

NOMBRE DE LA EMPRESA

Título de proyecto o documentación

Detalle

Subtítulo

ÍNDICE

Propósito	3
Documentos de análisis	4
Descripción de requerimiento	4
Reglas de negocio	5
Requerimientos funcionales	6
Especificación de capa semántica	7
Granularidad	8
Periodicidad	9
Dimensionamiento preliminar	9
Esquema de seguridad	10
Documento de diseño	11
Arquitectura	11
Modelo de datos	12
Diccionario de datos	13
Mapeo de datos	14
<i>Linaje de datos</i>	15
Diagrama de ejecución	16
Puntos de control y recuperación	19
Matriz de escalamiento	20
Manual de validación	20
Evidencia de publicación en repositorio	21
EVIDENCIAS	22
Verificación de artefactos para publicación a producción	22
Identificación de artefactos en Repositorio	22
Anexos Partner	22
Manual de Instalación	23
Aprobación por parte del cliente	23
Ran	23
Contacto	

Logo de la empresa

Subtítulo
Detalle 01
Detalle 02

PROPOSITO

El presente documento tiene el objetivo de describir integralmente la unidad o sistema que deberá pasar al ambiente productivo y ser gestionado por el área de Operación Big Data.

La información solicitada en el presente documento se utiliza para contextualizar el sistema que se está publicando al ambiente productivo de Big Data - BIMBO, de forma que los siguientes elementos clave para su operación son descritos y referidos para efectos de continuidad e independencia en la operación del equipo de desarrollo:

- Definición del sistema o unidad
- Entradas
- Procesos
- Salidas
- Proveedores de datos e información
- Clientes

Es requisito completar todas las secciones del documento. Y, en caso de que no aplique al sistema o unidad alguna de las secciones, deberá indicarse la justificación de negocio o técnica que permita omitirla, justo dentro de la sección correspondiente.

El presente documento respalda un recurso de gestión de datos e información que será publicado al ambiente productivo y que se sujetará a posibles cambios/actualizaciones dentro del ciclo productivo del ambiente Big Data.

DOCUMENTO DE ANÁLISIS

Descripción de análisis

Descripción de elementos clave del requerimiento original del sistema.

Reglas de negocio

Compartir las reglas de negocio utilizando Vocabulario de Negocio. Una regla de negocio no es un Query o una fórmula. Se trata de la expresión de

- Términos clave de negocio [Sustantivos]
- Relaciones entre los sustantivos de negocio
- Situaciones que regulan las relaciones para producir un resultado o acción esperada por el negocio.

Ejemplo:

RN1. Del método para calcular Venta Total por Producto

Se extrae de la tabla inventario_Movimiento_Detalle la clave de producto [campo CVEProd], cantidad [campo Qty] y precio Unitario [campo PriceU] para calcular mediante la formula

$$\text{Venta_total_prod} = \text{Qty} * \text{PriceU}$$

En la tabla Ventas_Semanal se inserta un registro para cada movimiento de orden identificado por el número de orden, número de cliente y venta total por cada línea del detalle de la venta [incluido la Venta Total]

RN2. Del cálculo de comisiones nivel A

Los datos correspondientes para el cálculo de comisiones correspondientes al NivelA se consiguen de la tabla Representante y de la tabla Tabulador Semanal considerando la semana vigente y la información almacenada correspondiente.

Un valor default de Comisión existe en la Configuración General para el caso de que no exista un valor específico para alguna Semana dentro del Tabulador.

Requerimientos funcionales

A continuación, explique y describa los elementos de datos e información que son parte integral del sistema o unidad que publica a producción.

