

# **Business Intelligence Aplicada a Servicios Ciudadanos**

**Alumno: Alejandro Planelles Soler**

Grado de Ingeniería Informática

**Consultor:**

**Humberto Andrés Sanz**

Fecha de entrega: 13 de Enero del 2016

## **GNU Free Documentation License (GNU FDL)**

Copyright © ANY Alejandro Planelles Soler

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License"

## **Agradecimientos**

El presente Trabajo no solo simboliza un Proyecto Final de Grado, sino que se trata del colofón final a muchos años de dedicación y esfuerzo focalizados en concluir los estudios de Ingeniería Informática.

En este sentido, me gustaría aprovechar la ocasión para dar las gracias a todos aquellos que han contribuido, de manera directa o indirecta, a que me encuentre tan cerca de acabar la carrera. Sin su confianza y apoyo, no estaría a punto de culminar una etapa formativa especialmente relevante en mi vida.

En primer lugar, dar las gracias a mi tutor de Proyecto, Humberto Andrés Sanz, quien ha resuelto todas las dudas que le he ido planteando durante la evolución del Trabajo. Su asesoramiento, me ha mostrado el camino hacia nuevos enfoques y contenidos, que han ayudado de manera significativa a la mejora de la calidad global del Proyecto. Espero que él también haya quedado satisfecho con el resultado.

Agradecer profundamente la confianza que mis padres han depositado siempre en mí. Asimismo, dar el merecido reconocimiento al esfuerzo económico que han realizado durante las diferentes etapas académicas que han transcurrido a lo largo de mi vida. Sin esa base educativa, no podría haber alcanzado los propósitos que me he planteado posteriormente.

A mis hermanos, por dedicarme las palabras de ánimo necesarias para transmitirme la sensación de que es posible lograr cualquier objetivo que me plantee sin importar la dificultad que pueda comportar.

Gracias a mi mujer, por otorgarme la valentía necesaria para volver al camino de los estudios años después de haber abandonado dicha senda. Pedirle perdón por privarla de muchos momentos de ocio en familia, y dedicar gran parte de mi tiempo libre a estudiar o finalizar prácticas. Agradecerle también que me proporcionara la moral necesaria en los momentos en los que más la he precisado. Por todo ello y por mucho más, gracias.

Finalmente dar las gracias a mi hijo Julen, por ayudarnos a crecer como familia. A sus 3 meses, me ha enseñado a valorar la importancia de muchos aspectos que anteriormente no era capaz de apreciar. Gracias también por haber nacido justo a tiempo para poder ver cómo me gradúo, y así compartir juntos ese momento tan especial de mi vida. Espero en un futuro ser yo quien le vea alcanzar sus propósitos personales.

## FICHA DEL TRABAJO FINAL

<b>Título del Trabajo:</b>	<b>Business Intelligence Aplicada a Servicios Ciudadanos</b>
<b>Nombre del autor:</b>	<i>Alejandro Planelles Soler</i>
<b>Nombre del consultor:</b>	<i>Humberto Andrés Sanz</i>
<b>Fecha de entrega:</b>	<i>01/2016</i>
<b>Área del trabajo final:</b>	<i>Business Intelligence</i>
<b>Titulación:</b>	<i>Grado de Ingeniería Informática</i>
<b>Resumen del Trabajo</b>	
<p>En el presente documento, se describe la memoria final del Trabajo Final del Grado de Ingeniería Informática. El desarrollo del mismo, consiste en poner en práctica las técnicas y herramientas derivadas del <i>Business Intelligence</i> dentro del ámbito de los Servicios Ciudadanos. La intención, es la de dotar de una base sólida a empresas dedicadas a la prestación de servicios ciudadanos, con la finalidad de poder transmitir una mayor transparencia hacia el cliente, y sobretodo, conceder mayor capacidad a la dirección de las empresas en la toma de decisiones que conciernen al funcionamiento de las mismas.</p> <p>Inicialmente, se valora el contexto sobre el que se desea implementar el Proyecto, con el propósito de considerar sus particularidades, y determinar en función de estas, cual es el mejor enfoque para alcanzar los objetivos propuestos.</p>	

Se realiza un análisis del BI en general, para explicar tanto el concepto como sus diferentes componentes, y mostrar como su aplicación puede dar soporte, en diversos aspectos, al ámbito corporativo.

Una vez aclarado el concepto del *Business Intelligence*, se elabora una lista de requisitos que precisan ser incluidos en los objetivos del proyecto. Para cubrir dichos requisitos, se definen una serie de indicadores clave de desempeño (*KPI, Key Performance Indicators*).

Una vez las necesidades han sido reconocidas, se decide qué tipo de software es la alternativa más viable, y se trabaja en la búsqueda de un producto que pueda dar la cobertura que se plantea, previa elaboración de un amplio análisis de las opciones existentes como alternativas reales.

Mediante un riguroso estudio de la plataforma escogida, se pretenden aplicar técnicas de *Business Intelligence*, capaces de medir la calidad de los servicios prestados, presentar la información en gráficas, y ayudar con el aporte de conocimiento extraído del análisis de los datos iniciales.

### **Abstract**

*Contemporary business requires relevant data to make the appropriate, effective and reliable decisions.*

This project shows how Business Intelligence delivers a rich set of benefits that drive to significant improvements in companies that need gathering relevant internal as well as external data which are often hardly acquirable. The present performance is referring specifically to Citizens' Services. The targets set are to enable users to make informed business decisions quickly and confidently by providing the query and reporting tools they need to find, share, manage, publish and analyze information, in a framework of absolute honesty and transparency with respect to the customers.

In order to achieve a suitable solution for the particular kind of companies where the platform can be implemented, the project expounds a comparative analysis of different tools with the capacity of fulfill the expectations.

The project also aims at describing processes of building Business Intelligence (BI) systems. Taking a BI system specific into consideration, the author presents a suggested methodology for the tool creation and implementation in a specific context. The considerations are focused on the objectives and functional areas.

