Universidad de Santiago de Compostela

Inteligencia de negocio

Luis Ardévol Mesa

Profesor: Joaquín Ángel Triñanes Fernández

Escola Técnica Superior de Enxeñaría Master en Tecnoloxías de Análise de Datos Masivos: Big Data

Curso 2024-2025

Contents

Contents							
1	Introducción a BI						
	1.1	Sistemas de información	1				
		1.1.1 Sistemas de información empresarial	2				
		1.1.2 Evolución de los sistemas de información	3				
		1.1.3 Tipos de sistemas de información	3				
	1.2	Inteligencia de negocio					
	1.3	3 Modelos de madurez					
		1.3.1 The Data Warehousing Institute (TDWI) Analytics MM	6				
		1.3.2 HP BI MM	7				
		1.3.3 Gartner BI MM	8				
	1.4	Metodología	9				
		1.4.1 Larissa Moss	9				
		1.4.2 CRISP-DM	11				

1 Introducción a BI

1.1 Sistemas de información

Un **sistema** es un conjunto de elementos entrelazados que interactúan entre sí en un entorno cambiante, con un cierto objetivo o realizando una función concreta. Se tienen los siguientes elementos básicos:

- Los componentes del sistema.
- Las relaciones entre los componentes, que determinan la estructura del sistema.
- El objetivo del sistema.
- El ambiente del sistema, que incluye todo lo que le rodea.
- Las fronteras del sistema, que lo separan del ambiente.
- Feedback: en muchos sistemas, la salida influencia la entrada del mismo.
- Tasa de rendimiento (*throughput*): procesos de transformación. Se trata de una medida de la cantidad de información que el sistema es capaz de procesar en un tiempo determinado.

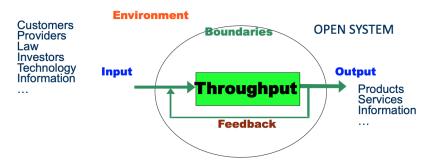


FIGURE I: Elementos de un sistema.

Un sistema de información consiste en un conjunto formal de procedimientos que, operando sobre un conjunto de datos estructurado según las necesidades de la organización, recopila, procesa, almacena y distribuye información necesaria para las operaciones diarias de la organización. Además, proporciona información a los niveles jerárquicos superiores para la toma de decisiones de acuerdo con la estrategia de negocio.

Es importante destacar que el objetivo de un IS es ayudar a la realización de actividades en todos los niveles de la organización, proveyendo la información correcta, con una **calidad suficiente** a la persona adecuada en el momento adecuado, en el formato más adecuado para el receptor.

Desde una perspectiva de negocio, un IS es una solución de organización y gestión basada en tecnología de la información, cuyo objetivo es tratar retos emergentes en el contexto del negocio. En cualquier nivel de la organización hay cambios en los objetivos, procedimientos, relaciones con el

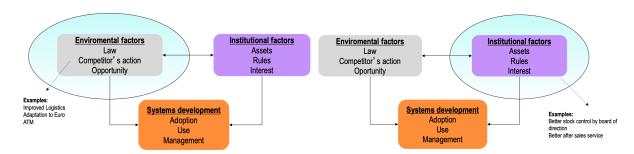


FIGURE II: Sistema de información en una organización.

entorno, etc. para conseguir una ventaja competitiva. El valor de un IS dependerá de su efectividad, alcance, aceptación, coste, calidad de la información que produce, etc.

Ejemplo 1.1.1. IS para la cría de mejillones. En este caso, algunos factores ambientales serían las regulaciones regionales y nacionales, el cambio climático, los avances en tecnología, las certificaciones, el mercado, etc. Factores institucionales sería la tecnología e infraestructuras, los recursos económicos y humanos, los intereses de maximizar beneficios y expandir el mercado, etc.

