

Universidad Internacional de La Rioja

Escuela Superior de Ingeniería y Tecnología

Grado en Ingeniería Informática

Dashboard de la Universidad de Valladolid en QlikView

|  |  |
| --- | --- |
| Trabajo fin de estudio presentado por: | Ana Lidia Herrera García |
| Director/a: | Jesús Pérez Melero |
| Fecha: | 17 de Julio del 2024 |
| Repositorio del código fuente: |  |

Resumen

El presente trabajo consiste en el diseño y la implementación de un Dashboard o Informe de Cuadro de Mandos para la Universidad de Valladolid, utilizando la herramienta de Business Intelligence *QlikView*. El proyecto surgió de la necesidad de mejorar la gestión y el análisis de los datos permitiendo una toma de decisiones más informada y eficiente.

El problema de partida se centró en la dispersión y el manejo manual de grandes volúmenes de datos provenientes de diversas fuentes de la universidad. Esta situación generaba ineficiencias, errores y limitaciones para obtener previsiones y estimaciones de los datos disponibles. La solución propuesta consistió en la creación de un Dashboard interactivo y centralizado que permitiera la visualización y análisis de estos datos de manera intuitiva y dinámica.

El Dashboard diseñado en QlikView integró datos académicos, administrativos y financieros, proporcionando a los usuarios una herramienta potente para monitorizar indicadores clave de rendimiento (KPIs), identificar tendencias, patrones y realizar análisis comparativos. La implementación de este sistema permitió optimizar procesos críticos como la gestión de matrículas, la planificación de cursos y la asignación de recursos.

Las principales conclusiones del trabajo indicaron que el Dashboard mejoró significativamente la eficiencia operativa y la transparencia en la toma de decisiones dentro de la universidad. Se reportó una mayor capacidad para acceder a información relevante y tomar decisiones basadas en datos actualizados y precisos. Además, se ensalzó la importancia del gobierno del dato entre los diferentes departamentos, mejorando la colaboración, integración y el uso de información estratégica.

En resumen, el proyecto ha demostrado ser una solución efectiva para los retos de gestión de datos de la Universidad de Valladolid, ofreciendo una plataforma robusta y flexible que potencia la inteligencia empresarial y la eficiencia institucional.

**Palabras clave:** Inteligencia Empresarial, Cuadro de mando, Indicadores clave de rendimiento, Análisis de datos, QlikView

Abstract

This work consists of the design and implementation of a Dashboard for the University of Valladolid, using the Business Intelligence tool *QlikView*. The project emerged from the need to improve data management and analysis, allowing for more informed and efficient decision-making.

The initial problem focused on the dispersion and manual handling of large volumes of data from various sources within the university. This situation generated inefficiencies, errors, and limitations in obtaining forecasts and estimates from the available data. The proposed solution consisted of creating an interactive and centralized Dashboard that allowed for intuitive and dynamic data visualization and analysis.

The Dashboard designed in QlikView integrated academic, administrative, and financial data, providing users with a powerful tool to monitor key performance indicators (KPIs), identify trends, patterns, and perform comparative analyses. The implementation of this system optimized critical processes such as enrollment management, course planning, and resource allocation.

The main conclusions of the work indicated that the Dashboard significantly improved operational efficiency and transparency in decision-making within the university. There was an increased ability to access relevant information and make decisions based on up-to-date and accurate data. Additionally, the importance of data governance among different departments was highlighted, improving collaboration, integration, and the use of strategic information.

In summary, the project has proven to be an effective solution to the data management challenges of the University of Valladolid, offering a robust and flexible platform that enhances business intelligence and institutional efficiency.

**Keywords**: Business Intelligence, Dashboard, Key performance indicators (KPIs), Data Analysis, QlikView

Índice de contenidos

[1. Introducción 10](#_Toc169472561)

[1.1. Motivación 10](#_Toc169472562)

[1.2. Planteamiento del trabajo 10](#_Toc169472563)

[1.3. Estructura del trabajo 11](#_Toc169472564)

[2. Contexto y Estado del Arte 13](#_Toc169472565)

[2.1. Análisis del Contexto 13](#_Toc169472566)

[2.1.1. Business Intelligence 13](#_Toc169472567)

[2.1.2. Diferencias entre Business Intelligence (BI), Business Analytics (BA) y Big Data (BD) 14](#_Toc169472568)

[2.1.3. Los datos, información y conocimiento 16](#_Toc169472569)

[2.1.4. Tecnologías 20](#_Toc169472570)

[2.1.4.1. Data Warehouse 21](#_Toc169472571)

[2.1.4.2. Dimensiones, métricas y tablas 24](#_Toc169472572)

[2.1.4.3. Análisis *OLAP* 25](#_Toc169472573)

[2.1.5. Modelado de datos dimensional 28](#_Toc169472574)

[2.1.5.1. Esquema Estrella 28](#_Toc169472575)

[2.1.5.2. Esquema Copo de Nieve 29](#_Toc169472576)

[2.2. Estado del arte 31](#_Toc169472577)

[2.2.1. Pentaho 31](#_Toc169472578)

[2.2.2. SAP BI 33](#_Toc169472579)

[2.2.3. MicroStrategy 34](#_Toc169472580)

[2.2.4. Power BI 35](#_Toc169472581)

[2.2.5. Tableau 36](#_Toc169472582)

[2.2.6. Qlik 37](#_Toc169472583)

[2.2.7. Diferencias entre las principales herramientas de Business Intelligence en el mercado 40](#_Toc169472584)

[3. Objetivos y metodología de trabajo 45](#_Toc169472585)

[3.1. Objetivo general 45](#_Toc169472586)

[3.2. Objetivos específicos 45](#_Toc169472587)

[3.3. Metodología de trabajo 46](#_Toc169472588)

[3.3.1. Metodología 46](#_Toc169472589)

[3.3.2. Tecnologías 48](#_Toc169472590)

[4. Solución del proyecto 49](#_Toc169472591)

[4.1. Análisis de Requisitos 49](#_Toc169472592)

[4.1.1. Visión General 49](#_Toc169472593)

[4.1.2. Alcance 49](#_Toc169472594)

[4.1.3. Stakeholders 51](#_Toc169472595)

[4.1.4. Arquetipos 52](#_Toc169472596)

[4.1.5. Historias de usuario 53](#_Toc169472597)

[4.1.6. Diagrama de casos de uso 53](#_Toc169472598)

[4.1.7. Requisitos funcionales 57](#_Toc169472599)

[4.1.8. Requisitos No funcionales 59](#_Toc169472600)

[4.1.9. Matriz de trazabilidad 60](#_Toc169472601)

[4.2. Diseño y modelado de datos 61](#_Toc169472602)

[4.2.1. Fuentes de datos 61](#_Toc169472603)

[4.2.2. Extracción, Transformación, Carga de BBDD 63](#_Toc169472604)

[4.2.3. Diseño del modelo 63](#_Toc169472605)

[4.3. Diseño del Dashboard 65](#_Toc169472606)

[4.4. Desarrollo del Dashboard y Visualización 70](#_Toc169472607)

[4.4.1. Tabla Simple 70](#_Toc169472608)

[4.4.2. Tabla Pivotante 70](#_Toc169472609)

[4.4.3. Gráfico Combinado 71](#_Toc169472610)

[4.4.4. Gráfico Tarta 72](#_Toc169472611)

[4.4.5. Gráfico Barras 73](#_Toc169472612)

[4.5. Calendario Maestro 73](#_Toc169472613)

[4.6. Set Analysis 75](#_Toc169472614)

[5. Conclusiones y trabajo futuro 77](#_Toc169472615)

[5.1. Conclusiones del trabajo 77](#_Toc169472616)

[5.2. Líneas de trabajo futuro 78](#_Toc169472617)

[Referencias bibliográficas 80](#_Toc169472618)

Índice de figuras

[Figura 1. Combinación de estrategias 15](#_Toc169472619)

[Figura 2. Dato, información, conocimiento 16](#_Toc169472620)

[Figura 3. Data Integration 17](#_Toc169472621)

[Figura 4. Dimensiones de la calidad de los datos 20](#_Toc169472622)

[Figura 5. Conceptos y tecnologías del Business Intelligence 21](#_Toc169472623)

[Figura 6. Contexto del Data Warehouse 23](#_Toc169472624)

[Figura 7. Dimensiones cubo OLAP 26](#_Toc169472625)

[Figura 8. Esquema estrella 29](#_Toc169472626)

[Figura 9. Esquema copo de nieve 30](#_Toc169472627)

[Figura 10. Arquitectura Pentaho 32](#_Toc169472628)

[Figura 11. Interfaz gráfica Pentaho Bi 33](#_Toc169472629)

[Figura 12. Interfaz gráfica de SAP BI 33](#_Toc169472630)

[Figura 13. Interfaz gráfica MicroStrategy 34](#_Toc169472631)

[Figura 14. Interfaz gráfica Power Bi 35](#_Toc169472632)

[Figura 15. Interfaz gráfica de Tableau 37](#_Toc169472633)

[Figura 16. Interfaz gráfica de QlikView 38](#_Toc169472634)

[Figura 17. Cuadrante Mágico de Gartner 42](#_Toc169472635)

[Figura 18. Arquitectura QlikView 44](#_Toc169472636)

[Figura 19. Cronograma del proyecto 47](#_Toc169472637)

[Figura 20. Diagrama de Gantt 47](#_Toc169472638)

[Figura 21. Diagrama de casos de uso 54](#_Toc169472639)

[Figura 22. Tablas SQL 61](#_Toc169472640)

[Figura 23. BBDD 62](#_Toc169472641)

[Figura 24. Proceso Ejecución Script 63](#_Toc169472642)

[Figura 25. Relaciones con Tablas Sintéticas 64](#_Toc169472643)

[Figura 26. Diagrama QlikView 64](#_Toc169472644)

[Figura 27. Diseño página Inicio 65](#_Toc169472645)

[Figura 28. Hoja Personal de Investigación 66](#_Toc169472646)

[Figura 29. Hoja Matrículas 66](#_Toc169472647)

[Figura 30. Hoja ayudas 67](#_Toc169472648)

[Figura 31. Hoja Ingresos 67](#_Toc169472649)

[Figura 32. Hoja Gastos 68](#_Toc169472650)

[Figura 33. Hoja Datos Capítulos 68](#_Toc169472651)

[Figura 34. Hoja Presupuestos con filtros 69](#_Toc169472652)

[Figura 35. Tabla Simple 70](#_Toc169472653)

[Figura 36. Tabla Pivotante contraída 71](#_Toc169472654)

[Figura 37. Tabla Pivotante expandida 71](#_Toc169472655)

[Figura 38. Gráfico Combinado 72](#_Toc169472656)

[Figura 39. Gráfico Tarta 72](#_Toc169472657)

[Figura 40. Gráfico Barras 73](#_Toc169472658)

[Figura 41. Gráfico Ayudas Universidad 75](#_Toc169472659)

Índice de tablas

[Tabla 1. Diferencias entre business Intelligence, business analytics y big data 15](#_Toc169472660)

[Tabla 2. Diferencias entre principales herramientas BI 40](#_Toc169472661)

[Tabla 3. Principales características del Dashboard 49](#_Toc169472662)

[Tabla 4. Historias de usuario 53](#_Toc169472663)

[Tabla 5. UC-01 Visualizar Datos 54](#_Toc169472664)

[Tabla 6. UC-02 Exportar Informe 55](#_Toc169472665)

[Tabla 7. UC-03 Buscar Información 56](#_Toc169472666)

[Tabla 8. Requisitos Funcionales 57](#_Toc169472667)

[Tabla 9. Requisitos No Funcionales 59](#_Toc169472668)

[Tabla 10. Matriz de Trazabilidad 60](#_Toc169472669)

# Introducción

La gestión eficiente de los datos ese ha convertido en un objetivo fundamental en cualquier institución académica. La Universidad de Valladolid con su vasto ecosistema de información proveniente de diversas fuentes académicas, administrativas y financieras, se ha enfrentado a grandes desafíos en la organización y en el análisis de estos datos. Este proyecto se centró en el diseño e implementación de un Informe de Cuadro de Mando utilizando la herramienta de Business Intelligence QlikView. Esta solución permitió a la Universidad de Valladolid visualizar y analizar sus datos de manera eficiente, optimizando así la toma de decisiones en todos los niveles de la institución.