**Palabras clave**

Business Intelligence, Key Performance Indicators, ETL, Dashboard

# Índice

1. **Introducción**
  - 1.1 **Contexto y justificación del Trabajo**
  - 1.2 **Objetivos Trabajo**
  - 1.3 **Enfoque y método seguido**
  - 1.4 **Planificación del Trabajo**
  - 1.5 **Breve resumen de los productos finales**
  - 1.6 **Breve descripción del contenido de la memoria**
2. **Business Intelligence**
  - 2.1 **Contexto y definición**
  - 2.2 **Componentes del Business Intelligence**
  - 2.3 **Beneficios del Business Intelligence**
3. **Requisitos y KPIs**
  - 3.1 **Requisitos funcionales**
  - 3.2 **Requisitos técnicos**
  - 3.3 **Requisitos económicos**
  - 3.4 **KPIs**
    - 3.4.1 **KPI: Calidad limpieza viaria**
    - 3.4.2 **KPI: Número de quejas ciudadanas**
    - 3.4.3 **KPI: Tiempo de respuesta a las emergencias**
    - 3.4.4 **KPI: Formación del personal**
4. **Selección de la solución**
  - 4.1 **Software libre o software propietario**
  - 4.2 **Análisis de soluciones**
    - 4.2.1 **Pentaho**
    - 4.2.2 **Jasper**
    - 4.2.3 **Palo**
    - 4.2.4 **Tabla comparativa: Selección de la plataforma**
5. **Instalación y configuración de Pentaho**
  - 5.1 **Descarga y configuración de componentes básicos**
  - 5.2 **El origen de los datos (Data Sources)**
  - 5.3 **Resto de componentes y plugins**
6. **Creación de herramientas BI**
  - 6.1 **ETL y Data Integration. Spoon (Kettle)**
  - 6.2 **Cubo OLAP y análisis dimensional. Schema Workbench (Mondrian)**
    - 6.2.1 **Visualización del Cubo OLAP**
  - 6.3 **CDE Dashboard**
    - 6.3.1 **Layout Panel**
    - 6.3.2 **Datasources Panel**
    - 6.3.3 **Components Panel**
7. **Conclusiones y líneas de futuro**
8. **Referencias**

## 1. Introducción

### 1.1 Contexto y justificación del Trabajo

El contexto de recesión económica global en el actualmente la sociedad se halla inmersa, está propiciando la búsqueda de soluciones por parte de las organizaciones, con independencia del sector al que estas pertenezcan. La finalidad de dicha búsqueda, es tratar de descubrir el modo de contribuir a prolongar, o incluso mejorar, su posición en el mercado. Las soluciones a las que se hace referencia, tienden a estrechar la relación existente entre los conceptos de “empresa” y “tecnología”, con la idea de establecer sinergias que permitan obtener una ventaja, que pueda llegar a ser determinante, frente al resto de organizaciones dedicadas a llevar a cabo el mismo tipo de actividad. Todo ello se sitúa dentro de un marco basado en la competencia feroz, donde las exigencias por parte de los clientes son cada vez mayores, y sin embargo, los presupuestos más ajustados que en épocas anteriores.

Ante esta situación, se ha originado una importante reacceleración del cambio tecnológico, investigando nuevas maneras de satisfacer el consumo mediante el uso de los datos como materia prima, y el alineamiento entre las TIC y el negocio como medio para alcanzar los objetivos planteados por parte de la dirección de las empresas.

Motivado por la creciente cantidad de fuentes de obtención de datos, que se suman a los silos de información ya existentes en las empresas, se presenta un panorama en el que se hace indispensable implementar un conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento. Frente a esta realidad, emergen las tecnologías asociadas a la inteligencia de negocio, también conocidas como *Business Intelligence*, con la voluntad de hacer posible que las organizaciones puedan mejorar notablemente su productividad, ya que colabora con la toma de decisiones, por lo que en la actualidad representa una de las tácticas estratégicas más importantes para lograr un verdadero potencial competitivo.

Estas tecnologías, consisten en el empleo de los principales componentes de orígenes de datos existentes en Business Intelligence (Datamart y Datawarehouse), para operar sobre alguno de los productos que mayor valor puede aportar a la organización (Cuadro de Mando Integral, Sistemas de Soporte a la Decisión y Sistemas de Información Ejecutiva).

Si extrapolamos lo anteriormente explicado, a la situación del sector de los servicios ciudadanos, en el que se desarrolla la actividad profesional del autor del presente Proyecto, se detecta la necesidad de establecer cambios en muchos aspectos que tiempo atrás era posible eludir. Se debe indicar que dichos servicios son generalmente contratados por parte de gobiernos municipales. Las circunstancias existentes durante periodos anteriores, permitían que muchos problemas que ahora conciernen a los Ayuntamientos, pudieran ser obviados, ya que se disponía de presupuestos más elevados que los actualmente destinados a las administraciones locales. En consecuencia, para conseguir los niveles de servicio deseados, simplemente había que incrementar el susodicho servicio, sin reparar demasiado en los gastos de personal o recursos que esta acción comportaba.

La situación actual, no se presta con facilidad a este tipo de incrementos presupuestarios, aunque pueda existir una notable necesidad de mejora sobre aspectos vinculados a los diferentes servicios que se proporcionan a los municipios. En el caso excepcional de que se decida llevar a cabo un aumento del valor inicial pactado con la empresa concesionaria, este debe estar perfectamente meditado y justificado, y por tanto, ser considerado de vital importancia. A esta dificultad, hay que añadir las elevadas exigencias de rendimiento por parte del cliente, los Ayuntamientos en este caso, en ocasiones tan elevadas, que si no se gestionan con extremado rigor pueden derivar en resultados donde el beneficio alcance un valor negativo.

Fruto del análisis realizado sobre el contexto en los párrafos anteriores, se aborda el punto de la justificación del proyecto. Dicha justificación se considera evidente teniendo en cuenta los factores anteriormente señalados.

