**INFORME DE RESULTADOS**

**Unidad 3 - Tarea 6**

**Peso: 10%**

**Equipo/Grupo : 3**

**Estudiantes :**

**- *Luis Felipe Gil Gómez***

* ***Gerson Gustavo Fernández Badillo***
* ***Davidson Harley Rave Buitrango***
* ***David Pabón García***

**Objetivo General**

**Diseñar un Gobierno de Datos y la estructura de un Proyecto de Big Data de envergadura para enfrentar la problemática en relación a la implementación de una Arquitectura de Plataforma de control y monitoreo de concentración de gases tóxicos en las fábricas de la empresa “Sustancias Locas”.**

Se requiere la convergencia de talento humano, planificación del trabajo, contratación de diversos servicios, instalación y configuración de aplicaciones y servidores locales y en la Nube, desarrollo de software, soporte técnico, entre otros.

Por lo tanto, es necesario la conformación de un gobierno de datos y la estructuración de un Proyecto de Big Data; así como una lista de verificación del Proyecto Big Data que incluya todos los ítems necesarios y los tiempos estimados para lograr el objetivo.

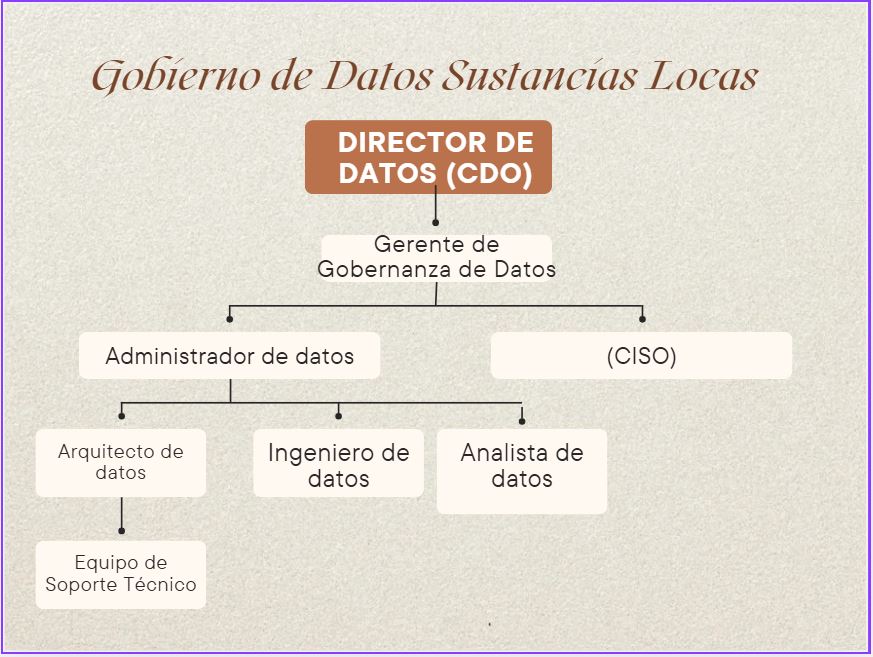
**1.-Diseño de un Gobierno de Datos**

*Después de analizar el enunciado y evaluar los requerimientos, proponga la estructura y roles que debe tener el gobierno de datos que se encargara del Proyecto Big Data*

**1.1- Roles del gobierno de datos**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Departamento o persona** | **Descripción** | **Rol, responsabilidades y/o funciones** |
| **1** | Director de Datos (CDO) | Miembro de la alta dirección o gerencia, responsable de la estrategia de datos general. | Liderar la estrategia de datos, coordinar roles, asegurar cumplimiento de políticas y alineación con objetivos corporativos. |
| **2** | Gerente de Gobernanza de Datos | Coordinador general de gobierno de datos. | Establecer políticas, procesos y estándares; coordinar personal. |
| **3** | Administrador de datos | Especialista en calidad y gestión de datos por área. | Supervisar la calidad de datos, garantizar metadatos correctos y coordinar con usuarios para mantener integridad de datos. |
| **4** | Arquitecto de datos | Diseñador de la infraestructura de datos. | Diseñar modelos de datos, arquitectura de almacenamiento, flujos de datos e integración con sistemas de terceros. |
| 5 | Ingeniero de datos | Especialista en extracción, transformación y almacenamiento de datos. | Desarrollar sistemas de datos, asegurar escalabilidad y disponibilidad del sistema de almacenamiento. |
| 6 | Analista de datos | Analista y modelador de datos. | Analizar las estadísticas para la detección y predicción. |
| 7 | Director de Seguridad de la Información (CISO) | Responsable de la seguridad y privacidad de los datos. | Definir políticas de acceso, cifrado, control de privacidad, cumplimiento legal y protección de datos sensibles. |
| 8 | Equipo de Soporte Técnico | Personal de infraestructura y soporte TI. | Instalación, configuración y mantenimiento de servidores, hardware y software. |

***1.2.- Estructura gráfica del gobierno de datos***



**2.-Diseño del Proyecto Big Data**

*Después de analizar el enunciado y evaluar los requerimientos, proponga la estructura y elementos que deben conformar el Proyecto Big Data*