## Motivación

La creciente cantidad de datos que generaba y manejaba la Universidad de Valladolid demandaba una herramienta que facilitara su análisis y utilización efectiva. Una gran cantidad de datos, la diversidad en cuanto a las fuentes y formatos, la necesidad de la unicidad del dato y el manejo manual habían llevado a ineficiencias, errores y limitaciones en la capacidad de obtener Insights valiosos. Esta situación ensalzaba la necesidad de un sistema centralizado e interactivo que no sólo almacenara y organizara los datos, sino que también permitiera su análisis dinámico y en tiempo real. La implementación de un Dashboard con QlikView respondió a esta necesidad, proporcionando una plataforma robusta para mejorar la transparencia, la eficiencia operativa y la toma de decisiones.

## Planteamiento del trabajo

El principal problema que abordó este trabajo es la dispersión y el manejo manual de grandes volúmenes de datos provenientes de diversas fuentes de datos dentro de la Universidad de Valladolid. De este modo se generaban ineficiencias operativas, errores humanos en la recopilación y el análisis de los datos, limitaciones a la hora de obtener previsiones y falta de transparencia y accesibilidad de la información para la toma de decisiones. La gestión eficiente de datos se volvió crucial para mejorar la toma de decisiones estratégicas y operativas.

Para resolver estos problemas, se propuso el desarrollo de un Dashboard interactivo y centralizado utilizando QlikView. Esta herramienta de business Intelligence permitió la integración de datos académicos, administrativos y financieros, proporcionando una plataforma potente para la visualización y análisis de datos. Permitió visualizar y analizar los datos de forma eficiente y en tiempo real, con una gran cantidad de informes y reportes accesibles de forma inmediata, flexible, exportables y personalizables gracias a los filtros. Esto permitió a los usuarios identificar tendencias, patrones y relaciones que antes eran difíciles de visualizar con métodos tradicionales.

Con esta solución, la universidad pudo mejorar sus procesos como la gestión de matrículas, gestión de ayudas, la asignación del personal de investigación, la planificación de los cursos y el control presupuestario de forma que todo ello implicó una toma de decisiones más informada y eficiente.

Este proyecto no solo buscó resolver los problemas actuales de gestión de datos, sino también establecer una base sólida para una cultura de datos robusta y sostenible en la Universidad de Valladolid.

## Estructura del trabajo

Este trabajo se estructuró en varios capítulos que abordaron diferentes aspectos de la gestión del informe del cuadro de mando de Business Intelligence para la Universidad de Valladolid.

A continuación, se detalla la organización de cada sección:

En el presente **Capítulo 1** se explicó la motivación, el planteamiento del problema a abordar y la estructura del documento.

En el **Capítulo 2** se especificó el análisis de contexto y los principales resultados encontrados en el Estado del Arte. Podemos destacar el amplio estudio que se efectuó de los elementos a tener en cuenta, las diversas herramientas de Business Intelligence y la comparativa para elegir que más se adecuaba a la resolución del problema.

En el **Capítulo 3** quedó detallado el objetivo general, los objetivos específicos, así como la metodología y las tecnologías que se utilizaron para el desarrollo.

En el **Capítulo 4** se explicó con detalle como se abordó el problema, se inició con el análisis de los requisitos, seguido del diseño, del modelado de datos y del diseño del Dashboard, y se concluyó con el desarrollo del informe del cuadro de mando y la visualización con diversas tablas y gráficas. También se explicó cómo se desarrolló el calendario maestro y expresiones mediante el set Analysis.

Finalmente, en el **Capítulo 5** se concluye con el análisis de las conclusiones del trabajo, así como las líneas de trabajo futuro identificadas del propio trabajo.

# Contexto y Estado del Arte

En este apartado se ha contextualizado el problema, se han detallado los conceptos y se han especificado las tecnologías relacionadas.

## Análisis del Contexto

Se ha analizado la problemática que se ha abordado, la magnitud del problema, sus causas y la necesidad que se ha identificado.

### Business Intelligence

La inteligencia empresarial,llamada también Business Intelligence (BI), ha sido reconocida como un software especializado en la gestión de datos empresariales. En términos más específicos, se ha denominado así al conjunto de estrategias, tecnologías, aplicaciones y métodos que se han empleado para recopilar, integrar, analizar y presentar información corporativa, así como para visualizar datos empresariales relacionados con el propósito de optimizar la toma de decisiones en una empresa (IBM, s.f.)

El propósito principal de la inteligencia empresarial ha radicado en la transformación de los datos en información relevante y en las prácticas que han facilitado a las compañías a comprender más a fondo sus actividades, descubrir tendencias, identificar patrones y detectar oportunidades, en resumen, que ha permitido respaldar la toma de decisiones estratégicas y operativas. Este proceso ha involucrado la utilización de herramientas y técnicas para recolectar datos provenientes de diversas fuentes, depurar y procesar esos datos, y posteriormente analizarlos para obtener conocimientos que puedan orientar hacia toma de decisiones más efectivas y fundamentadas.

Las tecnologías relacionadas con la inteligencia empresarial han abarcado sistemas de gestión de bases de datos, herramientas de análisis de datos, software de visualización de datos, paneles de control y cuadros de mandos, entre otros. Se ha demostrado que la inteligencia empresarial puede ser aplicada en distintos ámbitos dentro de una organización, tales como operaciones, finanzas, recursos humanos, marketing y ventas.

### Diferencias entre Business Intelligence (BI), Business Analytics (BA) y Big Data (BD)

Tanto el BI, BA y BD han permitido el análisis de datos con el fin de extraer la máxima información, no obstante, es importante tener presente las diferencias:

La Inteligencia de Negocio (Business Intelligence, **BI**) se enfoca en analizar el pasado de una empresa, utilizando informes para revisar su histórico y comprender su evolución. Mediante herramientas de BI, se accede a conjuntos de datos organizados y almacenados, donde se examina la información en busca de patrones analíticos. Este proceso implica el análisis de datos estructurados, que luego se almacenan en un servidor central para su posterior estudio y la obtención de conclusiones que ayuden a respaldar la toma de decisiones.

El enfoque de la Analítica de Negocio (Business Analytics, **BA**) se proyecta hacia adelante, proporcionando predicciones fundamentadas en modelos predictivos que orientan la formulación de nuevas decisiones y estrategias. A diferencia del enfoque del BI, el BA no se limita al análisis de estadísticas internas, sino que aprovecha una variedad de fuentes, como tendencias o indicadores económicos globales. Ambas disciplinas persiguen la optimización de procesos, mientras que el Business Intelligence se centra en corregir errores operativos presentes, el Business Analytics se dedica a prevenir dichos errores en el futuro.

Big Data (**BD**) se refiere a conjuntos de datos masivos y complejos que no pueden ser procesados de forma efectiva con herramientas de procesamiento de datos convencionales. Estos datos, provenientes de diversas fuentes, pueden ser tanto estructurados como no estructurados, y son tan vastos que requieren métodos especiales de procesamiento para extraer información útil. El término abarca tanto el almacenamiento masivo de datos como los métodos utilizados para identificar patrones repetitivos dentro de ellos. En lugar de almacenarse en un único servidor central, la información se clasifica y procesa en archivos distribuidos. El enfoque principal de Big Data está en la captura y procesamiento de datos, a diferencia de Business Analytics y Business Intelligence, que se centran en analizar esta información para mejorar la toma de decisiones.

(Big Data International Campus, s.f.)

Tabla 1. Diferencias entre business Intelligence, business analytics y big data

Tabla

Descripción generada automáticamente

*Fuente:* (Curto Díaz, 2016)

Figura 1. Combinación de estrategias

Tabla

Descripción generada automáticamente

*Fuente:* (Curto Díaz, 2016)

### Los datos, información y conocimiento

Los pilares fundamentales para la toma de decisiones se han basado en los datos, la información y el conocimiento. El proceso de transformar la información en conocimiento y su interacción ha ayudado a discernir y facilitar el proceso de toma de decisiones.

Figura 2. Dato, información, conocimiento

Diagrama

Descripción generada automáticamente

Fuente: (Madrid, 2008)

La transformación en las necesidades empresariales, la diversidad de los datos asociados a una empresa y la necesidad de su integración ha impulsado las herramientas de integración de datos hacia una mayor capacidad de limpieza y calidad de estos. Estas herramientas se basan en técnicas de aprendizaje automático supervisado y no supervisado (Curto Díaz, 2016).

El auge actual de los datos se debe a la enorme y creciente cantidad de información, influido por tecnologías clave que han alterado la manera en que se recopilan, almacenan, analizan y procesan los datos. Esta dinámica y los factores que la han impulsado han permitido obtener percepciones claras desde la profundidad de los datos, lo que a su vez ha conducido a la generación de nuevos conocimientos, la identificación de nuevas relaciones y la realización de nuevas predicciones (BSA y The Software Alliance, 2015).

Figura 3. Data Integration

Imagen que contiene Escala de tiempo

Descripción generada automáticamente

*Fuente:* (Curto Díaz, 2016)

#### Tipos de datos

Se han definido diferentes tipos de datos:

* **Estructurados**: son conjuntos de datos organizados y ordenados de manera sistemática, como en bases de datos relacionales o archivos de hojas de cálculo, donde cada elemento tiene una posición definida y unas características específicas (tamaño, formato, longitud, posibles relaciones) lo que facilita su manipulación y comprensión.
* **No estructurados**: son aquellos que carecen de un formato específico y pueden presentarse en diversos tipos de archivos, como documentos Word, imágenes, vídeos o correos electrónicos. Estos datos son más difíciles de procesar y suelen requerir técnicas especiales para su análisis.
* **Semiestructurados**: son una combinación de datos estructurados y no estructurados, que no tienen una organización clara, pero cuentan con metadatos que describen sus objetos y relaciones. Ejemplos de estos datos son los archivos HTML y XML, que poseen cierta estructura, pero también elementos que pueden variar en su formato y contenido.

(Trejo, 2020)

#### La calidad de los datos

La precisión de las decisiones en el Business Intelligence han dependido en gran medida de la calidad de los datos, lo que ha afectado directamente a la confiabilidad de los resultados, a la eficiencia en las operaciones, a la capacidad para detectar tendencias e identificar patrones, así como para impulsar el rendimiento empresarial.

Las dimensiones que se han de cumplir para los requisitos de la calidad de los datos:

* **Completitud**. Es esencial que los datos estén completos para asegurar la calidad y confiabilidad de los análisis, informes y decisiones en el ámbito de la inteligencia empresarial. La falta de datos completos puede provocar la pérdida de información valiosa o incluso la toma de decisiones basadas en datos sesgados o incompletos, lo cual puede ser crítico en ciertos procesos de la empresa si son datos necesarios.
* **Conformidad**. Los datos utilizados deben estar en un formato estándar, legible y válido. La conformidad es crucial para garantizar la integridad de los datos y reducir el riesgo de incumplimiento normativo, protegiendo así la reputación y la confianza de la organización.
* **Unicidad.**  Mantener la unicidad del dato en toda la organización es esencial para garantizar la consistencia, calidad y confiabilidad de la información en el ámbito de la inteligencia empresarial. Esto permite tomar decisiones más informadas y eficaces que conducen e impulsan al éxito de la empresa.
* **Precisión/ Exactitud.** La exactitud de los datos es fundamental para garantizar que representen con precisión el valor real o la fuente de referencia. Los datos precisos están libres de errores de entrada manual o de medición. La falta de precisión puede invalidar la utilidad de los datos.
* **Consistencia**. Los datos deben ser coherentes tanto dentro de un conjunto de datos como en relación con otros datos. La consistencia implica que los datos no deben contradecirse entre sí y que deben seguir las mismas convenciones y estándares en toda la organización. La consistencia es clave al cruzar información entre registros, se deben evitar incoherencias y contradicciones.
* **Integridad**. La integridad de los datos se refiere a su totalidad y nuevamente a su completitud. Los datos deben estar completos, actualizados y no deben faltar elementos importantes que podrían ser necesarios para su análisis. Esto asegura que toda la información relevante esté presente y lista para su uso (PowerData, 2014).