Se trata por tanto de desarrollar la implementación de una serie de herramientas basadas en el *Business Intelligence* que permitan, por una parte y a nivel externo, demostrar al cliente que las tareas que se llevan a cabo desde la organización son las acordadas en el pliego de condiciones firmado por ambas partes durante la negociación del mismo. De este modo, se asegura la transparencia entre empresa y cliente. Por otro lado, a nivel interno, proporcionar la posibilidad de realizar análisis exhaustivos sobre los sistemas de información transaccional vigentes, para que mediante la selección los correctos indicadores (*KPIs*) se puedan mejorar las prestaciones del servicio, sin necesidad de invertir más recursos de los estrictamente necesarios. Por tanto, la tecnología *Business Intelligence* permitirá extraer mayor cantidad y calidad de información, haciendo un apropiado uso de los datos, lo que supondrá un valor añadido para la empresa. A su vez, proporcionará un acceso más ágil a la información, que podrá ser utilizada para evaluar la posibilidad de volver a configurar servicios, favoreciendo la eficiencia de estos.

En definitiva, el proyecto constará de diversas partes con planteamientos muy distintos, que incluyen materia de carácter completamente teórico, pero también puntos en los que se tratan cuestiones especialmente de tipo técnico. En primer lugar se debe efectuar una introducción focalizada a aclarar en qué consiste el *Business Intelligence* (BI), y por tanto, en esta parte se describe de una manera comprensible los aspectos más interesantes que abarca la inteligencia de negocio. También se explica la aportación de este tipo de herramientas al mundo empresarial, así como los principales elementos que lo componen. Después de realizar un análisis comparativo sobre los dos grandes grupos de software que coexisten actualmente, el propietario y el libre, para decantarse por el que sea considerado más conveniente para el entorno sobre el que se plantea implantar la solución. Una vez decantados por uno de estos dos modelos, se valoran los principales productos disponibles, para posteriormente seleccionar el que mejor se adapte al cumplimiento de las expectativas del tipo de compañías sobre el que se desea implementar el BI.

Finalmente, se llevará a cabo la elaboración de la solución de inteligencia de negocio, teniendo en cuenta que indicadores son los más adecuados para el logro de los objetivos.

## 1.2 Objetivos del Trabajo

Los objetivos que se exponen a continuación, pertenecen por una parte a las metas de desarrollo personal vinculado al fruto de la realización del Trabajo, y por otro lado, a los objetivos que envuelven el proyecto por su propia naturaleza.

Desde una perspectiva personal, los objetivos primordiales que se pretenden alcanzar con la elaboración del actual Trabajo, son:

- El conocimiento, sin entrar en excesivo detalle, de algunas de las diferentes herramientas disponibles actualmente en el paradigma del *Business Intelligence*, para decidir cuál es la que mejor se adapta al caso concreto sobre el que será desarrollado el proyecto. Este objetivo, y los puntos asociados al mismo, serán considerados de tipo teórico.
  
- El entendimiento y correcto manejo del producto seleccionado. En este caso, se trata de profundizar en el uso del software de BI seleccionado, con el fin de comprender que posibilidades ofrece, y poder explicar de manera clara cada una de ellas. Esta meta es teórica, aunque en según qué fase, puede adquirir un carácter más práctico.
  
- Aprender a reconocer los *KPIs* más interesantes en función de las metas propuestas, desestimando aquellos que no aporten información relevante, y por tanto, focalizando el proyecto hacia los datos que posibiliten el logro de conocimiento. Muy importante en este sentido la investigación de proyectos de naturaleza similar para aprovechar los indicadores que puedan ser relevantes, y evitar aquellos que no aporten suficiente información.

- Adquirir las nociones suficientes para poder formar parte del desarrollo de un proyecto de implantación completo, partiendo de un producto que utilice herramientas de inteligencia de negocio.

Planteado desde el punto de vista del propio proyecto, los objetivos incluyen esencialmente:

- Estudio de las dos tipologías principales de software existente actualmente, libre y propietario, dentro del universo del *Business Intelligence*, con la finalidad de decidir cual se ajusta más a las necesidades del proyecto que se pretende desarrollar.

- Análisis de los principales productos disponibles para el tipo de software que haya sido seleccionado previamente. En este punto, se decreta cual va a ser el producto con el que se elaborará la parte teórica del proyecto.

- Valorar que indicadores (*KPIs*) deben ser tenidos en consideración. Para ello, es fundamental el estudio de los diferentes *KPIs* seleccionados en empresas de servicios, y como estos han podido conducir al éxito o fracaso de un proyecto BI.

- Implementar la inteligencia de negocios, dentro de organizaciones dedicadas a la prestación de servicios. Deben ser tenidos en cuenta los diferentes factores que envuelven el desarrollo de un proyecto de este tipo.

### **1.3 Enfoque y método seguido**

La planificación y desarrollo de los proyectos vinculados a las TIC, han madurado considerablemente en los últimos tiempos. Esto ha propiciado el aprendizaje y corrección de los errores que se cometían con frecuencia durante la implantación de proyectos en las organizaciones.

Mediante la recopilación de diversas lecciones ya aprendidas, y la identificación de las que pueden ser consideradas buenas prácticas, es factible obtener resultados óptimos en proyectos de tipo tecnológico.

Actualmente, se ofrecen diferentes metodologías que pueden ser utilizadas como guía para administrar y ejecutar proyectos, en función de las metas que previamente hayan sido definidas. Gracias a estas metodologías, es posible determinar el orden y los pasos necesarios a seguir para alcanzar de una manera eficaz los objetivos.

En este sentido, se debe saber que existen múltiples tendencias, filosofías, metodologías y herramientas que pueden ser aplicables en función del contexto del proyecto, y de la organización sobre la que se desea desarrollar.

Sin entrar en detalle, a continuación se dan a conocer algunas las metodologías más eficaces que existen actualmente, y cuál de ellas ha sido considerada, por parte del autor, como la más acertada para llevar a la práctica durante el desarrollo del presente Trabajo.