1.1.1 Sistemas de información empresarial

Para construir un sistema de información en una empresa, se debe tener en cuenta la estrategia de negocio, lo que resulta un factor clave. A continuación habría que considerar el sistema de información a partir del cual posteriormente se obtiene la información y las tecnologías de comunicación, como sensores, algoritmos, modelos, etc. Todo esto se retroalimenta, ya que estas tecnologías muestran si la estrategia planteada es o no viable.

Los EIS están diseñados para coordinar el flujo de información y los registros necesarios para que una empresa funcione de manera eficiente. El sistema debe estar alineado con la estrategia empresarial de la compañía, lo que significa que la gestión de la información debe apoyat los objetivos y la dirección general del negocio. Estos sistemas ayudan a gestionar distintos aspectos de las operaciones empresariales:

- Planificación de recursos empresariales (ERP): integra los procesos empresariales principales en un solo sistema para garantizar una gestión eficiente de los recursos y el flujo de información.
- Sistema de gestión de flujos de trabajo: automatizan y gestionan los procesos empresariales, asegurando que las tareas se completen en el orden correcto por las personas adecuadas.
- Sistemas de *Groupware*: herramientas que permiten a los empleados colaborar, compartir información y trabajar juntos en proyectos.
- Sistemas de comercio electrónico. Hay tres modelos principales: business-to-business (B2B), business-to-consumer (B2C) y consumer-to-consumer (C2C).
- Intercambio electrónico de datos (EDI): permite a las empresas intercambiar documentos comerciales en formato electrónico estandarizado.

La inteligencia de negocio ya existía antes del *Big Data*. Normalmente se liga a la minería de datos, pero con la IA y los modelos predictivos, se ha ampliado su alcance.

Decade	Required Information	Type of ISs	Objectives	
1950s	Basic Computational Data	Batch Processing Systems	Automate complex calculations and data processing tasks	
1960s	Transactional Data	TPS (Transaction Processing Systems)	Automate routine business operations and record-keeping	
1970s	Operational and Tactical Data	MIS (Management Information Systems)	Improve administrative control and reporting	
1980s	Analytical Data	DSS (Decision Support Systems)	Support decision-making with data modeling and analysis tools	
1990s	Strategic and Competitive Data	EIS (Executive Information Systems)	Enable real-time monitoring and strategic decision-making	
2000s	Enterprise-Wide Data	ERP (Enterprise Resource Planning Systems)	Integrate business functions for efficiency and process automation	
2010s	Big Data, Predictive Analytics	BI (Business Intelligence Systems)	Data-driven decision-making and predictive analytics	
2020s	Real-Time and AI-Enhanced Data	AI Systems and Cloud- Based Analytics	Automate decisions with AI, scalable cloud infrastructure	

TABLE 1.1: Evolution of Information Systems by Decade

1.1.2 Evolución de los sistemas de información

1.1.3 Tipos de sistemas de información

- Sistemas de nivel operacional. Monitorizan las actividades, operaciones y transacciones básicas de la empresa.
- Sistemas de gestión de informacion (MIS) y sistemas de soporte de decisión (DSS). Ayudan la monitorización, control y toma de decisiones y actividades administrativas a nivel gerencial.
- Sistemas de nivel estratégico. Apoyan la planificación a largo plazo a nivel gerencial para lograr una ventaja competitiva.
- Sistemas de nivel de conocimiento. Apoyan a los trabajadores de conocimiento e información dentro de la institución.

Ejemplo 1.1.2. *USC:*

- Operacional: inscripciones de estudiantes.
- MIS: informes mensiales sobre el desempeño estudiantil.
- DSS: análisis de tendencias en la inscripción de estudiantes.
- Estratégico: decisiones a largo plazo sobre nuevos grados y líneas de investigación.
- Conocimiento: repositorio digital con tesis, artículos e investigaciones.

Sistemas de información operacional

Las características principales de estos sistemas son:

• Ejecutan y registran operaciones rutinarias diarias necesarias para el funcionamiento de la empresa.