La información además ha de ser **accesible** sólo para quienes estén autorizados, garantizando así la **seguridad** de los datos. Además, debe estar disponible de manera oportuna, con cierta **puntualidad**, que tarde el tiempo mínimo para acceder a la información solicitada y que cuente con la rapidez suficiente para que la nueva información esté disponible para todas las fuentes. Si los procesos para almacenar los datos entrantes son complejos y lentos, puede afectar a la puntualidad y la efectividad del uso de la información y puede deberse a que los datos no estén bien almacenados, relacionados, organizados o tratados (Dataladder, 2021).

Figura 4. Dimensiones de la calidad de los datos

Forma

Descripción generada automáticamente con confianza media

Fuente: (Deloitte)

### Tecnologías

Se han listado las diferentes tecnologías que forman parte del Business Intelligence:

* *Data Warehouse.*
* *Reporting.*
* *Analisis OLAP (Online Analytical Processing).*
* *Análisis visual.*
* *Análisis predictivo.*
* *Cuadro de mando.*
* *Cuadro de mando integral.*
* *Minería de datos.*
* *Gestión del rendimiento.*
* *Previsiones.*
* *Reglas de negocio.*
* *Dashboards.*
* *Integración de datos (que incluye Extract, Transform and Load; ETL).*

(Curto Díaz, 2016)

Mediante el siguiente diagrama se han representado las diferentes tecnologías y aspectos que forman parte de la inteligencia de negocio.

Figura 5. Conceptos y tecnologías del Business Intelligence

Tabla

Descripción generada automáticamente

Fuente: (Forrester Research, 2008)

### 2.1.4.1. Data Warehouse

*Data Warehouse* es el repositorio de la información relevante para la organización, que proporciona una perspectiva global, común e integrada de los datos de la empresa. Dicho de otra forma, se trata de conjuntos de datos organizados por temas, que se encuentran integrados, que cuentan con variaciones en el tiempo y que no se alteran con facilidad, con el fin de respaldar la toma de decisiones.

Es de máxima importancia dentro de la inteligencia empresarial, ya que su diseño óptimo contribuye significativamente al éxito del rendimiento de la organización y a implementar un sistema eficaz de inteligencia empresarial. Con características de estabilidad, coherencia, fiabilidad y con histórico de información, almacena un gran volumen de datos.

El soporte técnico más utilizado para su almacenamiento suelen ser las bases de datos relacionales, aunque también se utilizan bases de datos orientadas a columnas, basadas en lógica asociativa o *appliances* especializadas.

En el contexto de un *Data Warehouse* se han considerado otros elementos que se combinan para poder responder a las necesidades de negocio:

* ***Data Warehousing***. Proceso de extraer y filtrar los datos de las actividades comunes de la organización, procedentes de los distintos sistemas expertos y/o de sistemas de información operacionales.
* ***Data Mart***. Se denomina al subconjunto de los datos del *Data Warehouse* que se centra en satisfacer a un determinado análisis, función o necesidad. Suelen tener una población de usuarios específica. Cubre las necesidades de un grupo de trabajo o departamento, por ejemplo, para la minería de datos o para el departamento de marketing. Los datos, que también son estructurados, pueden ser dependientes o no del *Data Warehouse*.
* ***Operational Data Store***. Es un tipo de almacén de datos que proporciona sólo los valores más recientes de los datos, sin conservar su historial. Generalmente se permite un pequeño desfase o retraso sobre los datos operacionales.
* ***Staging Area***. Se denomina así al sistema intermedio que permanece entre las fuentes de datos y el Data Warehouse para facilitar la extracción de datos desde fuentes de datos muy diversas y complejas. Funciona como una especie de “caché” de datos operacionales sobre el que posteriormente se realiza el proceso de almacenamiento mejorando así la calidad de los datos y permitiendo el acceso a información que no se encuentra en el Data Warehouse.
* ***Procesos ETL (Extracción, Transformación y Carga)*.** Son tecnologías de integración de datos que se utilizan para alimentar almacenes de datos como *Data Warehouse*, *Data Mart*, *Staging Area* y *ODS*.

Estos procesos se combinan con otras técnicas de consolidación de datos. En general, los almacenamientos de datos utilizan los procesos ETL para convertir y estandarizar la información procedente de diferentes fuentes antes de ponerla a disposición de las herramientas OLAP. Los datos se transforman, integran y almacenan con el fin de acceder a ellos para dar soporte en la toma de decisiones.

* **Metadatos**. Son datos estructurados y codificados que describen las características de las instancias de los datos conteniendo información para ayudar a identificar, descubrir, valorar y administrar las instancias descritas.

(Curto Díaz, 2016)

Figura 6. Contexto del Data Warehouse

Diagrama

Descripción generada automáticamente

*Fuente:* (Curto Díaz, 2016)

Se deben identificar las vistas para el análisis, las medidas cuantificables asociadas y la organización de los procesos de negocio donde se estructura el dato. En el diseño la información se puede almacenar de forma desnormalizada para optimizar las consultas y las diferentes relaciones.

### 2.1.4.2. Dimensiones, métricas y tablas

Se han identificado como importantes las siguientes definiciones:

* La *dimensión* se refiere a la vista que representa un proceso de negocio en el *Data Warehouse*. Este concepto está relacionado con el análisis multidimensional, que utiliza el procesamiento *OLAP* operando mediante tablas pivotantes o matrices.
* Las *métrica*s son los indicadores cuantificables y medibles de un proceso de negocio y se pueden clasificar en:
* ***Leading***. Métricas que miden el progreso de una actividad.
* ***Lagging****.* Métricas que recogen y evalúan los resultados de una actividad.
* Indicadores clave. Valores objetivos que proporcionan información sobre el rendimiento de una actividad o el logro de una meta.
  + *Key performance indicator (****KPI****).* Indicadores clave de rendimiento, son métricas del proceso. Definen un rango óptimo de rendimiento para alcanzar los objetivos.
  + *Key goal indicator (****KGI****).* Indicadores de metas, que suelen expresarse en términos de criterios de información y definen medidas para informar a la Dirección sobre el cumplimiento de los requisitos de negocio.

(Curto Díaz, 2016)

* Una ***Tabla*** *de hecho* es la representación en el *Data Warehouse* de los procesos de negocio. A través del ***cuadro de mando*** se muestra información ya consolidada, con una metodología principalmente gráfica y con cierto margen de interactividad. Esto facilita una comprensión rápida de la situación del negocio.

### 2.1.4.3. Análisis *OLAP*

El procesamiento analítico en línea (***OLAP, Online Analytical Processing***) es una tecnología de software que se utiliza para el análisis de datos empresarial con varias dimensiones. Realiza un análisis multidimensional de alta velocidad sobre un gran volumen de datos. Las organizaciones recopilan y almacenan datos de múltiples fuentes de datos, como sitios web, aplicaciones, medidores inteligentes y sistemas internos. *OLAP* combina y agrupa estos datos en categorías para proporcionar información procesable para la planificación estratégica.

En un almacenamiento de datos los conjuntos de datos se almacenan en tablas, cada una de las cuales puede organizar datos en sólo dos dimensiones a la vez. OLAP extrae datos de varios conjuntos de datos relacionales y los reorganiza en un formato multidimensional permitiendo un procesamiento mucho más rápido y un análisis más detallado.

La máquina subyacente que alimenta el sistema OLAP es el ***servidor OLAP***. Utiliza herramientas de Extracción, Transformación y Carga para transformar la información en las bases de datos relacionales y prepararlos para su uso en operaciones OLAP.

Una ***base de datos OLAP*** es una base de datos independiente que se conecta al almacenamiento de datos y que se puede utilizar para evitar que el almacenamiento de datos se vea sobrecargado por el análisis OLAP y para facilitar la creación de modelos de datos OLAP. (Amazon Web Service, s.f.)

El núcleo de la mayoría de los sistemas OLAP se denomina ***cubo OLAP*** o hipercubo. Este consiste en una base de datos multidimensional basada en matrices que posibilita el procesamiento y el análisis de varias dimensiones de datos de manera mucho más rápida y eficiente que una base de datos relacional tradicional. Una tabla única se ve ampliada con capas adicionales, cada una de las cuales agrega dimensiones adicionales, correspondiente al siguiente nivel en la "jerarquía de conceptos" de la dimensión.

Para explorar más a fondo la información dentro de un cubo OLAP se pueden realizar las siguientes *operaciones*:

* **Desglosar (Slide and Dice).** Esta operación crea un subcubo seleccionando una única dimensión del cubo OLAP principal.
* **Profundizar (Drill-Down).** Transforma datos menos detallados en datos más detallados, ya sea descendiendo en la jerarquía de conceptos o agregando una nueva dimensión al cubo.
* **Consolidar (Roll up).** La operación de consolidación es la función opuesta a la profundización, ya que agrega datos en un cubo OLAP ascendiendo en la jerarquía de conceptos o reduciendo el número de dimensiones.
* **Rotar (Pivot).** Esta operación permite cambiar la visualización actual del cubo para mostrar una nueva representación de los datos mediante vistas multidimensionales dinámicas, similar a una hoja de cálculo.

(IBM, s.f.)

Figura 7. Dimensiones cubo OLAP

Imagen que contiene Gráfico

Descripción generada automáticamente

*Fuente:* (IBM, s.f.)

Se han descrito diferentes tipos de análisis de datos multidimensionales:

* ***MOLAP*.** Es un tipo de procesamiento analítico en línea (OLAP) que trabaja directamente con un cubo *OLAP* multidimensional. Se denomina procesamiento analítico multidimensional en línea (*MOLAP*) e implica la creación de un cubo que representa datos multidimensionales y que almacena en el hipercubo los datos calculados previamente lo que permite proporcionar un análisis rápido.
* ***ROLAP***. Es un método de procesamiento analítico en línea que realiza análisis de datos multidimensionales operando directamente sobre datos en tablas relacionales, sin la necesidad de reorganizarlos en un cubo. Se denomina método de procesamiento analítico relacional en línea (*ROLAP*). Utiliza consultas SQL para buscar y recuperar información específica según las dimensiones requeridas.

El Procesamiento *ROLAP* es adecuado para analizar datos extensos y detallados. Tiene un rendimiento de consultas más lento que *MOLAP*, por lo que es mejor cuando la capacidad para trabajar directamente con grandes cantidades de datos es más importante que el rendimiento y la flexibilidad.

* ***HOLAP***. Este enfoque híbrido combina características de MOLAP y ROLAP para aprovechar lo mejor de ambas arquitecturas. Permite una recuperación rápida de resultados analíticos desde un cubo de datos, al mismo tiempo que permite extraer información detallada de bases de datos relacionales. Se denomina procesamiento analítico híbrido en línea (*HOLAP*).

Aunque ofrece mejor escalabilidad, puede experimentar ralentizaciones al acceder a fuentes de datos relacionales. Su arquitectura más compleja requiere más actualizaciones y mantenimiento, ya que debe manejar tanto datos de bases de datos relacionales como multidimensionales.

### Modelado de datos dimensional

Los sistemas de bases de datos relacionales llegaron en la década de los 80 enfocados en guardar, extraer y modificar datos. La técnica de modelado de Entidad Relación ha buscado eliminar la posible redundancia en el modelo de base de datos, ayudando a simplificar y mejorar el diseño de los sistemas transaccionales.

Estos avances han mejorado la eficiencia para insertar y modificar información en una base de datos transaccional y a su vez han complicado la extracción de la información cuando el modelado implica muchas entidades y aumenta el número de tablas.

Trabajar con fuentes de datos que no provienen de un único almacén de datos o *Data Marts,* se puede volver un proceso muy complejo. Un modelado de datos dimensional se ha convertido en una alternativa muy buena.

