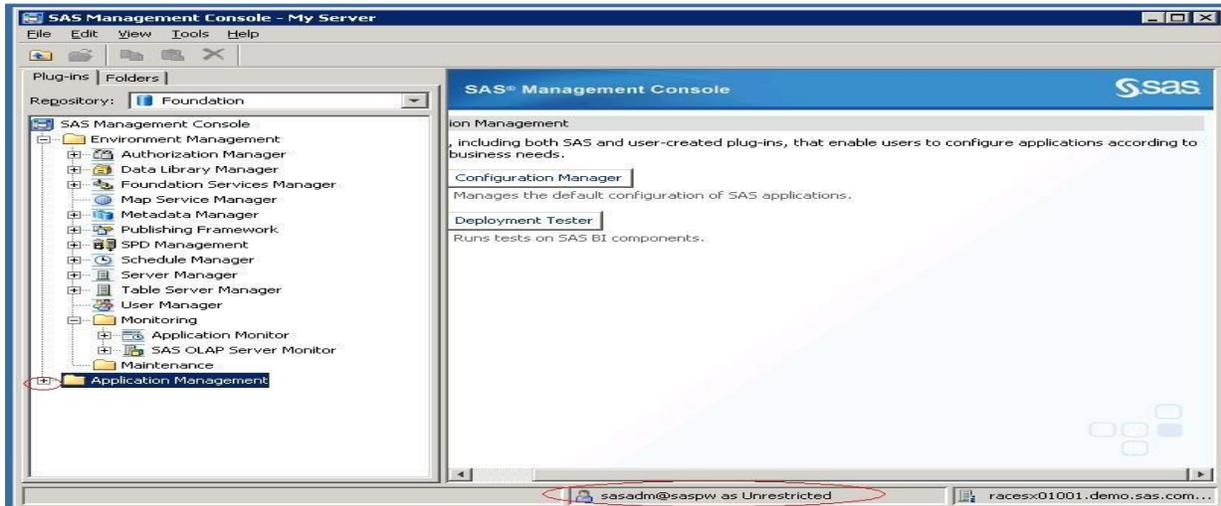


Fig. 3: SAS Management Consolé

Grupo Bancolombia tiene definidos buenas prácticas de codificación para que los procesos sean calendarizados, los resultados de esta fase fueron muy positivos lo que asegura la eficiencia de cada una de las consultas del modelo de conciliación. El tiempo de ejecución final oscila entre 20 y 30 minutos según la disponibilidad de los servicios cloud Bancolombia.

6.5. Pruebas y documentación

Durante la fase de pruebas se realizaron varias actividades como:

- Validar los cambios propuestos por el modelo de conciliación, esta prueba consiste en tomar una cantidad de registros manejable para verificar directamente en las fuentes de información que el cambio a realizar fuera correcto. Los aplicativos para validar la información son: SAP CRM y Oracle web, el aplicativo de clientes Valores es información exclusiva de Grupo Bancolombia.