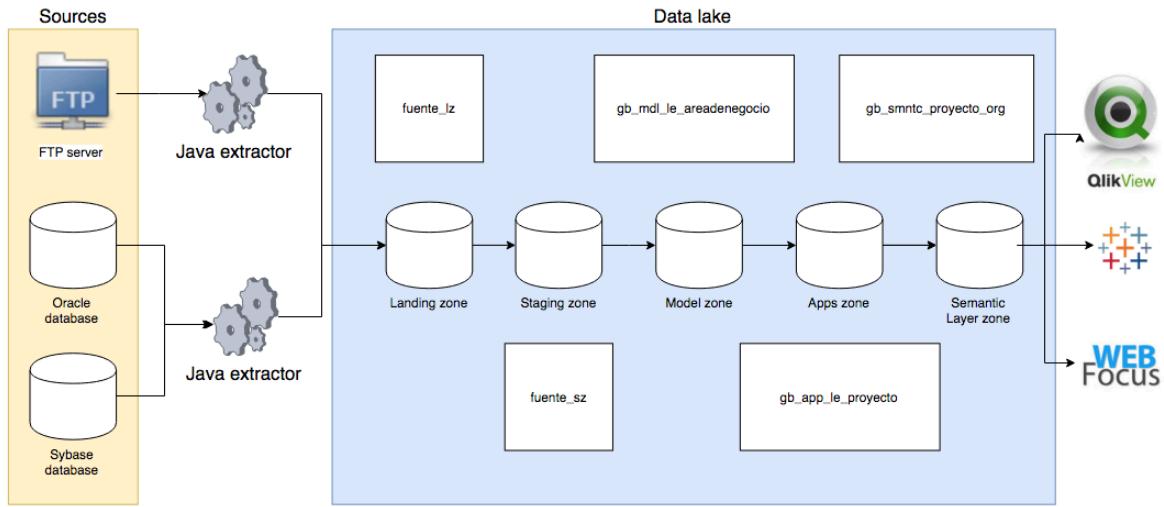
Describa las relaciones/dependencias en la siguiente tabla:

Ejemplo:

Sistema o Unidad	Relación o dependencia	Requiere configuración	Descripción
Extractor IC	Instalación y acceso al extractor	Definición de Script para consultar evento de cierre de agencia	El evento de cierre de agencia es el disparador de eventos en el sistema
Catálogo de Clientes BIMBO	Referencia para consulta y reporteo	Solicitar permisos de lectura a base de datos 089_stmn_ic tabla Clientes_IC	No realiza modificación de datos ni guardado Sólo referencia y consulta el catálogo

Especificación de capa semántica

De acuerdo con la arquitectura propuesta para el modelo de operación en el ambiente Big Data la capa semántica exhibe ante el usuario final el resultado de los procesos de extracción y transformación definiendo productos de datos e información (las salidas definidas del sistema o unidad) susceptibles de integrarse como insumos en nuevos procesos analíticos y de proceso de datos e información.



La capa semántica incluye los siguientes productos de datos:

Base de datos	Tabla	Descripción/Propósito

Ejemplo:

Base de datos	Tabla	Descripción/Propósito
gb_smntc_089_ventas	089_estadisticaventas_ic	Resumen ventas por Agencia – Inteligencia Comercial Cálculo y agregación de ventas a partir de IC por agencia Fuente para proceso analítico en Qlick View; todos los días a las 00:00 se procesan datos desde capa semántica

Granularidad

De las tablas que integran la capa semántica describa el nivel de mayor detalle que se puede consultar, esta información es para describir los registros que pueden ser consultados desde las tablas en la capa semántica.

Ejemplo

Tabla	Nivel de consulta
089_estadisticaventas_ic	Venta por Cliente Venta por Agencia Venta por Día Venta por Producto Venta por Ruta

Periodicidad

Indicar el ciclo con el que se va a lanzar los procesos que integran el sistema o unidad que se publica en producción, uno entre las siguientes opciones:

- Diario: lunes a sábado
 - Todos los días: lunes a domingo
 - Semana BIMBO
 - Mensual: Cada día 2 de mes
 - Sobre demanda

Dimensionamiento preliminar

Es importante que durante la fase de pruebas de su sistema o unidad se elabore una estimación de demanda para espacio en disco una vez que se publique a producción. Se trata pues de contar con una métrica objetiva de la demanda de espacio en disco, no constituye una solicitud exacta.