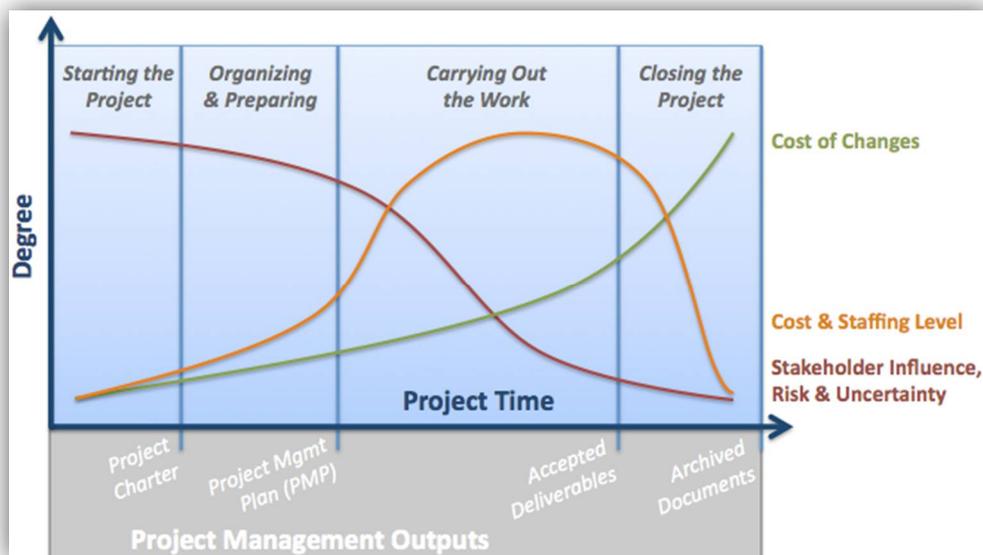
En primer lugar, se debe diferenciar cual es la finalidad de la aplicación de la metodología. Por un lado, si se tiene la intención de llevar a cabo la adecuada gestión del proyecto y la prestación de los servicios, las dos alternativas más aceptadas son las prácticas y recomendaciones del «PMBOK» e «ITIL» respectivamente. Estas metodologías están claramente focalizadas a la gestión. En cambio, si lo que se pretende se encuentra dirigido principalmente al desarrollo del Software, algunas de las opciones más interesantes son la «Métrica Versión 3» para la realización de grandes desarrollos y «Kanban» y/o «Scrum» como metodologías ágiles que precisen de algunos cambios evolutivos, o desarrollos y actuaciones de mantenimiento.

En el caso particular del actual Trabajo, se seguirán algunas de las reconocidas como buenas prácticas recogidas en el PMBOK (*Project Management Body of Knowledge*), es decir, una serie de normas voluntarias que marcarán las pautas a seguir durante el desarrollo del proyecto.

Señalar que PMBOK se constituye como estándar de administración de proyectos, y se encarga, entre otras cosas, de determinar cómo se debe administrar el alcance de un proyecto, elaborar un presupuesto y cómo gestionarlo, cómo hacer una evaluación y gestión de riesgos.

Según la Guía PMBOK, todos los proyectos, sin importar cuán pequeños o grandes sean, o el nivel de complejidad que comporten, pueden configurarse dentro de la siguiente estructura del ciclo de vida: Inicio, Organización y preparación, Ejecución del trabajo y Cierre.

A continuación, se muestra una figura que incluye estas fases, y el impacto a nivel de costes existente en cada una de las fases.



En el caso del proyecto concreto en que se desarrolla este Trabajo, en cada una de las fases se llevan a cabo las tareas que se especifican a continuación:

- Inicio: Determinar el área sobre la que se desea elaborar el Trabajo. En este punto, se deben tener en cuenta aspectos como que ámbitos se posee un mayor dominio, o que materia goza de un mayor interés desde el punto de vista del autor del proyecto. Es importante formarse una idea lo más precisa posible, que sirva para el desarrollo de las fases posteriores.

En el presente estudio, la propuesta inicial es la implantación de algunas de las tecnologías BI en empresas relacionadas con los servicios ciudadanos. Debido a que se trata de entornos relacionados la actividad laboral llevada a cabo por el autor del Trabajo, se parte de unos conocimientos sobre los procesos y los tipos de fuentes de datos. También se reconocen que puntos pueden ser mejorados mediante la aplicación de inteligencia de negocio.

- Organización y preparación: Enlazado con la fase previa, y con una idea ya correctamente establecida, se trata de valorar que tareas y en qué orden se tienen que efectuar durante el progreso del proyecto.

Se confecciona un diagrama de *Gantt* que contempla cada una de las principales partes del proyecto. En primer lugar, el análisis que conduce a la subsiguiente valoración de los productos asociados a tecnologías BI, con la meta de determinar cuál es el que mejor se adapta a los requerimientos de la empresa. Una vez tomada la decisión, es importante dedicar un tiempo a la exploración del software escogido, para adquirir un nivel lo suficientemente amplio del funcionamiento del mismo. Paralelamente, se van precisando los indicadores que pueden aportar más información. Después, se confeccionan las herramientas de BI y se hacen las pruebas pertinentes para detectar errores que puedan ser producto de los indicadores seleccionados. Finalmente, y cuando el funcionamiento de las herramientas es el esperado, se presentan los resultados. Como todo proyecto, la revisión periódica puede ayudar a mejorar el sistema, dando lugar a versiones más actualizadas y mejor adaptadas a las circunstancias del momento.

- Ejecución del trabajo: Tal y como se puede observar en la figura anteriormente expuesta, se trata de la fase más extensa del proyecto. Depende directamente de la etapa anterior.