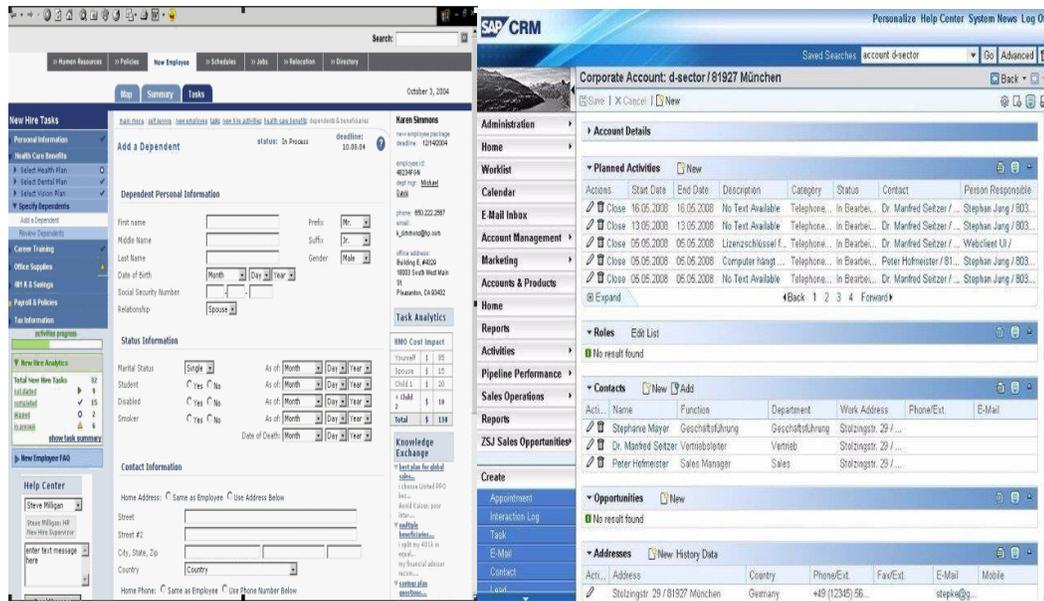


Fig. 4: Aplicativos de las fuentes.

- Monitorear diariamente los resultados del proceso de actualización, consiste en validar en los aplicativos que los cambios realizados fueran aplicados de manera satisfactoria. En caso de no ser así se considera un rechazo que debe ser tratado de forma manual por el responsable del proceso.
- Las pruebas de calendarización se realizaron en dos fases, la primera se ejecutó solamente con los nodos correspondientes a la construcción del proceso, luego se realizó otra prueba de calendarización con todo el modelo de conciliación completos. Grupo Bancolombia cuenta con un rol llamado **Usuario calendarizador** que son los encargados realizar estas pruebas y la calendarización final.

Este proceso debe monitorearse diariamente para verificar que las tres salidas del modelo de conciliación se exportan correctamente a la ruta compartida de la Vicepresidencia servicios para los clientes.

- A lo largo del informe se menciona los tiempos de ejecución de las fases del proyecto, esta es una prueba continua para validar que la ejecución diaria no se tarde más de 1 hora. El tiempo de ejecución es dependiente de la disponibilidad de los servicios cloud de Grupo Bancolombia, es posible que algunos días el proceso se ejecute más rápido o más lento.

6.6. Documentación

La documentación para una herramienta de usuario final (EUC), requiere de varios formatos para ser publicado dentro del catálogo de EUC.

6.6.1 Valoración de criticidad

Resultado: Criticidad media.

Este resultado indica que el proceso involucra datos de clientes alojados directamente en las fuentes de información.

La continuación del proceso de documentación consiste en comentar los nodos de SAS Enterprise, las líneas de código y consolidar la información requerida por el formato de herramientas de usuario final (EUC).

Grupo Bancolombia cuentas con SharePoint contenidos en el catalogo de EUC, como parte de la documentación es necesario generar el SharePoint con la información requerida

Para finalizar la documentación se realizaron dos videos que consisten en la gestión técnica del proceso en caso de que en un futuro se planteen modificaciones para la EUC.

El segundo video es la gestión manual que debe realizar el responsable del proceso con cada una de las salidas del modelo de conciliación.

7.RESULTADO FINAL

El modelo de conciliación cumple con los resultados planteados inicialmente en la propuesta del proyecto, además cumple con los resultados esperados por Grupo Bancolombia.

- Realizado el modelado de información y las consultas SQL necesarias para obtener la información requerida para llevar a cabo el proceso.
- Es bidireccional ya que actualiza las fuentes de información clientes Valores y clientes Bancolombia.
- Contiene buenas prácticas de programación para que las consultas realizadas sean eficientes, este resultado se respalda con la correcta calendarización del proceso dentro de la malla de servicio Bancolombia.
- Se le asigno un ID dentro del catalogo de herramientas de usuario final (EUC), lo cual indica que el proyecto cumple con las políticas definidas por la empresa para ser considerado una EUC.
- El proceso es escalable y dinámico en caso de que sea intervenido por otro colaborador de Grupo Bancolombia

8. DIFICULTADES

- Clientes Valores es una base datos tratada de manera muy rigurosa sin embargo ninguna fuente de información está libre del error humano o fallas tecnológicas. Por esta razón no es posible asegurar la calidad de la información al 100% Sin embargo los controles establecidos a lo largo del proceso buscan asegurar la mejor calidad posible.
- Si bien SQL es un motor de bases de datos bastante común en la ingeniería de sistemas, Grupo Bancolombia utiliza sus propias reglas y estructuras para realizar consulta en la nube, fue necesario aprender todos los lineamientos del Banco.
- Conciliar el modelo con el proceso de actualización, fue una dificultad acordar la forma en que iba a ser desplegado el proyecto, la solución consistió en establecer el esquema y el aplicativo a utilizar para consumir los resultados del proceso.
- El rol de aprendiz es limitado en cuanto a los permisos para el uso de los aplicativos, fue necesario rediseñar algunos procesos para que funcionaran dentro del rol de aprendiz. Un proceso importante a destacar es la exportación a la ruta publica de la Vicepresidencia.
- Es evidente que el Covid-19 es una inconveniente para la mayoría. Sin embargo, la virtualidad se tomó muy bien, se crearon buenas sinergias con los miembros del equipo a pesar de la virtualidad.