- Diseñados para aumentar la productividad.
- La inversión en estos sistemas es fácil de justificar ante la dirección, ya que sus beneficios son viables y palpables.
- Comúnmente, son el primer tipo de IS que se implementa en una empresa. Su uso inicial apoya los esfuerzos a nivel operacional.
- Son intensivos en entrada y salida de datos. Las operaciones que realizan generalmente implican cálculos y procesos simples y poco sofisticados.
- Proporcionan a los administradores informes y acceso en línea a registros históricos y diarios.
- Son los principales generadores de información para otros tipos de ISs dentro de la organización.

Como ejemplos de OIS se pueden comentar sistemas de facturación, sistemas de control de inventario o sistemas que calculan y procesan los pagos de los empleados.

Sistemas de soporte de decisión

Las características principales de estos sistemas son:

- Se incluyen tras haber implementado IS más relevantes, ya que reciben información de estos sistemas.
- Los cálculos suelen ser intensivos, mientras que las salidas son escasas.
- Combinan información cambiante con modelos analíticos sofisticados para apoyar la toma de decisiones no estructurada y semiestructurada.
- Tienden a ser interactivos, visuales y amigables, y están enfocados en el usuario final.
- No tienen la intención de ahorrar trabajo. Como resultado, la justificación económica para la inversión en estos sistemas es dificil, ya que no se conocen los beneficios directos del proyecto.
- Involucran modelos analíticos, análisis de "qué pasaría si", simulaciones y pronósticos.

Sistemas estratégicos

Las características principales de estos sistemas son:

- Incorporan información externa y obtienen información resumida de los sistemas operacionales y de soporte de decisión.
- Apoyan la introducción de productos y procesos dentro de la organización, ya que buscan obtener una ventaja sobre los competidores mediante innovación.
- Si función es lograr ventajas que los competidores no tienen, como ventaja de costos y servicios diferenciados para clientes y proveedores. En este contexto, los sistemas estatégicos crean barreras de entrada para el negocio. Por ejemplo los cajeros automáticos (ATM), la banca por internet y los sistemas de recomendación.
- Suelen desarrollarse de forma ad hoc dentro de la organización.
- Están orientados a alcanzar metas estratégicas a largo plazo (e.g. expandir la cuota de mercado, crear nuevos mercados).

Como ejemplos de SIS se pueden comentar pronósticos de ventas (con planificación de campañas de marketing), planificación de recursos de manufactura (MRP) enfocado en aumentar la productividad en un proceso de manufactura, descubrimiento y lanzamiento de productos en banca, como tipos de hipoteca, con el propósito de alcanzar objetivos comerciales tales como atraer nuevos clientes o fidelizar a los actuales, etc.

Sistema	Entrada	Procesos	Salida	Usuarios
Operacional	Transacciones	Almacenar, informar, unir	Informes, resúmenes, transacciones, facturas, nóminas.	Personal operativo, supervisor
DSS, táctico	Información operacional resumida, gran volumen de datos, modelos simples	Modelo, simulaciones, análisis, optimización, análisis de "qué pasaría si"	Informes analíticos, modelos de decisión	Personal técnico, analístas de datos
Sistemas estratégicos	Información agregada, información externa	Análisis estratégico, planificación de escenario, CI	Planes estratégicos, predicciones	Directiva

TABLE 1.2: Características de distintos IS.

1.2 Inteligencia de negocio

El término "inteligencia de negocio" (BI) se refiere a la gestión de la información en una empresa específica o área de negocio. Conciste en un conjunto de estrategias y herramientas focalizadas en la gestión del conocimiento mediante el análisis de los datos de la organización. BI se centra en:

- Marcar objetivos de negocio.
- Determinar la necesidad de datos, información y conocimiento para complir los objetivos.
- Hay que integrar todos esos daots con toda esa información, deben ser accesibles y fundamentar decisiones.

Como ya se vio antes, los OIS recogen y organizan los datos, los IS tácticos la analizan, resumen, transforman y visualizan, mientras que los SIS adquieren, descubren, evalúan y usan ese conocimiento generado a partir de los datos.