Este tipo de modelado está compuesto por una sola tabla de hechos, lo cual contiene una clave primaria y llaves adicionales que la ligan con las tablas de dimensión. Estas tablas contienen atributos y descripciones que añaden contexto a las métricas establecidas en las tablas de hechos. Las dimensiones suelen tener datos de varios niveles jerárquicos, “aplanados” dentro de una misma tabla. Contienen llaves para referenciar tablas de dimensión y datos de métricas para las tablas de hechos.

### 2.1.5.1. Esquema Estrella

Las tablas de hechos y dimensiones normalmente se combinan en un esquema tipo estrella. Es un esquema llamado así por su forma y es sencillo de entender. Los usuarios de negocio pueden fácilmente reconocer los nombres de las tablas y cómo se relacionan entre sí. La tabla de dimensión está desnormalizada.

Figura 8. Esquema estrella

Diagrama

Descripción generada automáticamente

*Fuente:* (Tunnell, s.f.)

### 2.1.5.2. Esquema Copo de Nieve

Es otro tipo de modelo dimensional en el que las tablas de dimensiones no están completamente aplanadas o desnormalizadas. Denominado también por la figura que conforma.

En un sistema de base de datos relacional, se suele utilizar este tipo de esquema para ahorrar espacio en disco ya que evita valores duplicados en la tabla de dimensión.

(García & Harmsen, 2013)

Figura 9. Esquema copo de nieve

Diagrama

Descripción generada automáticamente

*Fuente:* (Tunnell, s.f.)

La optimización de los datos y las relaciones entre las tablas han creado la forma de automatizar de forma más eficiente el modelo. Una vez que se ha elegido el modelado de esquema adecuado hará que las herramientas de BI funcionen más rápido ya que también se minimizarán las conexiones a las tablas y los tiempos.

Suele preferirse el esquema Estrella por ser más óptimo.

Simplificar y optimizar el modelo de datos también mejora el entendimiento y el mantenimiento.

El conocimiento de los sistemas de Reporting interno, de las fuentes de información y de la operativa se ha convertido en un punto fundamental para la integración de todos los datos al informe.

## Estado del arte

La implantación de los sistemas de información de inteligencia empresarial ha permitido la optimización de la toma de decisiones, que se ha traducido en mejores resultados para las organizaciones. A partir de los datos se obtiene información, se genera conocimiento y se generan nuevos datos. Como se ha detallado, entre las ventajas del *business Intelligence* se han encontrado:

* Ofrecer una visión integral, histórica, persistente, accesible y de alta calidad de la información relevante para la empresa.
* Desarrollar y gestionar métricas de rendimiento e indicadores clave para los objetivos corporativos.
* Fomentar la integración y colaboración entre departamentos y unidades de negocio, con el fin de asegurar la alineación con las estrategias organizacionales.
* Mejorar la comprensión, la documentación contextual dentro de la empresa, las percepciones y las demandas de los clientes, contribuyendo así a la competitividad.
* Maximizar la utilidad de la información relevante.
* Facilitar un acceso más rápido y confiable a la información.
* Agilizar el proceso de toma de decisiones.

El mercado del Business Intelligence se ha enriquecido con soluciones de código abierto (*open source OSBI*) como Pentaho, Tibco Jaspersoft, Birt y SpagoBI y hay otras soluciones como Power Bi, Qlik, Tableau...

Se han analizado las soluciones existentes más relevantes relacionadas con este problema u otros muy similares:

### Pentaho

En el mercado del BI de código abierto (*Open Source Business Intelligence*) una de las suites más conocidas es *Pentaho*. Surgió en el 2006 y ofrece 2 versiones: *Community*, de libre acceso, y la *Enterprise*, que brinda funcionalidades extendidas mediante una suscripción.

A continuación, se detalla el contexto:

* Metadata: Proporciona una capa de acceso a la información basada en términos de negocio facilitando la comprensión y la navegación de los datos.
* ETL: Incluye un servidor que utiliza la herramienta *Pentaho Data Integration* o *Kettle*, permitiendo la extracción, transformación y carga de datos de diversas fuentes.
* Análisis OLAP: Ofrece soporte para análisis multidimensional a través de *Mondrian* y para minería de datos mediante *Weka*, lo que permite explorar y descubrir patrones en los datos.
* Reporting: Admite la generación de informes estáticos, paramétricos y Ad Hoc, brindando flexibilidad en la presentación de datos según las necesidades del usuario.
* Cuadros de mando: En la versión *Community* se emplea *el Community Dashboard Framework*, mientras que en la *Enterprise* se cuenta con el *Dashboard Designer*, permitiendo visualizar y analizar los datos de manera intuitiva y personalizada.

Figura 10. Arquitectura Pentaho

Diagrama

Descripción generada automáticamente

*Fuente:* (Pentaho, s.f.)

Se trata de una herramienta altamente intuitiva, con una interfaz de usuario sencilla que puede ser extendida mediante código Javascript y puede integrarse con Hadoop. Sin embargo, su rendimiento es limitado y ofrece una cantidad reducida de componentes. Además, su soporte se basa en la colaboración de la comunidad de usuarios.

Figura 11. Interfaz gráfica Pentaho Bi

Interfaz de usuario gráfica, Gráfico, Aplicación

Descripción generada automáticamente

*Fuente:* (Fernandez, 2023)

### SAP BI

Esta plataforma de análisis empresarial de SAP permite la creación de paneles de control y visualizaciones de datos. Además, se integra con todos los productos analíticos de la suite de SAP.

Ofrece acceso a datos en tiempo real y una amplia capacidad para compartir informes o paneles de control.

Figura 12. Interfaz gráfica de SAP BI

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

*Fuente:* (Fernandez, 2023)

A pesar de contar con una interfaz de usuario muy intuitiva, esta herramienta tiene un coste elevado y ofrece integraciones limitadas con otros productos, así como un rendimiento poco optimizado.

### MicroStrategy

La plataforma *MicroStrategy Workstation* ofrece capacidades de análisis visual de datos de forma gratuita, sin necesidad de licencia. Se destaca por su alto nivel de personalización y la capacidad de realizar análisis de datos en tiempo real. Además, cuenta con funcionalidades de minería de datos y análisis predictivo.

Sin embargo, como inconveniente, presenta opciones de exportación limitadas y su plataforma de administración puede ser compleja de gestionar.

Figura 13. Interfaz gráfica MicroStrategy

Imagen de la pantalla de un celular con letras

Descripción generada automáticamente con confianza media

*Fuente:* (Fernandez, 2023)

### Power BI

La solución de inteligencia empresarial de Microsoft ofrece opciones de servicio en la nube *(SaaS)* o locales *(On Premise).* Es de fácil implementación y permite cargar, compartir y acceder a informes desde cualquier dispositivo. A través de *Microsoft Dynamics* permite el acceso a bases de datos, fuentes de datos locales y servicios en la nube.

Con *Quick Insights* es posible identificar correlaciones y patrones de datos, generando gráficos y gráficos personalizados. Permite extraer informes de forma autónoma y permite la integración del análisis avanzado a través de scripts y objetos visuales de *R, Microsoft Azure Machine Learning y Azure Stream Analytics.*

La herramienta centraliza la recopilación y almacenamiento de toda la información, tanto local como en la nube, en un único lugar de fácil acceso. Es ideal para proyectos transversales, ya que está diseñada para establecer relaciones con múltiples fuentes de datos.

Además, dispone de una gran capacidad de integración con otras soluciones de Microsoft (*Microsoft Dynamics, Excel, SQL Server*…) y soluciones de terceros (*Amazon, Salesforce, SAP,* Office 365, Teams).

(Marín, 2023)

Figura 14. Interfaz gráfica Power Bi

Aplicación

Descripción generada automáticamente

*Fuente:* (Fernandez, 2023)

La herramienta cuenta con una interfaz fácil de usar y amigable, diseñada para que los usuarios puedan utilizarla sin necesidad de tener experiencia técnica previa. Incorpora capacidades de inteligencia artificial y automatización de informes, lo que facilita la limpieza de datos y el análisis avanzado.

Cuenta con servicios de Power Platform como Power BI Desktop, Servicio Power BI, Power BI Mobile, Power Apps y Power Automate. Power Apps puede acceder a los datos de Power BI y con Power Automate se pueden construir flujos de trabajo para realizar determinadas acciones en Power BI.

Presenta compatibilidad con Windows, iOS y Android.

En cuanto a calidad/precio no sacrifica la calidad de su solución para lograr su modelo de precios asequible. Por ejemplo, el servicio en la nube de Power BI ofrece un conjunto de análisis mejorados y soluciones automatizadas de *Machine Learning* por un precio competitivo.

Sin embargo, las versiones empresariales pueden tener un costo elevado y presentar dificultades para migrar fuera del entorno de Microsoft.

### Tableau

Herramienta de BI que posibilita la exploración y el análisis de datos sin necesidad de tener conocimientos de programación. También ofrece la capacidad de conectarse a una multitud de fuentes de datos y proporciona una API para extraer datos de forma sistemática. Permite análisis avanzados y cálculos estadísticos minuciosos.

Permite realizar análisis avanzados y cálculos estadísticos detallados, así como comparar datos, aplicar filtros y establecer conexiones entre variables.

Facilita la exploración visual y la manipulación de datos, lo que permite a los usuarios empresariales acceder, preparar, analizar y presentar los resultados de sus datos de manera sencilla. Está diseñada para realizar análisis Ad-hoc a un nivel más profundo (Marín, 2023).

Figura 15. Interfaz gráfica de Tableau

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

*Fuente:* (Fernandez, 2023)

La productividad de Tableau permite a los usuarios vincular esta herramienta a una amplia variedad de recursos, así como importar y visualizar datos de manera sencilla. Entre los servicios que ofrece incluyen *Tableau Desktop, Tableau Server y Tableau Online.*

Es compatible con entornos web y dispositivos Android, Windows, Linux, iPhone/iPad y Mac. En cuanto al precio, es mucho más costoso que Power BI, con opciones de licencias por suscripción anual, pago único o basado en cotizaciones. Ofrece opciones tanto On Premise como Cloud.

En resumen, entre sus características destacadas se encuentran la creación de paneles interactivos, informes, notificaciones de datos, gestión de metadatos, comentarios en el tablero, alternancia de vista y la función de arrastrar y soltar para la visualización de datos.

### Qlik

Plataforma enfocada al análisis de datos y a aplicaciones interactivas para mejorar el proceso de acceso de datos, visualización y toma de decisiones del usuario.

La base de información para generar conocimiento se puede generar rápidamente gracias a su gran capacidad para la creación de scripts, así como el motor asociativo subyacente que permite realizar uniones naturales en tiempo real en función de las selecciones del usuario. Destaca las relaciones entre las entidades para el usuario.

Cuenta con una gran escalabilidad, ya que tiene una gran capacidad para trabajar con modelos de datos complejos y para la ejecución de cálculos avanzados. Cuenta con la posibilidad de crear soluciones verticales preconfiguradas.

Destaca también por la usabilidad. Es muy fácil hacer búsquedas directas e indirectas y tiene gran capacidad de comprensión asociativa. Fácil de usar por la variedad de filtros y muy intuitiva.

(García & Harmsen, 2013)

Figura 16. Interfaz gráfica de QlikView

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

*Fuente:* (Fernandez, 2023)

Permite integrar y combinar diversas fuentes de datos para crear informes y extraer métricas y conclusiones fácilmente. La herramienta Qlik es una solución basada en *Cloud* que usa *Machine Learning* para mejorar la analítica. También ofreces soluciones On premise y API abierta para integrarse con otras aplicaciones empresariales.

Qlik ofrece una analítica componible que permite integrar la analítica con los procesos de negocio mediante su API abierta y su Automatización de Aplicaciones. Gracias a esta funcionalidad, es posible integrar la analítica en las aplicaciones de negocio y ensamblar rápidamente flujos automatizados con un enfoque muy visual.

