



**UNIVERSIDAD  
DEL PACÍFICO**

**Escuela de  
Postgrado**

**“Creando valor con Datos & Analítica 4.0 en la industria de acero”**

**Trabajo de Investigación presentado  
para optar al Grado Académico de  
Magíster en Administración con Especialización  
en Innovación y Gestión de la Ciencia y Tecnología**

**Presentado por**

**Srta. Heidy Josefa Zelada Comeca  
Sr. Manuel Benjamin Ramírez Milla de León  
Sr. Rubén Dario Valenzuela Monge**

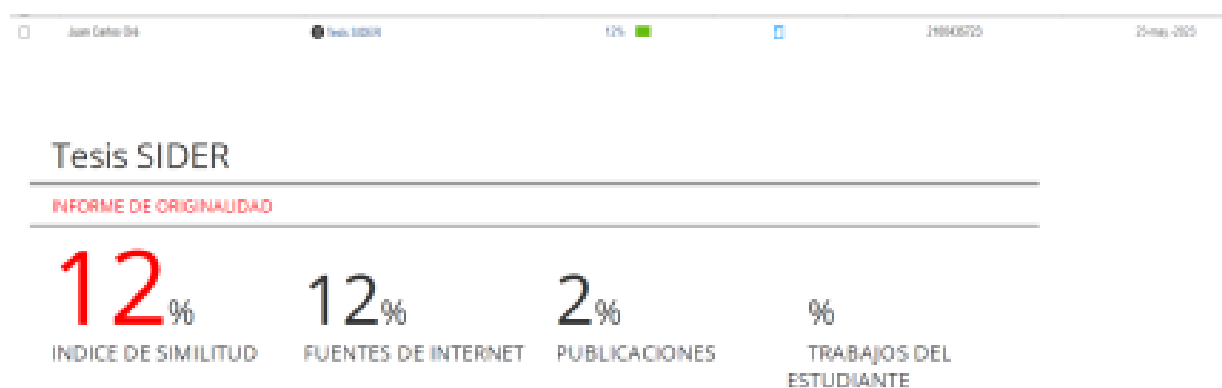
**Asesor: Juan Carlos Oré Oviedo**

**[0000-0001-9504-3932](tel:0000-0001-9504-3932)**

**Lima, Octubre 2022**

## REPORTE DE EVALUACIÓN DEL SISTEMA ANTIPLAGIO

A través del presente, JUAN CARLOS ORÉ OVIEDO deja constancia que el trabajo de investigación titulado “Creando valor con Datos & Analítica 4.0 en la industria de acero”, presentado por Doña Heidy Josefa Zelada Comeca, Don Manuel Benjamin Ramirez Milla de León y Don Rubén Dario Valenzuela Monge para, optar el Grado de Magister en administración con especialización en innovación y gestión de la ciencia y tecnología fue sometido al análisis del sistema antiplagio Turnitin del programa Blackboard el 22 de mayo 2023 dando el siguiente resultado:



Firma de asesor

Fecha: 23/05/2023

## TABLA DE CONTENIDO

<b>TABLA DE CONTENIDO .....</b>	<b>iii</b>
<b>ÍNDICE DE TABLAS .....</b>	<b>vi</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>vii</b>
<b>ÍNDICE DE ANEXOS.....</b>	<b>ix</b>
<b>CUERPO DEL TRABAJO .....</b>	<b>10</b>
<b>CAPÍTULO 1.- INTRODUCCIÓN.....</b>	<b>10</b>
1.1 Breve Historia de la Empresa .....	10
<b>CAPÍTULO 2.- PLANTEAMIENTO DEL PROBLEMA DE DATA &amp; ANALYTICS EN LA ORGANIZACIÓN.....</b>	<b>12</b>
2.1 Realidad problemática.....	12
2.1.1 Problemática General .....	12
2.1.2 Problemática: Datos y Analítica .....	13
2.2 Formulación del problema.....	13
2.3 Objetivos .....	13
2.3.1 Objetivos Generales.....	13
2.3.2 Objetivos Específicos .....	13
2.4 Justificación.....	14
2.4.1 Teórica.....	14
2.4.2 Práctica .....	14
2.4.3 Metodológica.....	16
2.4.4 Social.....	19
<b>CAPÍTULO 3.- ANTECEDENTES .....</b>	<b>21</b>
3.1 Internacionales .....	21
3.1.1 Tata Steel.....	21
3.1.2 Thyssenkrupp .....	22
3.1.3 Arcelormittal .....	22
3.1.4 Grupo <i>GERDAU</i> .....	23
3.2 Nacionales .....	24
3.2.1 Aceros Arequipa.....	24

<b>CAPÍTULO 4.- MARCO TEÓRICO.....</b>	<b>25</b>
4.1 Data & Analytics (D&A) .....	25
4.1.1 Analytics.....	25
4.2 Big Data .....	26
4.3 Internet de las cosas (IoT) .....	27
4.4 Inteligencia Artificial .....	28
4.5 Machine Learning.....	29
4.6 Industria 4.0 (avances de la comunicación y la conectividad) .....	31
4.7 Arquitectura Lambda.....	31
4.8 Data Management o Gestión de Datos .....	33
4.9 Data Governance o Gobierno de Datos .....	33
4.10 Data Warehouse .....	33
4.11 Data Lake .....	33
4.12 Activos Analíticos .....	34
<b>CAPÍTULO 5.- SITUACIÓN ACTUAL DE DATA &amp; ANALYTICS .....</b>	<b>35</b>
5.1 Estrategia de Data & Analytics actual .....	35
5.2 Ecosistema de datos.....	35
5.2.1 Fuentes de Datos .....	35
5.2.2 Microsoft Azure .....	36
5.2.3 Ecosistema Actual de Datos .....	37
5.3 Capital Humano.....	38
5.4 Activos analíticos .....	39
5.4.1 Modelos de Machine Learning .....	39
5.4.2 Dashboards y Reportes .....	39
5.5 Gobierno de Datos y Calidad de Datos .....	41
5.6 Modelo Operacional .....	44
5.7 Data Literacy / Culture .....	45
5.8 Data Compliance .....	45
5.9 Conclusiones de la situación actual.....	45
5.9.1 Conclusiones Generales.....	45
5.9.2 Nivel de madurez de la organización en data Analytics .....	46

<b>CAPÍTULO 6.- PLANTEAMIENTO DE LA SOLUCIÓN .....</b>	<b>47</b>
6.1 Definición de la estrategia de Data & Analytics futura .....	47
6.1.1 Estrategia .....	47
6.1.2 Indicadores de Medición de la Estrategia .....	49
6.2 Definición del Ecosistema de Datos .....	50
6.3 Definición del capital humano.....	51
6.4 Definición del portafolio de activos analíticos .....	54
6.5 Definición del plan de gobierno y calidad de datos .....	54
6.6 Presupuesto de la Solución.....	59
6.7 Revisión de la solución.....	60
6.7.1 Mapa Problema – Solución .....	60
6.7.2 Nivel de madurez objetivo para la organización en Data & Analytics .....	61
<b>CAPÍTULO 7.- PLAN DE IMPLEMENTACIÓN DE LA SOLUCIÓN .....</b>	<b>62</b>
7.1 Cartera de proyectos claves .....	62
7.1.1 Proyectos Fundacionales .....	62
7.1.1.1 CoE Data & Analytics .....	62
7.1.1.2 Data Lake Information Technology (IT) .....	64
7.1.1.3 Data Lake Operational Technology (OT).....	65
7.1.1.4 Data Governance .....	66
7.1.1.5 Data Literacy .....	67
7.1.2 Proceso de fabricación del acero de construcción .....	69
7.1.2.1 Optimización en uso de ferroaleaciones – U1 .....	70
7.1.2.2 Analítica Predictiva de Ventas – U2.....	71
7.1.2.3 Modelo de sobrecalentamiento – U3 .....	72
7.1.2.5 Modelo de visibilidad para el gruelo – U5 .....	75
7.2 Priorización y Definición de Precedencias .....	77
7.3 Evolución de indicadores de medición de la estrategia .....	77
7.4 Plan de Inversiones.....	80
<b>CAPÍTULO 8.- CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>81</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>82</b>
<b>ANEXOS     85</b>	

## ÍNDICE DE TABLAS

Tabla 1:Indicadores de la estrategia empresarial.....	11
Tabla 2:Objetivos específicos de la estrategia .....	13
Tabla 3:Uso de las metodologías .....	16
Tabla 4:The DAMA-DMBOK Framework.....	17
Tabla 5:Herbert Simon .....	29
Tabla 6:Modelos de ML Existentes.....	39
Tabla 7:Dashboards y Reportes .....	39
Tabla 8:Estrategia de D&A a tres años, 2023-2025 .....	47
Tabla 9:Roles de D&A .....	51
Tabla 10:D&A Head Count Ramp Up.....	53
Tabla 11:Portafolio de Activos Analíticos .....	54
Tabla 12:Proceso Unificado de Gobierno de Datos .....	56
Tabla 13:Presupuesto de la Solución .....	59
Tabla 14:Relación Problema, Objetivos y Proyectos.....	60
Tabla 15:Factores de Distribución de Costes a 3 años .....	62
Tabla 16:BC Optimización en uso de ferroaleaciones.....	63
Tabla 17:BC Data Lake Information Technology (IT) .....	65
Tabla 18:BC Data Lake Operational Technology (OT) .....	66
Tabla 19:BC Data Governance.....	66
Tabla 20:BC Data Literacy.....	68
Tabla 21:BC Optimización en uso de ferroaleaciones – U1 .....	71
Tabla 22:BC Optimización en uso de ferroaleaciones – U1 .....	73
Tabla 23:BC Mantenimiento Predictivo – U4.....	74
Tabla 24:BC Modelo de visibilidad para el gruero – U5.....	76
Tabla 25:Matriz de Beneficio y Complejidad por Proyecto.....	79
Tabla 26:Presupuesto General .....	80

## ÍNDICE DE FIGURAS

Figura 1:Organigrama .....	10
Figura 2:Modelo de madurez de gobierno de datos de IBM.....	17
Figura 3:AI Ladder Framework .....	18
Figura 4:Modelo de madurez de datos y analítica EIM Gartner .....	19
Figura 5:Las 4 Eras de la Analítica.....	26
Figura 6:Relación entre AI y ML.....	30
Figura 7:ML Life Cycle.....	30
Figura 8:Programación tradicional vs ML.....	30
Figura 9:Una Taxonomía de los Modelos de ML.....	30
Figura 10:Arquitectura Lambda.....	31
Figura 11:Bloques de la estrategia AS IS .....	35
Figura 12:Ecosistema Actual .....	37
Figura 13:Organigrama Actual de TITD .....	38
Figura 14:Nivel de Madurez de Data Governance Actual .....	41
Figura 15:IBM Data Governance Council Maturity Model .....	42
Figura 16:Gobierno de Datos - Nivel de Madurez Actual.....	43
Figura 17:Heatmap x Categoría.....	44
Figura 18:Heatmap x Fase.....	44
Figura 19:Why Maturity Matters? .....	45
Figura 20:Radar encuesta de madurez analítica – EIM Gartner.....	46
Figura 21:Buenas prácticas de D&A en la Estrategia de Negocios .....	49
Figura 22:Indicadores de Medición de la Estrategia.....	49
Figura 23:Ecosistema de Datos Propuesto .....	50
Figura 24:Organigrama de Data & Analytics.....	51
Figura 25:Data Governance AS IS vs TO BE .....	55
Figura 26:Proceso de Data Governance .....	55
Figura 27:Nivel de Madurez de Data Governance Objetivo.....	61

<b>Figura 28:Gráfico que compara la variación de madurez en DG AS IS vs TO BE .....</b>	<b>61</b>
<b>Figura 29:¿Por qué invertir en un Data Lake? .....</b>	<b>64</b>
<b>Figura 30:Gráfico proceso de producción del acero de construcción .....</b>	<b>69</b>
<b>Figura 31:Diagrama de Beneficio y Complejidad por Proyecto.....</b>	<b>77</b>



## ÍNDICE DE ANEXOS

<b>Anexo 1:Data Governance Maturity Model – AS IS .....</b>	<b>85</b>
<b>Anexo 2:Enterprise Information Management Maturity Self-Assessment – AS IS.....</b>	<b>98</b>

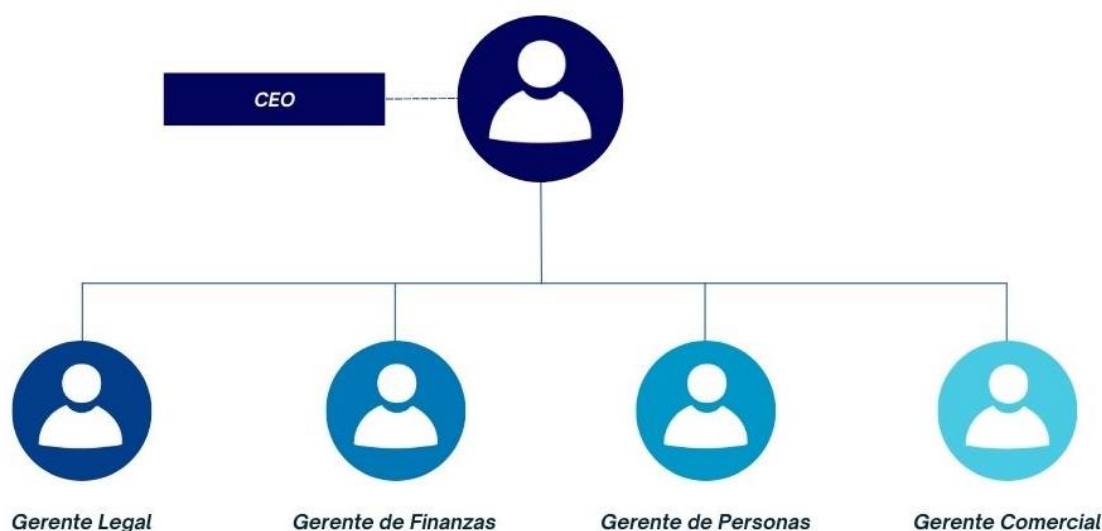
## CAPÍTULO 1.- INTRODUCCIÓN

### 1.1 Breve Historia de la Empresa

*SIDERPERU* es una empresa siderúrgica que inició su operación en 1956 en Chimbote – Perú, actualmente pertenece al grupo brasileño *GERDAU*. *SIDERPERU* produce y comercializa productos de acero tanto para el mercado peruano como para el extranjero. Sus productos atienden a los sectores de construcción civil, industria y minería.

*SIDERPERU* tiene una participación alrededor del 42% en el mercado peruano, su principal competidor es el otro fabricante local, Corporación Aceros Arequipa S.A. Estos dos fabricantes locales junto con empresas importadoras atienden la demanda del mercado peruano.

Figura 1: Organigrama



*SIDERPERU* al formar parte del grupo *GERDAU* está alineada a la evolución que ha venido experimentando, sin embargo, todavía está muy atrasada en temas de transformación digital y analítica avanzada.

*GERDAU* es la empresa productora de acero más grande de Brasil y una de las principales proveedoras de aceros largos en América y de aceros especiales en el mundo. “*GERDAU* está presente con operaciones industriales en 10 países y cuenta con más de 30 mil colaboradores directos e indirectos en el mundo. Tiene acciones cotizadas en las Bolsas de Valores de Sao Paulo, Nueva York y Madrid” *GERDAU* (2019, p. 12).

En el año 2019 *GERDAU* consolidó el propósito de la compañía en todos los niveles de la organización, este proceso se inició globalmente en el año 2018.

Los grandes desafíos que enfrentará el grupo *GERDAU*, para los próximos años, será el de mantenerse como una empresa sostenible y competitiva en la industria y convertirse en una organización cada vez más centrada en el cliente.

En los próximos 10 años *GERDAU* espera que el 20% de la facturación de la empresa se genere en nuevos negocios y productos.

Propósito *GERDAU*: “Empoderar personas que construyen el futuro” (GERDAU, 2019, p. 13) .

Aspiración *GERDAU*

“Ser en 10 años, una de las empresas de la cadena de valor de acero más rentables y admiradas en el mundo, y una de las más relevantes de las Américas” (SIDERPERU, 2022, p. 9).

Tabla 1: Indicadores de la estrategia empresarial

1.Seguridad	2.Colaboradores	3. EBITDA	4. Clientes	5. Sustentable
“Ambiente seguro”, cultura orientada a la seguridad.	Equipos altamente comprometidos y preparados.	Ser en 10 años una de las empresas de la cadena de valor más rentables y crecimiento de forma innovadora y sostenible.	Clientes satisfechos	Ser neutros en carbono para 2050 y ser referentes de ética y gobernanza.

*GERDAU* cuenta con una Política de Sostenibilidad, la cual establece las directrices para que todas sus unidades incluyendo *SIDERPERU* puedan conducir sus actividades considerando elementos económicos sociales, ambientales y de gobernanza, así como la evaluación de riesgos y oportunidades a ellos relacionados.

Un indicador clave para *SIDERPERU* es el EBITDA, este debe ser mayor a \$100,000.000 ó el equivalente al margen EBITDA mayor a 20%.

*GERDAU* cuenta con un desarrollo de capacidades de datos y analítica que viene mejorando los últimos años, principalmente en Brasil y Estados Unidos. La nube de *GERDAU* está en Amazon Web Services (AWS), la directiva vigente es la centralización de los datos administrativos y financieros de todos los países, pero el desarrollo independiente para los departamentos de Operaciones de las unidades de negocio.

## **CAPÍTULO 2.- PLANTEAMIENTO DEL PROBLEMA DE DATA & ANALYTICS EN LA ORGANIZACIÓN**

### **2.1 Realidad problemática**

#### **2.1.1 Problemática General**

##### **Segmentos de Mercado**

- Sector Construcción: Barras de Construcción, Fierro Habilitado, Rollo corrugado para construcción, Alambroón de Construcción, Alcantarillas y Guardavías, Alambre trefilado Clavos.
- Sector Industria: Tubos Electrosoldados laminados en caliente, laminados en frío y galvanizados), ángulos y perfiles laminados, Planchas Anti abrasivas, Mallas y Derivados.
- Sector Minería: Bolas de Molino y Barras Mineras para la molienda de minerales.

##### **Canales**

- Los productos de *SIDERPERU* se venden en todo el Perú a través de empresas de distribución (mayoristas) y fuerza de ventas directas. *SIDERPERU* tiene 3 centros de distribución ubicados en Chimbote, Lima y Arequipa.

##### **Competencia**

- La competencia directa de *SIDERPERU* es, Corporación Aceros Arequipa quien en estos últimos cinco años han incrementado su capacidad de producción cubriendo el 100% de la demanda de acero en el mercado peruano, esto genera que *SIDERPERU* busque nuevas formas de competir. La capacidad de producción de *SIDERPERU* es de 655 mil toneladas al año, el mercado total de acero peruano es de aproximadamente 1,6 millones de toneladas.

##### **Principales Riesgos**

- Concentración de ventas en un sólo producto, el 78% de las ventas es de barras de construcción.
- Dependencia de chatarra importada; no existe chatarra local suficiente para abastecer a los dos fabricantes locales.
- La informalidad del mercado peruano de la construcción (alineado con la informalidad en todas las demás actividades económicas) equivale al 70% del mercado total, generando como problemática la falta de alcance a la recopilación de datos del usuario o consumidor final.
- Incremento de participación de empresas importadoras en los tres últimos años, está entre 10%-13% aprox.

### 2.1.2 Problemática: Datos y Analítica

- No existe un departamento de datos y analítica, actualmente las actividades asociadas a datos son coordinadas por *TITD* con las unidades de negocio y proveedores.
- Las soluciones informáticas han generado silos de datos.
- No tienen implementado un gobierno de datos.
- No tienen implementado un proceso formal para la validación de la calidad de los datos.
- Existen múltiples requerimientos de data y analítica de las unidades de negocio.

### 2.2 Formulación del problema

- **Problema 1:** Generar una fuente de la verdad o repositorio centralizado, que almacene datos de calidad generando un ecosistema de datos para *SIDERPERU*, en el marco de un gobierno de datos, haciendo uso de la infraestructura, organización y recursos más apropiados.
- **Problema 2:** Hacer uso de Datos y Analítica para potenciar las capacidades de análisis, la validación de hipótesis, anticiparse y generar valor para *SIDERPERU*, a través de la generación de activos digitales.
- **Problema 3:** Contribuir a maximizar la rentabilidad de *SIDERPERU* haciendo uso de Datos y Analítica.

### 2.3 Objetivos

#### 2.3.1 Objetivos Generales

- Crear un marco de trabajo de Datos y Analítica en el marco de la estrategia de Datos y Analítica de *GERDAU – G1*.
- Generar una propuesta de valor para la industria del acero – *G2*.
- Generar una cultura del uso de datos en *SIDERPERU – G3*.

#### 2.3.2 Objetivos Específicos

Tabla 2: Objetivos específicos de la estrategia

Objetivo General	Cod.	Objetivo Específico
G1	E1	Crear el área de datos y analítica.
	E2	Implementar una plataforma tecnológica.
	E3	Implementar un gobierno de datos.
	E4	Implementar un modelo operativo.
G2	E5	Transmitir la relevancia y valor de los datos en <i>SIDERPERU</i> .
	E6	Desarrollar habilidades de datos y analítica.

G3	E7	Identificar los casos de uso de mayor relevancia y prioridad en el marco de los principios y objetivos de la Industria 4.0.
	E8	Reducción de labores operativas y foco en las labores de análisis.
	E9	Reducción de la merma.
	E10	Optimización en el uso de la materia prima.
	E11	Mantenimientos inteligentes, optimización de ciclos de mantenimiento.
	E12	Estimación de la demanda.

## 2.4 Justificación

### 2.4.1 Teórica

Las nuevas tecnologías conducen a prácticas innovadoras que otorgan a las empresas siderúrgicas una ventaja sobre sus competidores. Por ejemplo, la analítica predictiva puede mejorar la velocidad de producción y reducir los costos. Sin datos, las empresas no podrían generar conocimientos o identificar patrones para mejorar sus operaciones.

Comprender lo que esperan los clientes y tener la capacidad de anticipar el aumento de la demanda permite a los fabricantes de acero reducir el inventario, mejorar los márgenes y brindar un mejor servicio en general. Los análisis avanzados pueden segmentar la demanda en componentes individuales para hacer predicciones más precisas para cada categoría.

De acuerdo con un estudio de Boston Consulting Group sostiene que:

La Inteligencia Artificial (IA) está avanzando rápidamente, pero la mayoría de los fabricantes de acero aún no la ha aprovechado. Implementar IA requiere superar algunos desafíos claros en términos de datos y cultura organizacional, pero no requiere comenzar desde cero. La inversión necesaria para lanzar pilotos y comenzar a desarrollar capacidades es relativamente baja y, a menudo, se puede recuperar mediante aumentos en el rendimiento. La IA es una herramienta fácilmente disponible y las empresas siderúrgicas deben comenzar a utilizarlo. (Rodríguez et al., 2021)

Según el estudio (Rodríguez et al., 2021), el uso de la IA genera los siguientes beneficios:

- Reducir los costos de insumos de materia prima en más del 5%.
- Mejorar el rendimiento en los cuellos de botella en más del 6 %.
- Aumentar la productividad en más de un 15 %.

### 2.4.2 Práctica

De las entrevistas con los Stakeholders, se han podido identificar procesos que actualmente demandan recursos, tiempo y consecuentemente costos elevados para *SIDERPERU*.

## **General - Común a toda la empresa**

- Se vienen desarrollando múltiples iniciativas aisladas desde todas las áreas sin una coordinación que permita priorizarlas y evitar el traslape o redundancia de esfuerzos.
- Los proyectos hacen un uso ineficiente de los presupuestos al no ser parte de una estrategia general.
- No se desarrollan los proyectos en el marco de un gobierno de datos.
- No se desarrollan los proyectos bajo políticas o procesos que garanticen la calidad de los datos.

## **Departamento de Finanzas**

Compra de Divisas – Déficit de dólares americanos

Las ventas de *SIDERPERU* se cobran en dólares americanos y soles a sus clientes, pero sus obligaciones en una proporción significativa se pagan en dólares americanos, esto genera un déficit y oportunidad de optimización de costos.

## **Departamento de Operaciones**

Proyecto “Horarios Óptimos de Producción”

*SIDERPERU* ha desarrollado un proyecto en el departamento de operaciones que permite en función de los costes de la energía por hora establecer el plan óptimo de producción que reduce los costes de energía.

## **Departamento de Seguridad Industrial**

Proyecto “Safety Analytics”

*SIDERPERU* ha implementado una solución a partir de la información recopilada en el área de seguridad industrial a partir de diferentes fuentes (software y aplicativos en tablets y celulares para reportar las ocurrencias además de reportes en Excel entre otros) se han creado modelos de Inteligencia Artificial (IA) para prevenir riesgos de accidentes de trabajo a partir de datos históricos, con esta solución han minimizado la tasa de accidentabilidad así como la gravedad de los eventos no deseados (accidentes).

### 2.4.3 Metodológica

Tabla 3: Uso de las metodologías

Metodología	Evaluación de la Madurez Analítica	Evaluación de la Madurez del Gobierno de Datos	Adopción de la IA
DAMA Frameworks		X	
The IBM Data Governance Unified Process		X	
EIM Gartner	X		
AI Ladder Framework: the IBM Approach to Artificial Intelligence			X

#### **DAMA**

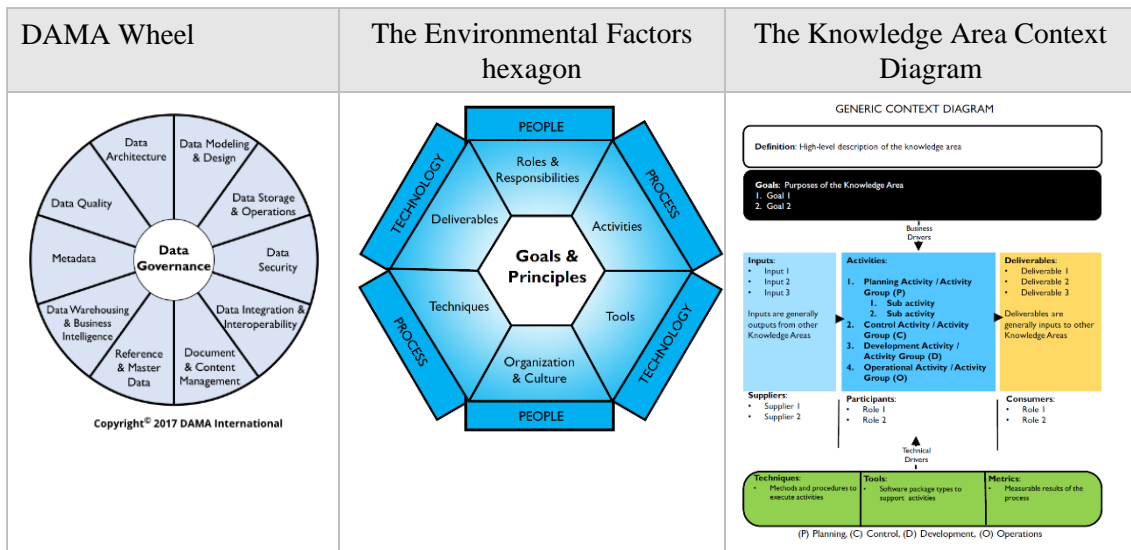
Es la principal organización internacional para profesionales de la gestión de datos. Su principal propósito es promover la comprensión, el desarrollo y la práctica de la gestión de datos e información como activos empresariales clave para apoyar a la organización (DAMA, 2022).