1.3 Modelos de madurez

Los modelos de madurez (MM) permiten evaluar el nivel de desarrollo y eficacia en el uso de soluciones, así como formas de mejora. Definen niveles de definición, eficiencia, manejabilidad y medición del entorno monitoreado. Generalmente todos siguen una estructura similar, y tienen las siguientes características comunes:

- Cada modelo tiene al menos 5 niveles de madurez; menos implica pasar de fases de forma confusa.
- Cada modelo empieza con una fase primitiva

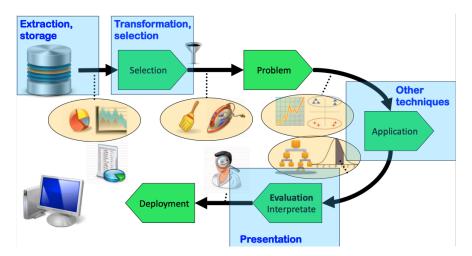


FIGURE III: Proceso de inteligencia de negocio. Para su aplicación se puede usar minería de datos u otras técnicas.

- Los modelos no se centran solo en la tecnología, se basan también en el involucramiento de las personas y los procesos.
- Proporcionan capacidades actuales, hojas de ruta para la mejora, comparación de referencias, alineación con los objetivos empresariales, mejoran la cultura basada en datos, optimizar la asignación de recursos y apoyar la planificación estratégica.

1.3.1 The Data Warehousing Institute (TDWI) Analytics MM

Este modelo consta de las siguientes dimensiones:

Madurez Organizacional

- Liderazgo
- Cultura
- Impacto
- Estrategia

Madurez de Recursos

- Financiamiento
- Talento / habilidades
- Roles / responsabilidades
- Capacitación

Madurez de la Infraestructura de Datos

- Diversidad, volumen y velocidad
- · Acceso a datos
- Integración y gestión de datos
- Arquitectura de datos

Madurez Analítica

- Alcance de capacidades
- Automatización / aumentar
- Enfoques de despliegue y entrega
- Innovación

Madurez de Gobernanza

- Procesos y herramientas de gobernanza de datos
- Procesos y herramientas de gobernanza del modelo
- Roles de gobernanza
- · Seguridad/privacidad

Entre la tercera y cuarta etapa, hay un abismo que muchas empresas no llegan o no quieren cruzar. Las fases del modelo son las siguientes:

Etapa 1: Naciente - Entorno pre-analítico. La cultura no está basada en datos y no se toman decisiones basadas en datos. No hay una infraestructura formal de BI. Los datos están dispersos y en silos. Calidad pobre de datos. IT y negocio no trabajan juntos. Informes en Excel o manuales.

Etapa 2: Temprana - Herramientas básicas de BI. Comenzando a entender el valor de la analítica. La organización se da cuenta de que se necesita cierta infraestructura de datos para apoyar la integración de datos para el análisis. Analítica rudimentaria pero en avance. Calidad de los datos inconsistente ya que nadie se ocupa de ellos.

Etapa 3: Establecida - Se ha implementado un data warehouse. Hay un grupo responsable de la analítica en IT. Se ha formado un equipo de gobernanza de datos. Se utilizan herramientas de visualización de datos. La solución no escala bien y no hay aún herramientas de fácil uso.

Etapa 4: Madura - Los usuarios finales integran la analítica en la toma de decisiones y operaciones. Cultura de innovación y colaboración entre IT y negocio. Uso de diversas fuentes, incluidos datos semi y no estructurados. La fuerza laboral incluye científicos de datos, ingenieros de datos, MLOps y roles como el Chief Analytics Officer.

Etapa 5: Avanzada / Visionaria - Analítica de autoservicio. La organización considera la analítica como un arma competitiva crítica. Infraestructura operativa, flexible y escalable (data lakes, DWH, nube), ML, AI se utiliza en NRT para la toma de decisiones.