Algunas organizaciones se preocupan por la dependencia de su proveedor con la nube, por ello valoran aún más un proveedor como Qlik, que les permite tener más flexibilidad y libertad de elección en su infraestructura

Cuenta con Dashboards interactivos, exportación de datos y Dashboards con formatos de imagen, Excel o PDF, acceso con API y dispositivos móviles. Es decir, es compatible con entornos web y dispositivos Windows, iPhone/iPad y Mac.

Como desventajas se presenta la curva de aprendizaje elevada para nuevos usuarios y el soporte empresarial (Fernandez, 2023).

Qlik cuenta con un amplio conjunto de tecnologías que brindan soporte a múltiples perfiles de negocio con amplias capacidades analíticas y de datos. Ofrece a los usuarios múltiples servicios mediante las herramientas Qlik Sense, Qlik Analytics Platform, Qlik Core, QlikView para crear espacios de trabajo y permitiendo a un usuario de negocio ser autosuficiente para realizar su propio análisis. Además, en función del producto elegido, encontramos varios modelos de precio: gratis, por mensualidades, pago único o pago basado en cotizaciones.

Un ejemplo de estas herramientas es QlikView, herramienta computacional desarrollada por QlikTech, fundada en Suecia en 1993, actualmente con sede en EEUU. Es una herramienta que permite el acceso a la información y posibilita el análisis de los datos, mejorando y optimizando el proceso de toma de decisiones de negocio y el desempeño de este.

### Diferencias entre las principales herramientas de Business Intelligence en el mercado

Una vez que se ha recopilado la información de las diferentes soluciones se exponen las diferencias entre las principales herramientas de BI:

Tabla 2. Diferencias entre principales herramientas BI

|  |  |  |  |
| --- | --- | --- | --- |
|  | **POWER BI** | **TABLEAU** | **QLIK** |
| Servicios | Power BI Desktop, Power BI y Power BI Mobile | Tableu Desktop, Tableu Server y Tableu Online | Qlik Sense, Qlik Analytics Platform, Qlik Core, QlikView |
| Integración | Microsoft, API | Varias fuentes | API, ETL |
| Compatibilidad | Windows, iOS y Android | Entornos web y dispositivos Android, Windows, Linux, iPhone/iPad y Mac | Entornos web y dispositivos Windows, iPhone/iPad y Mac |
| Precio | Asequible, versión empresarial costosa | Costoso | Opciones gratuitas |
| Alojamiento | En la nube (SaaS) o local (On Premise) | On Premise y Cloud | En la nube, On Premise o API abierta |
| Interfaz | Intuitiva | Menos intuitiva | Intuitiva |
| Informes | v | v | v |
| Cuadros de mando y reportes personalizables | v | v | v |
| Informes y análisis Ad hoc | v | v | v |
| Creación de informes interactivos | v | v | v |
| Cuadro de preguntas y respuestas | v |  | v |
| Indicadores de tendencias | v | v | v |
| Herramientas completas de informes y visualización de datos | v | v | v |
| Panel de navegación | v | v | v |
| Conjuntos de datos | v | v | v |
| Botones de ayuda y comentarios | v | v | v |
| Notificaciones de datos | v | v | v |
| Gestión de metadatos | v | v | v |
| Comentarios en el tablero |  | v |  |
| Alternar vista y arrastrar y soltar | v | v | v |
| Búsqueda natural | v |  | v |
| Plantillas de informes avanzados | v | v | v |
| Conectores predeterminados y personalizados | v | v | v |
| Colaboración segura en tiempo real | v | v | v |
| IA | v |  | v |

*Fuente: Elaboración propia*

Además de esta comparativa, se han analizado los datos del último Cuadrante Mágico de Gartner sobre plataformas analíticas y de Business Intelligence (ABI), informe anual del análisis del mercado que proporciona una visión objetiva de la posición de cada proveedor. Se pueden visualizar las principales herramientas de BI, concretamente ampliaremos Microsoft *Power BI* y *Qlik* en relación con sus novedades y últimas actualizaciones:

Figura 17. Cuadrante Mágico de Gartner

Escala de tiempo

Descripción generada automáticamente con confianza media

*Fuente:* (Gartner, 2023)

* ***Microsoft Power BI*** continúa posicionado como líder en el Cuadrante Mágico de Gartner para plataformas de análisis e inteligencia de negocios tras 16 años.

Power BI ha incorporado capacidad de seguimiento de métricas, que permite a los equipos colaborar visualmente para alinear objetivos y prioridades clave. También han presentado *Premium Gen 2*, una mejora que aumenta la relación precio/rendimiento y el valor gracias a sus capacidades de auto escalado. Por último, se han añadido *Marts* de datos de bajo código para simplificar el acceso a soluciones analíticas relacionales gestionadas de autoservicio.

* ***Qlik*** nombrado también líder en el Cuadrante Mágico de Gartner sobre plataformas de analítica y BI por décimo tercera vez consecutiva. Presenta novedades como la adquisición de *NodeGraph y Big Squid*, la presentación de *Qlik Forts* y el lanzamiento de *Qlik Application Automation***.** En 2011, Gartner creó una nueva subcategoría para herramientas de descubrimiento de datos, abanderando QlikView esta nueva categoría del Cuadrante Mágico de plataformas de Inteligencia de Negocios.

Ha integrado su adquisición de *Big Squid* para crear *Qlik AutoML*. Esta nueva herramienta proporciona previsión automática, agrupación y análisis de impulsores clave para científicos de datos. Además, ha presentado *Augmented Authoring*, una función que utiliza inteligencia artificial para crear cuadros de mando con entradas de lenguaje natural y diseños automatizados.

Finalmente, y por todo lo expuesto, ***QlikView*** ha sido la herramienta elegida ya que se centra en la exploración y la visualización de datos en tiempo real, lo que permite a los usuarios realizar análisis Ad hoc de manera más rápida. Además, facilita la conectividad con una amplia gama de fuentes de datos, por su motor asociativo, cognitivo y por tener una mayor flexibilidad en cuanto a los filtros y la visualización de información. Dispone de una interfaz intuitiva y además una versión gratuita para el desarrollo (García & Harmsen, 2013).

Figura 18. Arquitectura QlikView

Escala de tiempo

Descripción generada automáticamente

*Fuente:* (García & Harmsen, 2013)

# Objetivos y metodología de trabajo

En este capítulo se ha indicado el objetivo general del Trabajo de Fin de Grado, así como los objetivos específicos en los que se ha desagregado el propósito general.

## Objetivo general

Elaborar un Dashboard o informe de cuadro de mandos de los datos de la Universidad de Valladolid diseñado e implementado en la herramienta de Business Intelligence “QlikView”, que proporcione una solución de inteligencia empresarial que permita visualizar y analizar los datos de forma eficiente para optimizar la toma de decisiones.

## Objetivos específicos

El propósito general se ha dividido en objetivos intermedios mediante diferentes fases del proyecto. Los objetivos que se han detallado son específicos, medibles, alcanzables, relevantes y limitados en el tiempo.

* **Análisis de requisitos**

Realizar un estudio detallado de las especificaciones y requisitos del sistema, identificar a través de los *stakeholders* las necesidades y el alcance del proyecto.

* **Diseño y modelado de datos**

Diseñar y desarrollar los modelos de datos que han servido como base de datos para el análisis.

* **Desarrollo de Dashboards y visualizaciones**

Crear los Dashboards y las visualizaciones de datos que proporcionan información significativa a los usuarios finales.

* **Pruebas**

Implementar los Dashboards y realizar pruebas exhaustivas para garantizar su funcionalidad y rendimiento. Verificar que los datos que se obtienen son correctos y fiables.

* **Despliegue, usabilidad y mantenimiento**

Desplegar el informe de QlikView en el entorno de producción. Comprobar y verificar la usabilidad del informe.

* **Evaluación**

Verificar el cumplimiento de los requisitos de usuario, que la integración de los datos se ha realizado de forma correcta, que los datos son fiables y que el diseño es intuitivo y atractivo.

## Metodología de trabajo

Se ha definido la forma en la que se ha llevado a cabo el desarrollo.

### Metodología

Para este proyecto se ha utilizado un modelo de desarrollo en ***cascada***. Esta metodología es un enfoque secuencial para la gestión de proyectos en el que las fases se realizan en secuencia y cada fase debe completarse antes de pasar a la siguiente.

En proyectos más complejos y dinámicos, como los relacionados con el análisis de Big Data o con un número elevado de participantes podría ser más apropiado un enfoque más *ágil* o bien *iterativo incremental*.

Figura 19. Cronograma del proyecto

Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 20. Diagrama de Gantt

*Imagen que contiene Gráfico

Descripción generada automáticamente*

*Fuente: Elaboración propia*

### Tecnologías

La herramienta de business Intelligence que se ha utilizado es *QlikView*.

Para la extracción, transformación y carga, donde se incluye el proceso de recopilación, limpieza y transformación de datos a una estructura óptima desde diversas fuentes se ha realizado con *Power Query/Excel*. La gestión y carga de bases de datos mediante *Microsoft SQL Server Management Studio*.

El lenguaje de programación que ha utilizado es el *QlikView Scripting Language* (QSL). Es muy similar a *SQL* y se utiliza para extraer y transformar datos de diferentes fuentes, como bases de datos y archivos Excel. Además, QSL permite crear cálculos y expresiones para analizar los datos de manera más profunda.

# Solución del proyecto

## Análisis de Requisitos

Para poder dar solución al proyecto inicialmente se han identificado los requisitos del proyecto, se han definido los objetivos y el alcance del proyecto y se ha documentado los requisitos de datos y de funcionalidad del sistema.

### Visión General

El informe de los datos de la Universidad de Valladolid es un Dashboard destinado a los vicerrectores, rectores y directores del claustro de la Universidad que quieran llevar un control y seguimiento de la información a lo largo de los años de las matrículas, de las ayudas destinadas a las diferentes titulaciones, del personal de investigación, de los ingresos, los gastos y del cierre de los presupuestos.

El informe está dirigido sobre todo a personas que necesiten analizar los datos y cuyo objetivo sea optimizar la toma de decisiones. A diferencia de otros informes, al estar desarrollado en QlikView muestra mayor flexibilidad y versatilidad en la visualización de la información gracias a sus filtros y su motor asociativo y cognitivo.

### Alcance

Se han enumerado las principales características y capacidades del informe, que corresponden con las funcionalidades más importantes en la especificación de requisitos:

Tabla 3. Principales características del Dashboard

|  |  |  |
| --- | --- | --- |
| **Principales características** | | |
| **ID** | **Descripción** | **Prioridad** |
| **MF1** | El informe incluirá la visualización por año del número del personal de investigación por modalidad de estudio, tipo de personal, por dedicación, género, fecha de nacimiento y tipo de organismo | 2 |
| **MF2** | El informe mostrará el número de matrículas por curso académico, por titulación, nacionalidad, género, provincia y año de nacimiento | 1 |
| **MF3** | El Dashboard mostrará la evolución de la cantidad de las matrículas a lo largo de los cursos académicos | 1 |
| **MF4** | El informe mostrará el total de créditos matriculados y pendientes por titulación | 2 |
| **MF5** | El informe mostrará por año, código y descripción de la ayuda los importes concedidos y denegados de las ayudas | 1 |
| **MF6** | El sistema mostrará por año la cantidad de ayudas denegadas y concedidas y su evolución por género | 2 |
| **MF7** | El sistema mostrará por año el número de plazas de las ayudas por código y descripción y mostrará el total del importe del presupuesto inicial y ejecutado de la ayuda | 1 |
| **MF8** | El informe visualizará por año el total del importe de las matrículas por titulación, código y descripción de la ayuda | 1 |
| **MF9** | El Dashboard incluirá una tabla de ingresos por fecha, capítulo, artículo y concepto del ingreso con los importes totales segmentados por los importes del crédito inicial, modificaciones de ingreso, crédito total, derechos reconocidos y recaudación neta | 1 |
| **MF10** | El Dashboard incluirá una tabla de gastos por fecha, capítulo, artículo y concepto con los importes totales del crédito inicial, modificaciones de los gastos, crédito total, gastos comprometidos, obligaciones reconocidas y pagos netos | 1 |
| **MF11** | El informe mostrará la información, por fecha y por capítulo de ingresos y gastos, del importe de los créditos total e importes netos. | 2 |
| **MF12** | El informe mostrará de forma visual la información de los ingresos y de los gastos por fecha de cierre segmentado por el tipo de capítulo | 1 |
| **MF13** | El informe mostrará por fecha totalizadores de los créditos iniciales, modificaciones de los ingresos y gastos, derechos y obligaciones reconocidos, recaudación y pagos netos | 2 |
| **MF14** | El Dashboard mostrará por año el presupuesto detallado por el balance total del crédito asignado y de la recaudación total | 1 |
| **MF15** | El informe mostrará de forma visual por año de presupuesto e importe los créditos totales, la recaudación y los pagos netos | 1 |

*Fuente: Elaboración propia*

### Stakeholders

Se ha incluido una descripción de los *stakeholders* para facilitar la planificación y el desarrollo del Dashboard.