#### **DAMA Frameworks**

En su libro DMBOK segunda edición DAMA define la gestión de datos en 11 áreas de conocimiento y presenta cinco frameworks producto de las buenas prácticas y búsqueda de la excelencia. Los frameworks permiten a los gestores de datos en las empresas analizar en diferentes perspectivas la gestión de datos acorde a las necesidades específicas de sus organizaciones facilitando el ajuste de las estrategias, el desarrollo de roadmaps, el organizar equipos y alinear funciones.



Tabla 4: The DAMA-DMBOK Framework

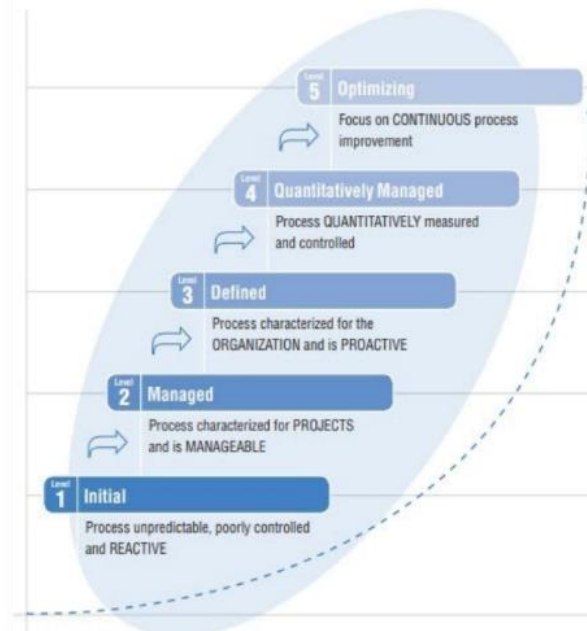


(Data Management Association, 2017, pp. 36-37)

### The IBM Data Governance Unified Process

El Modelo de Madurez del Consejo de Gobierno de Datos de IBM es una iniciativa innovadora diseñada con las aportaciones de un consejo de 55 organizaciones para crear consistencia y control de calidad en el gobierno a través de tecnologías empresariales probadas, métodos de colaboración y mejores prácticas.

Figura 2: Modelo de madurez de gobierno de datos de IBM

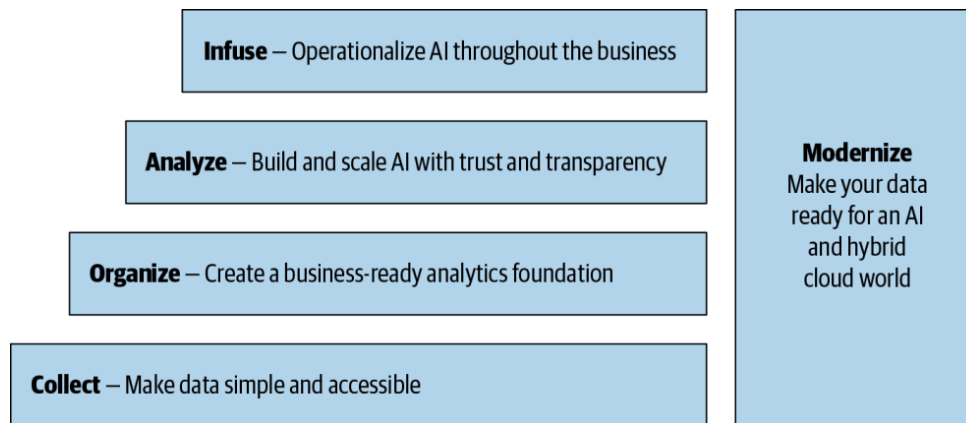


(Soares, 2012)

## AI Ladder Framework: the IBM Approach to Artificial Intelligence

“We often hear how clients struggle with their skills, and struggle with how to get a quick win with AI. According to Ritika Gunnar, VP of IBM Data and AI, the AI Ladder is a framework to help organizations build an information architecture, and ultimately determine where they are in their AI journey. It's a model for how we talk to clients about their data maturity, looking at how they collect, organize, and analyze data, and ultimately infuse AI throughout their organization” (Thomas & Zikopoulos, 2020)

Figura 3: AI Ladder Framework



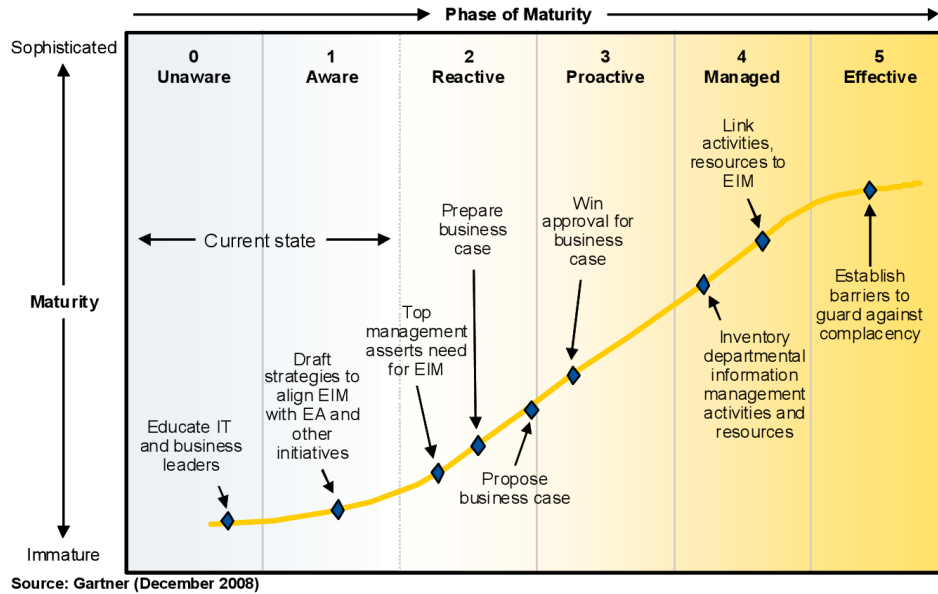
(Thomas & Zikopoulos, 2020)

## EIM Gartner

La idea detrás del Modelo de Información Empresarial de Gartner es utilizar los datos y la analítica para obtener los mejores resultados de negocio. Se ha desarrollado para que los líderes de datos y analítica puedan establecer una hoja de ruta para apoyar el desarrollo continuo.

Gartner considera que la gestión de la información empresarial no puede implementarse como un proyecto, sino más bien como un programa. Se considera que es transformacional y que se necesitan varios años para alcanzar el objetivo deseado. El EIM está apoyando la evaluación y la planificación del viaje de EIM. Con el modelo de madurez, los líderes pueden identificar la etapa de madurez y las acciones para alcanzar la siguiente etapa.

Figura 4: Modelo de madurez de datos y analítica EIM Gartner



(Laney, 2016)

#### 2.4.4 Social

##### Aporte de la estrategia de datos y analítica al propósito de SIDERPERU:

La estrategia de datos y analítica incluye el desarrollo de nuevas habilidades en los colaboradores de *SIDERPERU*, estas nuevas habilidades generarán un progreso individual que en una primera fase permitirá la promoción interna y en una segunda fase mayores y nuevas oportunidades laborales.

##### Actuación Social

La estrategia de datos y analítica alineada con la estrategia empresarial contribuirá en el cumplimiento de los objetivos financieros que permiten el sostenimiento de la estrategia de actuación social y su ampliación.

La estrategia de actuación social de *SIDERPERU*, contribuye para generar nuevas realidades, no sólo por medio del acero, sino principalmente por las personas. Las acciones de *SIDERPERU* en el marco social conectan con su propósito “empoderar a personas que construyen el futuro” (*SIDERPERU*, 2022), invierten en personas para que sean capaces de transformarse a sí mismas y a su entorno. Detallamos los 2 principales proyectos sociales de *SIDERPERU*.

##### Technical School

Es uno de los principales proyectos de responsabilidad social que tiene *SIDERPERU* que en alianza con el SENATI (Servicio Nacional de Adiestramiento en Trabajo Industrial) brinda formación gratuita a jóvenes egresados de secundaria, quienes reciben una formación teórica y práctica durante tres años, tras lo cual obtienen un título profesional acreditado.

*SIDERPERU* para atender este programa cuenta con una infraestructura adecuada y moderna, tiene talleres implementados con tecnología, su malla curricular está validada por el SENATI, cuenta también con una plana de profesores de esta misma entidad, garantizando una alta competitividad de sus egresados.

### **Orquesta Sinfónica Infantil y Juvenil**

*SIDERPERU* en alianza con dos centros culturales de la zona (Chimbote) contribuyen a la inclusión social promoviendo la educación a través de la música.

## **CAPÍTULO 3.- ANTECEDENTES**

La siderurgia es una de las mayores industrias manufactureras existentes desde que nació la primera revolución industrial. En las últimas décadas el sector ha experimentado transformaciones importantes ya que se han desarrollado procesos productivos y procesos altamente tecnológicos.

Según McKinsey(McKinsey, 2021) las empresas siderúrgicas han tenido un impacto positivo en lo digital y con la analítica. Esto se destacó en el Foro Económico Mundial cuando seleccionaron a 4 empresas siderúrgicas como los referentes de la industria 4.0 (POSCO South Korea , Bao Steel China, TATA Steel Netherlands, TATA Steel India,), además realizaron una encuesta a 30 empresas siderúrgicas a fin de comprender como están en su recorrido digital y analítica, concluyendo que hay 5 factores principales de éxito:

- Establecen objetivos y estrategias audaces.
- Las empresas invierten.
- Establecieron una arquitectura flexible de datos y tecnología.
- Desarrollaron un conjunto de habilidades.
- Establecieron un gobierno correcto detrás de los programas de datos y análisis.

### **3.1 Internacionales**

#### **3.1.1 Tata Steel**

Es una empresa siderúrgica ubicada en la India, cuya producción de acero anual es de 30 millones de toneladas métricas aproximadamente según el portal de estadística statista (Industria Siderúrgica: Mayores Productores de Acero Del Mundo En 2021 | Statista, n.d.)

#### **Modelo de sobrecalentamiento:**

El acero antes de adquirir la forma sólida es sometido a un proceso de sobrecalentamiento, con la finalidad de que adquiera la temperatura adecuada para la fundición. El acero se funde aproximadamente a 1,600 grados centígrados durante el sobrecalentamiento y tiene un rango de temperatura ideal. Si el acero sale demasiado caliente, los operadores del equipo deben ralentizar el paso de fundición. Si el acero no está lo suficientemente caliente, puede "congelarse" antes de que se haya fundido, lo que compromete su calidad.

Los científicos de datos de Tata Steel y McKinsey (McKinsey et al., 2021) utilizaron información histórica de la planta para construir y "entrenar" un modelo de optimización de sobrecalentamiento que examinaría los datos operativos en tiempo real y recomendaría puntos de ajuste del proceso conducentes a una mayor tasa de impacto, construyendo un modelo capaz de recomendar puntos de ajuste del proceso que elevarían la tasa de sobrecalentamiento al 90 por ciento de precisión.

El modelo de sobrecalentamiento produjo ahorros de costos anuales de \$ 4 millones.

### **3.1.2 Thyssenkrupp**

De acuerdo con (Industry 4.0 | Thyssenkrupp, n.d.), es una empresa siderúrgica ubicada en la Alemania que ha invertido en aplicaciones de Industria 4.0, detallamos algunas de ellas teniendo como referencia la información compartida por la misma empresa:

#### **Cadena de valor conectada (interactiva)**

Consiste en una red digital que rastrea toda la cadena desde el proveedor hasta los clientes, pasando por el fabricante de productos de acero seleccionados. Los procesos están conectados a una red digital en todo el ciclo de vida del producto. El sistema ciber físico le permite responder rápidamente a los requisitos del cliente en cuanto a plazos, incluso en relación con el material de partida ya en la etapa de producción de productos semiacabados (lote). Los clientes intervienen en la producción de productos. Los clientes realizan pedidos directamente en el sistema de TI de la planta y especifican cuándo se procesará su pedido. Los clientes pueden realizar cambios en las especificaciones, como el ancho y el grosor del producto, hasta que comience la producción. La cadena interactiva ahorra espacio y costos de almacenamiento de materiales y productos, liberando capital que puede ser utilizado para otros fines.

#### **Seguridad de datos**

En el espacio de datos industriales: sistemas (y dispositivos) para la transferencia rápida y segura de grandes cantidades de datos, tanto internamente como entre empresas; el resultado final es un espacio de datos industriales, con control total de sus datos por parte de los usuarios.

#### **Nuevos modelos comerciales con productos inteligentes**

Las empresas minoristas y los proveedores de servicios de logística están creando un ecosistema digital en el que todos están constantemente informados sobre la ubicación, el estado y la demanda de los productos para brindarles a los clientes todo lo que necesitan. Beneficios post-proyecto: ventaja competitiva gracias a nuevos modelos de negocio, productos y servicios inteligentes.

### **3.1.3 Arcelormittal**

Es una de las más grandes productoras de acero bruto del mundo, de acuerdo con información compartida por la misma empresa (Industry 4.0 | ArcelorMittal, n.d.) en sus plantas han instalado miles de sensores y dispositivos (hardware) para la recolección de grandes conjuntos de datos y su procesamiento. Estos son algunos proyectos:

- Creación de centros de excelencia digital cercanos a los sitios de producción en todo el mundo, permitiendo el crecimiento acelerado de las nuevas tecnologías desde el prototipo hasta la madurez.

- Recopilación de datos de decenas de miles de 'cosas' en la red de ArcelorMittal USA se utiliza buscando mejorar el rendimiento de las entregas. ArcelorMittal USA creó métricas de entrega para medir el desempeño de la entrega cliente por cliente. Esas métricas se personalizan para el cliente y cómo quieren recibir su material y se pueden ver desde prácticamente cualquier lugar.
- Los drones se están utilizando para mejorar la seguridad operativa, la eficiencia y la precisión.
- La automatización está conduciendo a una mayor productividad y calidad. Se están implementando depósitos de almacenamiento automatizados, vinculados a la programación de líneas y dispositivos de transporte, como grúas autónomas, y dan como resultado menos existencias y tiempos de entrega más cortos, dos importantes beneficios de la cadena de suministro.
- Los proyectos de inteligencia artificial (IA) el diseño de reconocimiento de imágenes y el modelo AI se utilizan para la toma de decisiones instantánea y automática para la liberación de soldadura en el laminador en caliente (planta de Canadá) o el reconocimiento de imágenes AI, uno para la medición del ancho de la bobina en frío y el otro para las emisiones de clasificación automática al medio ambiente (planta en Brasil). La IA ahorra tiempo y conduce a una mayor productividad.
- Modelos de hermanamiento digital utilizados para optimizar los recursos físicos y los procesos de producción utilizando datos recopilados de sensores.

#### **3.1.4 Grupo GERDAU**

En información compartida por la misma empresa GERDAU (2019) detallan algunas de sus iniciativas de transformación digital.

- Inteligencia artificial en la seguridad del trabajo: Son modelos de inteligencia artificial (IA) para prevenir accidente de trabajo, evitando que los riesgos se conviertan en accidentes, estos modelos han sido generados a partir del análisis de datos históricos.
- Inteligencia artificial en la producción del acero: Es un modelo de inteligencia artificial (IA) que permite optimizar el consumo de energía en los hornos de acería.
- Inteligencia artificial en la auditoría: Tienen auditores que programan en lenguaje Phyton, usan directamente el Data Lake de GERDAU para buscar los datos que requieren para realizar sus tareas.
- Inteligencia artificial en la gestión de la producción: Es un modelo de inteligencia artificial (IA) que fue desarrollado para dar soporte a los decisores y puedan equilibrar sus variables principales en la cadena de producción y comercialización.

## **3.2 Nacionales**

### **3.2.1 Aceros Arequipa**

Es una empresa siderúrgica peruana, ubicada geográficamente al sur del país, exactamente en el distrito de Paracas, provincia de Pisco, departamento de Ica.

De acuerdo con su memoria anual (“Memoria Anual Integrada 2021,” 2022), Aceros Arequipa impulsa las iniciativas digitales de todas las áreas de la compañía con la finalidad de alcanzar sus objetivos.

Durante el año 2021 desarrollaron más de 200 nuevos proyectos de digitalización e innovación, donde destacan dos proyectos:

#### **Aceros PRO:**

Es una herramienta digital (aplicación) que tiene como objetivo acelerar el proceso de compra y simplificar los procesos comerciales que se asocian a ello. El cliente en esta aplicación accede al catálogo de productos, puede ingresar y monitorear sus pedidos, puede acceder a su estado de cuenta y además puede colocar cualquier reclamo que tenga.

#### **Scrapcheck**

Es una herramienta digital (aplicación) de uso interno que reemplazó el uso de formatos en físicos, esto ha permitido agilizar el proceso de calificación para el ingreso de chatarra y su registro de la información en línea.

Cabe mencionar que desde el 2017 Aceros Arequipa ofrece al mercado de construcción formal productos con valor agregado:

- **ACEDIM BIM:** Es el desarrollo de la ingeniería de detalle en 3 dimensiones del acero de construcción, aplicando la metodología BIM soportado por la filosofía Lean Construction. Esto permite al proveedor y cliente la gestión de sus procesos de construcción.
- **Steel Track:** Es una plataforma que integra la todo el proceso de venta de los productos de la línea ACEDIM, ingeniería, comercial, planificación, planta, suministro, cliente, instalación.



## CAPÍTULO 4.- MARCO TEÓRICO

### Glosario

Acrónimo	Definición en Inglés	Definición en Español
SCADA	Supervisory Control and Data Acquisition	Supervisión, Control y Adquisición de Datos
MES	Manufacturing Execution System	Sistema de Ejecución de Manufactura
ERP	Enterprise Resource Planning	Planificación de Recursos Empresariales

#### **4.1 Data & Analytics (D&A)**

Data & Analytics se refiere a las formas en que se gestionan los datos para soportar todos los casos de uso y al análisis de los datos para impulsar mejores decisiones, procesos de negocio y resultados, tales como descubrir nuevos riesgos de negocio, desafíos y oportunidades. (Gartner, n.d.)

##### **4.1.1 Analytics**

La Analítica o Analytics, es todo el conjunto de soluciones que permite extraer y recolectar los datos de diferentes fuentes, ya sean sensores, máquinas, procesos industriales, sistemas, aplicaciones, entre otros, para prepararlos y puedan ser analizados y explotados, generando valor al negocio.

En la industria de aceros, las soluciones de analítica tienen tres objetivos principales:

- Extraer y almacenar los datos.
- Organizar la información de forma automatizada.
- Disponibilizar los datos, con políticas de seguridad, para que los usuarios los consuman a demanda y puedan tomar decisiones de negocio.

Según Thomas H. Davenport, un conocido gurú de Business Analytics, recientemente clasificó al Analytics en 4 eras de sofisticación, que van desde la analítica descriptiva hasta la inteligencia artificial. (Davenport, 2018)

##### **Analytics 1.0**

La era de la analítica descriptiva o “Business Intelligence”, fase que dominó la analítica en los negocios por décadas, donde el valor fue impulsado en gran medida por acciones reactivas en lugar de capacidades predictivas avanzadas o conocimientos estadísticos granulares. La mayoría de las tareas analíticas se realizaban de manera “artesanal”, con mucha mano de obra, donde se tomaba mucho tiempo para preparar la data y muy poco tiempo para analizarla.

##### **Analytics 2.0**

La era del Big Data, donde aparecieron potentes plataformas de gestión de datos (como Hadoop) y se empezó a innovar en ofertas basadas en información (Google, Facebook, LinkedIn), llevando al surgimiento de datos. El objetivo principal en este período cambió a “productos de datos”, construidos alrededor de datos y analítica para ser usados por los clientes.

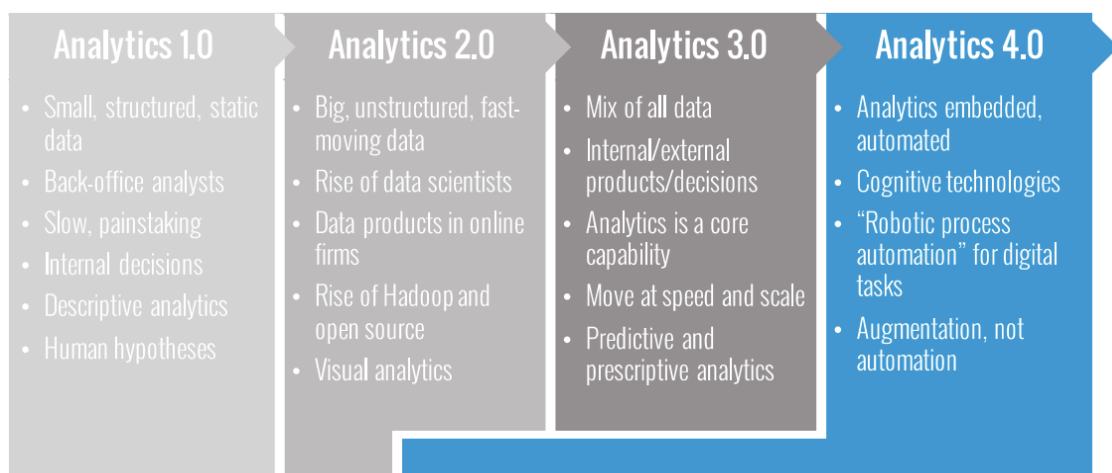
### Analytics 3.0

La era de las ofertas enriquecidas con datos, en la que las empresas de industrias tradicionales también adoptaron Big Data y Analytics. El sello de Analytics 3.0 es que las empresas que existen hace mucho tiempo transformen sus modelos de negocio y cultura con el uso extensivo de analítica. En esta era, las empresas de gran escala crean productos basados en datos e incorporan miles de modelos de aprendizaje automático (machine-learning).

### Analytics 4.0

La era de la inteligencia artificial, es el siguiente paso en la sofisticación analítica para las organizaciones.

Figura 5: Las 4 Eras de la Analítica



## 4.2 Big Data

De acuerdo con la definición de McKinsey (Manyika et al., 2011) Big data se refiere a aquellos conjuntos de datos cuyo tamaño excede la capacidad de las herramientas actuales de captura, almacenamiento, carga y análisis.

En el proceso de producción de las empresas industriales se capturan gran cantidad de datos, por ejemplo, a partir de los sensores, reportes, al usar aplicaciones, etc. Esto genera una gran cantidad de datos en un período determinado trayendo como consecuencia hacer inviable el procesamiento de la información con el software de uso común, a toda esta información podríamos denominarla Big Data.

Para muchos el Big Data es el combustible de la cuarta revolución industrial, y los tipos de datos que considera son:

- Estructurado: Son todos los datos que se pueden “almacenar, acceder y procesar en forma de formato fijo”, ejemplo: tabla de MS Excel de una planilla de pago.
- No estructurado: Cualquier dato con forma o estructura desconocida, ejemplo: email, archivos PDF, chats, etc.
- Semiestructurado: Estos tipos de datos pueden contener ambas formas de datos, ejemplo: archivo XML.

Big Data no es sólo tener un gran volumen de datos, los datos deben ser significativos y deben generar resultados beneficiosos para las organizaciones.

“Requirements of Big Data

Big Data is a term that is much used, both in the literature on health care and other areas of research. Although there is no universally accepted definition, the term generally refers to a large data set characterized by the ‘5 Vs’:

1 Volume, 2. Variety, 3. Velocity, 4. Veracity, 5. Value.”(Rumbold et al., 2020)

- El **volumen** se refiere a la cantidad de datos (volumen inmenso ya no en gigabytes y terabytes sino de petabytes de datos)
- La **variedad** se refiere a las variedades numéricas, nominales y ordinales de los datos.
- La **velocidad** se refiere a la rapidez con la que se pueden crear, almacenar o procesar los datos.
- La **veracidad** denota la fiabilidad y exactitud de los datos.
- El **valor** se refiere al valor de los datos.

Aunque las cinco V no son absolutas para el análisis de Big Data, el requisito básico es la necesidad de tecnologías y técnicas especiales para analizar un conjunto de datos complejo.

### **Beneficios del análisis de Big Data**

El mayor beneficio que se puede obtener del análisis del Big Data es la detección de patrones, permite tener una mejor comprensión de las correlaciones y dependencias, hacer modelos predictivos.

### **4.3 Internet de las cosas (IoT)**

Según Oracle Perú:

La Internet de las cosas (IoT) describe la red de objetos físicos (cosas) que llevan incorporados sensores, software y otras tecnologías con el fin de conectarse e intercambiar datos con otros dispositivos y sistemas a través de Internet. Estos dispositivos van desde

objetos domésticos comunes hasta herramientas industriales sofisticadas. (¿Qué Es El Internet de Las Cosas (IoT)? | Oracle Perú, n.d.)

### **El IoT y la Industria 4.0**

El IoT ó Industrial Internet of Things ha surgido debido a la digitalización y a la necesidad de fabricación inteligente, donde ha sido posible integrar OT y dominios de TI, ya que muchos procesos se están automatizando.