Esta organización proporciona un test para evaluar el nivel de madurez de una empresa u organización en BI, dando una visión holística: es de esperar que que las fases varíen para cada dimensión.

SCORE PER DIMENSION	STAGE	DIMENSION	SCORE	STAGE
≤ 5	Nascent	Organization	10	Early
6-10	Early	Resources	7	Early
11-15	Established	Data Infrastructure	11	Established
16-19	Mature	Analytics	4	Nascent
20	Visionary	Governance	7	Early

FIGURE IV: Test de madurez de TDWI: resultados.

1.3.2 HP BI MM

Etapa 1: Operación (Operar el negocio) - Implica soluciones ad-hoc enfocadas en la demanda local. No existe una estrategia formal de BI ni analítica.

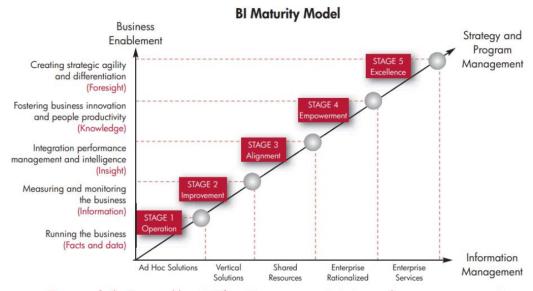
Etapa 2: Mejora (Medir y monitorear el negocio) - BI más estructurado. Se está utilizando la analítica. Los almacenes de datos (DW) aún están enfocados en unidades de negocio específicas.

Etapa 3: Alineación - BI alineado con los objetivos estratégicos. Integración de la información entre áreas temáticas. La calidad de los datos y la gobernanza están aumentando su importancia. La organización ha evolucionado de la gestión de proyectos de BI a la gestión de programas de BI.

Etapa 4: Empoderamiento - Se utilizan herramientas de autoservicio a todos los niveles. Se implementan análisis avanzados. Existe una única versión de la verdad en toda la organización. BI es fundamental.

Etapa 5: Excelencia (Cambiar el negocio) - Análisis predictivo para la mayoría de las decisiones empresariales. Arquitectura orientada a servicios (SOA) para la entrega de información. Los análisis se entregan rápidamente a los usuarios.

En este modelo de 5 etapas de madurez, el éxito se contempla como una función de cada una de las dimensiones: habilitación/preparación del negocio, gestión de la información, estrategia y programa de gestión.



Success = fn (business enablement, information management, strategy and program management)

FIGURE V: Modelo de madurez de BI de HP.

1.3.3 Gartner BI MM

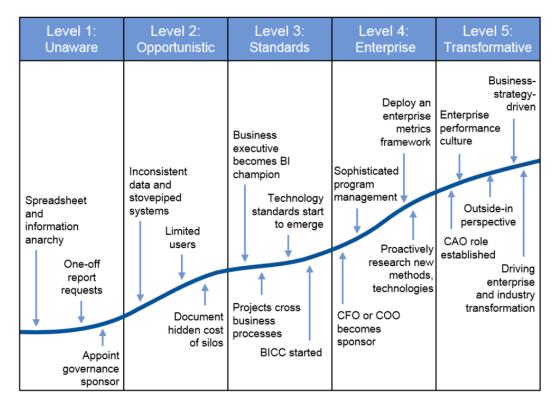


FIGURE VI: Modelo de madurez de BI de Gartner (BICC - BI competency center).

1.4. Metodología 9

Este se trata de un modelo popular, que tiene muy en cuenta el entorno de la organización a la hora de tomar decisiones (*external drivers*). En la primera etapa no se usa BI. En la segunda se empiezan a compartir datos. En la tercera no hay interacción entre las unidades de negocio. En la cuarta, existen políticas y estándares bien definidos y en última etapa, la información es confiable y se tiene una solución adaptable a las entidades de la empresa.

Nivel 1: Desconocido - BI y analítica ad-hoc. No hay un proceso formal de toma de decisiones. No hay infraestructura de TI.