**S1** Vicerrector. Usuario básico que puede utilizar la información de las matrículas, los créditos, el personal de investigación y las ayudas para un control y visualización por año académico. Controla y muestra el balance de los ingresos y los gastos para la optimización de la toma de decisiones. Tiene relación con todas las características del Dashboard [MF1-MF15]

**S2** Jefe de proyecto. Líder del proyecto de desarrollo de la herramienta de Business Intelligence. Define las líneas de trabajo, la metodología y la tecnología a utilizar.

**S3** Desarrollador. Programadores y analistas que participan en el desarrollo de la aplicación.

**S4** Administrador. Persona que tiene control total sobre las bases de datos, los permisos y el acceso a las licencias de la aplicación.

**S5** Ingeniero de pruebas. Persona que realizan las pruebas de validación.

### Arquetipos

Se han detallado diferentes arquetipos de usuario, todos vicerrectores, para ayudar a comprender las necesidades y las características de los usuarios:

* Arquetipo 1. Vicerrector de investigación.

Escenario: El Vicerrector quiere filtrar por varios años y que le aparezcan los datos del personal de Investigación de la Universidad de Valladolid segmentados por la dedicación y el género.

¿Qué necesita? Un sistema que le permita filtrar por fecha y muestre la información.

¿Cómo lo ayuda el sistema? La aplicación debe tener un sistema que permita filtrar por fecha y facilite un acumulado por hombre y mujer del personal de investigación según el tiempo de dedicación.

* Arquetipo 2. Vicerrector de asuntos económicos.

Escenario: El Vicerrector quiere poder buscar por año de cierre de presupuesto el balance de la recaudación total de los ingresos y los gastos.

¿Qué necesita? Un informe que le permita buscar por año y muestre el cierre del presupuesto de ese año.

¿Cómo lo ayuda el sistema? La aplicación debe contar con un buscador que cuando se introduzca el año debe mostrar en una tabla el acumulado del balance de recaudación total que corresponde con la diferencia entre los ingresos y los pagos netos del mismo año de cierre.

### Historias de usuario

Se han especificado las historias de usuario que indican la clase de usuario beneficiada por el requisito:

Tabla 4. Historias de usuario

|  |  |  |  |
| --- | --- | --- | --- |
| **CASO DE USO** | | **HISTORIA DE USUARIO** | |
| **UC-01** | Visualizar dato | **UH-01** | Como Vicerrectorquiero poder visualizar los datos de los informes con un reporte personalizable y flexible que me permita tomar mejores decisiones |
| **UC-02** | Exportar informe | **UH-02** | Como Vicerrector necesito poder exportar la información de los reportes que permita analizar los datos de forma precisa |
| **UC-03** | Buscar información | **UH-03** | Como Vicerrector necesito poder buscar un dato concreto a través de un buscador |

*Fuente: Elaboración propia*

### Diagrama de casos de uso

Diagrama de casos de uso:

Diagrama

Descripción generada automáticamenteFigura 21. Diagrama de casos de uso

*Fuente: Elaboración propia*

Documentación de los casos de usos que se han especificado en el diagrama:

Tabla 5. UC-01 Visualizar Datos

|  |  |  |
| --- | --- | --- |
| **Identificador** | UC-01 | |
| **Nombre** | Visualizar datos | |
| **Autores** | Lydia Mayo/24 | |
| **Prioridad** | Alta | |
| **Criticidad** | Alta | |
| **Fuente** | Dashboard de la UVA | |
| **Responsable** | Vicerrector | |
| **Descripción** | El vicerrector quiere visualizar datos de los reportes del Dashboard | |
| **Evento de activación** | El Vicerrector indica que quiere ver los datos actualizados de un reporte | |
| **Actores** | Vicerrector | |
| **Precondiciones** | Pre1. | El usuario debe tener abierto el Dashboard |
| **Escenario principal** | **Paso** | **Acción** |
| 1 | El usuario selecciona un año del filtro |
| 2 | El sistema actualiza los datos de los informes |
| 3 | El sistema muestra la información actualizada |
| **Postcondiciones** | Post1. | Todo el Dashboard muestra la información en base a ese filtro |
| **Resultado** | El vicerrector visualiza los datos solicitados | |
| **Escenarios de excepción** | 1a | El filtro de selección no tiene relación con los campos del reporte |
| 1a.1 | El reporte no muestra datos |
| 1b | No hay datos para ese filtro de selección |
| 1b.1 | El reporte no muestra datos y el filtro aparece en gris |

*Fuente: Elaboración propia*

Tabla 6. UC-02 Exportar Informe

|  |  |  |
| --- | --- | --- |
| **Identificador** | UC-02 | |
| **Nombre** | Exportar informe | |
| **Autores** | Lydia Mayo/24 | |
| **Prioridad** | Alta | |
| **Criticidad** | Alta | |
| **Fuente** | Dashboard de la UVA | |
| **Responsable** | Vicerrector | |
| **Descripción** | El vicerrector quiere exportar los datos de una tabla o gráfico a un Excel | |
| **Evento de activación** | El Vicerrector solicita exportar los datos | |
| **Actores** | Vicerrector | |
| **Precondiciones** | Pre1. | El usuario debe tener abierto el Dashboard |
| Pre2. | El usuario está visualizando los datos que quiere exportar |
| **Escenario principal** | **Paso** | **Acción** |
| 1 | El usuario selecciona sobre el botón XL para enviar a Excel |
| 2 | El sistema exporta el informe en formato .xlsx |
| **Postcondiciones** | Post1. | El usuario dispone de un archivo en Excel con los datos |
| **Resultado** | El vicerrector visualiza los datos solicitados en Excel | |
| **Escenarios de excepción** | 1a | El usuario hace click sobre el botón derecho del ratón sobre la tabla |
| 1a.1 | El usuario hace click sobre la opción Enviar a Excel |
| 1a.2 | El sistema exporta el informe en formato *.xlsx* |

*Fuente: Elaboración propia*

Tabla 7. UC-03 Buscar Información

|  |  |  |
| --- | --- | --- |
| **Identificador** | UC-03 | |
| **Nombre** | Buscar información | |
| **Autores** | Lydia Mayo/24 | |
| **Prioridad** | Alta | |
| **Criticidad** | Alta | |
| **Fuente** | Dashboard de la UVA | |
| **Responsable** | Vicerrector | |
| **Descripción** | El vicerrector quiere visualizar los informes asociados a un dato en concreto | |
| **Evento de activación** | El Vicerrector indica el dato a buscar | |
| **Actores** | Vicerrector | |
| **Precondiciones** | Pre1. | El usuario debe tener abierto el Dashboard |
| **Escenario principal** | **Paso** | **Acción** |
| 1 | El usuario selecciona el cuadro de buscar |
| 2 | El usuario introduce el dato por el que filtrar |
| 3 | El sistema muestra la información actualizada |
| **Postcondiciones** | Post1. | Todo el Dashboard muestra la información en base a ese filtro |
| **Resultado** | El vicerrector visualiza los datos solicitados | |
| **Escenarios de excepción** | 2a | El filtro de selección no tiene relación con los campos del reporte |
| 2a.1 | El reporte no muestra datos |
| 2b | No hay datos para ese filtro de selección |
| 2b.1 | El reporte no muestra datos y el filtro aparece en gris |
| 2c | El sistema localiza más de un campo con ese valor, por lo que muestra todos los campos asociados a ese valor en concreto por el que puede asociar el filtro |
| 2c.1 | El usuario selecciona el dato por el que desea filtrar |
| 2c.2 | El sistema muestra la información actualizada |

*Fuente: Elaboración propia*

### Requisitos funcionales

Los requisitos funcionales que se han detallado para describir la funcionalidad del sistema:

Tabla 8. Requisitos Funcionales

|  |  |
| --- | --- |
| RF-01 | Mostrar información actualizada en el Dashboard |
| Requisitos Asociados | UC-01 Visualizar datos |
| UC-03 Buscar información |
| Descripción | Si un usuario ha iniciado el Dashboard, cuando seleccione un filtro o bien introduzca un valor de búsqueda el sistema deberá mostrar los datos actualizados de los diferentes reportes y gráficas |

|  |  |
| --- | --- |
| RF-02 | Extraer la información de los reportes y gráficos |
| Requisitos Asociados | UC-01 Visualizar datos |
| UC-02 Exportar informe a Excel |
| UC-03 Buscar información |
| Descripción | El sistema deberá permitir exportar los datos de las tablas y gráficos en formato Excel u otros formatos |
|  |  |
|  |  |
|  |  |
| RF-03 | Integrar datos |
| Requisitos Asociados | UC-01 Visualizar datos |
| UC-03 Buscar información |
| Descripción | El sistema deberá permitir integrar datos de diferentes fuentes |
|  |  |
|  |  |
|  |  |
| RF-04 | Mostrar información en tablas |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá permitir mostrar la información en tablas simples y pivotantes que faciliten Drill-Down, Slice and Dice y aplicar formulas avanzadas |
|  |  |
|  |  |
|  |  |
| RF-05 | Mostrar información en gráfico de tarta |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá permitir mostrar la información en gráfico de tarta del personal de administración, del número de matrículas por género |
|  |  |
|  |  |
|  |  |
| RF-06 | Mostrar información en gráfico combinado |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá permitir mostrar la información en gráfico combinado de los presupuestos |
|  |  |
|  |  |
|  |  |
| RF-07 | Mostrar información en gráfico de barras |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá permitir mostrar la información en gráfico de barras del número de personal de investigación, de las matrículas, de las ayudas por género y estado y de los ingresos y gastos por capítulo |
|  |  |
|  |  |
|  |  |
| RF-08 | Mostrar información en tablas simples |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá ser capaz de mostrar la información en tablas simples del número de matrículas por curso académico |
|  |  |
|  |  |
|  |  |
| RF-09 | Mostrar información en tablas pivotantes |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El sistema deberá ser capaz de mostrar la información en tablas pivotantes de los datos del personal de investigación, de las ayudas, de los ingresos, de los gastos, de las matrículas y presupuestos |
|  |  |
|  |  |
|  |  |
| RF-10 | Mostrar pestañas con datos |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El Dashboard deberá ser capaz de mostrar la información del informe en pestañas separadas con accesos directos a la información |
|  |  |
|  |  |
|  |  |
| RF-11 | Mostrar inicio |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El Dashboard deberá ser capaz de mostrar un menú de inicio que permita acceder a todas las pestañas con la información de los reportes |

*Fuente: Elaboración propia*

### Requisitos No funcionales

Los requisitos no funcionales que se han detallado para especificar las condiciones en las que debe trabajar el sistema:

Tabla 9. Requisitos No Funcionales

|  |  |
| --- | --- |
| RNF-01 | Mostrar información de forma eficiente |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El Dashboard deberá ser capaz de mostrar la información en pocos segundos |
|  |  |
|  |  |
|  |  |
| RNF-02 | Mostrar información de forma intuitiva |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El Dashboard deberá mostrar la información de forma gráfica e intuitiva de modo que cualquier usuario podrá acceder a la información |
|  |  |
|  |  |
|  |  |
| RNF-03 | Mostrar información de forma accesible |
| Requisitos Asociados | UC-01 Visualizar datos |
| Descripción | El Dashboard deberá contar con un diseño accesible para poder realizar un análisis de los datos |