La estimación de requerimiento de espacio en disco permite al equipo de operación habilitar los recursos para un adecuado funcionamiento del sistema o unidad que se publica considerando que debe compartir e interactuar con otros ya dentro del ambiente.

En caso de que usted tenga una demanda notable de espacio en disco debe disponerse y garantizarse para la publicación y considerando el tiempo de vida para los próximos 12 meses.

En todo momento una modificación a la infraestructura Big Data debe sujetarse a un proceso de validación y autorización; de tal forma que el crecimiento se contempla para un periodo mínimo de un año y con la estimación de todos los módulos y sistemas que ya están publicados en producción.

Zona de almacenamiento	Dimensión inicial	Crecimiento mensual	Persistencia

Ejemplo

Zona de almacenamiento	Dimensión inicial	Crecimiento mensual	Persistencia
Landing Zone - datos crudos	15 Gb	2Gb	Permanente
Staging Zone	30 Gb	1Gb	Volátil, a cada ejecución
Model Zone	2 Gb	NA	Permanente
Semantic Zone	2 Gb	100 Mb	Permanente

Esquema de seguridad

Indique los criterios aplicables de seguridad y control de acceso a los datos de la zona semántica del sistema o unidad que se publica a producción.

El área responsable del proyecto suscribe el acuerdo de toda solicitud de acceso para consulta deberá someterse sin excepción a la Normativa de Seguridad – CAN Corporativo.

Cada solicitud de permisos debe apegarse a los formatos establecidos dentro de la Normativa de Seguridad – CAN Corporativo.

A continuación, se listan los contactos a participar en caso de control especial de acceso:

Contacto	Actividad	Correo	Teléfono

Ejemplo

Contacto	Actividad	Correo	Teléfono
Foo Bar	Autorizar solicitudes consulta	foo@grupobimbo.com	55 5666 6677
Zoo Bar	Líder	zoo@grupobimbo.com	55 3333 4444
Moo Bar	Gerente	moo@grupobimbo.com	55 2222 3333

DOCUMENTO DE DISEÑO

Arquitectura

El área de desarrollo deberá determinar cuáles son los documentos de arquitectura que se entregarán a producción, en el caso del área productiva este documento servirá para asignar paquetes de proyectos cuando se debe hacer algún ajuste.

Al área de operación le es indispensable conocer en resumen la arquitectura (general overview) que ayudó a dirigir el desarrollo del sistema o unidad que se publica a producción, con el objetivo de contextualizar el sistema o unidad dentro del ecosistema productivo Big Data.

Se solicita realizar una descripción ejecutiva de la arquitectura, considerando un ejercicio de explicar:

- Las fuentes de datos que se conectan y su método
- Los procesos o fases que contempla el sistema
- La ruta que siguen los datos
 - Datos Fuente
 - Datos Crudos
 - Datos Enriquecidos
 - Datos Limpios
 - Información de negocio
 - Métricas

Ejemplo:

Inteligencia Comercial para BRASIL se conecta a través de SQOOP a base de datos de Plataforma Comercial (Sybase) en TRES agencias. El evento disparador del proceso es el cierre de la agencia.

Todas las tablas de Plataforma Comercial se copian hacia la zona de datos crudos dentro de la base de datos tmp_ic. A partir de esta base de datos se aplican reglas de negocio que permiten el cálculo de métricas de ventas considerando su granularidad. Se crean registros finales en la tabla de la zona semántica desde donde se brinda el servicio a las áreas de negocio.