El IoT es la forma más adecuada de lograr eficiencia operacional, y tiene mucha más importancia en comunicación máquina a máquina y sistemas de misión crítica que tener un alto volumen de datos (Kebande, 2022).

Una de las principales aplicaciones del IoT en el sector industrial es la posibilidad de obtener información de los dispositivos hiperconectados, que ayuda a comunicar a los puestos de responsabilidad o de decisión estratégica el estado de los procesos. Esto también permite administrar en remoto los activos de la planta, por lo que el concepto de automatización se explota más que nunca (Las Principales Aplicaciones Del IoT En La Industria | Meinsa, n.d.).

### **4.4 Inteligencia Artificial**

Según la Real Academia Española se define como “Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” (RAE, 2022), en adelante simplemente IA.

IA es la inteligencia que pueden generar los computadores y máquinas robóticas controladas, el término “artificial” permite contrastar con la inteligencia natural que es producida por los seres humanos y algunos animales. En esencia, la inteligencia artificial es una ciencia que permite emular/imitar en escenarios específicos las funciones cognitivas de las mentes humanas esto incluye: comprender, expresar, percibir, calcular, recordar, organizar, razonar, imaginar, crear y resolver problemas.

El término Inteligencia Artificial fué acuñado por John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon en su artículo (McCarthy et al., 2006).

Pese a ser un concepto con más de 67 años, el cambio en las capacidades de captura y procesamiento de datos, así como los avances tecnológicos han colocado a la IA en la actualidad como la capacidad de mayor interés para las compañías a la búsqueda de innovar y diferenciarse. Como principales triggers identificamos:

### **Interés de las industrias**

Para poder entender la magnitud económica del interés de las industrias tenemos datos del reporte (HAI, 2022) que indica:

“Private investment in AI soared while investment concentration intensified. The private investment in AI in 2021 totaled around \$93.5 billion—more than double the total private investment in 2020, while the number of newly funded AI companies continues to drop, from 1051 companies in 2019 and 762 companies in 2020 to 746 companies in 2021. In 2020, there were 4 funding rounds worth \$500 million or more; in 2021, there were 15.”

#### 4.5 Machine Learning

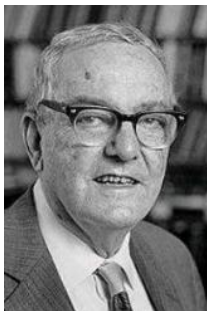
“Programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort.” (Samuel, 2000)

¿Qué es aprendizaje?

“Learning is any process by which a system improves performance from experience.” by Herbert Alexander Simon.

“Machine Learning is concerned with computer programs that automatically improve their performance through experience” by Herbert Alexander Simon.

Tabla 5: Herbert Simon



Premio Turing de la ACM (Association for Computing Machinery) junto con Allen Newell por haber dado “contribuciones fundamentales a la inteligencia artificial, a la psicología cognitiva y al tratamiento de las listas”, 1975.

Premio Nobel en economía “por su investigación pionera sobre el proceso de toma de decisiones en las organizaciones económicas”, 1978.

Podemos concluir que es el conjunto de métodos a través de los que los computadores son entrenados para aprender logrando el “aprendizaje automático”, la principal diferencia es que se aprende de los datos y no de reglas de decisión programadas por expertos en ciencias de la computación, en adelante simplemente ML. ML es un subconjunto de la inteligencia artificial.

Figura 6: Relación entre AI y ML



Por lo tanto, ML es la automatización de la extracción de conocimiento. Un programa que es capaz de aprender automáticamente un determinado comportamiento o patrón a partir de observaciones o ejemplos y una medida del rendimiento si su desempeño en las observaciones mejora con la experiencia.

Figura 7: ML Life Cycle



Figura 8: Programación tradicional vs ML

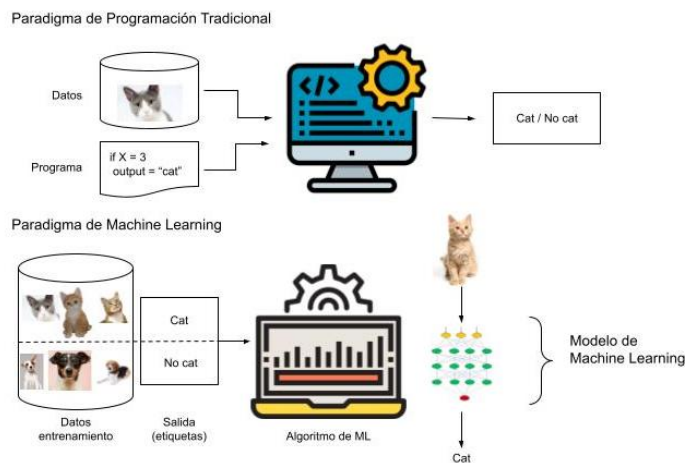
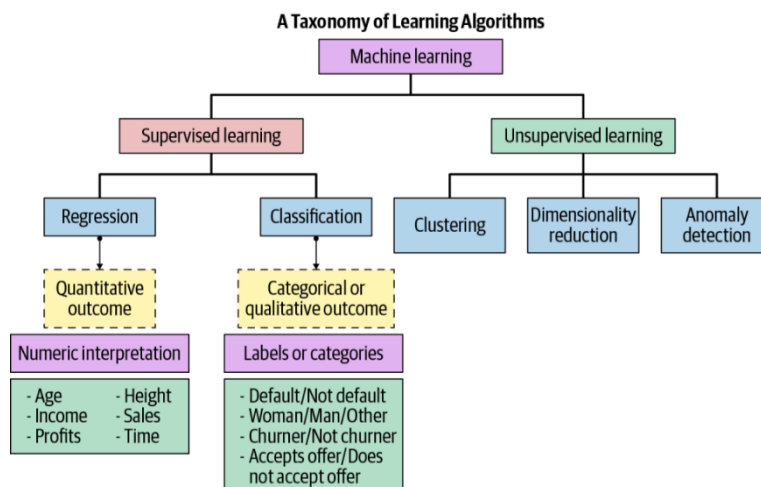


Figura 9: Una Taxonomía de los Modelos de ML



(Vaughan, 2020, p. 160)

#### 4.6 Industria 4.0 (avances de la comunicación y la conectividad)

Es la cuarta revolución industrial, que ha permitido que los fabricantes cambien la forma de operar sus negocios.

La industria 4.0 combina técnicas avanzadas de producción y operaciones con tecnologías inteligentes que se integran en las organizaciones, las personas y los activos; es decir es la aplicación masiva de nuevas tecnologías de comunicación e Internet a todos los procesos de fabricación, de modo que el funcionamiento de esta sea inteligente y eficiente. Se genera un mayor valor cuando los datos de las operaciones de producción se combinan con los datos operativos del ERP, la cadena de suministro, el servicio de atención al cliente y demás sistemas.

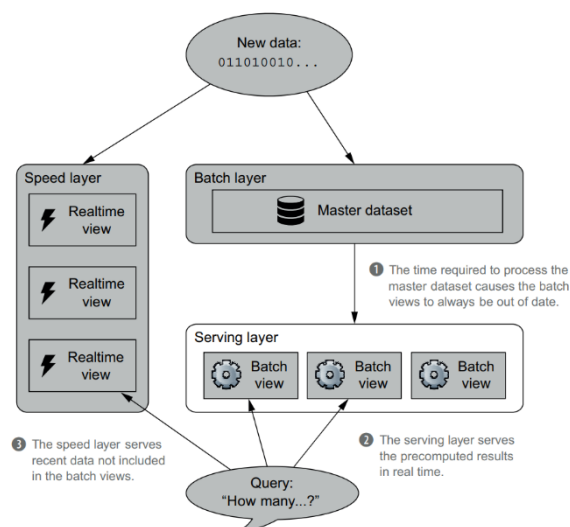
La recopilación y análisis de las grandes cantidades de datos de los sensores en la planta de producción, por ejemplo, pueden proporcionar herramientas para realizar mantenimiento predictivo y así minimizar el tiempo de inactividad de un equipo.

Por lo tanto, el core de la industria 4.0 es la digitalización continua de los procesos de producción; es decir se conecta el mundo de los dispositivos, la planta de producción, el área de TI y las oficinas, generando una gran cantidad de datos necesarios para optimizar la producción a partir de hacer utilizables los datos.

#### 4.7 Arquitectura Lambda

La arquitectura Lambda es una forma de procesar cantidades masivas de datos (es decir, "Big Data") que proporciona acceso a métodos de procesamiento por lotes y de procesamiento de flujos con un enfoque híbrido. La arquitectura lambda se utiliza para resolver el problema de la computación de funciones arbitrarias. La arquitectura lambda propiamente dicha se compone de 3 capas:

Figura 10: Arquitectura Lambda



(Marz & Warren, 2015)

### **Capa de lotes (Batch Layer)**

Los nuevos datos llegan continuamente, como una alimentación del sistema de datos. Se alimentan a la capa de lotes y a la capa de velocidad simultáneamente. Ésta examina todos los datos a la vez y finalmente los corrige en la capa de velocidad. Aquí podemos encontrar múltiples ETL y un almacén de datos tradicional. Esta capa se construye con un horario predefinido, normalmente una o dos veces al día. La capa de lotes tiene dos funciones muy importantes:

- Gestionar el conjunto de datos maestro.
- Precomputar/pre calcular las vistas por lotes.

### **Capa de servicio (Serving Layer)**

Las salidas de la capa de lotes en forma de vistas por lotes y las procedentes de la capa de velocidad en forma de vistas casi en tiempo real se envían a la capa de servicio. Esta capa indexa las vistas por lotes para que puedan ser consultadas en baja latencia sobre una base ad-hoc.

### **Capa de velocidad (Stream Layer)**

Esta capa se encarga de los datos que no se entregan en la vista por lotes debido a la latencia de la capa por lotes. Además, sólo se ocupa de los datos recientes con el fin de proporcionar una visión completa de los datos al usuario mediante la creación de vistas en tiempo real.

### **Ventajas de las arquitecturas lambda**

Estos son los principales beneficios de las arquitecturas lambda:

- Sin gestión de servidores: no es necesario instalar, mantener o administrar ningún software.
- Escalado flexible: su aplicación puede ser escalada automáticamente o escalada por el ajuste de su capacidad
- Alta disponibilidad automatizada: se refiere al hecho de que las aplicaciones sin servidor ya tienen incorporada la disponibilidad y la tolerancia a los fallos. Representa una garantía de que todas las solicitudes obtendrán una respuesta sobre si fueron exitosas o no.
- Agilidad de negocio: Reacciona en tiempo real a los escenarios cambiantes del negocio/mercado.

### **Desafíos de las arquitecturas lambda**

Complejidad: las arquitecturas lambda pueden ser muy complejas. Los administradores suelen tener que mantener dos bases de código separadas para las capas de lotes y de streaming, lo que puede dificultar la depuración.



#### **4.8 Data Management o Gestión de Datos**

La Gestión de Datos consiste en las prácticas, las técnicas de arquitectura y las herramientas para lograr un acceso consistente y la entrega de datos en todo el espectro de áreas sujetas de datos y tipos de estructuras de datos en la empresa, para cumplir con los requisitos de consumo de datos de todas las aplicaciones y procesos de negocio.(Data Management Association et al., 2017)

#### **4.9 Data Governance o Gobierno de Datos**

El Gobierno de Datos es la orquestación de políticas, estándares, procesos, roles y responsabilidades con el fin de permitir a la organización tomar ventaja de sus datos, manejándolos como activos estratégicos y poder cumplir objetivos de negocios, con eficiencia, efectividad y control.

#### **4.10 Data Warehouse**

Es una base de datos cuya información proviene de una o varias fuentes de datos, como por ejemplo sistemas transaccionales u otras bases de datos relacionales.

Los datos recopilados pueden ser estructurados, semiestructurados o no estructurados. Una vez que están integrados en esta base de datos son tratados y transformados. El objetivo de ello es que los usuarios puedan acceder a estos datos con la ayuda de herramientas de Business Intelligence, de cliente SQL o de tablas.

Para la creación de un Data Warehouse es necesario la ejecución de procesos ETL (Extracción, Transformación y Carga) a partir de los sistemas operacionales de una empresa.

##### **Extracción**

Recopilación de la información de las diferentes fuentes tanto internas como externas.

##### **Transformación**

Filtrado, limpieza, depuración, homogenización y agrupación de la información.

##### **Carga**

Organización y actualización de los datos y los metadatos en la base de datos.

#### **4.11 Data Lake**

Es un repositorio de datos centralizado diseñado para almacenar, procesar y gobernar grandes cantidades de datos estructurados, semiestructurados o no estructurados; es decir es el lugar donde se almacena los datos en bruto los cuales podemos procesar, ignorando los límites y tamaños.

De acuerdo con lo referido por IBM (Data Lakes I IBM, n.d.) los beneficios de un Data lake son:

##### **Mas flexibles**

Al no tener restricción con el almacenamiento de datos, los Data Lake son ideales para proyectos de análisis avanzado y machine learning.

### **Coste**

Los Data Lake no requieren de mucha planificación inicial para la recopilación de datos, la inversión es menor en recursos humanos. Los costos de almacenamiento son más bajos comparados con otros repositorios de almacenamiento.

### **Escalabilidad**

Los Data Lake proporcionan una Sandbox para que los trabajadores desarrollen POC exitosos. Cuando el proyecto ha demostrado valor a menor escala, es más fácil expandir ese flujo de trabajo a mayor escala por medio de la automatización.

### **Silos de datos reducidos**

Los silos de datos dentro de una organización se eliminan porque ya no hay un único propietario para un conjunto de datos determinado.

### **Experiencia del cliente**

Este beneficio no se ve de inmediato pero una prueba de concepto exitosa puede mejorar la experiencia general del usuario, permitiendo a los equipos comprender mejor y personalizar el proceso del cliente a través del análisis profundo y nuevos en la red.

#### **4.12 Activos Analíticos**

Son los datos transformados a partir de cualquier técnica utilizada y que son de utilidad para una empresa, pudiendo ser desde una visualización en un tablero hasta un modelo de ciencia de datos.

## CAPÍTULO 5.- SITUACIÓN ACTUAL DE DATA & ANALYTICS

### 5.1 Estrategia de Data & Analytics actual

En el marco de la estrategia empresarial de SIDERPERU viene desarrollando un proceso de transformación cultural desde el año 2015 y un proceso de transformación digital desde el año 2019.

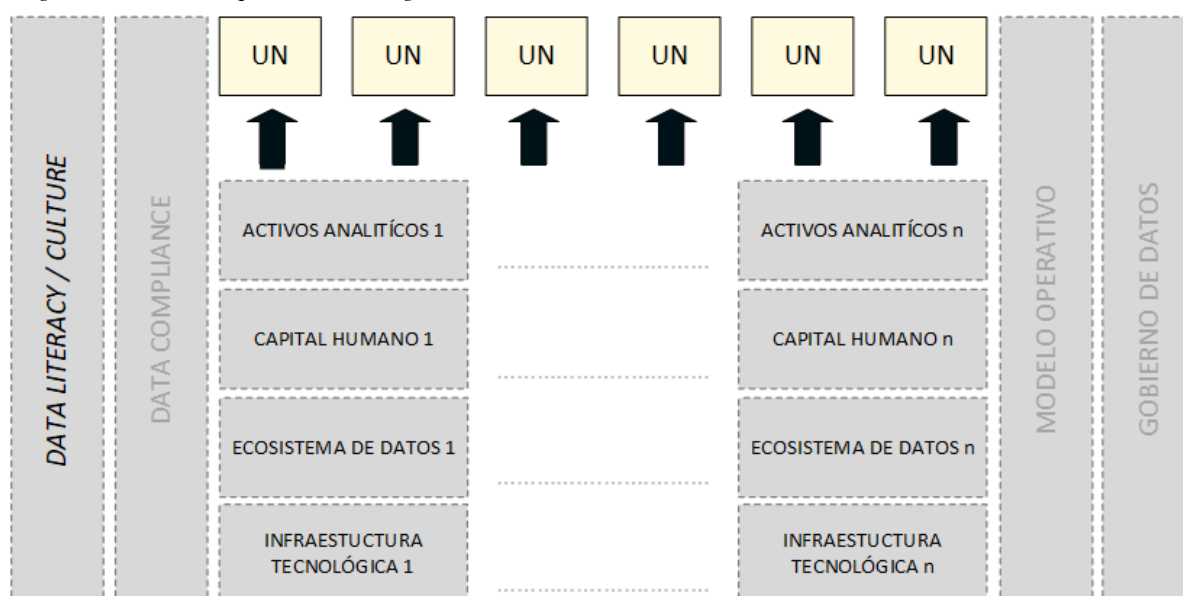
Luego de entrevistar a los Stakeholders y de analizar la documentación correspondiente, identificamos múltiples proyectos independientes de analítica, así como el interés de todos los niveles de la organización en desarrollar las capacidades de Datos y Analítica para tomar mejores decisiones.

Como consecuencia de lo anterior:

- Se tienen silos de datos.
- Cada implementación sigue un estándar propio.

Para describir mayores detalles a continuación se presentan nueve bloques temáticos.

Figura 11: Bloques de la estrategia AS IS



### 5.2 Ecosistema de datos

#### 5.2.1 Fuentes de Datos

##### ERP SAP

Es el ERP de SIDERPERU y estándar del grupo GERDAU, en este se registra la información de las órdenes de producción y que son transmitidas a la planta. En SAP registran el costo del producto, planeamiento, programación y control de la producción). SIDERPERU migró a este ERP el año 2010.

##### GMIS (GERDAU manufacturing integration system) y Nivel 2

Son sistemas MES (Sistema de Ejecución de Manufactura) que permite el control remoto de los diversos dispositivos de la planta de producción.

### **Sistemas SCADA**

Son sistemas directamente conectados a los equipos, los operadores monitorean el proceso de producción a través de los tres (3) sistemas independientes, parte de la información de estos sistemas se carga a SAP.

### **Fractal**

Sistema de gestión de las planillas.

### **DB Externas**

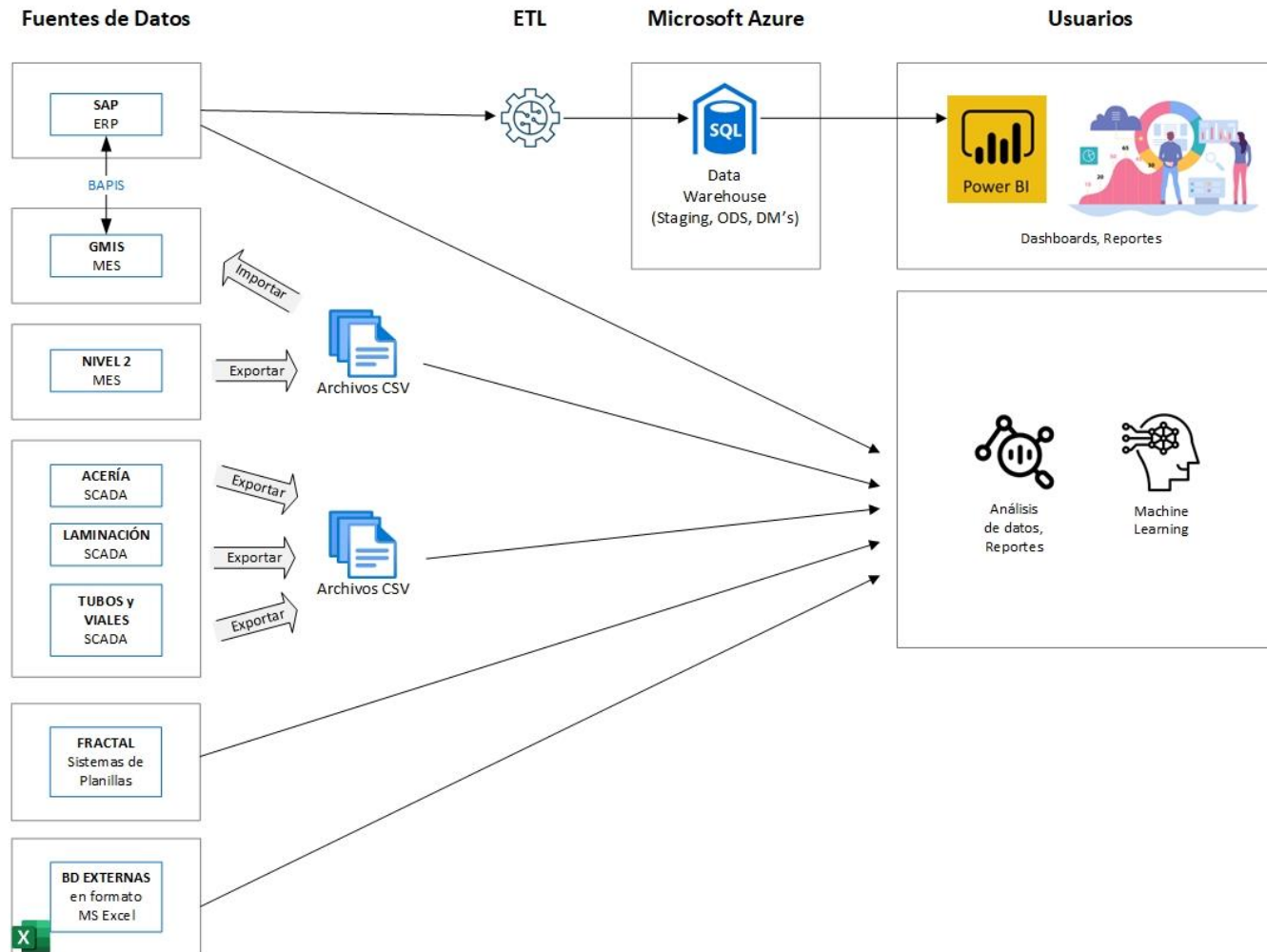
En determinados periodos del año se adquieren bases de datos externas en formato MS Excel para diversos propósitos.

#### **5.2.2 Microsoft Azure**

Es la plataforma Data Warehouse que permite disponibilizar los datos de SAP en cubos que a través de Power BI y son consultados para la generación de reportes y Dashboards.

### 5.2.3 Ecosistema Actual de Datos

Figura 12: Ecosistema Actual



### Observaciones

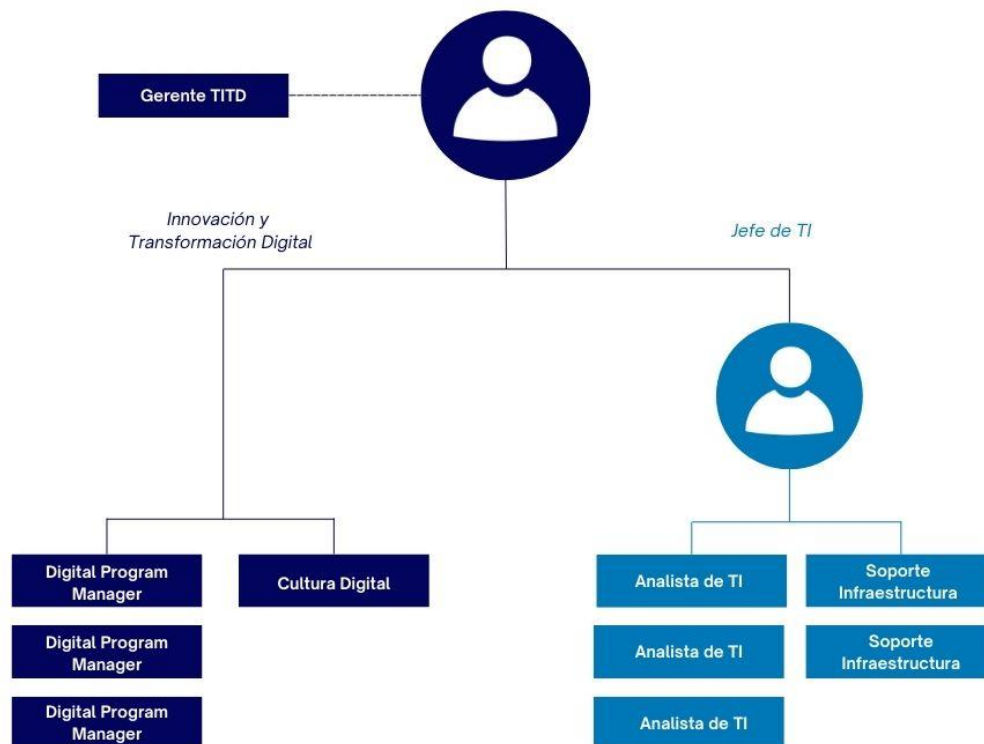
- El ERP SAP se encuentra en la red de GERDAU Brasil y se accede a través de un enlace MPLS desde SIDERPERU.
- Los sistemas MES se encuentran On Premise en el data center de Chimbote y tienen una arquitectura cliente servidor.
- Los sistemas SCADA se encuentran aislados, tienen un proyecto para montar un data center e interconectarlos.
- El Data Warehouse se encuentra en Microsoft Azure - cloud.

### 5.3 Capital Humano

SIDERPERU desde el 2021 formó la gerencia de Tecnología de la Información y Transformación digital TITD, con el objetivo de promover la incorporación de tecnologías innovadoras en búsqueda de la mejora del negocio y adaptación al cambio, alineando personas, procesos y sistemas. (“Memoria Anual 2021,” 2022)

Actualmente el equipo TITD está conformado por 10 personas más el gerente responsable.

Figura 13: Organigrama Actual de TITD



#### Transformación Digital (TD)

- Son los responsables de la comunicación de los temas de cultura.
- Son promotores y responsables del SIDERPERU TANK.
- Gestionan proyectos siempre con respaldo de TI.
- Son expertos en metodologías ágiles.

#### Tecnologías de la Información (TI)

- Da soporte técnico a todos los colaboradores/usuarios
- Ejecutan proyectos con el respaldo de TD.
- Supervisan y asesoran a las áreas que adquieren un servicio informático externo.

- Entregan los servicios electrónicos (Laptos, PCs, celulares, correos, software, etc).

## **SQUAD Industria 4.0**

Dentro de las iniciativas de transformación digital *SIDERPERU* cuenta con un equipo denominado Squad industria 4.0, el foco de este equipo son las iniciativas industriales. El Squad industria 4.0 es un equipo multidisciplinario conformado por representantes de diferentes áreas:

- Sponsor: Gerente Industrial/jefe de Tecnología de la Información
- Project manager: Un representante del área de Tecnología de la Información.
- Líder técnico: Un representante del área de Tecnología de la Información.
- Equipo de ejecución: Representantes de las áreas Acería (1), Laminación (2), Tubos y viales (1), Ingeniería (2) y suministros (1).