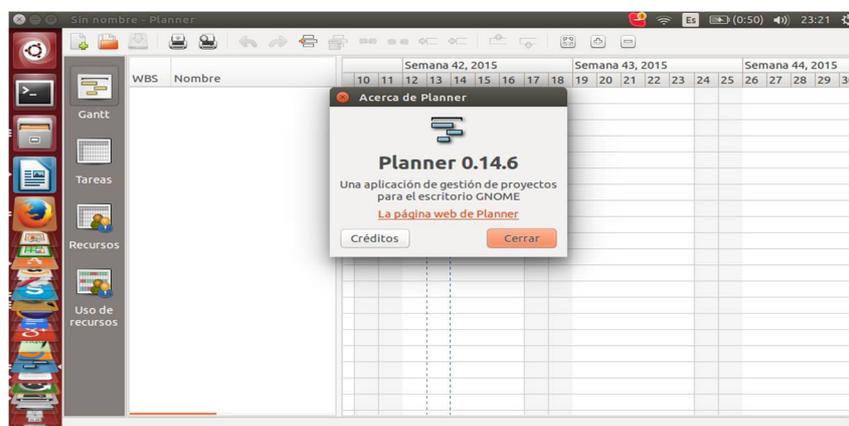
Si se ha organizado correctamente el Trabajo, solo hay que seguir las pautas establecidas para cada momento concreto. En definitiva, consiste en ejecutar las tareas especificadas durante la fase de «Organización y preparación».

- Cierre del proyecto: Presentar un trabajo completo, que cumpla con los objetivos planteados en la fase inicial.

#### 1.4 Planificación del Trabajo

Con la finalidad de crear una planificación precisa del proyecto, se emplea la elaboración del diagrama de *Gantt*. El diagrama de *Gantt*, o gráfica de *Gantt* es una herramienta gráfica cuyo objetivo es mostrar el tiempo de dedicación previsto para diferentes tareas o actividades a lo largo de un tiempo total determinado. A pesar de que, en principio, el diagrama de *Gantt* no indica las relaciones existentes entre actividades, la posición de cada tarea a lo largo del tiempo hace que se puedan identificar dichas relaciones e interdependencias.

El software escogido para realizar el diagrama de *Gantt* es «Planner». A continuación se muestra una imagen de la interfaz de la aplicación:



La decisión ha sido tomada tras valorar diferentes opciones que han sido desarrolladas para realizar diagramas de *Gantt* de manera sencilla e intuitiva en Linux.

«Planner» es una herramienta sencilla, que permite planear, programar y realizar el seguimiento de proyectos. Es una aplicación GTK+ escrita en C y licenciada bajo GPL.

Como información adicional, comentar que «Planner» fue creado originalmente por *Richard Hult* y *Mikael Hallendal*. En la actualidad está siendo desarrollado por GNOME.

El programa permite:

- Almacenamiento en XML o BD *Postgresql*
- Gestión de calendarios
- Gestión de recursos
- Seguimiento del avance del proyecto
- Enlazar tareas
- Exportación a diferentes formatos (PDF, HTML)

Destacar que siguiendo la filosofía de la guía PMBOK, las tareas se muestran en waterfall (cascada). En un desarrollo de este tipo, se dispone de una fecha final de desarrollo medianamente realista porque está completamente planificado, teniendo siempre presente que cualquier desvío en las tareas retrasa la salida porque no se dispone de un producto entregable. Esto se pone de manifiesto cuando se observa que la mayor parte de las tareas, parten de otra distinta que ha sido realizada previamente.

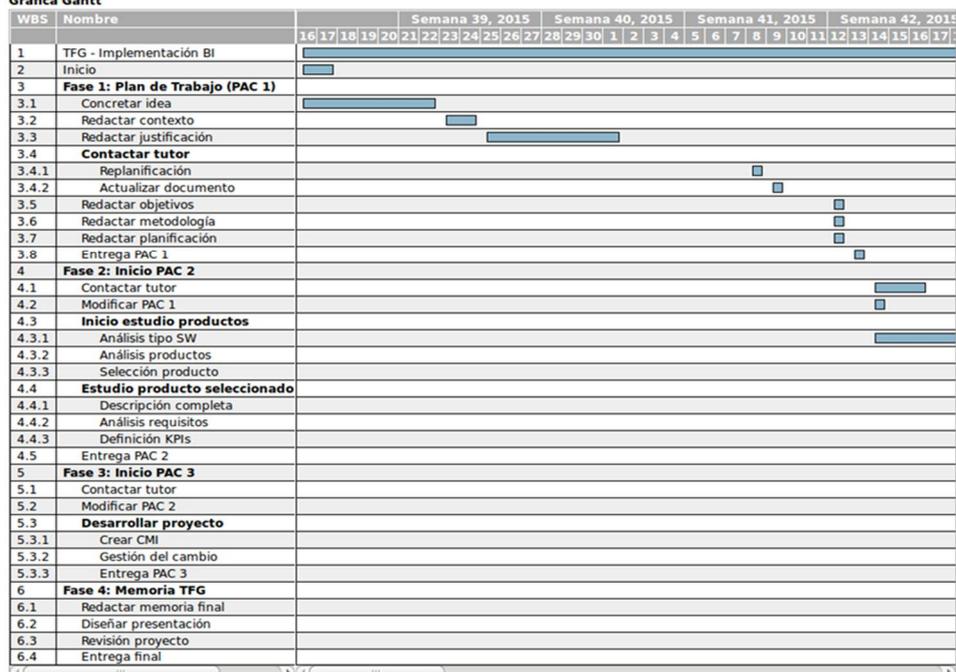
A continuación, se muestra una imagen completa en la que aparece el diagrama de Gantt del presente proyecto. Para mostrarlo por completo, ha sido necesaria la conversión del archivo a formato HTML, y posteriormente se ha capturado con *Fireshot*, un complemento de *Firefox Mozilla* que permite ver páginas web enteras.

Debido a la amplitud del archivo, es complicada su visualización. Se han incorporado dos imágenes adicionales que ayudan a distinguir con mayor facilidad la distribución temporal de las tareas.