Nivel 2: Oportunista - Las unidades de negocio usan BI de forma individual. Cada proyecto/unidad tiene su propia infraestructura.

Nivel 3: Estandarizado - La coordinación mejora. Proyectos que abarcan múltiples procesos de negocio. Se crea un centro de competencia de BI. Empiezan a surgir estándares tecnológicos para la infraestructura, DW y plataformas de BI.

Nivel 4: Empresarial - Los altos ejecutivos patrocinan BI. Se han definido métricas de rendimiento que guían la estrategia. BI apoya los procesos de decisión en toda la empresa.

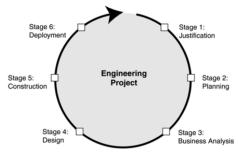
Nivel 5: Transformador - BI y la analítica son una iniciativa estratégica gestionada y apoyada por el negocio y TI. Se utilizan para generar ingresos. Usuarios en todos los niveles, incluidos clientes y socios.

1.4 Metodología

El primer paso en la implementación de BI es desarrollar una declaración clara del problema. Todo proceso debe ser repetible (esto hará más fácil su asimilación). Los procesos, creados por expertos, ayudan a la planificación y gestión de procesos. Además, reduce el miedo inicial a afrontar el problema, ya que existe un proceso estandarizado y reduce la dependencia en personal específico. Como ejemplos de metodologías se pueden mencionar:

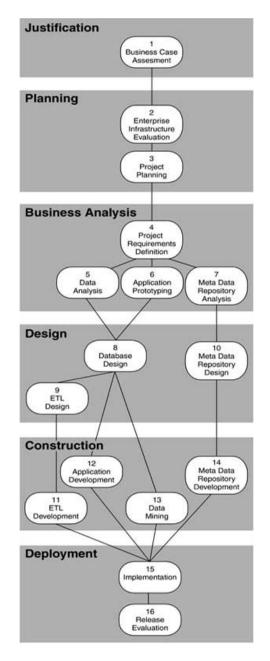
- Gestión: PMBok, Agile, ...
- Gestión y desarrollo: Larissa, Kimball, Inmon, SAFE, QPM (QlikView), ASAP (SAP), ...
- CRISP-DM: tareas genéricas para minería de datos (DM).
 - Procesos estandarizados intersectoriales para DM.
 - Propuestos por un consorcio (SPSS, NCR, AG, OHRA).
 - Otras metodologías centradas en DM podrían ser: 5A'S, critikal, CAT (Clementine Application Template), SEMMA (SAS, Sample, Explore, Modify, Model, Assess), ...

1.4.1 Larissa Moss



Este modelo define una hoja de ruta con 6 etapas iterativas y 16 pasos.

Define etapas, pasos, cargos, estándares y entregable. Es ágil y flexible, además de promocionar subproyectos, ya que fragmenta los proyectos grandes. Veamos un desglose de las etapas y los pasos a seguir en cada una de ellas:



- i. Justificación: evaluar la necesidad empresarial que da lugar al nuevo proyecto de ingeniería.
 - Paso 1. Evaluación del caso de negocio.
 - Define el problema u oportunidad empresarial y propone una solución de BI.
 - Cada lanzamiento de aplicación de BI debe estar justificado en costos y debe definir claramente los beneficios.
- Planificación: desarrollar planes estratégicos y tácticos, que establezcan cómo se llevará a cabo y desplegará el proyecto de ingeniería.
 - Paso 2. Evaluación de la infraestructura empresarial.
 - Técnico: hardware, software, redes, ...
 - No técnico: procedimientos, metodologías,
 - Paso 3. Planificación del proyecto: alcance, personal, presupuesto, tecnología, representantes empresariales, ...
- iii. Análisis de negocio: realizar un análisis detallado del problema u oportunidad empresarial para obtener una comprensión sólida de los requisitos del negocio para una solución potencial (producto).
 - Paso 4. Definición de requisitos del proyecto.
 - Gestión y especificación del alcance, necesidades del usuario, ...
 - Se entrega el *Service Level Agreement* (SLA).
 - Paso 5. Análisis de datos, disponibilidad, calidad,
- Paso 6. Prototipado de la aplicación. Ayuda en la definición de requisitos y evita riesgos.
- Paso 7. Análisis del repositorio de metadatos.
 - Los metadatos técnicos deben ser mapeados a los metadatos del negocio.
 - Los metadatos describen una organización en términos de sus actividades comerciales, los objetos del negocio y las reglas bajo las cuales se realizan las actividades comerciales (e.g. definiciones, unidades, relaciones, fuentes, ...)
- iv. Diseño: concebir un producto que resuelva el problema empresarial o permita la oportunidad empresarial.
- v. Construcción: construir el producto, el cual debe proporcionar un retorno de inversión dentro de un periodo de tiempo predefinido.