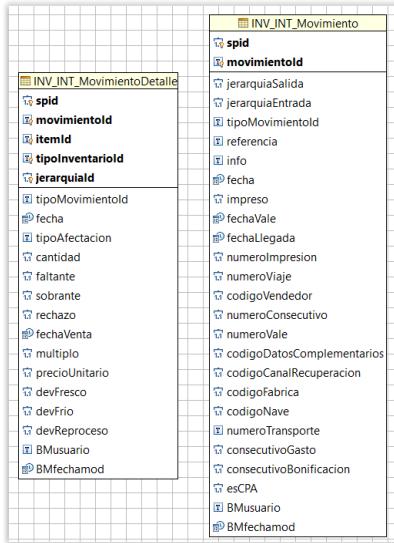
Modelo de datos

Proporcionar el modelo físico de datos. Y los scripts DDL correspondientes al modelo como ANEXOS al presente documento.

Un diagrama ER sería la herramienta ideal para describir el modelo de datos, aunque si no está disponible es necesario describir cada una de las Entidades Clave del modelo de datos a partir de la siguiente Ficha

Nombre de la ENTIDAD/TABLA					
Campo	Tipo dato fuente	Tipo dato semántico	Longitud	Valor default	¿Es Llave?
Se relaciona con:					

Ejemplo:



Diccionario de datos

Complemento natural del modelo de datos. Principalmente de los datos que intervienen en la construcción de tablas del área semántica.

Considere como referencia al menos la siguiente ficha para especificar sus datos:

Nombre de la ENTIDAD/TABLA					
Campo Fuente	Descripción	Tipo dato semántico	¿Es calculado?	Fórmula	¿Es Llave?

--	--	--	--	--

Formato:

Diccionario de
datos.xlsx

Ejemplo:

Diccionario de
datos ejemplo.xlsx

Base: GLOBO Tabla: SALE		
<i>Campo</i>	<i>Tipo de dato</i>	<i>Descripción</i>
sale_point_id	INT	Número del area de negocio
Bonification	FLOAT	Bonificación diferencia entre el precio base y el precio de venta que se aplicó a la partida
closing_sales	STRING	Si ya se encuentra cerrado y en corte Z
average_cost	FLOAT	Costo promedio al momento de la operación en la partida
last_cost	DOUBLE	Costo ultimo al momento de la operación
property_account	STRING	Cuenta predial (Solo facturación electrónica)
discount	DOUBLE	Porcentaje de descuento que se aplicó al producto

Mapeo de datos

Proporcione información sobre los datos en la capa semántica con dos perspectivas: tabla y campo

El linaje de datos construye a base de niveles los antecedentes y consecuentes, desde el punto de vista de los datos para pasar desde un conjunto de datos fuente por datos corregidos enriquecidos y procesados hasta tener información clave de un proceso de negocio al contar con la aplicación de las correspondientes reglas de negocio.

Es posible que un nivel del linaje de datos sea terminal por que entrega resultado o por que se encuentra una excepción.

A nivel de tablas, indique el linaje de los datos considerando el siguiente formato:

Linaje de datos

Nivel	Base de datos	Tabla	Proceso	Siguiente	Terminal

Ejemplo:

Nivel	Base de datos	Tabla	Proceso	Siguiente	Terminal
1	Microleaks_data	Inv_movimiento	Ingesta	3	No
2	Microleaks_data	Inv_movimientodetalle	Ingesta	3	No
3	Microleaks_data	Agencia	Consolidación de Faltantes	6	No
4	Microleaks_data	Bus_cal	Consolidación de Faltantes	6	No
5	Microleaks_dta	Producto Inv_movimientodetalle	Consolidación de Faltantes	6	No
6	Microleaks_data	Microleaks_ds	Publicación de Distribución secundaria		SI

Formato para mapeo de datos:

Mapeo modelo
semántico.xlsx

De forma complementaria a nivel de campos es necesario para los campos en todas las tablas de la capa semántica indicar el mapeo entre la fuente y el campo correspondiente:

Campo fuente	Descripción	Campo semántico	Observaciones

Ejemplo:

XALDIGITAL

Reglas de negocio

Se debe hacer un cruce entre la tabla GLOBO_MOVEMENT y GLOBO_MOVEMENT_REASON con los campos MOVEMENT_REASON_ID
 Se debe hacer un cruce entre la tabla GLOBO_MOVEMENT y GLOBO_MOVEMENT_TYPE con los campos MOVEMENT_TYPE_ID
 Se debe hacer un cruce entre la tabla GLOBO_MOVEMENT y GLOBO_TRANSFER_TYPE con los campos TRANSFER_TYPE_ID
 Se debe filtrar el campo MOVEMENT_REASON_ID de la tabla GLOBO_MOVEMENT con el valor 24

SMNTC_LAYER_MYAPP.TABLA1						
TABLA FINAL	COLUMNNA-TablaFinal	TIPO DATO	TABLA ORIGEN	COLUMNNA	TIPO DATO	OPERACIÓN [Aplicada para generar campo destino]
SMNTC_LAYER_GLOBO_INVENTORIES	SALE_POINT_ID	INT	GLOBO_MOVEMENT	SALE_POINT_ID	INT	
SMNTC_LAYER_GLOBO_INVENTORIES	PRODUCT_ID	INT	GLOBO_MOVEMENT	PRODUCT_ID	STRING	
SMNTC_LAYER_GLOBO_INVENTORIES	PRODUCT_NAME	STRING	GLOBO_PRODUCTS	PRODUCT_DESC	STRING	
SMNTC_LAYER_GLOBO_INVENTORIES	EXISTENCE	INT	GLOBO_MOVEMENT	EXISTENCE_INI	DOUBLE	Se obtiene el valor absoluto de este campo.
SMNTC_LAYER_GLOBO_INVENTORIES	UNITS	STRING	GLOBO_PRODUCTS	PRODUCT_UNIT	STRING	
SMNTC_LAYER_GLOBO_INVENTORIES	PRICE	DECIMAL(9,2)	GLOBO_MOVEMENT	PRICE	DECIMAL(9,2)	
SMNTC_LAYER_GLOBO_INVENTORIES	COST	DECIMAL(9,2)	GLOBO_MOVEMENT	MOVEMENT_COST	DECIMAL(9,2)	
SMNTC_LAYER_GLOBO_INVENTORIES	SALE_POINT_NAME	STRING	GLOBO_SALE_POINT	SALE_POINT_NAME	STRING	
SMNTC_LAYER_GLOBO_INVENTORIES	PACKING_CONTENT	INT	N/A	N/A	INT	Constante con valor 1.
SMNTC_LAYER_GLOBO_INVENTORIES	TOTAL_COST	DECIMAL(9,2)	GLOBO_MOVEMENT	EXISTENCE_INI	DOUBLE	Se obtiene el valor absoluto del resultado de la multiplicación entre estos dos campos.
SMNTC_LAYER_GLOBO_INVENTORIES	PARTITIONKEY	STRING	GLOBO_MOVEMENT	MOVEMENT_COST	DECIMAL(9,2)	
SMNTC_LAYER_GLOBO_INVENTORIES	PARTITIONKEY	STRING	GLOBO_MOVEMENT	PARTITIONKEY	STRING	