**TFG - Implementación de un CMI**

**Empresa:** UOC  
**Encargado:** Alejandro Planelles Soler  
**Inicio:** 16 septiembre, 2015  
**Fin:** 13 enero, 2016  
**Fecha de informe:** 13 octubre, 2015

**Gráfica Gantt**



**Tareas**

WBS	Nombre	Inicio	Fin	Trabajo	Terminado	Coste	Asignado a
1	TFG - Implementación BI	Sep 16	Jan 13	86d	0%		
2	Inicio	Sep 16	Sep 17	2d	0%		
3	<b>Fase 1: Plan de Trabajo (PAC 1)</b>	Sep 16	Oct 13	18d			
3.1	Concretar idea	Sep 16	Sep 22	5d	0%		
3.2	Redactar contexto	Sep 23	Sep 24	2d	0%		
3.3	Redactar justificación	Sep 25	Oct 1	5d	0%		
3.4	<b>Contactar tutor</b>	Oct 8	Oct 9	2d			
3.4.1	Replanificación	Oct 8	Oct 8	1d	0%		
3.4.2	Actualizar documento	Oct 9	Oct 9	1d	0%		
3.5	Redactar objetivos	Oct 12	Oct 12	1d	0%		
3.6	Redactar metodología	Oct 12	Oct 12	1d	0%		
3.7	Redactar planificación	Oct 12	Oct 12	1d	0%		
3.8	Entrega PAC 1	Oct 13	Oct 13	1d	0%		
4	<b>Fase 2: Inicio PAC 2</b>	Oct 14	Nov 11	25d			
4.1	Contactar tutor	Oct 14	Oct 16	3d	0%		
4.2	Modificar PAC 1	Oct 14	Oct 14	1d	0%		
4.3	<b>Inicio estudio productos</b>	Oct 14	Oct 28	11d			
4.3.1	Análisis tipo SW	Oct 14	Oct 20	5d	0%		
4.3.2	Análisis productos	Oct 21	Oct 27	5d	0%		
4.3.3	Selección producto	Oct 28	Oct 28	1d	0%		
4.4	<b>Estudio producto seleccionado</b>	Oct 29	Nov 10	9d			
4.4.1	Descripción completa	Oct 29	Nov 2	3d	0%		
4.4.2	Análisis requisitos	Nov 3	Nov 6	4d	0%		
4.4.3	Definición KPIs	Nov 9	Nov 10	2d	0%		
4.5	Entrega PAC 2	Nov 11	Nov 11	1d	0%		
5	<b>Fase 3: Inicio PAC 3</b>	Nov 12	Dec 9	23d			
5.1	Contactar tutor	Nov 12	Nov 12	1d	0%		
5.2	Modificar PAC 2	Nov 12	Nov 12	1d	0%		
5.3	<b>Desarrollar proyecto</b>	Nov 12	Dec 9	21d			
5.3.1	Crear CMI	Nov 12	Dec 8	19d	0%		
5.3.2	Gestión del cambio	Nov 12	Nov 12	1d	0%		
5.3.3	Entrega PAC 3	Dec 9	Dec 9	1d	0%		
6	<b>Fase 4: Memoria TFG</b>	Dec 10	Jan 13	40d			
6.1	Redactar memoria final	Dec 10	Dec 30	15d	0%		
6.2	Diseñar presentación	Dec 10	Jan 6	20d	0%		
6.3	Revisión proyecto	Jan 7	Jan 12	4d	0%		
6.4	Entrega final	Jan 13	Jan 13	1d	0%		

**Recursos**

Nombre	Nombre corto	Tipo	Grupo	Correo-e	Coste

## 1.5 Breve resumen de los productos finales

Acto seguido a la valoración sobre el impacto que las herramientas de *Business Intelligence* pueden tener aplicado en el negocio de los servicios ciudadanos, la conclusión que se obtiene, es que el uso de inteligencia de negocio puede ayudar sustancialmente a la obtención de información representativa, y ser presentada de una manera sencilla, tanto al cliente al que se le prestan los servicios contratados, como a los responsables en la toma de decisiones de la propia empresa.

Se ha determinado que la implantación de herramientas basadas en la inteligencia empresarial, es el modo más adecuado para poder llevar a cabo la aportación de conocimiento a las compañías. Estas herramientas, contienen una serie de indicadores (*KPIs*), que permiten mostrar la información de manera segura y fiable en tiempo real, sobre aquellos parámetros que presenten un interés significativo para el sector.

## 1.6 Breve descripción del contenido de la memoria

Paralelamente a la evolución del Trabajo, se describen las fases que se van realizando hasta llegar a la implementación de las diversas herramientas de BI.

En el primer punto, se ha argumentado el contexto, desde una perspectiva genérica, que justifica la selección de la materia sobre la que se realizará el Trabajo. Asimismo, se hace referencia a los objetivos generales y a la metodología empleada para materializarlos. Además, se muestra el diagrama de *Gantt*, en el que se precisa la planificación temporal del proyecto.

El segundo punto, considerado fundamentalmente teórico, explica en que consiste el *Business Intelligence*, y sobre que entornos es adecuado su puesta en práctica. En este apartado, el contexto que se expone es específico, y se ajusta a la materia que concierne al ámbito del proyecto que se pretende desarrollar.

Es en esta fase, en la que también se hace alusión a los principales componentes del *Business Intelligence*, y se explica cada uno de ellos, sin entrar en un nivel de detalle excesivo. Posteriormente, se citan los tipos de beneficios que aportan las soluciones de BI, y el impacto específico que tiene su aplicación sobre el caso estudiado.

En el tercer punto, se hace un análisis de requisitos, tanto técnicos como funcionales y económicos. A partir de este, es necesario definir los KPIs que mayor información puedan aportar a la empresa. Este punto se considera crítico, ya que en función de la selección, se puede, o no, conseguir el propósito inicial de la implementación de inteligencia de negocio en las compañías dedicadas a ofrecer servicios ciudadanos.

El cuarto punto, consiste en decidir cuál es la solución más apropiada para llevar a cabo la parte de implementación del proyecto. Con esta finalidad, se elabora un análisis previo. Dicho análisis, compara en primer lugar, y desde una óptica global, las circunstancias que hacen que sea más apropiado el decantarse por un tipo de software libre o propietario, teniendo en cuenta el contexto sobre el que se desea desarrollar la implementación. Después, las diferentes soluciones candidatas son enumeradas, y se procede a la valoración minuciosa de cada una de ellas. Cuando se toma la decisión considerada más acertada, se procede a mostrar aspectos de la instalación y configuración de la misma. En conclusión, el cuarto punto presenta algunos matices relacionados con la parte práctica del Trabajo.