1.4. Metodología

vi. Despliegue: implementar o vender el producto terminado, y luego medir su efectividad para determinar si la solución cumple, excede o no cumple con el retorno de inversión esperado.

1.4.2 CRISP-DM

Esta metodología es libre e independiente de aplicación, contexto y herramientas. Se trata de una guía que considera el problema y las técnicas, y está basado en la experiencia. Hay 2 documentos principales: el modelo de referencia, que describe las fases, tareas y salidas, y la guía de usuario, con consejos de aplicación práctica, listas, etc. CRISP-DM es el proceso más usado para proyectos de ciencia de datos.

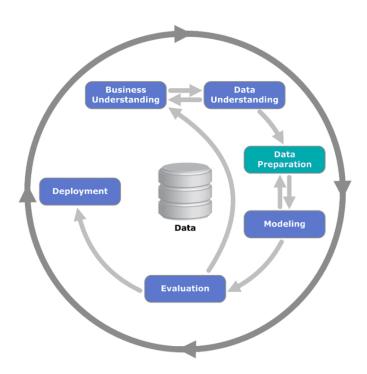


FIGURE VII: Fases de CRISP-DM.

Hay 4 niveles de jerarquía: las fases se dividen en tareas generales, estas se fragmentan en tareas específicas que a su vez de dividen en instancias de proceso. La metodología se compone de 6 fases bien diferenciadas; las 3 primeras representan de forma general el 80% de un proyecto de ciencia de datos (30% entre las fases 1 y 2 y 50% para la fase 3):



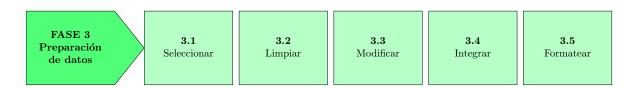
Un gran fallo es considerar que comprender el negocio es un objetivo secundario; se trata de uno de los aspectos más importantes. Se tienen las siguientes tareas:

- 1.1 Definir los objetivos del negocio.
 - Conocer lo que quiere el cliente desde la perspectiva del negocio: ¿mejorar la gestion de inventario?, ¿reducir el tiempo de respuesta?, ¿mejorar la calidad de los productos?, ...

- Mostrar los factores que afecten al proyecto.
- Establecer el criterio de éxito.
- 1.2 Evaluación de viabilidad. Es común usar análisis DAFO (debilidades, amenazas, fortalezas, oportunidades) para evaluar la viabilidad del proyecto.
 - Indicar recursos (personal, datos, *software*, ...), restricciones, suposiciones, requerimientos y otros factores como el hecho de si tenemos permitido usar los datos.
 - Evaluación del costo-beneficio.
 - Riesgos: no disponibilidad de los datos, conocimiento no disponible, herramientas
- 1.3 Definir los objetivos de minería.
 - Objetivos específicos del problema: traducir las necesidades del cliente en objetivos de minería mediante segmentación, patrones secuanciales, ...
 - Establecer objetivos técnicos: traducir las metas en parámetros de salida, como el ratio de éxito, predicciones de fracaso, coste de error, etc.
- 1.4 Crear un plan de proyecto.
 - Establecer los pasos: la duración, los recursos necesarios, las entradas y las salidas. Organizar y gestionar riesgos.
 - Elección inicial de herramientas y técnicas.