## **5.4 Activos analíticos**

### **5.4.1 Modelos de Machine Learning**

Tabla 6: Modelos de ML Existentes

Proyecto	Descripción	Arquitectura Tecnológica
Safety Analytics	Es un modelo que ayuda a reducir eventos no deseados y dar continuidad a la operación libre de accidentes.	-On Premise -Aplicación Web
Modelo de optimización de planes de producción	Es un modelo para reducir costes de energía eléctrica, genera un plan de producción que optimiza los costes en función de las ordenes de producción y los costes de la energía por hora. Ha sido desarrollado internamente por un Data Scientist que finalizada la implementación dejó <i>SIDERPERU</i> .  Se ha contratado a un consultor local para el desarrollo del Frontend Web del proyecto.	-On Premise -Aplicación Web
Tu Salud	Es un modelo para detectar de manera temprana síntomas de COVID-19 de todos los colaboradores de <i>SIDERPERU</i> , permitiéndoles tomar acciones frente a dicha enfermedad.	-On Premise -Aplicación Web

### **5.4.2 Dashboards y Reportes**

Tabla 7: Dashboards y Reportes

Dashboard	Usuarios	Descripción	Arquitectura Tecnológica
Sistema IBA (Historiador)	Operaciones	<p>Los principales equipos de acería están en este sistema, en el cual se visualiza en tiempo real las variables del proceso.</p> <p>Son 1,300 señales aproximadamente que se pueden monitorear:</p> <ul style="list-style-type: none"> <li>• La información recolectada es a través de sensores.</li> <li>• Es una herramienta de visualización e historia, que permite hacer autopsias de eventos subestándar. Esta herramienta almacena señales continuas, discretas y videos.</li> </ul>	<p>-Cloud Azure</p> <p>-Power BI</p>
Precios internacionales-Platts	Planeamiento	<p>Visualización del comportamiento de precios internacionales de la barra de construcción, palanquilla, chatarra y bobina caliente. La actualización de esta información es diaria de lunes a viernes.</p>	<p>-Cloud Azure</p> <p>-Power BI</p>
Programación de importaciones	Planeamiento	<p>Visualización de las importaciones realizadas por <i>SIDERPERU</i>. (fechas de llegada, producto, cantidad, costo, etc)</p>	<p>-Cloud Azure</p> <p>-Power BI</p>
Contraloría Planeamiento GLP	Finanzas	<p>Visualización de la rentabilidad de <i>SIDERPERU</i> por:</p> <p>Producto, línea de negocio, jefatura de ventas, cliente, mes, etc. La actualización de esta información es una (1) vez al mes, posterior al cierre contable.</p>	<p>-Cloud Azure</p> <p>-Power BI</p>

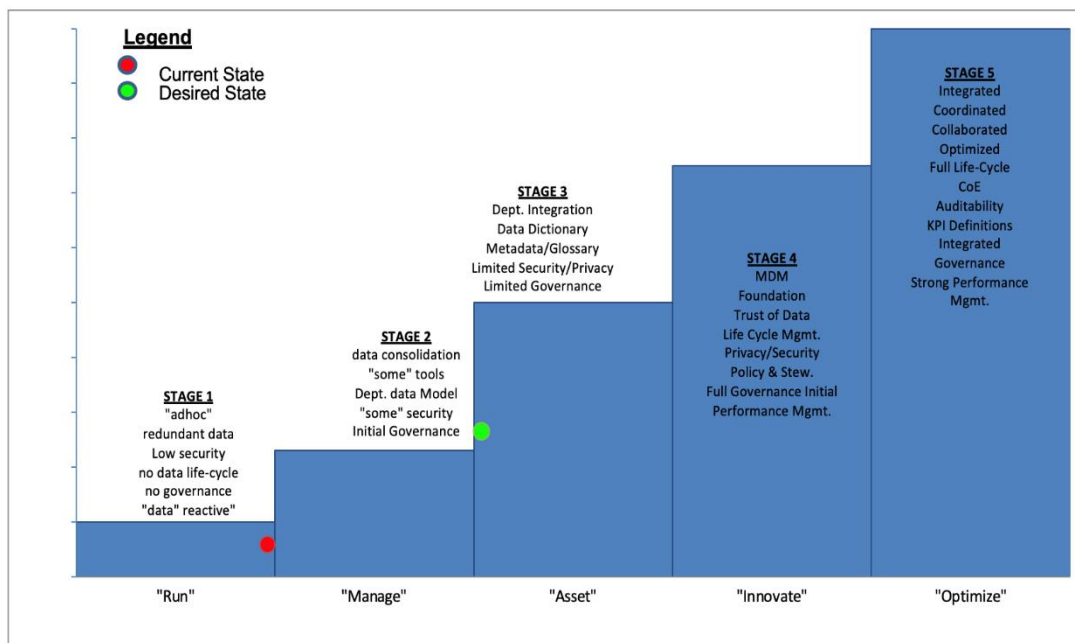


Inteligencia de Mercado	Comercial	Visualización del comportamiento de las ventas de <i>SIDERPERU</i> por: Producto, línea de negocio, jefatura de ventas, cliente, mes. La actualización de esta información es tres (3) veces al día de lunes a domingo.	-Cloud Azure -Power BI
-------------------------	-----------	---	---------------------------

## 5.5 Gobierno de Datos y Calidad de Datos

Empleando la metodología de medición de la madurez de gobierno de datos de IBM, se ha desarrollado la encuesta a los Stakeholders de *SIDERPERU* obteniendo los resultados detallados en el Anexo “Data Governance Maturity Model”.

Figura 14: Nivel de Madurez de Data Governance Actual



Según esta evaluación, el AS-IS se encuentra en la fase inicial “Stage 1: Run”, caracterizado por:

- La existencia de datos redundantes.
- No existen procesos formales para el seguimiento de los datos.
- No existe un proceso formal de gobierno de datos.
- La gestión de datos es ad-hoc y reactiva.
- Tiene un bajo nivel de seguridad.

### Madurez por categoría

El modelo de madurez de gobierno de datos de IBM evalúa 12 diferentes categorías. Cada categoría se puede evaluar de forma individual y ayuda a las organizaciones a obtener una idea de cómo establecer

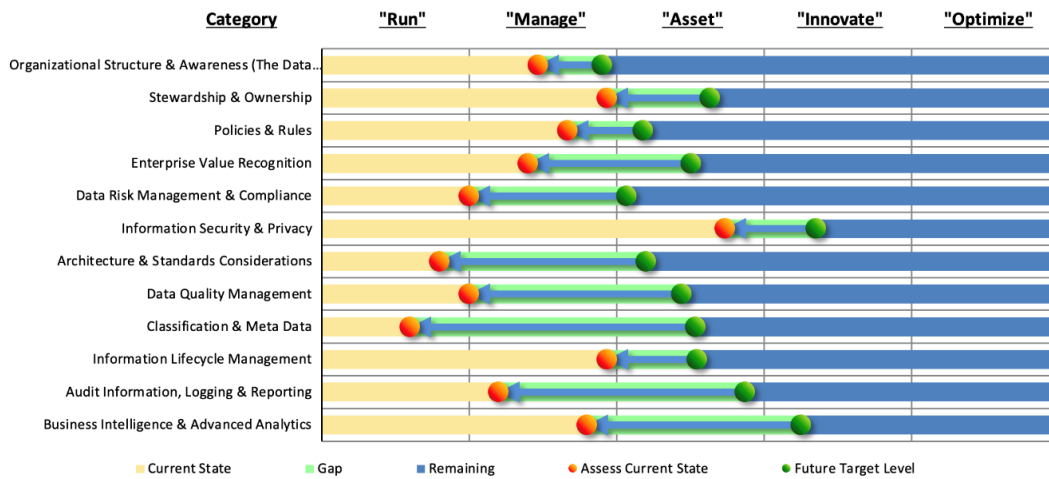
un plan integral de gobierno de datos. Cada categoría tiene 5 etapas de madurez. (“The IBM Data Governance Council Maturity Model\_ Building a”, 2007)

Figura 15: IBM Data Governance Council Maturity Model

Category	Description
<b>1 Organizational Structures &amp; Awareness</b>	Describes the level of mutual responsibility between business and IT, and recognition of the fiduciary responsibility to govern data at different levels of management.
<b>2 Stewardship</b>	Stewardship is a quality control discipline designed to ensure custodial care of data for asset enhancement, risk mitigation, and organizational control.
<b>3 Policy</b>	Policy is the written articulation of desired organizational behavior.
<b>4 Value Creation</b>	The process by which data assets are qualified and quantified to enable the business to maximize the value created by data assets.
<b>5 Data Risk Management &amp; Compliance</b>	The methodology by which risks are identified, qualified, quantified, avoided, accepted, mitigated, or transferred out.
<b>6 Information Security &amp; Privacy</b>	Describes the policies, practices and controls used by an organization to mitigate risk and protect data assets.
<b>7 Data Architecture</b>	The architectural design of structured and unstructured data systems and applications that enable data availability and distribution to appropriate users.
<b>8 Data Quality Management</b>	Methods to measure, improve, and certify the quality and integrity of production, test, and archival data.
<b>9 Classification &amp; Metadata</b>	The methods and tools used to create common semantic definitions for business and IT terms, data models, types, and repositories. Metadata that bridge human and computer understanding.
<b>10 Information Lifecycle Management</b>	A systematic policy-based approach to information collection, use, retention, and deletion.
<b>11 Audit Information, Logging &amp; Reporting</b>	The organizational processes for monitoring and measuring the data value, risks, and efficacy of governance.

El siguiente gráfico muestra el nivel de madurez actual y deseado por cada una de las doce categorías:

Figura 16: Gobierno de Datos - Nivel de Madurez Actual



### Mapa de calor por categoría

El siguiente gráfico es un mapa de calor que muestra cuáles son las etapas más fuertes o las más desarrolladas dentro de cada categoría. Cada categoría se analiza de forma independiente para poder definir planes de acción en cada una de ellas, de acuerdo con las prioridades de gobierno de datos de la *SIDERPERU*.

Por ejemplo, podemos ver que en la categoría de “Policies & Rules” la fase 1 es la más fuerte con 87% de desarrollo, mientras que las fases 4 y 5 son las más débiles con 0% de desarrollo.

Es importante mencionar que la madurez no necesariamente es descendente en cada fase, por ejemplo, en la categoría de “Data Risk Management & Compliance” vemos que la fase 1 está menos desarrollada que la fase 2.

Figura 17: Heatmap x Categoría

Category	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	%
Organizational Structure & Awareness (The Data)	53%	53%	40%	0%	0%	29%
Stewardship & Ownership	73%	60%	20%	0%	40%	39%
Policies & Rules	87%	40%	40%	0%	0%	33%
Enterprise Value Recognition	60%	40%	0%	40%	0%	28%
Data Risk Management & Compliance	20%	60%	20%	0%	0%	20%
Information Security & Privacy	53%	73%	67%	60%	20%	55%
Architecture & Standards Considerations	20%	20%	40%	0%	0%	16%
Data Quality Management	40%	40%	20%	0%	0%	20%
Classification & Meta Data	20%	20%	20%	0%	0%	12%
Information Lifecycle Management	60%	73%	20%	20%	20%	39%
Audit Information, Logging & Reporting	60%	0%	40%	0%	20%	24%
Business Intelligence & Advanced Analytics	60%	40%	40%	40%	0%	36%
	35%	30%	21%	9%	6%	

### Mapa de calor por fase

El siguiente gráfico es un mapa de calor que muestra cuáles son las categorías más fuertes o las más desarrolladas en cada fase. Este análisis sirve para determinar cuáles son las categorías que se necesitan reforzar para pasar a una siguiente fase de madurez.

De acuerdo con el resultado de nuestra encuesta, *SIDERPERU* se encuentra en la Fase 1 de madurez y desea pasar a una Fase 3, por lo tanto, debemos desarrollar cada una de las categorías para que se encuentren en verde en las fases previas a la fase objetivo.

Figura 18: Heatmap x Fase

Stage	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	%
Organizational Structure & Awareness (The Data)	53%	53%	40%	0%	0%	29%
Stewardship & Ownership	73%	60%	20%	0%	40%	39%
Policies & Rules	87%	40%	40%	0%	0%	33%
Enterprise Value Recognition	60%	40%	0%	40%	0%	28%
Data Risk Management & Compliance	20%	60%	20%	0%	0%	20%
Information Security & Privacy	53%	73%	67%	60%	20%	55%
Architecture & Standards Considerations	20%	20%	40%	0%	0%	16%
Data Quality Management	40%	40%	20%	0%	0%	20%
Classification & Meta Data	20%	20%	20%	0%	0%	12%
Information Lifecycle Management	60%	73%	20%	20%	20%	39%
Audit Information, Logging & Reporting	60%	0%	40%	0%	20%	24%
Business Intelligence & Advanced Analytics	60%	40%	40%	40%	0%	36%
	35%	30%	21%	9%	6%	

## 5.6 Modelo Operacional

*SIDERPERU* para las solicitudes de nuevos activos analíticos ha tenido dos fases, una primera en la que cada unidad solicitaba al departamento de TITD de forma independiente y acorde con sus necesidades el desarrollo y elaboración de Dashboards, herramientas digitales y modelos de ML. En una segunda fase estos nuevos activos analíticos provienen del evento *SIDER TANK*, donde las diferentes áreas presentan sus proyectos de mayor impacto con el objetivo de que el comité evaluador apruebe la inversión requerida. Sólo para el caso del área industrial cuentan con el Squad industria 4.0 quienes son los encargados de identificar las necesidades y supervisar la implementación de sus proyectos ganadores.

## 5.7 Data Literacy / Culture

SIDERPERU tiene un programa anual de capacitaciones referidas a automatización, digitalización, valor de los datos y su potencial para generar una ventaja competitiva. Estas capacitaciones carecen de una estrategia de comunicación. Actualmente la participación a dichas capacitaciones es a nivel de notificación general generando poca asistencia.

Quedan pendientes muchas otras acciones para desarrollar nuevas habilidades e impactar en la cultura de la toma de decisiones empleando datos.

## 5.8 Data Compliance

Se tiene como parte de las políticas de protección de datos y confidencialidad medidas generales de seguridad, accesos restringidos respecto de los sistemas, reportes y documentos en los servidores.

Estas políticas no están vinculadas a dominios de datos, al estar los datos en silos se encuentran repetidos y no necesariamente sincronizados, como consecuencia es muy complejo implementar cambios en los accesos.

## 5.9 Conclusiones de la situación actual

### 5.9.1 Conclusiones Generales

- No cuenta con una estrategia de datos y analítica, pero está enfocada en el desarrollo de proyectos de transformación digital e iniciativas aisladas de datos.
- No tiene un repositorio de datos unificado e integrado.
- Dado el nivel de madurez actual no tienen un ciclo de vida (DataOps) del dato adecuado, no cuentan con un gobierno de datos ni con un proceso de calidad de datos.
- Tiene personal acorde con la estrategia empresarial vigente, existe personal con capacidades analíticas que puede ser parte de la propuesta de datos y analítica.
- No tiene capacidad de poder identificar alertas tempranas producto del análisis de los datos.
- Existe una gran oportunidad de optimizar tiempos y reducir costes haciendo uso de datos y analítica.

En el siguiente gráfico de Forbes Canadá 2017 podemos apreciar las grandes ventajas de las empresas que desarrollan capacidades de datos y analítica (Leading) vs las que no (Lagging).

Figura 19: Why Maturity Matters?



(Forbes, 2017, p. 7)

### 5.9.2 Nivel de madurez de la organización en data Analytics

Para poder conocer el nivel de madurez analítica se ha desarrollado la encuesta según la metodología EIM Gartner con el apoyo de los Stakeholders, con los siguientes resultados:




EIM Dimension:	Overall	Vision	Strategy	Metrics	Governance	Org./Roles	Life Cycle	Infrastructure
EIM Maturity Index:	<b>1,84</b>	1,54	2,01	1,80	2,65	1,50	1,59	1,83
EIM Maturity Level:	<b>Reactive (-)</b>							
----- Range of weighted responses (1-5) for each EIM maturity dimension -----								
EIM Maturity Level Description:	Su organización es similar a aproximadamente el 30% de las que nos encontramos. Dichas organizaciones generalmente operan en un modo reactivo y centrado en la aplicación, a menudo esperando hasta que los problemas relacionados con la información se manifiesten en pérdidas comerciales significativas o falta de competitividad antes de abordarlos.							
EIM Program Balance:	Somewhat Balanced							

Figura 20: Radar encuesta de madurez analítica – EIM Gartner



Ver detalles adicionales en el Anexo 2 Enterprise Information Management Maturity Self-Assessment – AS IS.

El score general de *SIDERPERU* es 1,84 (REACTIVE -) correspondiente a empresas reactivas en relación con el desarrollo de iniciativas de datos y analítica o como consecuencia de perjuicios relevantes.

En el gráfico se aprecia que sólo en la categoría Governance se supera el nivel 2. *SIDERPERU* viene ejecutando un plan de difusión del valor de los datos y de las capacidades derivadas de su correcta gestión. El proceso de transformación cultural está en una fase temprana, pero existe consciencia a nivel gerencial de la importancia de estas capacidades.

## CAPÍTULO 6.- PLANTEAMIENTO DE LA SOLUCIÓN

### 6.1 Definición de la estrategia de Data & Analytics futura

#### 6.1.1 Estrategia

Tabla 8: Estrategia de D&A a tres años, 2023-2025

	AS IS	TO BE
DIFERENCIAS		
MGD <sup>1</sup>	Stage1 (Run)	Stage3 (Asset)
MD&A <sup>2</sup>	Reactiva (-)	Proactiva (-)
CoE <sup>3</sup>	No se tiene un área de datos, se tiene un departamento de TI y de Trans. Digital.	Si
DLIT <sup>4</sup>	No	Si – Un programa de alfabetización en niveles distintos de la estructura.

<sup>1</sup> MGD: Madurez del Gobierno de Datos.

<sup>2</sup> MD&A: Madurez de Datos y Analítica.

<sup>3</sup> CoE: Centro Integrado de Excelencia de Ciencia de Datos.

<sup>4</sup> DLIT: Data Literacy.

IT <sup>5</sup>	Business Intelligence, Microsoft Azure	Un Data Lake IT en AWS y Un Data Lake OT On Premise
ED <sup>6</sup>	No se tiene un ecosistema, se tienen silos de datos.	Un ecosistema con una plataforma para procesos tiempo real (OT) y otra para procesos offline (IT) interrelacionadas en la medida de lo necesario e integradas por el gobierno de datos, el modelo operacional y la coordinación del CoE.

---

<sup>5</sup> IT: Infraestructura Tecnológica.

<sup>6</sup> ED: Ecosistema de Datos.



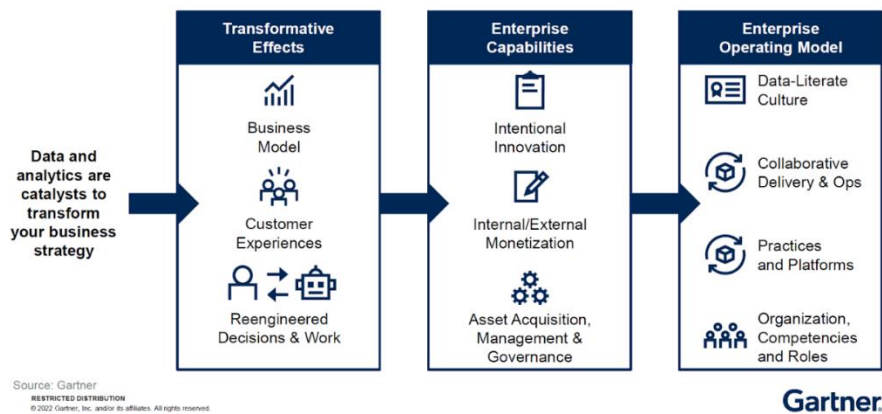
## Beneficios

- Incrementar de las capacidades de datos y analítica de *SIDERPERU* para potenciar los procesos y la toma de decisiones basadas en datos. Esto contribuye al cumplimiento de la estrategia empresarial como resultado de la monetización de los activos analíticos.
- Fortalecer sus capacidades dinámicas generando mayores oportunidades de innovar, permitiendo generar valor y ventajas diferenciales.

Según Gartner datos y analítica fungirá de catalizador para transformar o llevar la estrategia de negocio como muestra en el siguiente gráfico:

Figura 21: Buenas prácticas de D&A en la Estrategia de Negocios

### Best Practice: Infuse Business Strategy With D&A



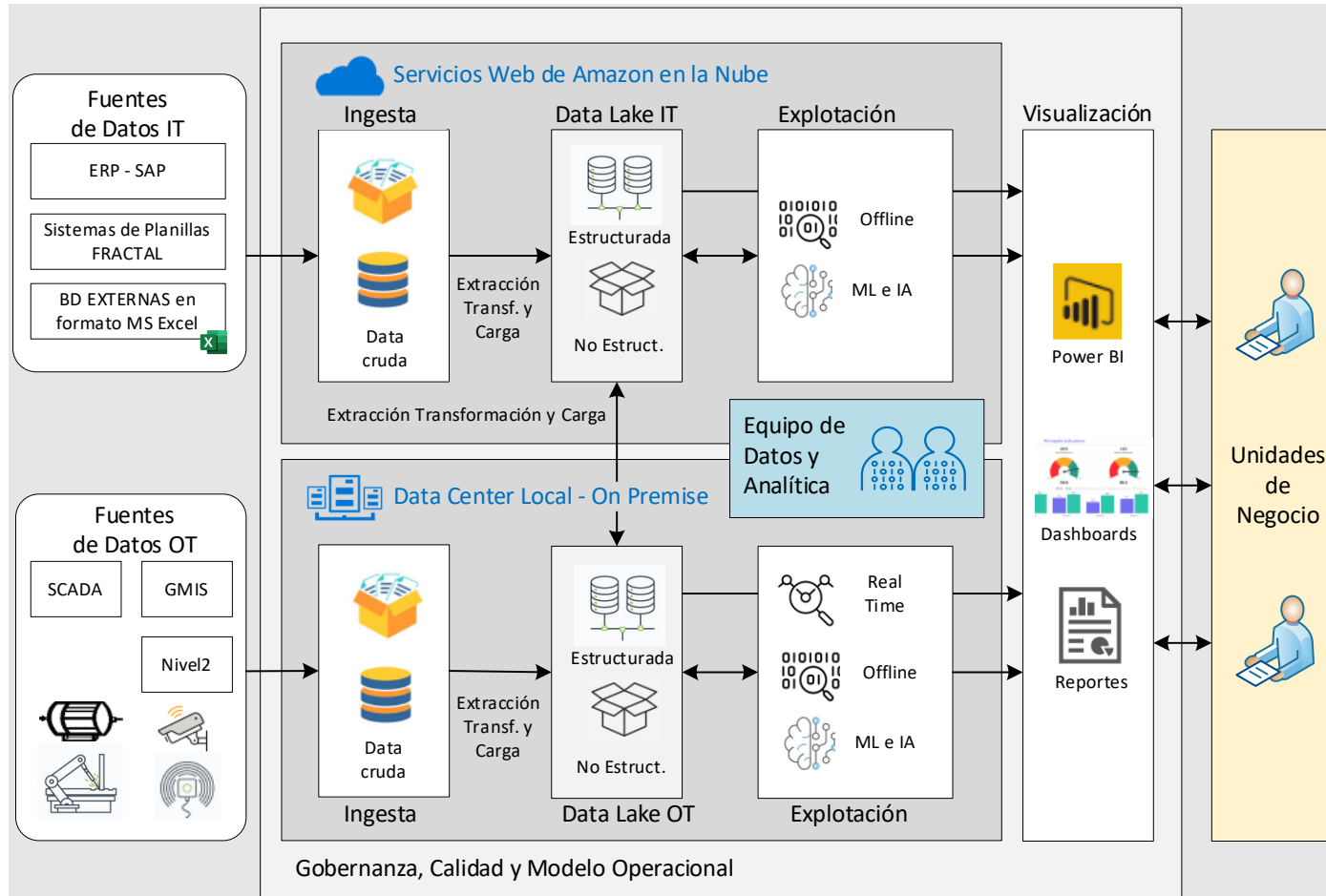
## 6.1.2 Indicadores de Medición de la Estrategia

Figura 22: Indicadores de Medición de la Estrategia

Categoría	Indicador	Meta a 3 años	Periodicidad de Medición
Estrategia	Índice de Madurez de Datos y Analítica	Proactiva (-)	Semestral
Estrategia	Índice de Madurez de Gobierno de Datos	Stage 3	Semestral
Gobierno de Datos	Datos Gobernados, de los dominios relevantes para los activos analíticos.	100%	Trimestral
Gobierno de Datos	Datos No Gobernados, de los dominios relevantes para los activos analíticos.	0%	Trimestral
Activos Analíticos	Casos de Uso en Producción	6	Trimestral

## 6.2 Definición del Ecosistema de Datos

Figura 23: Ecosistema de Datos Propuesto



### Observaciones

#### Data Lake IT

- Fuentes de datos administrativas.
- Integra datos operativos según la necesidad, para un análisis *offline*.
- Implementado en la infraestructura cloud empresarial AWS.

#### Data Lake OT

- Fuentes de datos operativos.
- Implementado On Premise para cumplir con los requisitos de baja latencia y analítica en *tiempo real*.
- Integra datos administrativos según las necesidades.