El quinto punto hace referencia a aspectos puramente técnicos. Se procede a la instalación, configuración y desarrollo de la solución. La idea es mostrar mediante explicaciones y alguna captura de pantalla, como se trabaja desde el inicio hasta el desarrollo de las herramientas sobre la plataforma seleccionada.

El sexto punto, presenta las herramientas de inteligencia de negocio. Este punto, como el anterior, es puramente técnico, y muestra algunos ejemplos de resultados obtenidos tras las diferentes pruebas llevadas a cabo.

## 2. Business Intelligence

### 2.1 Contexto y definición

En la actualidad, las empresas manejan un flujo de información el cual era inconcebible hace apenas unos años. La explosión de las redes sociales, las *smarts cities*, el *big data*, los dispositivos móviles o los sensores, entre otras muchas y variadas fuentes de datos, brindan nuevas oportunidades a las organizaciones. Este incremento exponencial del volumen de datos generado continuamente, hace imprescindible el uso de sistemas que sean capaces de analizarlos y convertirlos en información útil.

Es en este contexto es donde entra en juego uno de los conceptos más mencionados últimamente en el área de TI: *Business Intelligence*. La inteligencia de negocio, hace referencia a la práctica de un conjunto de herramientas que pueden ayudar a las corporaciones a adquirir un mejor entendimiento sobre ellas mismas. Esto es debido, gracias a la capacidad de explotar su información, con la intención de poder manipularlos de una manera más sencilla y entender el objetivo del negocio. También favorece la predicción de sucesos, lo que facilita el planteamiento de futuros escenarios, y por tanto, la correcta toma de decisiones.

Según *Gartner*, consultora internacional especializada en Tecnologías de Información: *"BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un Datawarehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones"*.

En resumen, BI es un conjunto de metodologías, aplicaciones y tecnologías, dedicadas a reunir, depurar y transformar datos de sistemas transaccionales e información desestructurada (tanto interna como externa a la organización), con el propósito de aportar inteligencia a los procesos organizativos, dando soporte a la toma de decisiones.

## 2.2 Componentes del Business Intelligence

### - Las fuentes de información.

Internas:

- Sistemas operacionales o transaccionales, incluyendo las aplicaciones desarrolladas a medida.
- Sistemas de información departamental, principalmente presentada en hojas de cálculo.

Dichas fuentes de información, serán las que alimentarán el *Datawarehouse*.

Externas:

- Principalmente estudios de mercado o extraída de fuentes ajenas a la organización (como el Instituto Nacional de Estadística).

En esta fase, el punto clave es identificar las fuentes más apropiadas de las cuales recuperaremos la información, deberemos analizar los formatos, la disponibilidad y la calidad de la información.

- El proceso ETL de extracción, transformación y carga de los datos en el Datawarehouse. Mediante este proceso, se pretende que los datos sean transformados, limpiados, filtrados y redefinidos, ya que por lo general, la información obtenida de sistemas transaccionales no está preparada para ser aplicada en la toma de decisiones.

Los procesos para completar esta fase son:

- Extracción: Se recuperan los datos en bruto de las distintas fuentes de información.
- Limpieza: Se recuperan los datos en bruto para determinar su calidad. Se eliminan los datos duplicados y si es posible, se corrigen los problemas de valores equivocados o vacíos. En este momento disponemos de datos más limpios y de mayor calidad respecto a los extraídos inicialmente.
- Transformación: Se recuperan los datos ya limpios de la fase anterior, y son estructurados y se recapitulan en diversos modelos de análisis. El resultado de este proceso es la obtención de datos limpios, consistentes, indexados y útiles.

- Integración: Primero se comprueba la validez de los datos que cargamos en el *Datawarehouse*. Especialmente la consistencia con las definiciones y formatos del *Datawarehouse*. Después los integra en los distintos modelos de las distintas áreas de negocio que han sido definidas en el almacén de datos.
- Actualización: Los nuevos datos al *Datawarehouse*.

- El *Datawarehouse* o almacén de datos, con el Metadata o Diccionario de datos. Se busca almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración.

Al recuperar la información de los distintos sistemas, tanto transaccionales como departamentales o externos, y almacenándolos en un entorno integrado, el *Datawarehouse* permite realizar el análisis de la información en un contexto relacionado con la propia organización. Además de los *Datawarehouse*, hay que mencionar los *Data Mart*, más pequeños que estos, y dirigidos a una comunidad de usuarios dentro de la organización, que puede estar formada por los miembros de un departamento, o grupo de trabajo multidisciplinario que persiga la misma meta. Los *Data Mart* pueden ser independientes o dependientes. Los primeros son alimentados directamente de los orígenes de información, mientras que los segundos se alimentan desde el *Datawarehouse* corporativo. Un componente crítico de un *Datawarehouse* es el Metadata. El Metadata es el repositorio central de información de la información. Nos da el significado de cada uno de los componentes y sus atributos que residen en el *Datawarehouse* (o *Data Mart*). La información que contiene el Metadata es útil para los departamentos de tecnología y los propios usuarios. Puede incluir definiciones de negocio, descripciones detalladas de los tipos de datos, formatos y otras características.

- El motor OLAP, que nos debe proveer capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de escenarios en grandes volúmenes de datos.

La idea es que los usuarios puedan hacer unos análisis multidimensionales al máximo nivel de agregación o al máximo nivel de detalle.

- Las herramientas de visualización, que nos permitirán el análisis y la navegación a través de los mismos.