*Fuente: Elaboración propia*

### Matriz de trazabilidad

Las relaciones de los requisitos funcionales y los de usuario se han representado mediante una matriz de trazabilidad:

Tabla 10. Matriz de Trazabilidad

|  |  |  |  |
| --- | --- | --- | --- |
|  | CASOS DE USO | | |
| REQUISITOS DE USUARIO | UC-01 Visualizar datos | UC-02 Exportar informe | UC-03 Buscar información |
| RF-01 Mostrar información actualizada en el Dashboard | X |  | X |
| RF-02 Extraer la información de los reportes y gráficos | X | X | X |
| RF-03 Integrar datos | X |  | X |
| RF-04 Mostrar información en tablas | X |  |  |
| RF-05 Mostrar información en gráfico de tarta | X |  |  |
| RF-06 Mostrar información en gráfico combinado | X |  |  |
| RF-07 Mostrar información en gráfico de barras | X |  |  |
| RF-08 Mostrar información en tablas simples | X |  |  |
| RF-09 Mostrar información en tablas pivotantes | X |  |  |
| RF-10 Mostrar pestañas con datos | X |  |  |
| RF-11 Mostrar inicio | X |  |  |

*Fuente: Elaboración propia*

## Diseño y modelado de datos

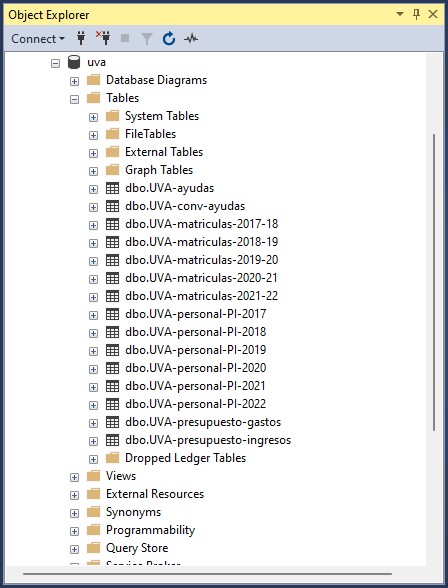
Se han identificado las fuentes de datos, se han tratado los datos mediante la extracción, transformación y carga de los datos y se han diseñado los modelos de datos que representan la estructura y las relaciones entre las entidades.

### Fuentes de datos

Los datos se han obtenido de la página web del Gobierno de España del Ministerio de la Transformación digital y de la función pública (Datos.gob.es, s.f.) que contiene un catálogo de datos abiertos. Las fuentes de datos son el catálogo de datos del Ministerio de Universidades (Ministerio de ciencia, innovación y universidades, s.f.) y los Dataset de UniversiData (UniversiData, s.f.).

Se han creado las bases de datos a través del gestor de Microsoft SQL Server Management Studio:

Figura 22. Tablas SQL



*Fuente: Elaboración propia*

Figura 23. BBDD



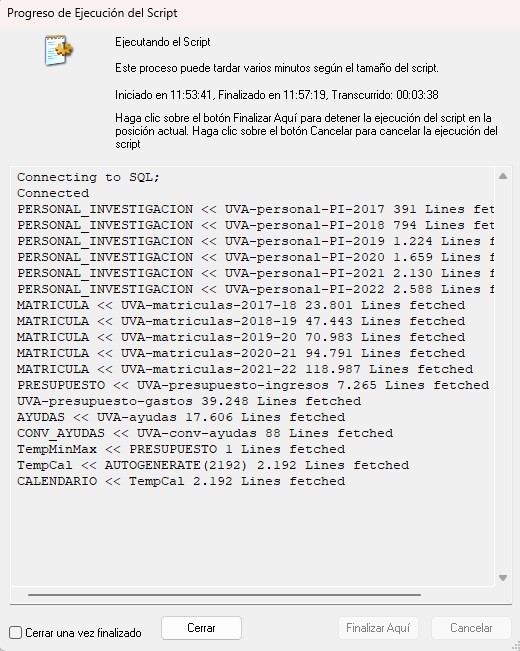
*Fuente: Elaboración propia*

### Extracción, Transformación, Carga de BBDD

Se han tratado los datos revisando las duplicidades y eliminando los datos de las columnas anonimizadas.

A través del script de QlikView se realizan las conexiones ODBC a SQL y se cargan las BBDD.

Figura 24. Proceso Ejecución Script

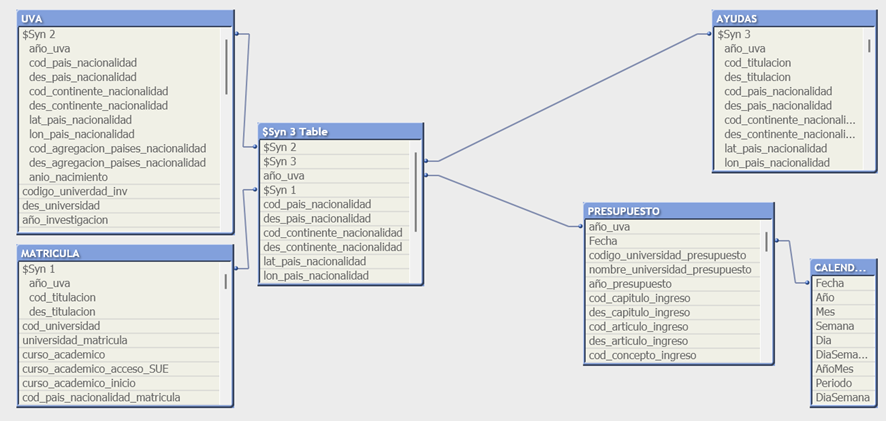


*Fuente: Elaboración propia*

### Diseño del modelo

Se ha procedido a renombran algunos campos en función de las relaciones que se han establecido para evitar la creación de tablas sintéticas, las cuales provocan ralentización en el documento e inexactitud de los datos.

Figura 25. Relaciones con Tablas Sintéticas



*Fuente: Elaboración propia*

Se han optimizan las relaciones entre las entidades de forma que la asociación se produce por un único campo clave común.

Figura 26. Diagrama QlikView

Imagen que contiene Texto

Descripción generada automáticamente

*Fuente: Elaboración propia*

## Diseño del Dashboard

Una vez que se han diseñado los modelos de datos que representan la estructura y las relaciones entre las entidades se ha procedido a realizar el diseño del Dashboard.

La página de inicio se ha desarrollado con botones que dirigen a cada una de las pestañas. Cada una de las pestañas también tienen acceso directo.

Figura 27. Diseño página Inicio

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

*Fuente: Elaboración propia*

Desde cada pestaña se puede realizar la selección de cualquiera de las demás pestañas o se puede seleccionar a la izquierda el botón que corresponde a la pestaña a la que queremos dirigirnos.

El botón que se corresponda con la pestaña seleccionada quedará marcado en otro color, mediante la función GetActiveSheetId() y variables definidas:

if(GetActiveSheetId()='Document\Personal de Investigación'*,$(vColorFondoActivo),$(vColorFondo*))

Cada una de las hojas del diseño del Dashboard:

Figura 28. Hoja Personal de Investigación

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 29. Hoja Matrículas

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 30. Hoja ayudas

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 31. Hoja Ingresos

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 32. Hoja Gastos

Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

Figura 33. Hoja Datos Capítulos

Tabla

Descripción generada automáticamente con confianza baja

*Fuente: Elaboración propia*

Cada uno de los filtros temporales permiten modificar el tiempo real la selección de los datos y las expresiones calculadas. Los filtros se aplican a los datos de todas las hojas donde haya relación. El buscador también permite hacer selecciones sobre los diferentes campos de la base de datos.

Figura 34. Hoja Presupuestos con filtros

Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

*Fuente: Elaboración propia*

## Desarrollo del Dashboard y Visualización

Como se ha definido en el punto anterior, el Dashboard cuenta con una pestaña de Inicio desde la que se puede hacer click sobre los diferentes botones que dirigen a las correspondientes pestañas, o bien se puede seleccionar directamente la pestaña sobre la que se quiere obtener la información actualizada.

Contiene además diferentes elementos fundamentales para el desarrollo, se han utilizado tablas simples, tablas pivotantes, gráficos de tarta, de barras y combinados para representar los datos y los requisitos del sistema:

### Tabla Simple

Es una tabla que puede contener dimensiones y expresiones calculadas. Una tabla simple es una tabla regular.

En la tabla *Matrículas* se ha definido como dimensión el campo correspondiente al año del curso académico y como expresión se contabiliza el número de matrículas académicas.

Figura 35. Tabla Simple

Tabla

Descripción generada automáticamente

*Fuente: Elaboración propia*

### Tabla Pivotante

Las tablas Pivotantes ofrecen más flexibilidad cuando trabajamos con múltiples dimensiones. Inicialmente muestra todas las dimensiones contraídas y sólo está visible la primera. Se puede con el botón derecho seleccionar la opción de expandir o contraer todo, o bien se puede utilizar el icono dispuesto al lado de cada celda de la dimensión.

En este ejemplo de tabla Pivotante se muestra como dimensiones el tipo de estudio, la titulación y el curso académico. Las expresiones que se calculan son el número de las matrículas, la cantidad de créditos matriculados y el número de los créditos pendientes. Además, se muestra un total de cada una de las expresiones por cada dimensión.

Figura 36. Tabla Pivotante contraída



*Fuente: Elaboración propia*

Figura 37. Tabla Pivotante expandida

****

*Fuente: Elaboración propia*

### Gráfico Combinado

Combina el gráfico de líneas con el gráfico de barras.

Para un gráfico con dimensión de los años del presupuesto y con expresiones que calculan el total del crédito de ingreso y de gasto de forma lineal, y la recaudación y los pagos netos con diagrama de barras, se muestra para una selección del filtro de los años comprendidos entre 2027 y 2022 los siguientes datos del presupuesto:

Figura 38. Gráfico Combinado

Gráfico

Descripción generada automáticamente

*Fuente: Elaboración propia*

### Gráfico Tarta

El gráfico de tarta, denominado así por su forma, permite establecer límites de dimensión en el caso de que disponga de muchos valores.

El gráfico en forma de Tarta está midiendo el número de personal de investigación que hay en función del año o años que se seleccionen en el filtro. Como dimensión se ha configurado el género y como expresión contabiliza el número en cada caso.

Figura 39. Gráfico Tarta

Gráfico, Gráfico circular

Descripción generada automáticamente

*Fuente: Elaboración propia*

### Gráfico Barras

El gráfico de barras establece una representación bidimensional donde el conjunto de datos se corresponde con las barras o rectángulos.

Las dimensiones son el año de presupuesto y el tipo de capítulo del ingreso y la expresión corresponde a la suma de la recaudación neta del ingreso. El filtro tiene aplicado varios años, pero se visibiliza y analiza mejor por año.

Figura 40. Gráfico Barras

Gráfico

Descripción generada automáticamente

*Fuente: Elaboración propia*

## Calendario Maestro

Para crear tablas, listas y gráficos se ha creado un calendario Maestro para poder completar las fechas que faltan en el filtro o selector. De esa forma cuando se utilice la dimensión fecha los datos que se muestren quedarán reflejados de forma uniforme y entendible.

En el script se coge la fecha de la tabla *PRESUPUESTO*, por lo que de esta variable se calcula la mínima fecha, la máxima y se definen las variables.

TempMinMax:  
  
**load**  
Min(Fecha) as MinDate,  
Max(Fecha) as MaxDate  
  
Resident PRESUPUESTO;  
  
**Let** vMinDate= Peek('MinDate',0,'TempMinMax');  
**Let** vMaxDate= Peek('MaxDate',0,'TempMinMax');  
**Let** vToday= ***$(vMaxDate)***;

Para crear el calendario Temporal se han utilizado las funciones de *Peek* y *AutoGenerate*, de esta forma se crean todas las fechas desde la variable mínima hasta la máxima.