### 6.3 Definición del capital humano

Figura 24: Organigrama de Data & Analytics

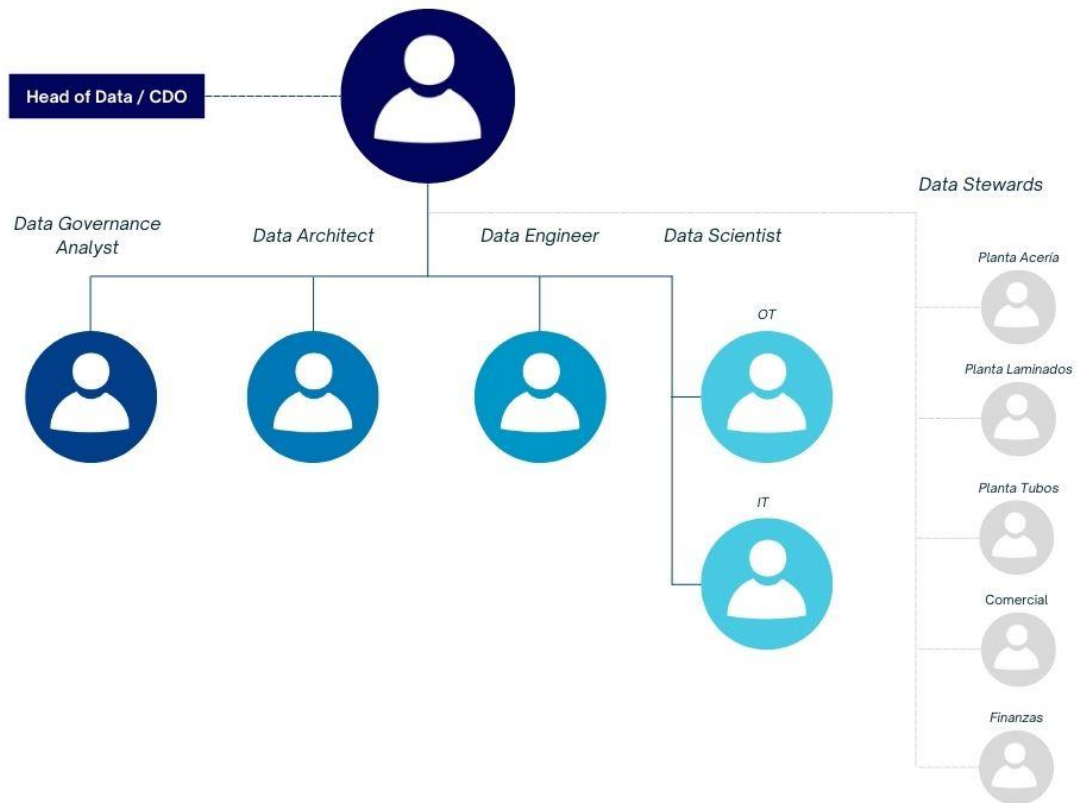


Tabla 9: Roles de D&A

Rol	Funciones Principales
Head of Data & Analytics o CDO	Es el ejecutivo sénior que tiene la responsabilidad de la estrategia de información y datos de toda la empresa, el gobierno, el control, el desarrollo de políticas y la explotación efectiva de la empresa. El CDO tiene la responsabilidad por la protección y privacidad de la información, el gobierno de la información, la calidad de los datos y la gestión del ciclo de vida de los datos, junto con la explotación de los activos de datos para crear valor empresarial.
Data Architect	Los Data Architects o Arquitectos de Datos son empleados senior que traducen los requisitos de negocio en requisitos tecnológicos y definen los estándares y principios de datos. El arquitecto de datos es responsable de visualizar y diseñar el marco de gestión (framework) de datos empresariales de una

	<p>organización. Este framework describe los procesos utilizados para planificar, especificar, habilitar, crear, adquirir, mantener, usar, archivar, recuperar, controlar y depurar datos.</p>
Data Scientist	<p>Los Data Scientist o Científico de Datos son expertos en datos analíticos que utilizan la ciencia de datos para descubrir información a partir de cantidades masivas de datos estructurados y no estructurados para ayudar a dar forma o satisfacer necesidades y objetivos específicos de negocio. Los científicos de datos se están volviendo cada vez más importantes en los negocios, ya que las organizaciones confían más en el análisis de datos para impulsar la toma de decisiones y se apoyan en la automatización y el aprendizaje automático como componentes centrales de sus estrategias de TI.</p> <p>El rol del científico de datos es fundamental para las organizaciones que buscan extraer información de los activos de información para iniciativas de "big data" y requiere una amplia combinación de habilidades. Por ejemplo, se requiere colaboración y trabajo en equipo para trabajar con las partes interesadas del negocio para comprender los problemas del negocio. Se requieren habilidades analíticas y de modelado de decisiones para descubrir relaciones dentro de los datos y detectar patrones. Se requieren habilidades de gestión de datos para construir el conjunto de datos relevante utilizado para el análisis.</p>
Data Engineer	<p>Los Data Engineers o Ingenieros de Datos diseñan, crean y optimizan sistemas para la recopilación, el almacenamiento, el acceso y el análisis de datos a escala. Crean "data pipelines" utilizados por científicos de datos, aplicaciones centradas en datos y otros consumidores de datos.</p> <p>Este rol de TI requiere un conjunto significativo de habilidades técnicas, incluido un conocimiento profundo del diseño de bases de datos SQL y múltiples lenguajes de programación. Los ingenieros de datos también necesitan habilidades de comunicación para trabajar entre departamentos y comprender lo</p>

	que los líderes empresariales quieren obtener de los grandes conjuntos de datos de la empresa.
Data Steward	Los Data Stewards pueden ayudar a las organizaciones a alcanzar los objetivos de mejora de la calidad de los datos. Los Stewards actúan como fideicomisarios de los datos, tienen un profundo conocimiento de los procesos de negocio y el uso de los datos y, en última instancia, son los responsables del resultado de un programa de mejora de la calidad de los datos.
Data Governance Analyst	El objetivo principal del Data Governance Analysts o Analista de Gobierno de Datos es ayudar al equipo de gobierno de información/datos en la formación y ejecución del framework, la política y los estándares de gobierno de datos. Esta posición ayuda en la implementación de un programa de gobierno de datos empresariales.

Tabla 10: D&A Head Count Ramp Up

Tarea	Año1	Año2	Año3	Comentarios
Head of Data & Analytics	1	1	1	
Data Governance Analyst	1	1	1	
Data Architect	1	1	1	
Data Scientist	1	2	2	Desde el año 2 se consideran: Un DS experto en la industria y procesos operacionales. Un DS experto en modelos asociados a datos administrativos y financieros.
Data Engineer	1	1	1	
Data Steward	5	5	5	

## 6.4 Definición del portafolio de activos analíticos

Tabla 11: Portafolio de Activos Analíticos

Cod.	Nombre	Descripción	Data Lake
U1	Optimización en uso de ferroaleaciones	Analizará la composición química y los aditivos que se requieren para la fabricación del acero líquido de manera automática, evitando los desperdicios actuales por adición manual y a veces excesiva.	OT
U2	Predictiva de Ventas	A partir de la analítica de datos, se gestionará la venta de la construcción formal, identificando por tipo de obra el mix de diámetro requerido consiguiendo una previsión de ventas en concordancia con la producción.	IT
U3	Modelo de sobrecalentamiento	Modelo de optimización de sobrecalentamiento que examinaría los datos operativos en tiempo real y recomendaría puntos de ajuste del proceso conducentes a una mayor tasa de impacto.	OT
U4	Mantenimiento Predictivo	Permitirá detectar con anticipación la posible pérdida de calidad de servicio que este entregando un equipo.	IT
U5	Modelo para visibilidad del guero	<i>SIDERPERU</i> requiere implementar un sistema de cámaras para puntos ciegos, con estas imágenes se propone hacer un modelo ML que identifique estos puntos ciegos y de asertividad para la colocación de las palanquillas.	IT

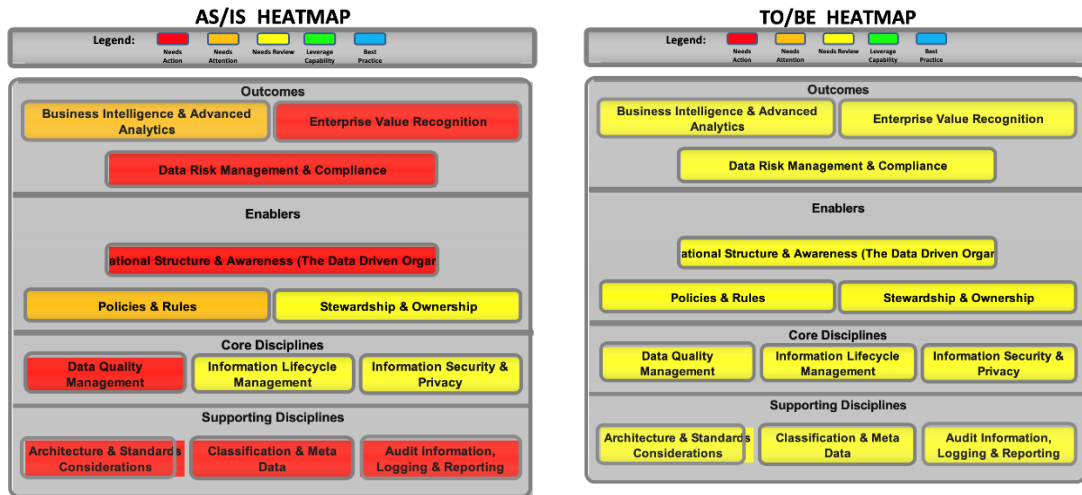
## 6.5 Definición del plan de gobierno y calidad de datos

En el Gráfico 15 vimos que *SIDERPERU* se encuentre en una Fase 1 (reactiva) y desea progresar hacia una Fase 3 (proactiva) en un periodo de 3 años.

El siguiente gráfico muestra un mapa de calor más detallado de la situación actual y la situación deseada en cada una de las 12 categorías, agrupadas en cuatro grupos principales:

- Outcomes (resultados)
- Enablers (habilitadores)
- Core Disciplines (disciplinas fundamentales)
- Supporting Disciplines (disciplinas de apoyo)

Figura 25: Data Governance AS IS vs TO BE



Se puede apreciar que en la situación actual (AS-IS) que hay muchas categorías que requieren acción inmediata (en rojo) y que serán desarrolladas con el plan propuesto. La situación deseada (TO-BE) aún no será un escenario ideal, ya que este es un proceso iterativo y que irá evolucionando en el tiempo, sin embargo, este será un punto de partida.

La metodología de Gobierno de Datos de IBM nos sugiere un proceso unificado de gobierno, el cual consiste en 14 pasos (10 requeridos y 04 opcionales), tal como se muestra en el siguiente gráfico:

Figura 26: Proceso de Data Governance

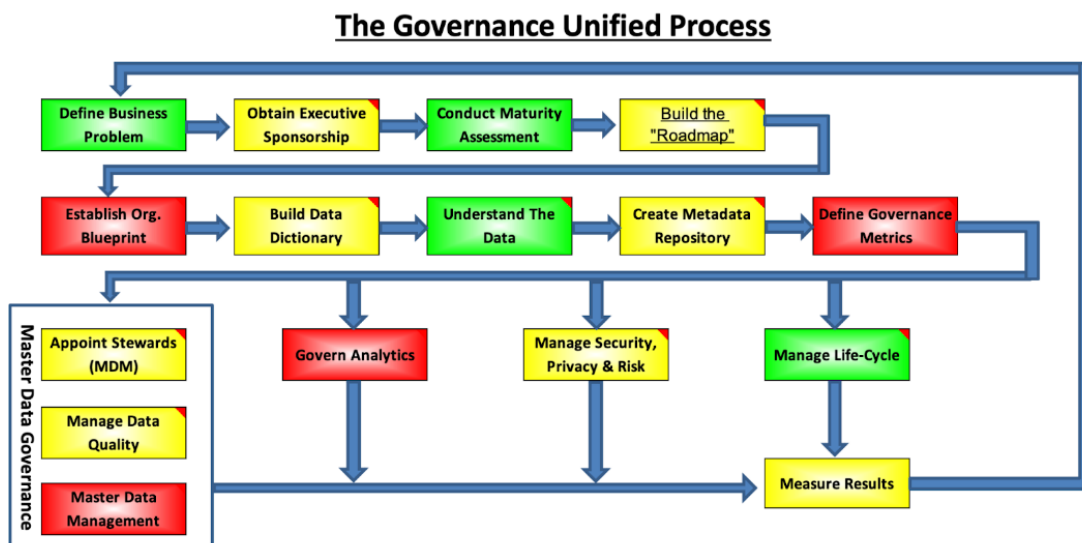


Tabla 12: *Proceso Unificado de Gobierno de Datos*

Paso	Descripción
1	<b><u>Definir el problema de negocio</u></b>
2	<p><b><u>Obtener el patrocinio ejecutivo</u></b></p> <p>Es importante obtener el sponsorship de ejecutivos clave del negocio y de TI para el programa de Gobierno de Datos.</p> <p>Elegir casos de negocio y buscar “victorias rápidas”.</p> <p>Crear un equipo de trabajo de Gobierno de Datos.</p> <p>Identificar un dueño para el Gobierno de Datos.</p>
3	<p><b><u>Realizar una evaluación de madurez</u></b></p> <ul style="list-style-type: none"> <li>• Ya se realizó con la herramienta de IBM.</li> </ul>
4	<p><b><u>Crear una hoja de ruta (roadmap)</u></b></p> <ul style="list-style-type: none"> <li>• Resumir los resultados de la evaluación de madurez.</li> <li>• Enumerar las iniciativas clave de personas, procesos y tecnología necesarias para cerrar las brechas destacadas en la evaluación.</li> <li>• Crear el roadmap basado en una priorización de iniciativas.</li> </ul>
5	<p><b><u>Establecer un modelo organizativo (blueprint)</u></b></p> <ul style="list-style-type: none"> <li>• Definir la estructura organizacional para el Gobierno de Datos.</li> <li>• Establecer el consejo de Gobierno de Datos.</li> <li>• Establecer el grupo de trabajo de Gobierno de Datos.</li> <li>• Identificar a los Data Stewards.</li> <li>• Llevar a cabo reuniones periódicas del consejo y grupos de trabajo de Gobierno de Datos.</li> </ul>
6	<p><b><u>Crear un diccionario de datos</u></b></p> <ul style="list-style-type: none"> <li>• Seleccionar un dominio de datos.</li> <li>• Asignar Data Stewards para mantener los términos clave del negocio.</li> <li>• Identificar elementos de datos críticos.</li> <li>• Iniciar el diccionario de datos con un glosario de términos existente.</li> <li>• Rellenar el diccionario de datos.</li> <li>• Vincular términos de negocio con términos técnicos a través de metadata.</li> </ul>



	<ul style="list-style-type: none"> <li>• Compatibilidad con los requisitos de auditoría, generación de informes y registro de Gobierno de Datos.</li> <li>• Integrar el diccionario de datos con el entorno de aplicaciones (diccionario de negocio).</li> </ul>
7	<b><u>Comprender los datos</u></b>
8	<b><u>Crear un repositorio de metadatos</u></b> <ul style="list-style-type: none"> <li>• Combinar metadatos de negocio del diccionario de datos y metadatos técnicos del proceso de descubrimiento.</li> <li>• Asegurar el linaje de datos apropiado.</li> <li>• Realizar un análisis de impacto.</li> <li>• Administrar metadatos operativos.</li> </ul>
9	<b><u>Definir métricas de gobierno</u></b> <ul style="list-style-type: none"> <li>• Comprender los Key Performance Indicators (KPIs) del negocio.</li> <li>• Definir KPIs impulsados por el negocio para el Gobierno de Datos.</li> <li>• Definir KPIs técnicos para el Gobierno de Datos.</li> <li>• Establecer un tablero (dashboard) para la evaluación de la madurez de Gobierno de Datos.</li> </ul>
10	<b><u>Opcional: Gobernar los datos maestros</u></b> <ul style="list-style-type: none"> <li>• Designar a los Data Stewards</li> <li>• Administrar datos de calidad</li> <li>• Implementar MDM (Master Data Management)</li> <li>• Designar al líder de Data Stewards</li> <li>• Determinar la configuración del programa de Data Stewards.</li> <li>• Identificar sponsors ejecutivos para cada dominio de datos.</li> <li>• Establecer políticas de calidad de datos.</li> <li>• Construir los datos de negocio</li> <li>• Reclutar Data Stewards para cada dominio de datos.</li> <li>• Empoderar al consejo de Gobierno de Datos para supervisar al programa de Data Stewards.</li> <li>• Limpiar los datos.</li> <li>• Supervisar la calidad de los datos a lo largo del tiempo.</li> <li>• Definir las áreas temáticas de los datos maestros.</li> <li>• Identificar los sistemas y procesos de negocio que consumen los datos.</li> </ul>

	<ul style="list-style-type: none"> <li>• Identificar las fuentes de datos actuales.</li> <li>• Definir los atributos de datos del sistema de registro.</li> <li>• Designar data Stewards para cada sistema de registro.</li> <li>• Establecer políticas para el gobierno de datos maestros.</li> <li>• Implementar una consola de administración de datos para intervención manual y monitoreo.</li> <li>• Administrar posibles tareas superpuestas.</li> <li>• Unificar la data duplicada de una misma fuente o múltiples fuentes para crear nuevo registro maestro.</li> <li>• Vincular registros de múltiples fuentes.</li> <li>• Administrar relaciones.</li> <li>• Administrar jerarquías.</li> <li>• Administrar agrupaciones.</li> <li>• Diseñar la solución de MDM.</li> </ul>
11	<b><u>Opcional: Gobernar la analítica</u></b>
12	<b><u>Opcional: Administrar la seguridad y privacidad</u></b>
13	<b><u>Opcional: Gobernar el ciclo de vida de la información</u></b>
14	<b><u>Medir los resultados</u></b>

## 6.6 Presupuesto de la Solución

Todos los valores están expresados en miles de dólares americanos, los flujos incluyen signo.

Tabla 13: Presupuesto de la Solución

Proyecto	Año 0	Año 1	Año 2	Año 3	VAN Años [1 a 3]
CoE D&A	0	-275	-362	-362	-791
Opex	0	-275	-362	-362	-791
Capex	0	0	0	0	0
Data Lake IT	-60	-156	-312	-468	-781
Opex	0	-156	-312	-468	-721
Capex	-60	0	0	0	-60
Data Lake OT	-370	-283	-301	-320	-1 090
Opex	0	-38	-56	-75	-132
Capex	-370	-245	-245	-245	-958
Data Governance	0	-169	-179	-179	-421
Opex	0	-109	-119	-119	-276
Capex	0	-60	-60	-60	-144
Data Literacy	0	-22	-24	-24	-56
Opex	0	-22	-24	-24	-56
Capex	0	0	0	0	0
Mantenimiento Predictivo	-50	0	0	0	-50
Capex	-50	0	0	0	-50
Sub total OPEX	0	-599	-873	-1 047	-1 976
Sub total CAPEX	-480	-305	-305	-305	-1 213
Total	-480	-904	-1 178	-1 352	-3 189

### **Inversión - CAPEX**

Esta compuesta en el Año 0 por las implementaciones de:

- Data Lake IT.
- Data Lake OT.
- El caso de uso 4 – Mantenimiento Predictivo, implementado por una consultora.

En los siguientes años:

- Licencias de software.
- Ampliación/Upgrade de hardware y los servicios de implementación asociados.

### **Egresos Recurrentes/Costes Operacionales - OPEX**

Está compuesto por:

- Costes de personal.
- Costes de arrendamiento de infraestructura cloud.
- Costes de soporte y gestión de las plataformas.
- Costes de iniciativas/acciones de entrenamiento.

## 6.7 Revisión de la solución

### 6.7.1 Mapa Problema – Solución

Tabla 14: Relación Problema, Objetivos y Proyectos

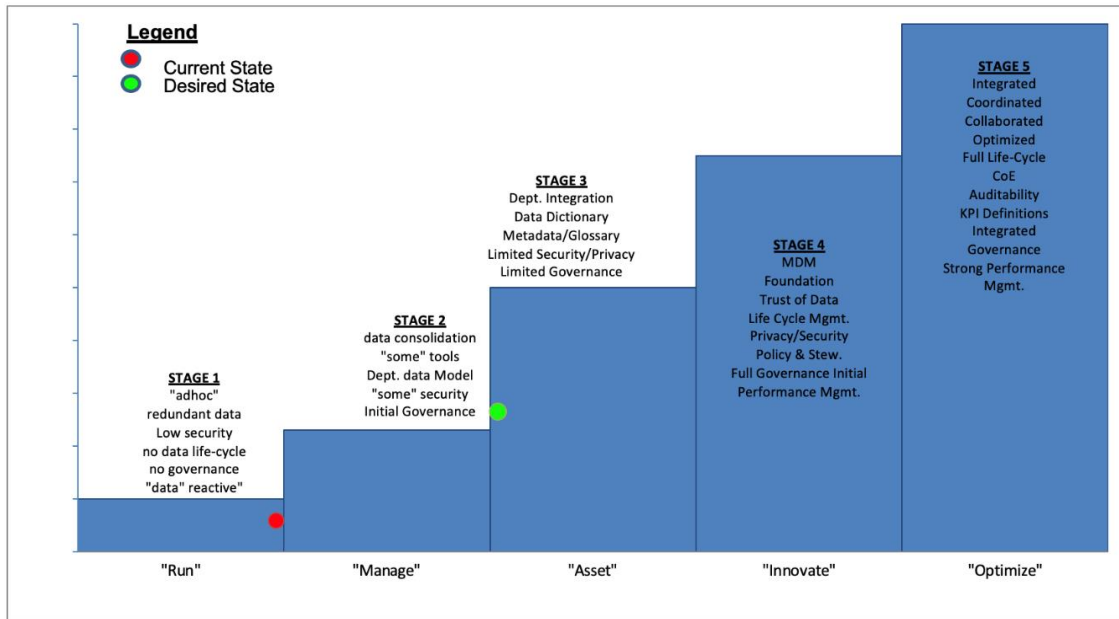
			Solución									
			P1	P2	P3	P4	P5	U1	U2	U3	U4	U5
<b>Problema</b>			CoE D&A	Data Lake IT	Data Lake OT	Data Governance	Data Literacy	Optimización en uso de ferreoaleaciones	Análítica Predictiva de Ventas	Modelo de sobre calentamiento	Mantenimiento Predictivo	Modelo de visibilidad para el grueso
<b>Problema1:</b> Generar una fuente de la verdad o repositorio centralizado, que almacene datos de calidad generando un ecosistema de datos para <i>SIDERPERU</i> , en el marco de un gobierno de datos, haciendo uso de la infraestructura, organización y recursos más apropiados.			x	x	x	x						
<b>Problema2:</b> Hacer uso de Datos y Analítica para potenciar las capacidades de análisis, la validación de hipótesis, anticiparse y generar valor para <i>SIDERPERU</i> , a través de la generación de activos digitales.			x			x	x	x	x	x	x	x
<b>Problema3:</b> Contribuir a maximizar la rentabilidad de <i>SIDERPERU</i> haciendo uso de Datos y Analítica.								x	x	x	x	x
Objetivo General	Cod.	Objetivo Especifico	P1	P2	P3	P4	P5	U1	U2	U3	U4	U5
G1	E1	Crear el área de datos y analítica.	x									
	E2	Implementar una plataforma tecnológica.		x	x							
	E3	Implementar un gobierno de datos.				x						
	E4	Implementar un modelo operativo.	x									
G2	E5	Transmitir la relevancia y valor de los datos en <i>SIDERPERU</i> .					x					
	E6	Desarrollar habilidades de datos y analítica.	x				x					
G3	E7	Identificar los casos de uso de mayor relevancia y prioridad en el marco de los principios y objetivos de la Industria 4.0.	x									
	E8	Reducción de labores operativas y foco en las labores de análisis.	x			x				x	x	x
	E9	Reducción de la merma.						x		x		x
	E10	Optimización en el uso de la materia prima.						x		x		x
	E11	Mantenimientos inteligentes, optimización de ciclos de mantenimiento.									x	
	E12	Estimación de la demanda.							x			

### 6.7.2 Nivel de madurez objetivo para la organización en Data & Analytics

El nivel objetivo de madurez analítica objetivo empleando la metodología EIM Gartner es Proactive (-).

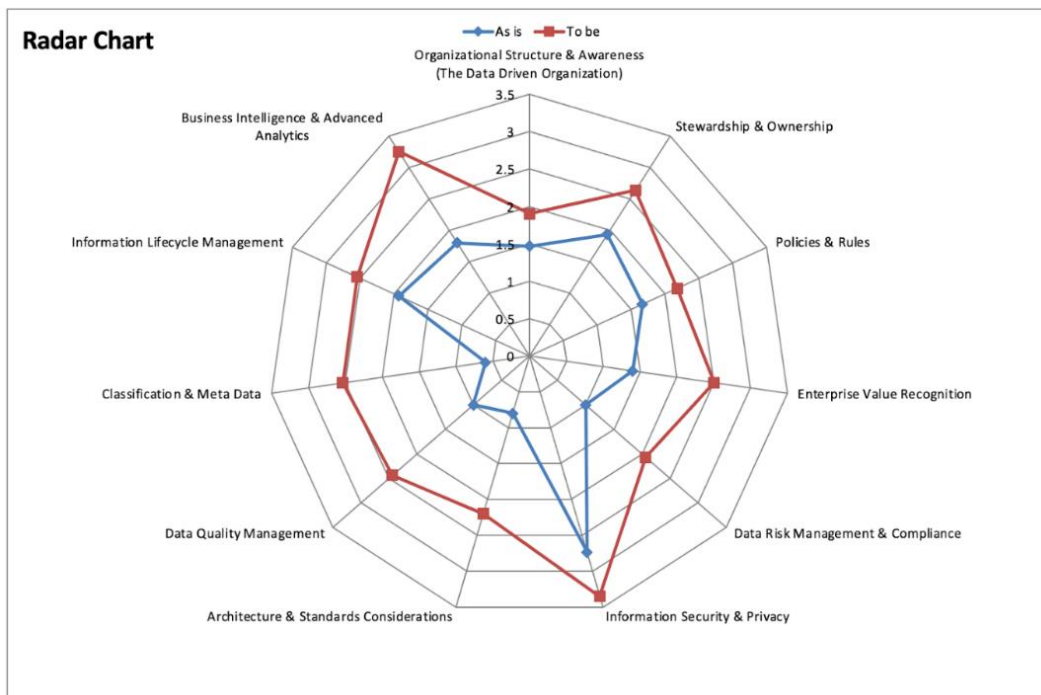
De igual forma el nivel de madurez objetivo del gobierno de datos en la metodología de IBM es “Stage 3: Asset”.

Figura 27: Nivel de Madurez de Data Governance Objetivo



Comparación entre el AS IS y el TO BE

Figura 28: Gráfico que compara la variación de madurez en DG AS IS vs TO BE



## CAPÍTULO 7.- PLAN DE IMPLEMENTACIÓN DE LA SOLUCIÓN

Todos los valores están expresados en miles de dólares americanos, los flujos incluyen signo.