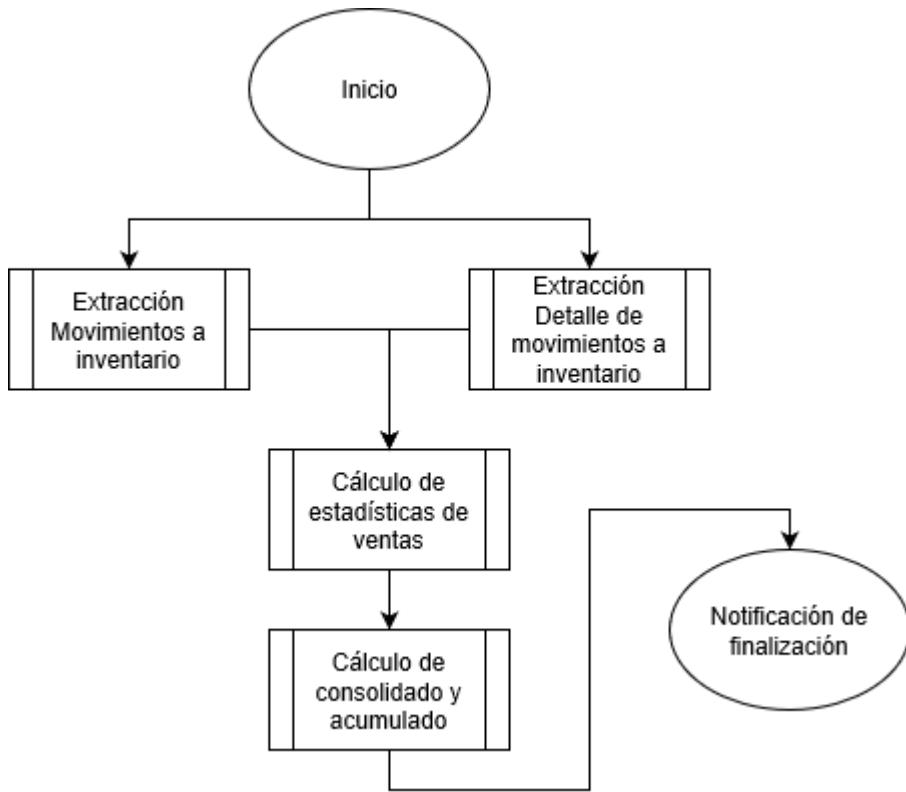
DIAGRAMA DE EJECUCIÓN

Indicar a nivel de fases utilizando un sencillo diagrama de flujo – puede incluso documentarlo a través de un diagrama de interacción tipo BPMN.

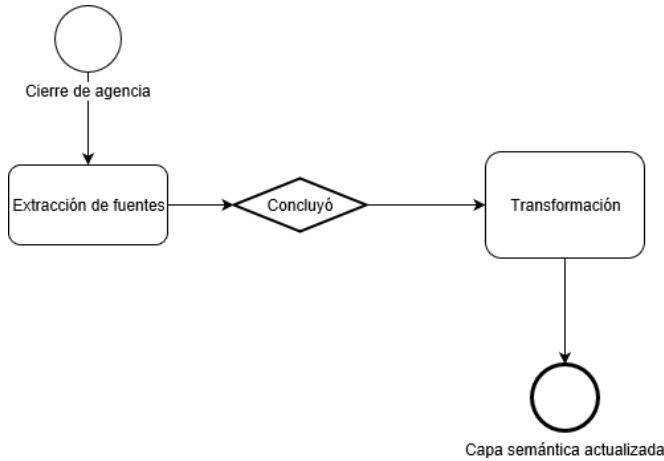
Ejemplo:

XALDIGITAL

Data Experts



Ejemplo



Nota: En el ejemplo se entiende que pueden ser ejecutados en paralelo, pero es necesario ser claro por el tema de borrado y recargas de tablas, países, etc.

En este diagrama de ejecución es posible identificar:

- Actores
- Precedencias
- Ejecución simultánea
- Dependencia con otros sistemas o unidades

Asegúrese de documentar esta información en su diagrama o si así lo requiere utilice un formato como el siguiente (Indicar si es necesario que la información enviada tenga la precedencia, el tipo de ejecución paralela o excluyente, etc).

ID Fase	Proceso	Tipo de ejecución	Resultado esperado	Tiempo estimado	¿Es final?

Ejemplo

ID Fase	Proceso	Tipo de ejecución	Resultado esperado	Tiempo estimado	¿Es final?
1	Extracción			45 mins	No
	Movimientos a inventario	Línea de comando	Tabla en zona Landing con datos	30 mins	
	Movimientos a detalle de inventarios	Línea de comando	Tabla en zona Landing con datos	45 mins	
2	Cálculo de estadísticos de ventas	Línea de comando	Tabla semántica actualizada	90 mins	No
3	Cálculo de consolidados	Línea de comando	Tabla semántica actualizada	60 mins	Si

PUNTOS DE CONTROL Y RECUPERACIÓN

En concordancia con el diagrama de ejecución indique para cada fase de la ejecución a donde deberá continuar en caso de la ocurrencia de un fallo o problema que provoque la interrupción de ese proceso.