TempCal:  
  
**Load**   
 Date(***$(vMinDate)*** +RowNo() -1) as TempDate  
AutoGenerate   
 ***$(vMaxDate)*** - ***$(vMinDate)*** +1;  
  
**drop** Table TempMinMax;

Para crear el calendario Maestro renombramos las fechas para crear el vínculo y se crean variables que pueden ser útiles como la de Año, Mes, Semana, Día, mediante funciones ya predefinidas en QlikView:

CALENDARIO:  
  
**Load**  
TempDate as Fecha,  
Year(TempDate) as Año,  
num(Month(TempDate),'00') as Mes,  
Week(TempDate) as Semana,  
Day(TempDate) as Dia,  
num(WeekDay(TempDate)) as DiaSemana1,  
year(TempDate)&num(Month(TempDate),'00') as AñoMes,  
Autonumber(year(TempDate)&num(Month(TempDate),'00')) as Periodo,  
WeekDay(TempDate) as DiaSemana  
  
  
Resident TempCal  
WHERE year(TempDate)>='2015'  
Order by TempDate ASC;  
  
**Drop** Table TempCal;

## Set Analysis

Herramienta muy potente en Qlikview “**{<[ ]>}”** que permite hacer selecciones, por ejemplo, en la pestaña *Ayudas,* el gráfico *Nº Ayudas,* permite contar el número de ayudas concedidas o denegadas según el género mediante las siguientes expresiones:

Count({<des\_estado={'Concedida'}>}cod\_genero\_ayuda)

Count({<des\_estado={'Denegada'}>}cod\_genero\_ayuda)

Figura 41. Gráfico Ayudas Universidad

Gráfico, Gráfico de barras

Descripción generada automáticamente

*Fuente: Elaboración propia*

En las dimensiones también se permite hacer tratamiento de los datos como es el caso del tratamiento del código de las convocatorias de ayuda si es nulo:

if (isNull(cod\_convocatoria\_ayuda),'Sin Código',cod\_convocatoria\_ayuda)

# Conclusiones y trabajo futuro

Una vez concluido el trabajo se han analizado las conclusiones que se han obtenido al conseguir cada uno de los puntos objetivos.

## Conclusiones del trabajo

En el *análisis de requisitos* se han detallado las especificaciones y los requisitos del sistema, objetivo que se ha conseguido tal y como hemos especificado en el punto 4.1 donde se han identificado los requisitos del sistema, el alcance, los requisitos de los datos junto con las métricas y la funcionalidad del sistema.

En el *diseño y modelado de datos* se han creado las bases de datos como se ha mostrado en el punto 4.2.1 donde se han identificado las fuentes de datos. Uno de los desafíos principales fue la integración de diversas fuentes de datos y poder asegurar la calidad y consistencia de la información, por ello, se han implementado soluciones de limpieza y normalización de datos. En el punto 4.2.2 se ha llevado a cabo la extracción, transformación y carga de las bases de datos y se han diseñado los modelos de datos y optimizado las relaciones, por lo que este objetivo también se ha alcanzado.

En el *desarrollo del Dashboard* se ha visualizado el informe de QlikView, en el apartado 4.3 se han mostrado alguna de las hojas, y en el punto 4.4 donde se muestran algunos elementos fundamentales del informe como tablas y gráficos. Por tanto, este objetivo también se ha conseguido.

Se han realizado *pruebas* exhaustivas para comprobar que los datos son correctos y fiables. Se han contrastado con los registros de las bases de datos y con los totales de cada una de las tablas para garantizar que son correctos y fiables. Las expresiones en las que se realizan cálculos también se han verificado de forma individualizada para garantizar la funcionalidad y la consistencia de los datos mediante consultas en SQL en las bases de datos.

Se ha verificado la usabilidad del informe con el propio funcionamiento del Dashboard y por ello se puede concluir también con el apartado de *Evaluación*, donde todos los objetivos se han conseguido.

Por todo ello, el objetivo general se puede dar por alcanzado con éxito:

*“Se ha creado un Dashboard o Informe de cuadro de mandos de los datos de la Universidad de Valladolid diseñado e implementado en la herramienta de Business Intelligence QlikView, que proporciona una solución de inteligencia empresarial que permite visualizar y analizar los datos de forma eficiente para optimizar la toma de decisiones”.*

El Dashboard ha proporcionado una herramienta centralizada que facilita el acceso rápido y eficiente a datos críticos permitiendo la mejora de la toma de decisiones basados en información actualizada y precisa.

La visualización clara y comprensible de los datos ha ayudado a identificar tendencias y patrones, lo que ha permitido tomar decisiones más informadas y proactivas.

## Líneas de trabajo futuro

Como posibles mejoras al Dashboard de los datos de la Universidad de Valladolid se han establecido las siguientes líneas de trabajo:

* Posibilidad de integrar nuevas bases de datos de otras Universidades nacionales e internacionales, o incluso con fuentes externas relevantes como datos gubernamentales.
* Integrar el informe con software de sistemas académicos que permitan mostrar información actualizada y en tiempo real.
* Integrar flujos de trabajo en tiempo real para reflejar el estado de la matriculación o eventos universitarios.
* Incluir tasas de predicción mediante inteligencia artificial. Integrar análisis predictivo y machine learning para proporcionar estimaciones, predicciones y recomendaciones proactivas.
* Crear un Dashboard que aprenda de las consultas más frecuentes de los usuarios, analizando el feedback de rendimiento y usabilidad, de forma que permita mejorar continuamente el Dashboard, asegurando que evolucione para satisfacer las necesidades de los usuarios.
* Implementar un control de acceso de usuarios en función de los perfiles y las licencias para asegurar que solo los usuarios autorizados puedan ver o modificar datos específicos.
* Crear una automatización para la actualización de la información de los datos de la Universidad de Valladolid para cursos académicos sucesivos.
* Establecer un menú de ayuda con documentación detallada y recursos como tutoriales, videos y guías de usuario que faciliten el uso del Dashboard, así como los diferentes usos de los filtros para visualizar la información y gestión de los reportes que permita la mejora de la toma de decisiones.

# Referencias bibliográficas

*Aitana*. (15 de 5 de 2023). Obtenido de Tecnología Innovación Compromiso: https://blog.aitana.es/2022/03/31/power-bi-y-qlik-mejores-soluciones-de-bi-segun-gartner/

*Amazon Web Service*. (s.f.). Obtenido de https://aws.amazon.com/es/what-is/olap/

Arias Pérez, J., & Aristazábal Botero, C. (20 de Enero de 2011). *Semestre Económico, Revista Científica*. Obtenido de https://revistas.udem.edu.co/index.php/economico/article/view/370/329

*Big Data International Campus*. (s.f.). Obtenido de https://www.campusbigdata.com/big-data-blog/item/148-diferencias-entre-big-data-business-analytics-y-business-intelligence

*BSA y The Software Alliance*. (2015). Obtenido de https://www.bsa.org/es/informes/por-que-son-tan-importantes-los-datos

Curto Díaz, J. (2016). *Introducción al business intelligence.* Barcelona: Editorial UOC.

*Dataladder.* (17 de 12 de 2021). Obtenido de https://dataladder.com/es/dimensiones-de-la-calidad-de-los-datos-10-metricas-que-debe-medir/#timeliness

*Datos.gob.es*. (s.f.). Obtenido de Gobierno de España. Ministerio para la Transformación Digital y de la Función Pública.: https://datos.gob.es/es

*Deloitte.* (s.f.). Obtenido de chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www2.deloitte.com/content/dam/Deloitte/uy/Documents/strategy/Calidad\_de\_datos.pdf

*Evotic*. (s.f.). Obtenido de https://evotic.es/business-intelligence-bi/power-bi-vs-qlikview/#:~:text=Power%20BI%20se%20basa%20en,hoc%20de%20manera%20m%C3%A1s%20r%C3%A1pida.

Fernandez, O. (9 de 10 de 2023). *AprenderBigData*. Obtenido de https://aprenderbigdata.com/herramientas-bi/

*Forrester Research.* (2008).

García, M. Á., & Harmsen, B. (2013). *QlikView 11 para Desarrolladores.* Birmingham-Mumbai: Packt Publishing Ltd.

*Gartner*. (Enero de 2023). Obtenido de https://www.gartner.es/es/metodologias/magic-quadrant

*Gartner Peer Insights*. (s.f.). Obtenido de https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms

Gonzalez Farran, X., Guitart, I., & Rodriguez, J. R. (2016). *¿Cómo planificar un proyecto de inteligencia de negocio?* Barcelona: UOC.

*IBM*. (s.f.). Obtenido de https://www.ibm.com/es-es/topics/business-intelligence

*IBM*. (s.f.). Obtenido de https://www.ibm.com/es-es/topics/olap#:~:text=OLAP%20(del%20ingl%C3%A9s%20Online%20Analytical,de%20datos%20unificado%20y%20centralizado.

*Introducción al Business Intelligence*. (s.f.). Obtenido de https://reader.digitalbooks.pro/content/preview/books/99484/book/OEBPS/ch\_2.html

*IQlik*. (s.f.). Obtenido de https://iqlik.wordpress.com/

Madrid, E. P. (2008). *Redalyc.* Obtenido de https://www.redalyc.org/articulo.oa?id=63501110

Magallanes, V. A. (6 de Abril de 2024). *Guia completa: Tipos de dimensiones en modelos dimensionales*. Obtenido de https://www.linkedin.com/pulse/guia-completa-tipos-de-dimensiones-en-modelos-juan-magallanes-j4xwf/

Marín, L. (9 de Noviembre de 2023). *Arbentia Empowering Consulting*. Obtenido de https://www.arbentia.com/blog/comparativa-de-herramientas-de-business-intelligence-microsoft-power-bi/

*Microsoft Power Bi*. (2024). Obtenido de https://powerbi.microsoft.com/en-us/blog/

Miguel, R. d. (27 de Diciembre de 2023). *Revista Byte*. Obtenido de https://revistabyte.es/comparativa/plataformas-de-business-intelligence-bi/

*Ministerio de ciencia, innovación y universidades*. (s.f.). Obtenido de www.universidades.gob.es

*PackPub*. (s.f.). Obtenido de https://www.packtpub.com/

*Pentaho*. (s.f.). Obtenido de https://www.hitachivantara.com/pentaho/pentaho-plus-platform.html?source=pentaho-redirect

*PowerData*. (5 de 3 de 2014). Obtenido de https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/368790/las-6-dimensiones-de-la-calidad-de-los-datos

*Qlik Community*. (s.f.). Obtenido de https://community.qlik.com/

QlikTech. (2024). *QlikView Help*. Obtenido de https://help.qlik.com/es-ES/

*Reddit*. (s.f.). Obtenido de https://www.reddit.com/r/BusinessIntelligence/

Redmond, S. (s.f.). *Qlik Tips*. Obtenido de https://www.qliktips.com/

Salesforce. (2024). *Tableau*. Obtenido de https://www.tableau.com/

TIC, E. d. (24 de Abril de 2023). *CLEFormación*. Obtenido de https://cleformacion.com/aplicaciones-olap-bases-de-datos-dimensionales/

*Transtats*. (s.f.). Obtenido de https://www.transtats.bts.gov/

Trejo, D. (Abril de 2020). *Data Governance literacy.* Obtenido de https://www.researchgate.net/publication/366214271\_Gobierno\_de\_datos\_conceptos\_basicos\_Data\_Governance\_literacy

Tunnell, S. (s.f.). *Qlikview Developer: Certificado Desarrollador QlikView*. Obtenido de https://www.udemy.com/course/qlikview-developer-certificado-desarollador-qlikview/learn/lecture/5820286?start=0#overview

*UniversiData*. (s.f.). Obtenido de El portal de los datos abiertos sobre la educación superior: https://www.universidata.es/universidades/universidad-de-valladolid-uva