### 7.1 Cartera de proyectos claves

#### **Definición, Justificación, Business Case y Precedencias de Proyectos Claves**

Para elaborar los Business Cases de cada proyecto hemos seguido como metodología la asignación de costes CAPEX y OPEX, según un factor de reparto directamente proporcional al valor capturado. A continuación, se presenta el factor de distribución.

Tabla 15: Factores de Distribución de Costes a 3 años

Código	Use Case	Valor Capturado Años [1 a 3]	Factor Distribución a 3 años
U1	Optimización en uso de ferreoaleaciones	1 644	23,4%
U3	Modelo de sobrecalentamiento	870	12,4%
U2	Analítica Predictiva de Ventas	0	0,0%
U4	Mantenimiento Predictivo	1 644	23,4%
U5	Modelo de visibilidad para el gruero	2 863	40,8%

#### 7.1.1 Proyectos Fundacionales

Estos proyectos representan los pilares para el desarrollo e implementación de D&A en SIDEPERU, estos proyectos preceden a todos los demás y un grupo de ellos no tienen fecha de término como Data Governance y Data Literacy, porque su ejecución continua es un requisito para garantizar el éxito de todos los proyectos y/o casos de uso posteriores.

##### 7.1.1.1 CoE Data & Analytics

El CoE de Data y Analytics es un proyecto fundamental para cualquier organización que busque transformar su negocio. El COE es un equipo o un grupo de expertos que brinda mejores prácticas, liderazgo, investigación, capacitación o apoyo para un tema de interés.

A medida que las organizaciones se vuelven más complejas, los equipos a menudo trabajan en silos, sin compartir su conocimiento. El COE busca construir una base más amplia de conocimiento y comprensión de data y analítica. Gartner menciona que un CoE ayuda a las empresas a enfocar y alinear sus recursos y experiencia actuales en torno a una capacidad específica para lograr y mantener un rendimiento y valor de clase mundial.

El objetivo principal del CoE es garantizar que los datos sean accesibles y se utilicen en todo *SIDERPERU* para generar valor. Entregar valor es la última medida del éxito.

Entre los beneficios que el CoE aportará a *SIDERPERU* se incluyen:

- Definir y difundir políticas y mejores prácticas para el gobierno de datos, las operaciones de datos y la gestión de programas que agilicen los procesos y faciliten la creación de valor.
- Uso eficiente de los recursos
- Provisión de productos y servicios de calidad a los clientes.
- Reducción de los costos operativos mediante la eliminación de prácticas ineficientes y la reducción del tiempo de implementación de nuevas habilidades y tecnologías.
- Los expertos del CoE pueden ayudar a los empleados a adaptarse para mejorar su forma de trabajar, compartir técnicas efectivas entre sí, capturar técnicas viables y apoyar el aprendizaje individual y en equipo.
- Fomentar una cultura de Data & Analytics más profunda mediante la creación de un lenguaje común para los equipos de tecnología y negocios y la alineación de incentivos en torno a la entrega de valor al negocio.
- Mejorar la toma de decisiones basada en datos.
- Mejorar las habilidades de Data & Analytics y asignar recursos en toda la organización.
- Entregar valor al negocio a través de Data & Analytics.

## **Business Case**

Tabla 16: BC Optimización en uso de ferroaleaciones

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
CoE D&A	0	-275	-362	-362	-791
Opex	0	-275	-362	-362	-791
Capex	0	0	0	0	0
Captura de Valor	0	0	0	0	0

Periodo	Indicador Financiero		
	VAN	TIR	Relación Beneficio - Coste
Años [1 a 3]	-791	n/a	n/a

## **Comentarios**

- Al ser un proyecto fundacional tiene asociado el coste de su implementación y de operación.
- No se presentan ingresos, porque los efectos positivos se presentan a través del valor capturado por los Use Cases.

### 7.1.1.2 Data Lake Information Technology (IT)

El mundo ha cambiado y, hoy, la diferencia entre el éxito y el fracaso depende en cómo las organizaciones usan su información. La información es poder, es por eso que el Data Lake IT es un proyecto fundamental.

La infraestructura de TI actual no les permite a los usuarios conseguir sus metas personales y de negocio. Muchos usuarios de negocio crean sus propias soluciones, duplicando esfuerzos y aumentando las posibilidades de error.

Los Data Lakes habilitan la captura de toda la información sobre el negocio. Los datos se recopilan de múltiples fuentes, tanto internas como externas, y se transfieren al Data Lake en su formato original. Además, permite importar cualquier cantidad de datos que puedan venir en tiempo real. Este proceso le permite escalar a datos de cualquier tamaño, mientras ahorra tiempo en la definición de estructuras de datos, esquemas y transformaciones.

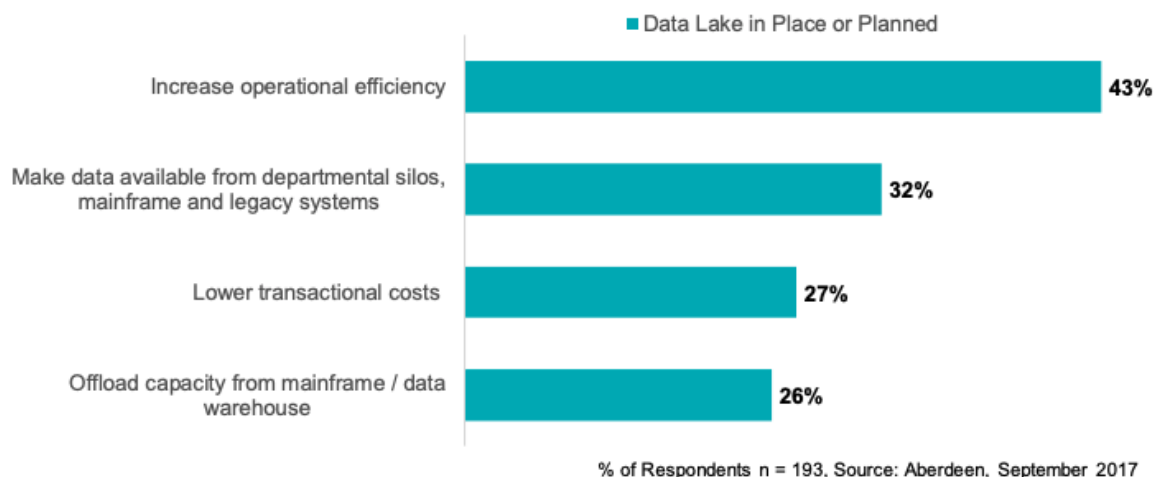
Uno de los grandes beneficios de un Data Lake es que pone la información a disposición de muchos empleados para hacer a la organización más inteligente, más ágil y más innovadora. Todos los usuarios extraen la información de la misma fuente de datos, lo que garantiza su fidelidad y precisión. Además, permite a los usuarios de cualquier parte de la organización ver toda la información que necesitan, sin necesidad de bases de datos adicionales ni sistemas duplicados.

El Data Lake puede colocar a *SIDERPERU* como líder por el valor generado a partir de la explotación de sus datos. De acuerdo a un artículo publicado por Aberdeen (Lock, 2017), las empresas líderes que implementan Data Lakes obtienen 24% en crecimiento de ingresos orgánicos, versus las empresas seguidoras que obtienen sólo 15%.

El mismo estudio de Aberdeen muestra cómo la eficiencia operacional puede incrementarse en 43% en las empresas que tienen un Data Lake implementado.

*Figura 29: ¿Por qué invertir en un Data Lake?*





## Business Case

Tabla 17: BC Data Lake Information Technology (IT)

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Data Lake IT	-60	1 272	1 830	1 674	3 726
Opex	0	-156	-312	-468	-721
Capex	-60	0	0	0	-60
Captura de Valor	0	1 428	2 142	2 142	4 507
Valor actual del FCE	-60	1 136	1 459	1 192	
VAN	-60	1 076	2 535	3 726	

Periodo	Indicador Financiero		
	VAN	TIR	Relación Beneficio - Coste
Años [1 a 3]	3 726	2160%	5,77

## Comentarios

- La plataforma IT que es una solución cloud tiene una inversión baja, que es recuperada durante el primer año con los primeros modelos.
- El proyecto tiene una TIR muy alta por la gran diferencia entre los ingresos proyectados y CAPEX + OPEX.

### 7.1.1.3 Data Lake Operational Technology (OT)

Este es otro proyecto fundamental, ya que las empresas industriales a menudo tienen dificultades para extraer valor de sus datos desde la tecnología de operaciones, principalmente porque suele ser tecnología legada, lo cual crea silos de información y resulta complejo su uso general por toda la empresa para analizarlos y obtener valor. Es por esto que un Data Lake OT es un paso importante para la transformación digital.

Crear un repositorio centralizado de datos operacionales, tanto de datos históricos, como también de nueva información en tiempo real, es un paso indispensable para implementar herramientas modernas de TI, tales como analítica avanzada e inteligencia artificial.

Los Data Lakes estandarizan los datos independientemente del sistema, lo cual permite enviarlos a prácticamente cualquier aplicación. Esta estandarización es útil cuando se tiene una amplia variedad de sensores o equipos que deben operar con la máxima eficiencia para reducir el tiempo de inactividad de las plantas, además de garantizar la seguridad y maximizar la producción.

Otro beneficio importante de implementar un Data Lake OT es que los usuarios de múltiples áreas funcionales podrán ver dashboards en tiempo real desde cualquier lugar y podrán monitorear el rendimiento de los equipos para anticiparse a cualquier falla o emergencia.

Finalmente, el Data Lake OT agilizará la colaboración y la toma de decisiones en *SIDERPERU*, sin necesidad de reemplazar equipos antiguos

Los equipos de TI y OT deberán trabajar juntos para romper los silos de información en toda la organización y poder tener la data a disposición de todos los usuarios que la necesiten para crearle el máximo valor a *SIDERPERU*.

## **Business Case**

Tabla 18: BC Data Lake Operational Technology (OT)

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Data Lake OT	-370	-283	1 210	1 520	1 425
Opex	0	-38	-56	-75	-132
Capex	-370	-245	-245	-245	-958
Captura de Valor	0	0	1 512	1 840	2 515
Valor actual del FCE	-370	-252	965	1 082	
VAN	-370	-622	343	1 425	

Periodo	Indicador Financiero						
	VAN	TIR	Relación Beneficio - Coste	Payback	año(s)	mes(es)	día(s)
Años [1 a 3]	1 425	97%	2,31	1,64	1	8	22

## **Comentarios**

- La plataforma OT que es una solución On Premise tiene una inversión alta, que es recuperada durante el segundo año de generación de valor de los proyectos.
- La plataforma OT por su naturaleza requiere de inversión a medida que se demandan más capacidades de procesamiento.

### **7.1.1.4 Data Governance**

## **Business Case**

Tabla 19: BC Data Governance

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Data Governance	0	-169	-179	-179	-421
Opex	0	-109	-119	-119	-276
Capex	0	-60	-60	-60	-144
Captura de Valor	0	0	0	0	0

Periodo	Indicador Financiero			
	VAN	TIR	Relación Beneficio - Coste	Payback
Años [1 a 3]	-421	n/a	n/a	n/a

### **Comentarios**

- Al ser un proyecto fundacional tiene asociado el coste de su implementación y de operación.
- No se presentan ingresos, porque los efectos positivos se presentan a través del valor capturado por los Use Cases.

#### **7.1.1.5 Data Literacy**

La alfabetización en datos es un proyecto fundamental que tiene como objetivo desarrollar los planes y programas formativos que permitan acortar las brechas de conocimiento asociadas a datos. La alfabetización en datos es una nueva competencia empresarial.

### **Definición**

Data Literacy (alfabetización en datos) según Gartner:

"La capacidad de leer, escribir y comunicar datos en su contexto, incluyendo las fuentes y las estructuras de datos, los métodos, técnicas analíticas, técnicas aplicadas y la capacidad de describir la aplicación del caso de uso y el valor resultante" — Gartner IT Glossary.

### **Relevancia**

Acorde con las conclusiones de la quinta encuesta anual 2020 a los CDOs realizada por Gartner, la falta de alfabetización en datos es uno de los tres principales factores que impiden el éxito de los equipos de D&A.

### **Proyecto**

- El proyecto inicia con definición de un equipo transversal en la empresa a cargo del proyecto de Data Literacy. El equipo estará conformado el CDO, un responsable del equipo de Recursos Humanos y los responsables de líneas de negocio.

- El equipo tendrá como primer objetivo el identificar los roles que requieren alfabetización de datos y el nivel de alfabetización necesario. De igual forma se deben de establecer los objetivos de aprendizaje por rol.
- Se diseñarán cursos de formación personalizados por rol y línea de negocio, los tópicos de mayor relevancia y competencias a alcanzar.
- Colaborar con los responsables de RRHH y las líneas de negocio para hacer evaluaciones de las competencias actuales, diseñar hojas de ruta para la mejora de las competencias, crear un plan de estudios y determinar los KPI's de la formación.
- Un insight de este proyecto es la actualización de los perfiles de los empleados, para introducir como requisito niveles mínimos de alfabetización de datos para ser contratados en la empresa.

### **Business Case**

Tabla 20: BC Data Literacy

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Data Literacy	0	-22	-24	-24	-56
Opex	0	-22	-24	-24	-56
Capex	0	0	0	0	0
Captura de Valor	0	0	0	0	0

Periodo	Indicador Financiero			
	VAN	TIR	Relación Beneficio - Coste	Payback
Años [1 a 3]	-56	n/a	n/a	n/a

### **Comentarios**

- Al ser un proyecto fundacional tiene asociado el coste de su implementación y de operación.
- No se presentan ingresos, porque los efectos positivos se presentan a través del valor capturado por los Use Cases.

### 7.1.2 Proceso de fabricación del acero de construcción

Antes de detallar los casos de uso, daremos un alcance respecto al proceso de fabricación del acero de construcción.

*SIDERPERU* produce acero vía el uso de un horno eléctrico, para ello usa el acero reciclado (chatarra). La fabricación consta de cinco procesos:

Paso 01: Industrialización de la chatarra

*SIDERPERU* tiene proveedores nacionales e internacionales de chatarra, en este proceso la chatarra es cortada a una medida adecuada y se quitan todas las impurezas para luego pasar al proceso del horno eléctrico.

Paso 02: Horno eléctrico

El horno eléctrico transforma la chatarra en acero líquido, yendo de una temperatura de ambiente a 1,600°C aproximadamente. Este acero líquido es depositado en unos recipientes llamados “colada”.

Paso 03: Horno Cuchara

Se encarga de tomar la composición química del acero y llevarla a la composición química requerida. Es decir, toman muestras, observan los resultados y agregan los aleantes necesarios para tener la composición química deseada para cada tipo de barra de acero.

Paso 04: Colada continua

Recibe el acero líquido a una temperatura de 1,600°C y se encarga de solidificar el acero con un proceso de enfriamiento que pasa por unos moldes de cobre y por enfriamiento con agua. En este proceso se obtiene un semiproducto denominado “palanquilla”.

Paso 05: Laminación

Es un proceso que toma la palanquilla y se encarga de transformarla mediante un proceso de reducciones donde una serie de rodillos van girando y aplastando la palanquilla hasta obtener las barras de construcción (producto final).

*Figura 30: Gráfico proceso de producción del acero de construcción*



Los datos utilizados en el desarrollo de los casos de uso son referenciales.

### 7.1.2.1 Optimización en uso de ferroaleaciones – U1

#### Contexto

Este caso de uso se ejecutará en el paso 03 del proceso de producción de acero (planta acería). Para que el acero de construcción obtenga las propiedades físicas y mecánicas es necesario que cuando el acero líquido este en el horno cuchara se agregue ferroaleaciones.

Actualmente la dosificación de las ferroaleaciones se realiza de forma manual, generando un consumo elevado de este.

El costo por tonelada de las ferroaleaciones es aproximadamente el 4% respecto del costo por tonelada de acero de construcción producido.

#### Justificación

Como resultado de la adopción de la Industria 4.0, los costos de producción podrían disminuir entre un 10 y un 30 % según (Bauernhansel et al., 2016), así mismo, de acuerdo con un estudio de las empresas manufactureras alemanas estas refieren mejoras en los costos de producción por ahorro en materia prima está entre el 5% - 8% (Rüßmann et al., 2015).

Con este modelo buscamos reducir el consumo de ferroaleaciones y tener un consumo adecuado que garantice lograr las propiedades requeridas para el acero de construcción, la reducción esperada es 8%, siendo el ahorro anual de US\$ 1'140,000.

## **Business Case**

Tabla 21: BC Optimización en uso de ferroaleaciones – U1

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Optimización en uso de ferroaleaciones	-661	0	916	1 011	789
Opex			-129	-129	-219
Capex	-661	0	0	0	-661
Captura de Valor	0	0	1 045	1 140	1 644
Valor actual del FCE	-661	0	730	719	
VAN	-661	-661	69	789	

Periodo	Indicador Financiero						
	VAN	TIR	Relación Beneficio - Coste	Payback	año(s)	mes(es)	día(s)
Años [1 a 3]	789	54%	1,87	1,90	1	11	26

## **Comentarios**

- A tres años, es un proyecto con un VAN (US\$ 789,000) y un TIR (54%) que hacen financieramente viable su implementación. Se tiene un Payback durante el segundo año.
- El primer año el modelo no captura valor y el segundo año desde febrero.

### **7.1.2.2 Analítica Predictiva de Ventas – U2**

#### **Contexto**

El mercado del acero de construcción este compuesto por dos tipos de mercado, el 70% corresponde a la construcción y el 30% a la construcción formal. La autoconstrucción proviene del crecimiento de la informalidad lo que complica su identificación, la construcción formal deja una huella digital de las obras que se ejecutarán en el país haciendo que su identificación sea predecible.

La venta de SIDERPERU está en la misma proporción del mercado, por lo que la actual referencia para construir el mercado del acero de construcción es el cemento. El cemento al ser un producto perecible hace que el comportamiento de sus despachos dé como referencia un crecimiento o contracción del sector construcción.

Depender de esta información trae como consecuencia lo siguiente:

- No tienen una adecuada segmentación del mercado.
- Tienen dificultad en la generación de nuevos productos y servicios.
- Les falta rapidez y certeza en la toma de decisiones.

- Se les presenta inconsistencias entre el comportamiento del acero y el cemento que deben ser analizadas y sustentadas con otras variables (tipo de obra, incremento de la informalidad, índice de pobreza, etc)

### **Justificación**

Con la implementación de este modelo buscamos incrementar la participación de SIDERPERU en el mercado construcción formal a 42% a partir del año ..., permitirá además que con las obras plenamente identificadas realizar un análisis de acuerdo con el tipo de obra (colegios, hospitales, puentes, edificios multifamiliares, etc), ubicación, inicio de obra, etc y vincularlo con la previsión de ventas logrando una optimización de los niveles de stock por cada diámetro de barra.

### **Business Case**

- Luego de un análisis de las necesidades y la revisión de las buenas prácticas de la industria, en este modelo sabemos que se debe de conseguir, está pendiente el definir el como hacerlo, pero será consecuencia directa de la madurez que los proyectos anteriores generen, así como el asesoramiento de expertos.
- Este modelo tiene un potencial muy alto de beneficios que en función de la calidad de las estimaciones es muy variable.
- De la revisión del numeral 6.6 de presupuesto general, se tiene un VAN para el segundo año de US\$ 1'961,000, esto nos permite definir que el proyecto "Analítica Predictiva de Ventas" no puede iniciar antes del cierre del segundo año y nos delimita el gasto en función de la estrategia a definir con el apoyo de los Stakeholders.

#### **7.1.2.3 Modelo de sobrecalentamiento – U3**

##### **Contexto**

Este caso de uso se ejecutará en el paso 02 del proceso de producción de acero (planta acería). Es en el horno eléctrico que el acero se funde a una temperatura aproximada de 1,600 grados centígrados. Durante el sobrecalentamiento el acero líquido tiene un rango de temperatura ideal. Si el acero sale demasiado caliente, los operadores del equipo deben ralentizar el paso de fundición. Si el acero no está lo suficientemente caliente, puede "congelarse" antes de que se haya fundido, lo que compromete su calidad. Cuando el acero se congela la producción se pierde y este material se convierte en chatarra.

El proceso de sobrecalentamiento actualmente se ejecuta en función de la experiencia del operario y de las indicaciones de los sistemas de control.

La pérdida de producción por no llegar a la temperatura ideal genera una merma o pérdida con un costo aproximado del 8% con respecto al costo de la producción de acero líquido.



## Justificación

Como resultado de la adopción de la Industria 4.0, los costos de producción podrían disminuir entre un 10 y un 30 % según (Bauernhansel et al., 2016), así mismo, de acuerdo con un estudio de las empresas manufactureras alemanas estas refieren mejoras en los costos de producción por ahorro en materia prima está entre el 5% - 8% (Rüßmann et al., 2015).

Con este modelo, esperamos reducir en 5% la merma generada, siendo el ahorro anual esperado de US\$ 700,000.

## Business Case

Tabla 22: BC Optimización en uso de ferroaleaciones – U1

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Modelo de sobrecalentamiento	-350	0	394	627	410
Opex			-73	-73	-123
Capex	-350	0	0	0	-350
Captura de Valor	0	0	467	700	700
Valor actual del FCE	-350	0	314	446	
VAN	-350	-350	-36	410	

Periodo	Indicador Financiero				año(s)	mes(es)	día(s)
	VAN	TIR	Relación Beneficio - Coste	Payback			
Años [1 a 3]	410	52%	1,48	2,08	2	1	29

## Comentarios

- A tres años, es un proyecto con un VAN (US\$ 410,000) y un TIR (52%) que hacen financieramente viable su implementación. Se tiene un Payback durante el tercer año.
- El primer año el modelo no captura valor y el segundo año desde mayo.

### 7.1.2.4 Mantenimiento Predictivo – U4

## Contexto

La industria siderúrgica tiene un proceso continuo en el que un paro de producción conlleva a incurrir en un costo alto, riesgos de seguridad e impactos medioambientales. Por lo que el mantenimiento es vital para tener un proceso seguro y buscar tener la mayor productividad posible.

Existe un concepto denominado tiempo de inactividad de la producción y se refiere al período de tiempo en el que la producción se detiene sin producir, este tiempo se clasifica en dos categorías:

1. Tiempo de inactividad planificado: De acuerdo con mantenimiento programado, en este período no se produce normalmente, este tiempo se utiliza principalmente para evitar interrupciones repentinas.

2. Tiempo de inactividad no planificado: Este tiempo es contrario al planificado y se refiere a la cantidad de tiempo que la producción está parada.

El tiempo de inactividad no planificado es más costoso que el planificado.

Actualmente en SIDERPERU la proporción de estas dos categorías en costos de mantenimiento son: 45% en tiempo de inactividad planificada y 55% en tiempo de inactividad no planificada.

Esto conlleva a tener los siguientes inconvenientes:

- Falta de identificación de fallas en equipos críticos.
- Mantenimiento reactivo cuando se presentan fallas en los equipos.
- Lubricación deficiente de equipos críticos.
- Equipos obsoletos dificultad de repuestos.
- Tuberías, equipos y máquina obsoletos.

### **Justificación**

En el informe de McKinsey (Bradbury et al., 2018), afirman que el mantenimiento predictivo puede ayudar a las empresas reducir sus costos de mantenimiento entre un 18-25 %.

Con este modelo esperamos un ahorro del 3% en "Tiempo de inactividad planificado" y 5% en "Tiempo de inactividad no planificado", siendo el ahorro esperado de US\$ 781,000.

Algunos beneficios adicionales son:

- Permitirá tener un monitoreo remoto de los equipos de las plantas, prediciendo las fallas.
- Permitirá tener acciones correctivas preventivas.
- Disminución del tiempo de inactividad de los equipos.
- Evitar fallas de equipos críticos de cada planta controlando sus variables y reducir las paradas que estos producen.
- Controlar adecuadamente el proceso.
- Evitar incurrir en mayores horas de producción evitando paradas.

En conclusión, se conseguirá tener "Trabajo de mantenimiento más eficiente"

### **Business Case**

Tabla 23: *BC Mantenimiento Predictivo – U4*

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Mantenimiento Predictivo	-74	428	688	688	1 346
Opex		-93	-93	-93	-224
Capex	-74	0	0	0	-74
Captura de Valor	0	521	781	781	1 644
Valor actual del FCE	-74	382	548	490	
VAN	-74	308	856	1 346	

Periodo	Indicador Financiero				año(s)	mes(es)	día(s)
	VAN	TIR	Relación Beneficio - Coste	Payback			
Años [1 a 3]	1 346	625%	5,52	0,19	0	2	10

### **Comentarios**

- A tres años, es un proyecto con un VAN (US\$ 1'346,000) y un TIR (625%) que hacen financieramente viable su implementación. Se tiene un Payback durante el primer año.
- El primer año el modelo captura valor desde mayo.

### **7.1.2.5 Modelo de visibilidad para el guero – U5**

#### **Contexto**

Este caso de uso se ejecutará en el paso 04 del proceso de producción de acero y tendrá impacto en la producción y en la seguridad de la planta de acería.

El acero líquido es depositado en un recipiente denominado "colada", esta es trasladada a través de una grúa hasta el horno cuchara.

Del horno cuchara el acero líquido es vaciado a unas lingoteras que son unos moldes que darán la forma rectangular a la sección del producto terminado denominado palanquilla.

En este proceso la compañía tiene inconvenientes tales como:

- Falta de visión para el guero en el cambio de lingoteras.
- Limitaciones para la operación de puentes grúa debido a la poca visibilidad, capacidad de carga, manipulación de carga.
- Falta de iluminación para los operadores grúa y en general.

#### **Justificación**

Como resultado de la adopción de la Industria 4.0, los costos de producción podrían disminuir entre un 10 y un 30 % según (Bauernhansel et al., 2016), así mismo, de acuerdo con un estudio de las empresas manufactureras alemanas estas refieren mejoras en los costos de producción por ahorro en materia prima está entre el 5% - 8% (Rüßmann et al., 2015).