**2.1- Elementos del Proyecto Big Data**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Elemento** | **Descripción** | **Observaciones** |
| **1** |  |  |  |
| **2** |  |  |  |
| **3** |  |  |  |
| **..** | … | … | … |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **N** |  |  |  |

**2.2- Lista de verificación del Proyecto Big Data**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Elemento** | **Estado** | **Tiempo** |
| **1** |  |  |  |
| **2** |  |  |  |
| **3** |  |  |  |
| **..** | … | … | … |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **N** |  |  |  |

**3.- Conceptos y elementos de Big Data**

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Componente** | **Descripción (breve cita bajo Norma APA)** | **Referencia (APA)** |
| **1** | Arquitectura de Sistema de Información |  |  |
| **2** | Hadoop |  |  |
| **3** | Spark |  |  |
| **4** | PostgreSQL |  |  |
| **5** | Amazon Kinesis FireHose |  |  |
| **6** | Amazon QuickSight |  |  |
| **7** | Amazon Glue |  |  |
| **8** | HTTP API |  |  |
| **9** | Boto3 |  |  |
| **10** | Amazon RedShift |  |  |
| **11** | Amazon Athena |  |  |
| **12** | Amazon CloudWatch |  |  |
| **13** | Boto3 |  |  |
| **14** | Formato de datos JSON |  |  |
| **15** | Formato de datos CSV |  |  |
| **16** | Formato de datos Parquet |  |  |
| **17** | Microcontrolador Arduino |  |  |
| **18** | Microcontrolador ESP8266 |  |  |
| **19** | Sensor MQ-135 |  |  |
| **20** | Benceno |  |  |

**4.- Diseño de la estructura de la hoja de cálculo del microcontrolador “lecturas”**

*Colocar en esta sección los datos que se generan en el sensor y se almacenan en la hoja de cálculo local “lecturas” en el microcontrolador. El mismo formato de hoja de cálculo se utiliza en el servidor local.*

**4.1. - Diseño de la estructura de la hoja de cálculo para almacenar lecturas del microcontrolador**

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Nombre Columna** | **Descripción** | **Observaciones** |
| **1** |  |  |  |
| **2** |  |  |  |
| **3** |  |  |  |
| **..** | … | … | … |
| **N** |  |  |  |

**4.2. – Poblamiento de la hoja de cálculo del microcontrolador**

*Colocar en esta sección veinte (20) registros con datos ficticios. Nota: puede colocar el “pantallazo” de la hoja de cálculo o incrustar “copia” y “pega”. Las columnas son los datos*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **#** | **Dato Columna 1** | **Dato Columna 2** | … | **Dato Columna “N”** |
| **1** |  |  | … |  |
| **2** |  |  | … |  |
| **3** |  |  | … |  |
| **..** | … | … | … | … |
| **20** |  |  | … |  |

**5.- Diseño de la estructura de la tabla de “lecturas”**

*Colocar en esta sección la estructura de la tabla de “lecturas” dónde se almacenan los datos de las diferentes hojas de cálculo que provienen de los sensores. Esta tabla se encuentra dentro de la base de datos “monitoreo-produccion”*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tabla** | **lecturas** | |
| **#** | **Nombre Campo** | **Tipo de dato** | **Tamaño (Bytes)** |
| **1** |  |  |  |
| **2** |  |  |  |
| **3** |  |  |  |
| **..** | … | … | … |
| **N** |  |  |  |
| **Tamaño de un (1) registro en bytes** | | |  |

**6.- Diseño y creación de la Base de Datos “monitoreo-produccion”**

*Colocar en esta sección el diagrama de Entidad-Relación sin atributos y el Diccionario de datos de la base de datos de gestión. Esta es la base de datos que contiene, además de la tabla lecturas,*

**6.1.- Diagrama Entidad-Relación (entidades, relaciones, atributos y cardinalidades)**

**6.2- Diccionario de Datos (tipos de dato, tamaño, claves primarias y foráneas)**

**6.3- Scripts de creación de las tablas de la Base de Datos “monitoreo-produccion”**

**7.- Poblamiento de la base de datos “monitoreo-produccion”**

*Colocar en esta sección un pantallazo de los resultados de DML (INSERTS) en pgAdmin4 realizados en todas las tablas de la base de datos.*

**7.1.- Scripts de inserción de registros en todas las tablas de la base datos (INSERTS)**

**7.2- Pantallazos de consultas SELECT de las tablas pobladas**

**8.- Conclusiones.**

*Elabore las conclusiones planteando la importancia y utilidad de esta tarea y su relación con el contenido de la asignatura. ¿Cuáles eran sus expectativas al inicio de la asignatura y en este momento final? ¿Cómo cree Ud. que el conocimiento aprendido en la asignatura afectará sus oportunidades laborales y desempeño profesional una vez obtenga su título en el Pascual Bravo?*

*Adicionalmente a estas conclusiones generales del grupo de trabajo, cada estudiante debe expresar sus propias conclusiones.*

**9.- Video de sustentación:**

*Elabore un video de sustentación con la participación de todos los integrantes (si es en equipo). Este vídeo debe informar sobre las actividades realizadas en general . Coloque el enlace en esta sección*