9. CONCLUSIONES

- Destacar el aporte del modelo a la sección servicios calidad de datos, debido a que el objetivo principal de esta sección es asegurar un índice de calidad mayor al 90% y que debe ir ascendiendo continuamente. Las propuestas de cambio indicadas por el proceso en forma masiva es parte de los diferentes procesos realizados por la sección para mejorar el umbral.
- El modelo de conciliación en general tiene un concepto sencillo, sin embargo, su complejidad radica en todas las fuentes de información que maneja Grupo Bancolombia, por ejemplo, no todas las entidades pertenecientes al grupo manejan las mismas reglas para guardar la información.
- Las herramientas de programación visual como SAS Enterprise Guide son útiles para asociar información de otros aplicativos como las conexiones a las bases de datos, el uso de SQL para la extracción de la información necesaria, códigos Python para automatización, insumos de Excel y programas para la calendarización del proceso.
- Desde el punto de vista profesional, la practica académica es una oportunidad para mejorar en la aplicación de los conocimientos adquiridos a lo largo del proceso como estudiante, en el desarrollo de la practica académica fue evidente el uso de programación, planteamiento de arquitecturas de software y análisis y diseño de sistemas desde diferentes perspectivas como el uso de Backlogs para el cumplimiento de las tareas, diagramas de componentes y diagrama de actividades
- El proceso de prácticas académicas es un apoyo para conocer mejor la marca personal de cada uno. Hay Valores éticos que deben ser clave como la responsabilidad, integridad, honestidad entre otros, sin embargo, la marca personal viene definida por otro tipo de cualidades. En mi caso particular destaca ser un profesional analítico, preciso y correcto en cada uno de su proceso.
- El aporte de la Universidad de Antioquia para el proceso de prácticas académicas es evidente a lo largo del pregrado, la universidad se esfuerza por educar profesionales que cumplan con el ser, el hacer y el saber. Todo esto se trabaja a lo largo de la carrera desde diferentes puntos de vista como, por ejemplo: Impulsar el trabajo en equipo, dictar diferentes asignaturas relacionadas a las habilidades blandas que hoy día son una parte fundamental para que las empresas consideren a un profesional además de los diferentes retos que se presentan a lo largo del pregrado para pulir la parte del hacer y el saber.
- Destacar la asignación de los asesores internos por parte de la universidad, la experiencia de los encargados de cada uno de los practicantes aporta al proceso profesional. En mi caso fue la primera vez que me enfrente a un proyecto real, el asesor interno fue de gran ayuda para entender que la planeación del proyecto puede tener muchos cambios a lo largo de la ejecución y las diferentes formas de enfrentar adelantos o retrasos de un proyecto.

10. REFERENCIAS

- [1] [Xiaopeng Li](#) ; [Wenli Zhou](#), “Performance Comparison of Hive, Impala and Spark SQL” 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, August 2015
- [2] Rosa Gini, Martijn Schuemie “Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies” , eGEMs , February 2016
- [3] Patrick M. Kelly, T. Michael Cannon, and Donald R. Hush "Query by image example: the comparison algorithm for navigating digital image databases (CANDID) approach", Proc. SPIE 2420, Storage and Retrieval for Image and Video Databases III, (23 March 1995); <https://doi.org/10.1117/12.205289>
- [4] Yang, Seungyeon, Park, Seok-Cheon , Lee, Jinhee “Design of table shape comparison Algorithm for Improvement of Transformation Managing Process in Meta Data management System”, koreascience,August 2013
- [5] David J. Hand, Niall M. Adams, “Data Mining”, Wiley, June 2015; <https://doi.org/10.1002/9781118445112.stat06466.pub2>
- [6] SAS,SAS Enterprise Guide [Online].Available: <https://support.sas.com/en/software/enterprise-guide-support.html>