De forma que el operador identifique el siguiente paso y se concluya con éxito la ejecución.

Ejemplo:

ID Fase	Proceso	Limpieza en caso de fallo	Regresa a Fase
1	Extracción		
	Movimientos a inventario	Eliminar partición correspondiente al día actual	1
	Movimientos a detalle de inventarios	Eliminar partición correspondiente al día actual	1
	Fase exitosa:	No hay mensajes de error en LOG	
2	Cálculo de estadísticos de ventas	Eliminar datos correspondientes al día actual	2
	Fase exitosa:	Sin mensajes de error en LOG Con registros en tabla semántica	
3	Cálculo de consolidados	Eliminar datos correspondientes al día actual	3
	Fase exitosa:	Sin mensajes de error en LOG Con registros en tabla semántica	

MATRIZ DE ESCALAMIENTO

En caso de fallos dentro del proceso es necesario contactar a personal clave en la línea de negocio que permitan dar continuidad al proceso.

Persona	Contacto	Tipo de contacto	Urgencia

Ejemplo

Persona	Contacto	Tipo de contacto	Tiempo de respuesta
Foo Bar	foo@grupobimbo.com 55 6666 7788	Responsable de BD fuente	Inmediata
Moo Bar	moo@grupobimbo.com 55 7788 9900	Delivery CAN	2-4 horas
Zoo Bar	zoo@grupobimbo.com 55 9988 7755	Analista de negocio	6-8 horas
Roo Bar	roo@grupobimbo.com 55 4466 3322	Responsable desarrollo	Día siguiente

MANUAL DE VALIDACIÓN

Incluir evidencias de las pruebas y validaciones realizadas en ambiente de desarrollo. Considerando aquellas que tienen relación directa con la capa semántica.

Es importante proporcionar al Operador una guía paso a paso para identificar la correcta ejecución del proceso (queries, cifras control o similares que permitan a la operación hacer la validación de registros transferidos, cargas, validaciones (fechas, cierres, etc.))

Es importante incluir los posibles errores y soluciones que el sistema pudiera presentar.

EVIDENCIA DE PUBLICACIÓN EN REPOSITORIO

Todos los códigos fuente deben incorporarse al repositorio oficial al cual se dará acceso total al equipo de desarrollo hasta que se declare la publicación en producción, donde se retira el acceso.

Debe incorporar en el presente documento evidencia de la incorporación TOTAL de fuentes en el repositorio asignado.

EVIDENCIAS

Proporcione evidencias (capturas, screenshot, etc)

VERIFICACIÓN DE ARTEFACTOS PARA PUBLICACIÓN A PRODUCCIÓN

Para ser llenado durante el proceso de pase a producción

- Creación de base de datos
- Creación de tablas
- Creación de Usuarios
- Asignación de permisos
- Generación de ligas dinámicas y referencias
- Programas y scripts necesarios publicados en los directorios correspondientes

IDENTIFICACIÓN DE ARTEFACTOS EN REPOSITORIO

Proporcione el URL dentro del repositorio donde se encuentra la última versión de Ejecutables

ANEXOS PARTNER

Proporcione los detalles de la herramienta que se está utilizado ya sea de AWS, GOOGLE, MICROSOFT, etc.

MANUAL DE INSTALACIÓN

Especificar paso a paso el como y donde se debe instalar el sistema

APROBACIÓN POR PARTE DEL CLIENTE

Incluir aquí el documento de liberación de proyecto

Carta de
Liberación_Proyecto

RAN

Incluir el documento

Logo de la empresa

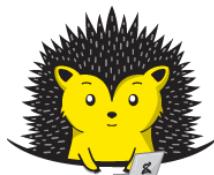
Subtítulo
Detalle 01
Detalle 02

CONTACTO

Nombre:

Correo:

Teléfono:



GRACIAS!