Con este modelo esperamos reducir en un 5% la merma generada actualmente, siendo el ahorro anual esperado de US\$ 1'361,000; esto se conseguirá por que el modelo propondrá tener un adecuado cambio de las lingoteras, garantizando que estas no fallen evitando que se genere

pérdida de acero líquido (merma), además evitará paradas de producción y minimizará el riesgo de exposición de los trabajadores por colocación inadecuada de las lingoteras.

### **Business Case**

Tabla 24: BC Modelo de visibilidad para el guero – U5

Proyecto	Año 0	Año 1	Año 2	Año 3	Años [1 a 3]
Modelo de visibilidad para el guero	-129	596	1 050	1 050	1 988
Opex		-311	-311	-311	-747
Capex	-129	0	0	0	-129
Captura de Valor	0	907	1 361	1 361	1 361
Valor actual del FCE	-129	532	837	747	
VAN	-129	404	1 241	1 988	

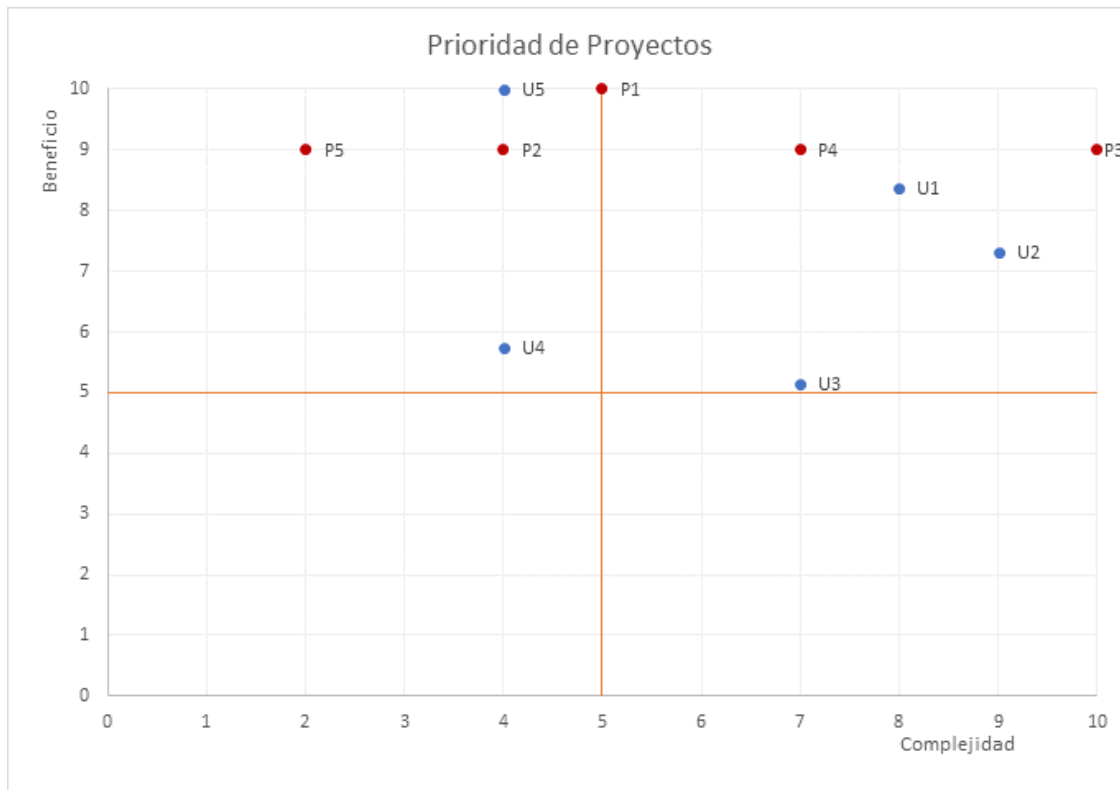
Periodo	Indicador Financiero				año(s)	mes(es)	día(s)
	VAN	TIR	Relación Beneficio - Coste	Payback			
Años [1 a 3]	1 988	517%	1,55	0,24	0	3	27

### **Comentarios**

- A tres años, es un proyecto con un VAN (US\$ 1'988,000) y un TIR (517%) que hacen financieramente viable su implementación. Se tiene un Payback durante el primer año.
- El primer año el modelo captura valor desde mayo.

## 7.2 Priorización y Definición de Precedencias

Figura 31: Diagrama de Beneficio y Complejidad por Proyecto



### Comentarios

- El beneficio de los proyectos se ha calculado en función del valor capturado por mes en una escala de 1 a 10, siendo el de mayor beneficio 10.
- De P1 a P5 (color rojo en el diagrama) son los proyectos fundacionales que inician con máxima prioridad, P1 a P3 desde el primer mes y P4 a P5 desde el segundo mes.
- Los casos de uso 4 y 5 por su alto impacto y complejidad media inician antes que los demás, El caso de uso 4 a cargo de una consultara y el caso de uso 5 a cargo del equipo de D&A.

## 7.3 Evolución de indicadores de medición de la estrategia

Como parte del proceso de control y medición del progreso en la ejecución de la estrategia de D&A, se definieron los indicadores en el numeral 6.1.2 así como la periodicidad de medición. Efectuadas las actualizaciones se desarrollaran reuniones de trabajo para evaluar cualquier desviación y planes concretos para volver a estar en cronograma o para ajustar las expectativas.

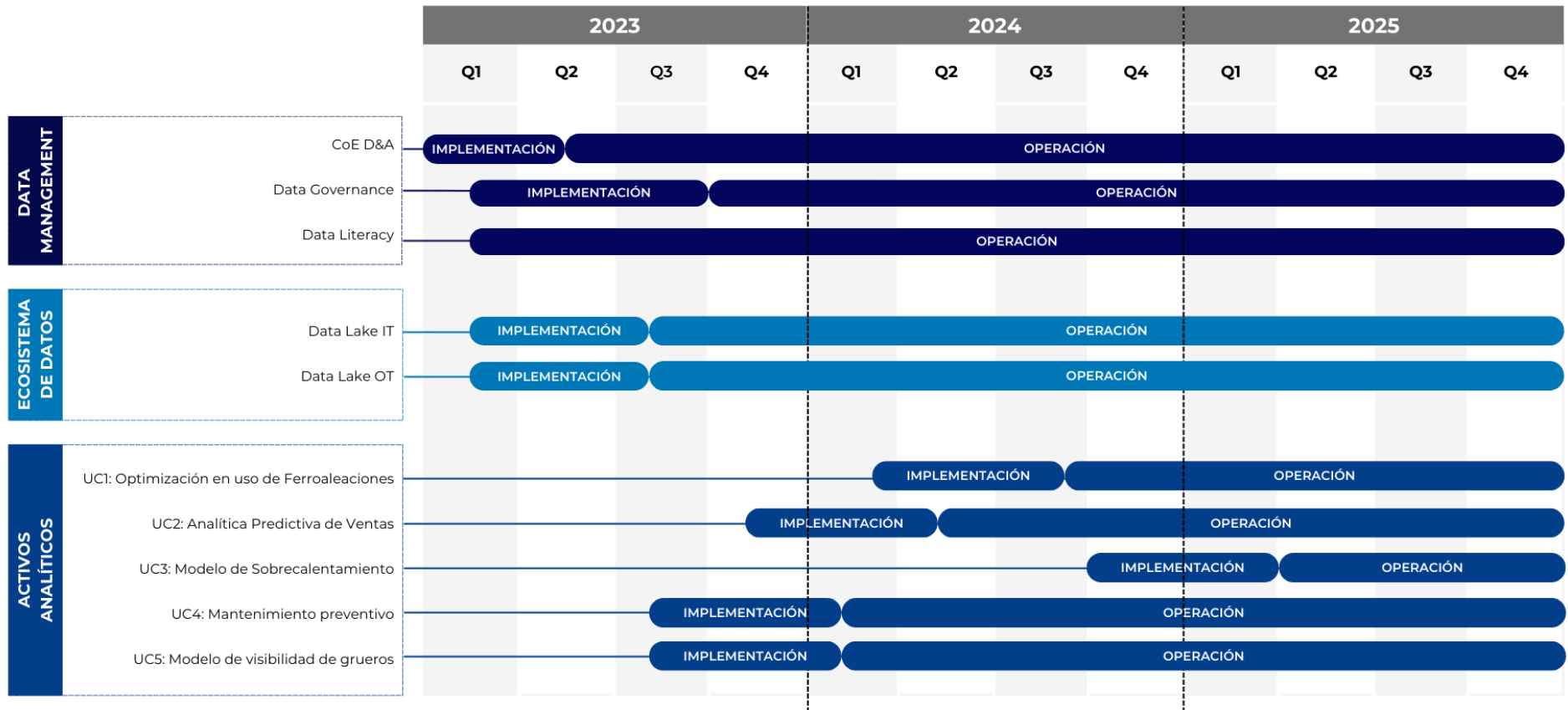


Tabla 25: *Matriz de Beneficio y Complejidad por Proyecto*

Código	Caso de Uso	Prioridad	Mes Inicio	Mes Fin	Project Go-Live	Complejidad Eje X	Beneficio Eje Y
P1	CoE D&A	0	1	36	1	5	10
P2	Data Lake IT	0	1	36	1	4	9
P3	Data Lake OT	0	1	36	1	10	9
P4	Data Governance	0	2	36	2	7	9
P5	Data Literacy	0	2	36	2	2	9
U1	Optimización en uso de ferreoaleaciones	4	14	36	18	8	8,4
U2	Analítica Predictiva de Ventas	3	11	36	15	9	7,3
U3	Modelo de sobrecalentamiento	5	17	36	21	7	5,1
U4	Mantenimiento Predictivo	1	5	36	9	4	5,7
U5	Modelo de visibilidad para el gruero	1	5	36	9	4	10,0

### Comentarios

- El beneficio de los proyectos se ha calculado en función del valor capturado por mes en una escala de 1 a 10, siendo el de mayor beneficio 10.
- Los proyectos P1, P2, P3, P4 y P5 son los proyectos fundacionales que inician con máxima prioridad, P1 a P3 desde el primer mes y P4 a P5 desde el segundo mes.
- Los casos de uso 4 y 5 por su alto impacto y complejidad media, inician en el mes 5 luego implementados el proyecto P2.
- El caso de uso 4 a cargo de una consultara y el caso de uso 5 a cargo del equipo de D&A.

## 7.4 Plan de Inversiones

Todos los valores están expresados en miles de dólares americanos, los flujos incluyen signo.

Tabla 26: Presupuesto General

		Año 0	Año 1	Año 2	Año 3	Año 4	Año 5	Años [1 a 3]	Años [1 a 5]
Ingresos [A]	Captura de Valor	0	1 428	3 654	3 982	3 982	3 982	7 022	11 812
Egresos [B]	Opex	0	-599	-873	-1 047	-1 066	-1 085	-1 976	-3 269
Inversión [C]	Capex	-480	-305	-305	-305	-305	-305	-1 213	-1 579
	VAN							3 833	6 963
FCO = [A] + [B]	Flujo de Caja Operativo	0	829	2 781	2 935	2 916	2 897		
FCE = [A] + [B] + [C]	Flujo de Caja Económico	-480	524	2 476	2 630	2 611	2 592		
	Valor actual del FCE	-480	468	1 974	1 872	1 659	1 471		
	VAN	-480	-12	1 961	3 833	5 493	6 963		

Periodo	Indicador Financiero						
	VAN	TIR	Relación Beneficio - Coste	Payback	año(s)	mes(es)	día(s)
Años [1 a 3]	3 833	222%	2,20	1,01	1	0	2
Años [1 a 5]	6 963	233%	2,44	1,01	1	0	2

### Comentarios

- La inversión inicial es de US\$ 480 mil dólares americanos.
- Se ha efectuado a nivel financiero un análisis a 3 y 5 años concluyendo que en ambos casos el proyecto tiene un VAN muy conveniente con tasas internas de retorno de 222% a 3 años y 233% a 5 años. La relación beneficio/coste nos indica que a 3 años por cada US\$ el proyecto genera US\$ 1,20 y a 5 años por cada US\$ el proyecto genera US\$ 1,44.
- En ambos casos el Payback es a 1 año y 2 días.
- El general está compuesto por 38% CAPEX y 62% OPEX.



## CAPÍTULO 8.- CONCLUSIONES Y RECOMENDACIONES

### CONCLUSIONES

1. Desde la perspectiva financiera la ejecución de la propuesta de estrategia de D&A es muy conveniente:

Al 100% de cumplimiento de previsiones de ingresos.	Al 60% de cumplimiento de previsiones de ingresos.
<ul style="list-style-type: none"><li>– A tres años tiene un VAN de US\$ 3'833,000.</li><li>– Una TIR de 222%.</li><li>– Una relación Beneficio-Coste de 2,20 y un Payback durante el segundo año.</li></ul>	<ul style="list-style-type: none"><li>– A tres años tiene un VAN de US\$ 1',024,000.</li><li>– Una TIR de 78%.</li><li>– Una relación Beneficio-Coste de 1,32 y un Payback durante el segundo año.</li></ul>

2. La industria del acero administra grandes volúmenes de producción, el impacto de la optimización de procesos haciendo uso de D&A generará beneficios económicos muy significativos.
3. SIDERPERU desarrollará una ventaja competitiva haciendo uso de D&A para optimizar sus procesos y toma de decisiones.
4. La implementación de capacidades de D&A no es producto de un “Big Bang”, es consecuencia de la ejecución de una estrategia de medio y largo plazo de desarrollo constante, en la que es determinante: un cambio cultural, la gestión del cambio (en Latinoamérica la industria del acero es muy tradicional), incremento de capacidades y la tecnología.

### RECOMENDACIONES

1. Los expertos de los diferentes dominios de negocio deben involucrarse activamente en la creación de los activos analíticos para garantizar la calidad de los datos.
2. La cantidad de datos que se generan hoy por hoy está en constante crecimiento, así como también aumentan la cantidad de fuentes de datos. Por lo tanto, es importante diseñar una arquitectura de datos con capacidad de escalabilidad
3. Seleccionar como primeros proyectos para la implementación los de complejidad media o baja que representen un menor riesgo, permitiendo así el ajuste de procesos y la captura de valor en el menor tiempo posible. Consecuencia de lo anterior se fortalecerá la credibilidad y la motivación para sumarse a la estrategia de Data & Analytics de la empresa.

## REFERENCIAS BIBLIOGRÁFICAS

- Bauernhansel, T. y Krüger, J. y Reinhart, G. y Schuh, G. (2016). WGP-Standpunkt Industrie 4.0. *WGP-Standpunkte, 1*.
- Bradbury, S. y Carpizo, B. y Gentzel, M. y Horah, D. y Thibert, J. (2018). Digitally enabled reliability: Beyond predictive maintenance. *McKinsey & Company, October*, 1–2.
- DAMA. (2022). *Sitio web DAMA*. <https://www.dama.org/cpages/home>
- Data Lakes I IBM*. (s/f). Recuperado el 21 de octubre de 2022, a partir de <https://www.ibm.com/es-es/topics/data-lake>
- Data Management Association y Henderson, D. y Earley, S. (2017). *DAMA-DMBOK : data management body of knowledge*.
- Davenport, T. H. (2018). From analytics to artificial intelligence. *Journal of Business Analytics, 1(2)*, 73–80. <https://doi.org/10.1080/2573234X.2018.1543535>
- Forbes. (2017). *DATA & ADVANCE ANALYTICS IN CANADA - HIGH STAKES, HIGH REWARDS:*
- Gartner. (s/f). *What Is Data and Analytics?* Gartner. Recuperado el 21 de octubre de 2022, a partir de <https://www.gartner.com/en/topics/data-and-analytics>
- GERDAU. (2019). *RELATO INTEGRADO GERDAU 2019*.
- HAI, H.-C. A. I. (2022). *Artificial Intelligence Index Report 2022*.
- Industria siderúrgica: mayores productores de acero del mundo en 2021 | Statista*. (s/f). Recuperado el 21 de octubre de 2022, a partir de <https://es.statista.com/estadisticas/600708/productores-de-acero-a-nivel-mundial--por-volumen-de-produccion/>
- Industry 4.0 | ArcelorMittal*. (s/f). Recuperado el 21 de octubre de 2022, a partir de <https://corporate.arcelormittal.com/media/case-studies/industry-4-0>
- Industry 4.0 | thyssenkrupp*. (s/f). Recuperado el 21 de octubre de 2022, a partir de <https://www.thyssenkrupp.com/en/company/innovation/industry-4-0>
- Kebande, V. R. (2022). Industrial internet of things (IIoT) forensics: The forgotten concept in the race towards industry 4.0. *Forensic Science International: Reports, 5*. <https://doi.org/10.1016/j.fsir.2022.100257>

- Laney, D. (2016). *The Gartner Enterprise Information Management Maturity Model*. Gartner Webinar.
- Las principales aplicaciones del IoT en la industria | Meinsa.* (s/f). Recuperado el 26 de octubre de 2022, a partir de <https://meinsa.com/2020/11/las-principales-aplicaciones-del-iot-en-la-industria/>
- Lock, M. (2017). *ANGLING FOR INSIGHT IN TODAY'S DATA LAKE*.
- Manyika, J. y Chui Brown, M. y B. J., B. y Dobbs, R. y Roxburgh, C. y Hung Byers, A. (2011). Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute, June*.
- Marz, N. y Warren, J. (2015). *Big Data: Principles and best practices of scalable realtime data systems &gt; Chapter 1. A new paradigm for Big Data : Safari Books Online*. Manning Publications.
- McCarthy, J. y Minsky, M. L. y Rochester, N. y Shannon, C. E. (2006). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*.
- McKinsey. (2021). *How digital and analytics can unlock full potential in steel*.
- McKinsey y Kumar, R. y Dubashi, N. y Gupta, R. y Singh, K. (2021). *How a steel plant in India tapped the value of data-and won global acclaim - McKinsey & Company*.
- Memoria Anual 2021. (2022). *SIDERPERU, Febrero*, 16–17. <https://ri.gerdau.com/a-gerdau/siderperu/>
- Memoria Anual Integrada 2021.* (s/f).
- ¿Qué es el Internet de las cosas (IoT)? | Oracle Perú.* (s/f). Recuperado el 26 de octubre de 2022, a partir de <https://www.oracle.com/pe/internet-of-things/what-is-iot/>
- RAE. (2022). *Inteligencia Artificial - RAE*. <https://dle.rae.es/inteligencia#2DxmhCT>
- Rodriguez, J. y Kothiyal, M. y Kalvenes, J. y Wolfgang, M. y Lukic, V. y Nath, G. (2021). Strengthening the Steel Industry with AI. *BCG*. <https://www.bcg.com/publications/2021/value-of-ai-in-steel-industry>
- Rumbold, J. M. M. y O’Kane, M. y Philip, N. y Pierscionek, B. K. (2020). Big Data and diabetes: the applications of Big Data for diabetes care now and in the future. En *Diabetic Medicine* (Vol. 37, Issue 2). <https://doi.org/10.1111/dme.14044>

- Rüßmann, M. y Lorenz, M. y Gerbet, P. y Waldner, M. y Justus, J. y Engel, P. y Harnisch, Michael. (2015). Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. *Boston Consulting Group*, 9(1). [https://www.bcg.com/publications/2015/engineered\\_products\\_project\\_business\\_industry\\_4\\_future\\_productivity\\_growth\\_manufacturing\\_industries](https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries)
- Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 44(1–2). <https://doi.org/10.1147/rd.441.0206>
- SIDERPERU. (2022). *Reporte de Sostenibilidad 2021*. <https://ri.gerdau.com/gerdau/siderperu/>
- Soares, S. (2012). The IBM data governance unified process: driving business value with IBM software and best practices. *ISACA JOURNAL*, 4.
- The IBM Data Governance Council Maturity Model\_ Building a. (2007). *IBM Corporation*, 8–10.
- Thomas, R. y Zikopoulos, P. (2020). *AI Ladder*. O'Reilly Media, Incorporated.
- Vaughan, D. (2020). *Analytical Skills for AI and Data Science*. O'Reilly Media, Incorporated.

## ANEXOS

### Anexo 1: Data Governance Maturity Model – AS IS

DATA GOVERNANCE MATURITY MODEL--		ACHIEVE
	VERDADERO	INVEST
	FALSO	IMPROVE
	PARTIAL	MAINTAIN
ACHIEVE, meaning you have to do it and will invest whatever it takes		
INVEST, meaning the client is willing to put more time, effort, funding into improving the condition		
IMPROVE, meaning that the client is not yet prepared to make a solid investment but wants to identify that somehow, through shifting responsibilities or changing policies, they will improve on that particular issue		
MAINTAIN, meaning that they won't can't do anything about this now		

1 Organizational Structure & Awareness	AS IS	TO BE
Comments - Maturity levels of organizational awareness of Information Governance will establish themselves by the level of teamwork between the business and IT. A strong awareness of the value of data and the risks around that data become more prominent and creates effective control structures that enable persistent measurement and refinement of governance behaviors across the enterprise.		
High-Level Questions:  1) To what extent do you have Management interest in and support of Information Governance across your lines of business?  2) To what extent do you have an Information Governance Organization (IGO) resourced and structured to address the enterprise's Information Governance needs?  3) To what extent do you have a Communication plan designed to convey Information Governance objectives and value throughout the organization?		
Managers use basic transactional data to measure and report on business results.	VERDADERO	ACHIEVE
There are some policies and practices in place for the use of data within the Enterprise.	FALSO	IMPROVE
Awareness of the value of data is present and is the subject to some form of improvement process.	PARTIAL	IMPROVE
There is growing awareness as to the importance of data and its role in the quality of decisions.	PARTIAL	IMPROVE
While data is generally departmentalized, there is an increasing need/awareness for cross Enterprise collaboration. data.	VERDADERO	ACHIEVE

There exists a data blueprint, flow or map that IT uses to understand how data originates and is used.	FALSO	MAINTAIN
A business case exists for advancing Data Governance and it is tied to financial benefits.	PARTIAL	IMPROVE
The risk(s) associated with bad data or incomplete information are understood and assessed.	PARTIAL	IMPROVE
There is a written plan on how data and information is managed within the Enterprise.	FALSO	IMPROVE
Departmental data requirements are shared for purposed of improving security, privacy, and "single view".	FALSO	IMPROVE
A Information Governance Competency Center exists and is actively delivering significant, measurable value.	FALSO	IMPROVE
Enterprise wide data transformation capabilities exist, MDM is implemented and analytics are effectively used.	n/a	n/a
Information Governance is a foundation for effective Enterprise Performance Management.	n/a	n/a
The organization uses information as a competitive advantage and sees data as a primary contributor.	n/a	n/a
Information Governance is funded and fully integrated into Enterprise Governance.	n/a	n/a

2 Stewardship	AS IS	TO BE
Comments - Stewardship is at the foundation of strong Data Governance practice. Key to understanding Stewardship is to understand how committed the Enterprise is to individual, departmental and eventually, Corporate or Enterprise wide data quality, ownership.		
High-Level Questions: 1) How is data stewardship managed in your organization? 2) Provide answers separately for SAP and non-SAP environments? 3) Is the organization ready to make the investment in a Stewardship Program in order to improve data quality?		
Some departments or areas are responsible for quality of data within the Enterprise and have assigned Stewards.	PARTIAL	INVEST
There is some understanding at the Departmental level as to the "sources and uses" of data.	VERDADERO	ACHIEVE
Changes to a Data definition during projects are often documented/shared with other analysts.	PARTIAL	IMPROVE
Data definitions are often managed by Stewards across the Enterprise resulting in measurable improvement.	PARTIAL	IMPROVE

Conflicting data definitions are generally documented and analyzed and resolution ownership exists.	PARTIAL	IMPROVE
Executive Management recognizes a need for data stewards and is committed to an investment.	PARTIAL	INVEST
The Data Steward role includes meta data management within domain, data exception handling, and security.	n/a	IMPROVE
Stewardship processes, procedures, & standards exist and are "assigned", measured and communicated.	PARTIAL	IMPROVE
The appropriate Data Steward resource is in place and the role/responsibility has been effectively communicated.	n/a	IMPROVE
Data stewards have direct authority and influence development of information systems.	FALSO	IMPROVE
Stewardship responsibility includes data Life-Cycle Management across the Enterprise.	FALSO	IMPROVE
Stewards directly contribute to Performance Management for business processes and data relationships.	n/a	n/a
Stewards promulgate the use of data standards/quality, and directly interact with stakeholders/senior mgmt.	n/a	n/a
Stewards collaborate with process owners on business changes where data & data quality impact KPIs.	PARTIAL	IMPROVE
Stewards contribute SLAs related to Data Quality, define data standards and monitor enterprise performance.	PARTIAL	IMPROVE

3 Policy	AS IS	TO BE
<p>Comments - Includes corporate policies and standards around Information Governance. Levels of maturity will be based on creation and formalization of Information Governance policies and standards and then by full adoption and incorporation into the basic fabric of an organization.</p> <p>Examples of Information Governance policies include the following:</p> <ol style="list-style-type: none"> <li>1) Retention periods for specific document types (e.g., bank statements, invoices to customers, inventory records, safety records).</li> <li>2) Electronically Stored Information should be treated as part of Records Management policy.</li> <li>3) Any record that is relevant to any pending or anticipated litigation, claim, audit, agency charge, investigation or enforcement action shall be retained at least until final resolution of the matter.</li> <li>4) Enterprise data is owned by the enterprise and not by individual lines of business.</li> </ol>		

5) Enterprise data may be shared with named external data providers provided the data is anonymized and is not used to compete with the business.		
6) HR will create job descriptions for Data Stewards who will report into the Lines of Business and be responsible for Enterprise Data.		
High-Level Questions:		
1) To what extent do you have a strategy for managing, maintaining, developing and deploying enterprise Information Governance policies aligned with corporate objectives?		
2) To what extent do you have a strategy for tracking overall Information Governance compliance with appropriate emphasis on accountability?		
3) To what extent do you have processes to design, implement, and manage operational best practices in the area of Information Governance?		
4) To what extent do you have routine and consistent periodic reviews of Information Governance policies and procedures?		
There are currently rudimentary policies in place for the management of Enterprise data.	VERDADERO	ACHIEVE
Data policy is evident with at least one Governance area (Data Quality, Security/Privacy, Risk) Management).	VERDADERO	ACHIEVE
Reviews and updates to data policies and procedures are evident.	PARTIAL	IMPROVE
Enterprise Data and Information related policies are collaborated across many or several departments.	PARTIAL	IMPROVE
Some data "best practices", related to policy, have been identified and shared.	PARTIAL	IMPROVE
Influence for Data policy is evident at both the enterprise as well as the departmental levels.	FALSO	IMPROVE
Data Policies directly relate to the quality of data, building trust and confidence with end users.	PARTIAL	IMPROVE
Data Policy include data Life-Cycle-Management (in addition to Security and Privacy).	FALSO	IMPROVE
Data Policy results in improvements in data quality, recognized at the Departmental level.	PARTIAL	IMPROVE
Data Policy is under the domain of a formal Enterprise Information Governance initiative, program or CoE.	FALSO	IMPROVE
Policies support processes that design, implement and manage Information Governance "best practices".	FALSO	IMPROVE