- 2.1 Recolección: definir las fuentes, integración (suele ser la parte más complicada), localización, problemas y soluciones, etc.
- 2.2 Descripción: formato, claves, cantidad, etc.
- 2.3 Exploración: consultas, visualización, informes, distribución de valores, agregaciones de datos, estadísticas, etc.
- 2.4 Verificación de calidad: ¿todos los tipos de datos?, ¿todas las clases representadas?, ¿suficientes datos?, ¿datos completos?, ¿valores nulos o anormales?, etc.

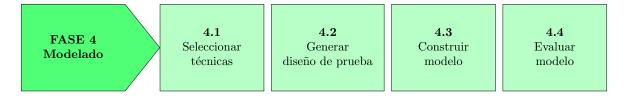


Esta fase supone un 50% del tiempo total del proyecto.

- 3.1 Selección de datos.
 - Seleccionar tablas, atributos y filas.

1.4. Metodología

- Volumen adecuado para las herramientas.
- Selección: dividir el conjunto de datos en training, test y validation.
- Muestra de datos (samples).
- 3.2 Limpiar daots incompletos o erróneos.
 - Mejorar la calidad de los datos.
 - Seleccionar un subconjunto de datos, reemplazar valores nulos.
 - Completar, eliminar o ignorar valores.
- 3.3 Transformación.
 - Elegir los atributos más relevantes.
 - Derivar las nuevas características más significativas a partir de las originales.
 - Discretización, mapear, normalizar, ...
- 3.4 Integrar: almacén de datos (datawarehouse).
- 3.5 Formatear: si es necesario para las técnicas posteriores. Conocido como ETL (extracción, transformación y carga).
 - Normalmente usa un modelo gráfico. Involucra la monitorización de la ejecución, el grabado de los datos y excepciones y errores.



Para la descripción de esta fase, se asume una vista minable: una tabla que contenga los atributos relevantes, etiquetados como entradas y salidas.

- 4.1 Técnicas de selección, considerando:
 - Adecuación al problema: clasificación, predicción, agrupamiento, asociación, dependencias, ...
 - · Con datos adecuados.
 - Cumpliento con los requisitos del problema.
 - Ejecutable a tiempo.
 - Conocimiento de la técnica.
- 4.2 Diseños de test.
 - Establecer training, testing y validation.
 - Establecer criterios para la bondad del modelo.
- 4.3 Crear el modelo: fijar los parámetros y ejecutarlo.
- 4.4 Evaluar el modelo: ¿cumple con la calidad esperada?



- 5.1 Interpretar los resultados: ¿queda resuelto el problema?, ¿es apropiada la respuesta?, ¿es válida?, ¿está bien definido el objtivo del negocio?, ¿es un conocimiento nuevo?, ¿es útil el modelo?, ¿es mejor que lo que se tenía antes?, ¿hay muchos patrones?, ...
- 5.2 Revisar el proceso: ¿algún error desde el punto de vista técnico?, ¿se ha pasado algo por encima?
- 5.3 Definir los siguientes pasos: ¿iterar?, ¿desplegar?, ¿reconstruir?, ...



- 6.1 Planificación del despliegue: ¿quiénes son los usuarios?, ¿cómo y cuándo será usado el modelo?, ¿cómo será desplegado?, ¿cómo una herramienta?, ¿es necesario un programa de ordenador?, ...
- 6.2 Planificación del seguimiento y mantenimiento: ¿se está usado?, ¿se está usado de forma adecuada?, ¿modelos actualizables?, ...
- 6.3 Generar un informe final.
- 6.4 Revisar el proyecto: debilidades y fortalezas, aspectos a mejorar, ...