New Data Governance Policies or the ability to change Policy is performed in a systematic and accountable process.	FALSO	IMPROVE
Improvements to Data Governance policies are part of a broader employee metrics.	n/a	n/a
Data Governance policies are integrated with Enterprise/Operational Policies for continuous improvement.	n/a	n/a
Data Governance policies support a complete Total Quality Metrics (TQM) capability for all Enterprise data and information.	n/a	n/a

4 Enterprise Value Creation	AS IS	TO BE
Comments - Business Outcomes define how an organization realizes returns on investment in the collection, production, and use of data. Returns may take the form of increases in operational efficiency, worker productivity, time-to-market, sales and revenue, as well as reductions in costs. Costs take many forms and may be realized as process failures, opportunity costs, scrap and rework costs, as well as the cost of failure to act upon the information presented.		
High-Level Questions:  1.) Discuss how you see a Data Governance program monitoring and measuring performance against clearly defined enterprise business objectives and the corporate charter?  2.) How will, in your opinion, Data Governance give you an ability to measure strategic initiatives by the results they generate?  3.) Can you describe the methods to identify the intrinsic value associated with your organization's data?  4.) Can you discuss with us the ability to quantify the costs of acquiring and maintaining its information?		
The Enterprise has a process/methodology for measuring performance against defined business objectives.	PARTIAL	INVEST
Business imperatives and strategic initiatives are acknowledged, acted upon and quantifiably measured.	PARTIAL	INVEST
The quality and value of data is an important factor to value creation and business outcomes.	PARTIAL	INVEST
Monitoring/measuring performance against business objectives is evident.	PARTIAL	INVEST
Management recognizes the need for data consistency and "single-view" may improve performance/outcomes.	PARTIAL	IMPROVE
The Enterprise is considering or has underway an SOA implementation for managing processes.	n/a	n/a
Methods to identify the financial value of enterprise data & business measures exist.	n/a	n/a
There is a formal Data Governance Center of Excellence (CoE) established.	FALSO	MAINTAIN

The Enterprise is considering or has underway a Master Data Mgmt. (MDM) implementation for managing data.	FALSO	INVEST
Data Governance provides direct value to business results and drives a formal ROI that relates directly to data quality.	PARTIAL	INVEST
The ability to measure through KPI's the success of all strategic initiatives by quantifiable results is evident.	PARTIAL	IMPROVE
A formal process is in place to link Data Governance & Business Outcomes, and adjust for Continuous Improvement.	FALSO	IMPROVE
Management implements data driven initiatives that provide reasonable value over an acceptable ROI threshold.	FALSO	IMPROVE
Intrinsic value of data is leveraged to effectively prioritize initiatives and proactively prevent breaches.	FALSO	IMPROVE
Data is related to outcomes as a "competitive advantage" and integrated, considered critical to outcomes/results.	n/a	n/a

<b>5 Data Risk Management &amp; Compliance</b>	<b>AS IS</b>	<b>TO BE</b>
Comments - Data Risk Management and Compliance seeks to understand how a organization or enterprise treats risk through the use of strong data management practices. Many of the elements of risk are intangible and often unassignable but concurrently have a major impact on the business. This section seeks to flush out how the organization deals with such costs and exposure, while highlighting the importance of compliance.		
High-Level Questions: 1.) Discuss how you see risk associated with the way the enterprise manages data? 2.) Does the enterprise take "compliance" seriously and if so, how does that manifest itself in specific business requirements? 3.) Can you site some examples of how tracability is being used or customer service being monitored to improve the quality of business outcomes? 4.) To what extent is Information Governance tied to corporate risk management Goals?		
Risk assessments for information management are conducted on a periodic basis.	FALSO	INVEST
The Executive Team views risk/exposure of data as a security priority.	PARTIAL	IMPROVE
Managing data risk is performed at the application level and is somewhat consistent across lines of business.	FALSO	IMPROVE
Risk remediation processes exist but may not be incorporated into formal Information Governance.	PARTIAL	IMPROVE
Risk assessment Audits are "consistent" and performed by each department.	PARTIAL	INVEST

Information risk, audit, and compliance requirements are understood and partially applied across the Enterprise.	PARTIAL	IMPROVE
The enterprise has defined processes for risk remediation and is aligned with Information Governance.	FALSO	IMPROVE
Information Risk Management and Audit is somewhat integrated across the enterprise.	PARTIAL	IMPROVE
Information Management goals include measurement and risk assessment that results in continuous improvement.	FALSO	IMPROVE
A strategic Information Governance framework, including risk/compliance, has been started.	FALSO	IMPROVE
Risk remediation is somewhat effective and proactive related to Information Governance.	FALSO	IMPROVE
External Audits are part of the CIO's Scorecard and includes "performance against audits".	FALSO	IMPROVE
The C Level and the IG Program Executive work in a coordinated fashion to minimize risk exposure.	FALSO	MAINTAIN
Periodic risk assessments for Governance are routinely conducted & provide insight regarding risk exposure.	FALSO	INVEST
Audits extend across Information Governance and 3rd Party Business Partners.	FALSO	IMPROVE

6 Information Security & Privacy	AS IS	TO BE
Comments - Information Security & Privacy seeks to understand how the enterprise deals with protecting data and information as a vital asset. While security has to do with the ability to protect information and data, privacy has to do with who is authorized to see what. The set of questions below help facilitate a understanding of the maturity of the organization to deal with both issues.		
High-Level Questions: 1.) Discuss how you see security and privacy as "important" in maintaining integrity (employees, clients, reporting, etc.) within the business? 2.) Does the organization take "security" seriously and if so, how does that manifest itself in specific requirements for handling data/information? 3.) Can you site some examples of how privacy is being implemented? 4.) How do you access the cost of unsecure data within the organization?		
Management is aware of the storage of personal information in the IT system and data related legislation.	PARTIAL	IMPROVE
Some policies/tools exist to enforce data security in most systems.	FALSO	IMPROVE
Employee access is controlled/restricted to specific "authorized" information.	VERDADERO	ACHIEVE
The Enterprise enforces to some degree, data security and privacy both internally and externally.	VERDADERO	ACHIEVE

Information Security is considered in transmissions of data and (personal/customer/supplier) over network(s).	PARTIAL	IMPROVE
Specific Policies, Practices or Procedures are in place to secure sensitive Personal Health Information (PHI)	PARTIAL	INVEST
The Organization is typically compliant with government regulations with regard to data security & privacy.	VERDADERO	ACHIEVE
Security processes consider the restriction of information access based on an employee/supplier/user profiles.	VERDADERO	ACHIEVE
The Organization can immediately deny system access to employees who are terminated or are not in compliance.	FALSO	IMPROVE
The C Level Officer is responsible for effectively enforcing security and privacy across the Enterprise.	PARTIAL	IMPROVE
IT, with regard to security, has an audit trail for access to all Data Bases (including DBA access).	PARTIAL	IMPROVE
IT within the Organization, does not use live/actual data in development or test systems.	PARTIAL	INVEST
Systems are in place to predict security and privacy breaches and proactively resolve them prior to occurrence.	FALSO	IMPROVE
Patterns in data access and system activities concerning security/privacy are analyzed and acted upon.	FALSO	IMPROVE
Data element level encrypted security and privacy is implemented consistently across the Enterprise.	PARTIAL	IMPROVE

7 Architecture & Standards	AS IS	TO BE
Comments - Information/Data Architecture and Standards provides the foundation for high quality data and data		
High-Level Questions:		
The architecture and standards are documented and communicated/approved across the enterprise	FALSO	IMPROVE
Standards for interfacing between two applications are decided by a CoE on application interfaces.	FALSO	IMPROVE
The interface between two applications have standards for data elements and often include data definitions.	PARTIAL	IMPROVE
Standards are consistent across department and LOB.	FALSO	IMPROVE
IT Enterprise Architecture considers data architecture & standards as a fundamental component of design.	PARTIAL	INVEST

Business rules are separated from applications and apply to multiple applications throughout the enterprise.	FALSO	IMPROVE
IT development efforts subordinate departmental requirements to specific Enterprise Level standards.	PARTIAL	INVEST
Management understands Master Data Management and information Life Cycle Management within IT architecture.	PARTIAL	IMPROVE
The Organization has data "standards compliance" reporting capability in place.	FALSO	IMPROVE
Specific data standards are helping reduce collaboration time between application developers.	n/a	INVEST
Data standards are included for end-to-end business processes and complement an existing Enterprise Architecture.	n/a	IMPROVE
Organizational roles are funded and staffed to develop and maintain data standards.	n/a	INVEST
The established standards are used across the enterprise by all major stakeholders.	n/a	IMPROVE
Data standards are linked to project business cases for continued improvement in Data Governance.	n/a	IMPROVE
Data architecture is driving performance improvements, and facilitating collaboration with all major stakeholders.	n/a	INVEST

8 Data Quality Management	AS IS	TO BE
Comments - Data Quality Management means having the tools, practices and procedures in order to establish and		
High-Level Questions:		
Management can quantify the "cost impact" of poor data quality and is motivated to do something about it.	PARTIAL	IMPROVE
The Organization views data integration and migration projects as an opportunity for data quality improvement.	PARTIAL	INVEST
The Organization's reference data is synchronized, eliminating redundancy and manual reconciliation.	FALSO	INVEST
Data quality is part of the Organization's development planning and has "a standard definition".	PARTIAL	INVEST
The Organization has chosen a Data Quality tool(s), in order to facilitate quality improvements on a consistent basis.	FALSO	IMPROVE
Data quality problems are analyzed and documented, reflecting the impact of data quality on the business.	PARTIAL	IMPROVE
Cost avoidance gained from data quality has been introduced as business case for any IT projects.	FALSO	IMPROVE

Data Quality includes analysis, impact, source/use, matching, deduping and "cleansing" and TQM drives improvement.	PARTIAL	INVEST
A formal methodology exists that includes the analysis, rule development, implementation and testing of Data Quality.	FALSO	INVEST
Most/all key projects have a Data Quality component with clearly defined objectives and business case.	FALSO	IMPROVE
A common platform and tool set for Data Quality improvement exists across the enterprise.	FALSO	IMPROVE
The Enterprise can now manage all unstructured data with advanced storage, search and retrieval capabilities.	FALSO	IMPROVE
Data Quality is a standard part of the CIO scorecard and quality metrics are integrated with Performance metrics.	FALSO	INVEST
Data Quality is a standard part of business planning inside the enterprise as well as with business partners.	FALSO	IMPROVE
A strong set of controls at partner onboarding and channel deployment are used to prevent data quality problems.	FALSO	IMPROVE

9 Classification & Meta Data	AS IS	TO BE
Comments - Classification & Meta Data means "the creation of the data, purpose of the data, time and date of creation,		
High-Level Questions:		
The Organization is aware of the importance of meta data relating to data governance (quality, life-cycle, security).	PARTIAL	IMPROVE
The Organization has standard definition "meta data" using common business glossary and consistent taxonomy.	FALSO	IMPROVE
Full time equivalents (FTE) are responsible for developing and managing meta data.	FALSO	IMPROVE
The Enterprise has requirements identified for meta data repository and plans to use it beyond a "data dictionary".	n/a	INVEST
There is a physical data dictionary and a messaging catalog available.	n/a	INVEST
The Organization has appropriate staffing (or has at least identified the need) for developing meta data repository.	PARTIAL	INVEST
A methodology driven approach to meta data is evident and is populated at the Dept. level and consolidated.	n/a	IMPROVE
The Organization has multiple meta data repositories and has identified the overlaps and gaps.	PARTIAL	IMPROVE
There are meta data repositories used for design and production activities and been approved by all stake holders.	n/a	IMPROVE

The meta data repository is methodology driven and approved by all stake holders.	FALSO	INVEST
The Enterprise has consolidated/integrated its meta data repositories across business/physical levels.	FALSO	INVEST
The Organization has mapped its meta data to business processes and users.	n/a	INVEST
Meta data directly supports performance improvements and business value.	n/a	INVEST
A meta data repository to drive requirements, analyze work/programs/projects, and provide collaboration exists.	n/a	INVEST
Meta data repository that drives design/production exists and has costs aligned by Business Unit.	n/a	INVEST

10 Information Lifecycle Management	AS IS	TO BE
Comments - Information Lifecycle Management		
High-Level Questions:		
Management is aware of the need for Lifecycle Data Management as storage demands and costs increase.	PARTIAL	IMPROVE
The higher value data is favored over other particular data elements. Data value is prioritized.	PARTIAL	IMPROVE
Specific data/information is periodically and systematically destroyed, retired or archived.	PARTIAL	IMPROVE
Senior Management looks at data/information as an asset with tangible value.	VERDADERO	ACHIEVE
Data/information has the appropriate level of security associated with its importance.	PARTIAL	IMPROVE
Departments understand a correlation between the value of data and how it is stored/retrieved.	PARTIAL	IMPROVE
Data "sources and uses" are defined and flowcharted (workflow, production & ad hoc)	FALSO	IMPROVE
Processes for Information Life Cycle are well defined, executed and measured.	FALSO	IMPROVE
Information/data within the enterprise is "owned" at the C Level for accountability and understanding.	PARTIAL	IMPROVE
Unstructured data/information is given the same priority as structured data/information.	FALSO	IMPROVE
A system exists that tracks meta data associated with use, changes and maturity levels (of data).	FALSO	IMPROVE
Processes, rules, systems and procedures are in place to archive data/information and store it securely.	PARTIAL	IMPROVE
The enterprise uses data/information for competitive advantage and places high value on analytics.	PARTIAL	IMPROVE

Single view, "Customer Insight" and "360 degree view" are successfully integrated, real-time and valued.	FALSO	IMPROVE
Data/Information Lifecycle Management has specific KPIs that include sourcing, usage, retirement, search & retrieval.	FALSO	IMPROVE

11 Audit Information, Logging & Reporting	AS IS	TO BE
Comments - Audit Information, Logging & Reporting represent methods, processes and tools by which enterprises enable		
The Organization has a framework to audit data changes for critical data.	PARTIAL	INVEST
A Data Management strategy is in place that considers audit, logging and reporting.	PARTIAL	IMPROVE
Data Governance processes exist for certification of financial reports and regulatory compliance.	PARTIAL	INVEST
Data audit and data change management are consistent across the lines of business.	n/a	IMPROVE
Localized Data Governance addresses out-of-process changes in each Line-Of-Business.	n/a	IMPROVE
The sources of data are known at the "Enterprise" level and are considered important in audit and reporting.	FALSO	INVEST
Corporate reports rolled up to the C Level are consistent across LOB's.	PARTIAL	INVEST
Certification, with confidence, is evident in financial reporting (consistency may still be a problem)	PARTIAL	INVEST
A framework for auditing all data changes to critical data across all lines of business has been developed.	FALSO	IMPROVE
A framework for auditing all data changes to critical data has been defined & communicated across the enterprise.	n/a	INVEST
A comprehensive strategy for reducing the number of out-of-process changes across all lines of business exists.	n/a	IMPROVE
The Organization utilizes a centralized method for managing and tracing data sources and flows across the Enterprise.	FALSO	INVEST
The auditing framework is fully operational, and management has confidence in the ability to audit all data changes.	n/a	IMPROVE
Reports, KPI's, & metrics are used to improve Data Management processes across the enterprise.	PARTIAL	INVEST
The ability to trace all data flows is evident across the enterprise. All data can be easily traced from any source.	FALSO	INVEST



12 Business Intelligence & Advanced Analytics	AS IS	TO BE
Comments - Advanced Analytics are a direct result of outstanding "data governance". With analytics, an organization may		
The Organization provides "some" consolidated views of performance data across the enterprise.	PARTIAL	INVEST
There is a range of BI capabilities (reporting, analysis, dashboards, scorecards & alerts) with basic analytics.	PARTIAL	INVEST
IT delivers performance information to the business and somewhat customized to their needs.	PARTIAL	INVEST
The Organization can analyze planning requirements and "what if" scenarios to anticipate the correct course of action.	PARTIAL	IMPROVE
Finance can deliver planning, budget and forecasting to each business unit.	PARTIAL	IMPROVE
Unstructured text can be analyzed and trends assessed for actionable activities.	FALSO	MAINTAIN
BI and Analytics are used with specific business rules to provide "insight" into opportunities and drive outcomes.	PARTIAL	INVEST
BI and Analytics are cost-effectively embedded in operational systems within the enterprise (ERP, CRM, SCM etc.).	PARTIAL	INVEST
Analytics are used to identify relationships between entities in order to detect fraud, collusion or identity theft.	FALSO	IMPROVE
Management sees the value of predictive analytics and is willing to invest in strategic capabilities.	PARTIAL	INVEST
Advanced streaming analytics are used to achieve higher customer related investments for upsell/cross sell.	FALSO	INVEST
The need for scalable, parallel, real-time data analysis is required for certain situations.	PARTIAL	INVEST
Business imperatives are related to Performance Management and advanced data analytics.	FALSO	IMPROVE
Advanced data flow/event driven analytics are used to accelerate and optimize the business.	FALSO	INVEST
Predictive Analytics (e.g., regression, logistic regression, neural networks) effectively predicts outcomes.	FALSO	INVEST

**Anexo 2: Enterprise Information Management Maturity Self-Assessment – AS IS**

VISION Indicators	Not at all	Somewhat	Mostly	Definitely
Information assets are linked and leveraged across several programs	x			
Information is a central component of business strategy and architecture	x			
Business management encourages cross-functional information accessibility to improve responsiveness to the business, customers and marketplace	x			
Exogenous data sources have begun to be integrated for enhanced analytics		x		
Data fiefdoms (pockets of ownership) have begun to disband	x			
Information is a source of power, but managed in silos			x	
There is general acknowledgment that information management (or lack thereof) is a serious problem		x		
Customers and partners directly influence information vision	x			
Information is recognized as a corporate asset, competitive differentiator, source of transformation and even as a product itself	x			
Information is defined primarily by the value it brings, not by its structure or other characteristics	x			
Progress is hampered by culture, contradictory incentives, organizational barriers and lack of leadership		x		
People spend time arguing about whose data is correct and who owns it instead of seeking uniform availability		x		
IT attempts to formalize objectives for information availability to achieve targeted operational needs	x			
Necessary, valued and prioritized information is leveraged across all programs and investments	x			
Information is viewed as an indispensable fuel for enterprise performance and innovation to be shared seamlessly	x			
Senior business executives champion and communicate information-related best practices	x			

STRATEGY Indicators	Not at all	Somewhat	Mostly	Definitely
Information is hoarded by departments and individuals as a source of power and influence, or is unknown altogether				x
An information management organization may be in formative stages, but sponsorship is nonexistent		x		
Information strategy is no longer a separate work task but is embodied in how the business operates	x			
Information is seen merely as application-specific			x	
Most EIM capabilities and resources are in place and functioning				
Business units are committed and involved		x		
The information strategy considers the organization's extended ecosystem of partners, suppliers and customers	x			
Different information/content types are treated and managed with a common or consistent approach	x			
Information management resources and technologies start to become pooled and shared across projects	x			
Business informational needs and risks are met proactively			x	
Strategy definition is shifting from a static, annual process toward more of a dynamic "living document"		x		
The office of the CDO is empowered to drive EIM vision in support of the business needs				
A high-level sponsor (e.g., CDO) is named to define an enterprisewide information strategy and coordinate a broad agenda, including funding and roadmap				
An EIM organization emerges to establish and control standards, and improve information availability while reducing expenses, but the main focus is on technology				
Data and analytics leadership has a say in corporate strategy as information is deemed an actual corporate asset	x			
A well-funded and well-led information program addresses most enterprise needs (current and planned)	x			
Business units recognize the broader value of information and begrudgingly share it on cross-functional projects			x	

<b>METRICS Indicators</b>	<b>Not at all</b>	<b>Somewhat</b>	<b>Mostly</b>	<b>Definitely</b>
A portfolio approach to EIM investments and risks is adopted where in business cases are aligned, interrelated and regularly maintained	x			
Information-related metrics correlate to business value metrics	x			
There is a proliferation of overlapping and sometimes inconsistent nonfinancial metrics, such as those for operational efficiency or customer satisfaction	x			
Some information metrics are linked to business initiatives, and business cases are linked	x			
Any goals and measures for information management and delivery are purely subjective and rarely, if at all, tracked		x		
Qualitative measures emerge for information management that may not necessarily link well to business KPIs		x		
Feedback loops reporting information value and investment yield emerge	x			
Information management financial justifications for each investment are typically related to expense savings, or embedded in other projects	x			
Simple (often predisposed) cost/benefit models justify independent information management investments, or are merely part of IT-business projects themselves	x			
Information management is not a budget item, and priorities are based on influence peddling and failure prevention	x			
Information valuations and yield (e.g., ROI) models drive investments in information, technology and business innovation	x			
Priorities are based on user surveys, minimizing each information management program expense, and infrastructure performance/scale	x			
Some data profiling is done for tactical purposes, usually only for specific data quality needs	x			
As EIM becomes a budget item of its own, various measures of information value (e.g., quality, top/bottom line) and risk are developed, tracked and communicated	x			

GOVERNANCE Indicator	Not at all	Somewhat	Mostly	Definitely
Information governance priorities are based on business need, not IT demands			x	
Policies evolve into a full set of precepts (e.g., principles, guidelines, policies, standards, procedures), which are well-communicated and enforced		x		
All information assets, including external sources, are inventoried			x	
Information security and risk are now linked to the same information governance process		x		
Information owners are assumed but not formalized, and upstream data quality is performed as needed			x	
Business process improvement is now part of information governance		x		
Policy adherence procedures for key information assets are developed and monitored by information owners and stewards who have limited authority		x		
Information ownership mainly takes the form of territoriality to buttress against change			x	
An enterprise information governance organization is functioning (e.g., within the context of the office of the CDO) and carries sway on all IT and business projects	x			
Formal data quality/integration, metadata and master data management (MDM) programs emerge, but tend to focus on small subsets of business data or business outcomes to be improved		x		
Few official policies exist for the handling or use of most information, other than those required by law or industry regulations			x	
Efforts to align governance and stewardship of different kinds of data, spanning content and structured data, start to emerge		x		
Information stewards have become information advocates, focused more on fostering information value generation now that data quality has become part of the culture		x		
Ad hoc data quality efforts and a lack of data definitions result in low data trust and usage		x		
Enterprise information governance is encoded into an automated information asset management system	x			
Data quality is largely automated		x		
Policies, mostly for information silos, have emerged for information management and use, but are not monitored and are regularly circumvented			x	

<b>ORGANIZATION and ROLES Indicators</b>	<b>Not at all</b>	<b>Somewhat</b>	<b>Mostly</b>	<b>Definitely</b>
Business people typically are resigned to source and manage their own data, or must join the IT backlog			x	
Specialty roles — such as big data infrastructure/architecture specialists, data scientists, information strategists, information architects, information product managers and data curators — become prevalent				
EIM and analytics move outside of IT as a CDO is installed to lead a separate enterprise information service organization				
Business users are engaged in information-related activities mostly to resolve issues, rather than in upfront design and planning		x		
Formal information and content management organizations materialize within IT and governance councils as well as stewardship bodies in the business	x			
Data-related meetings have become business-centric				
Information-related responsibilities are resourced on an application-by-application and project-by-project basis			x	
The CDO oversees and has authority and budget for most aspects of the information life cycle				
Information-related competency centers have emerged under the CDO for core analytics, data modeling, metadata and master data				
IT also houses business intelligence analysts and data integration specialists			x	
Projects are set up and staffed one at a time, but tend to lack a plan for organizational continuity or intraenterprise synergy		x		
Pooled or centralized database administrators, data administrators and data modeling resources have emerged, all of whom are strictly part of the IT department	x			
An information product management function develops and facilitates new revenue streams				
This information services organization supports the entire LOB-customer-partner-supplier information ecosystem				
At least one data scientist is hired by a business unit	x			

LIFE CYCLE Indicators	Not at all	Somewhat	Mostly	Definitely
Information architecture is not yet formalized or embedded in the EIM program	x			
Valiant attempts at enterprise metadata management (EMM) are less than successful	x			
There are still no enterprise policies or procedures for information disposal or archival	x			
Semantically consistent and important information assets are shared across all needed programs and investments	x			
Metadata management is mostly manual (e.g., spreadsheets) and remains focused on individual data assets				
Enterprise procedures are in place for the defensible disposal of information assets based on value, risk and compliance modeling				
Information life cycles and metadata are documented, implemented and aligned — and all differences are culled as a matter of course				
Information life cycles are seen as business processes, not as IT workflows or tasks				
Information architects are embedded (and key players) in the EIM program				
Only some information governance policies are encoded as procedures	x			
Information value and governance are measured, monitored and optimized across critical information life cycle paths				
Information is governed outside the control of any given business application	x			
Data is deleted early due to lack of infrastructure, resulting in potential regulatory violations		x		
The focus, if there is one, is on local efficiencies and use of individual datasets			x	
Usually data integration involves making data extracts, and is outside the purview of information governance			x	
End-of-life procedures for information assets are established only for those subject to industry regulations			x	
Technical efficiency is deemed more important than the business efficacy from shared data		x		
Data is kept and maintained in silos, and IT does its best to integrate data as requested			x	
Information flows are well-documented but not maintained	x			
There is no understanding of information having its own life cycle			x	
Data is retained well beyond its usefulness and may introduce risks	x			
Metadata standards, tools and procedures have emerged				
Data integration is effective in linking disparate data, but efforts to semantically align and form shared procedures across silos are sporadic	x			
EMM and MDM are ongoing initiatives that help coordinate and enable business initiatives				
New information uses build on previously deployed and documented models spanning the information life cycle	x			
Information architects are often involved in EIM, but are less than central to or sufficiently influential on such efforts				