Las principales herramientas de visualización en *Business Intelligence* son:

- Generadores de informes: Son utilizadas para crear informes estándar por parte de desarrolladores profesionales.
- Herramientas de consultas e informes: Son utilizadas para crear informes por parte de los usuarios finales.
- Herramientas OLAP: Posibilitan tratar la información multidimensionalmente.
- Herramientas de *Dashboard* y *Scorecard*: Se emplean para visualizar la información relacionada con el rendimiento por parte de los usuarios finales.
- Herramientas de planificación, modelización y consolidación: Mediante su uso, es posible crear planes de negocio, presupuestos y las previsiones. Proveen a los *Dashboard* y los *Scorecard* con los objetivos y los umbrales de las métricas.
- Herramientas de minería de datos: Permiten a estadísticos o analistas de negocio crear modelos estadísticos de las actividades de los negocios.

### **2.3 Beneficios del Business Intelligence**

Los beneficios que se pueden obtener mediante el uso de la inteligencia de negocio, se pueden dividir en: beneficios tangibles, beneficios intangibles y beneficios estratégicos. La fusión de estos tipos de beneficios, es lo que se pretende conseguir cuando se decide integrar una solución BI en una empresa. Teniendo en cuenta la naturaleza de la organización sobre la que se desea implementar la solución, se espera que la aplicación de BI proporcione:

- Beneficios intangibles: Disponer de la información necesaria para decidir qué acciones llevar a cabo en tiempo real, y predecir las que será necesario realizar en un futuro. Los beneficios intangibles tendrán una repercusión directa en el mantenimiento de los acuerdos vigentes, y en la adquisición de futuros contratos.

- Beneficios tangibles: Especialmente se pretende, mediante la optimización de los procesos del negocio, conseguir la rentabilidad deseada. Obviamente, este tipo de beneficio tiene un importante impacto sobre el aspecto económico.

- Beneficios estratégicos: En el caso sobre el que se trabaja, este tipo de beneficios pueden ayudar a puntualizar que servicios han dado mejores resultados con los clientes, y por tanto, valorar si su aplicación es lo más correcto cuando se tomen futuras decisiones.

### **3. Requisitos y KPIs**

El propósito principal de este punto, es detallar los requisitos del Proyecto, con la finalidad de cubrir las necesidades del tipo de organización donde es posible su implementación.

En este sentido, se definen que *KPIs* son aplicables para alcanzar el propósito planteado mediante los requerimientos.

Con los requisitos fijados y los indicadores concretados, el siguiente paso es el de encontrar el producto que mejor pueda adaptarse a las pretensiones expuestas.

#### **3.1 Requisitos funcionales**

Se conciben como requerimientos funcionales, aquellas declaraciones de los servicios que proveerá el sistema, relacionado con la manera en que éste reaccionará a entradas particulares.

En el caso del presente proyecto, los requerimientos funcionales son los siguientes:

- Acceso a la visualización de la información en tiempo real.
- Generar informes de rendimiento.
- Simular escenarios futuros mediante la modificación de variables.
- Gestión de accesos para usuarios internos y clientes.

### 3.2 Requisitos técnicos

- Interfaz gráfica intuitiva.
- Rendimiento suficiente para una adecuada experiencia de usuario (UX).
- Acceso desde cualquier ubicación, multiplataforma y dispositivo.
- Sistema escalable en cuanto a número de usuarios y procesamiento.

### 3.3 Requisitos económicos

- Inversión mínima en costes de Licencias Software.
- Implantación rápida a través de recursos propios, sin depender de terceros.

### 3.4 KPIs

A modo de introducción, solo mencionar que las siglas *KPI* se corresponden con las iniciales de *Key Performance Indicators*. En español se traduce como “Indicadores Clave de Desempeño”. Estos *KPIs* consisten en métricas que nos ayudan a cuantificar el rendimiento de procesos en función de unas metas y objetivos planteados para las distintas actividades que se lleven a cabo dentro de una organización.

El proyecto actual, demanda tener en cuenta los requisitos formulados anteriormente, y en función de estos, se deben encontrar aquellos que puedan servir de ayuda a alcanzar los objetivos del conjunto del Trabajo.

La idea del caso concreto sobre el que se trabaja, es determinar que *KPIs* pueden ser los más interesantes, para ayudar a organizaciones a mejorar su rendimiento sobre los servicios que se ofrecen al ciudadano. Estos indicadores serán pocos pero relevantes, y en función de la evolución tras la puesta en marcha, se irán incluyendo nuevos con la intención de mejorar la precisión de la información obtenida.

#### **3.4.1 KPI: Calidad limpieza viaria**

Para medir la calidad de la limpieza viaria, se utilizan datos mensuales de diferentes fuentes, como puedan ser las reuniones de calidad que periódicamente se acostumbran a realizar en los distintos Ayuntamientos. Por tanto, estos datos en general, son fruto de la valoración por parte de los inspectores de los departamentos de Medio Ambiente.

#### **3.4.2 KPI: Número de quejas ciudadanas**

Se trata de un *KPI* realmente objetivo, ya que dichas quejas quedan registradas en las oficinas de atención ciudadana (OAC) que hay presente en la mayoría de ciudades, y que frecuentemente se envían directamente al departamento de producción de la empresa.

#### **3.4.3 KPI: Tiempo de respuesta a las emergencias**

Los datos de tiempo de respuesta a las incidencias consideradas de alta importancia, son accesibles en muchos casos mediante software independiente dedicado a la gestión de incidencias (*SGI*, Sistema Gestor de Incidencias). Mediante su uso, las empresas que disponen de estos productos, logran generar informes en diferentes formatos. Es relativamente sencillo filtrar los datos más interesantes para explotar este indicador de manera correcta.

#### **3.4.4 KPI: Formación del personal**

En apariencia se trata de un indicador poco trascendente, no obstante, las formaciones que se realizan, tanto en materia medioambiental, como en sistemática de trabajo, han permitido mejorar el rendimiento de los trabajadores de muchas de las empresas dedicadas a la prestación de servicios ciudadanos.

En el presente proyecto, se espera que con su presencia como indicador, se pueda mostrar de manera objetiva el impacto que tiene en el resto de *KPIs*.

#### **4. Selección de la solución**

##### **4.1 Software libre o software propietario**

La intención de este apartado, es examinar las características más relevantes del software libre y del software propietario. En ningún caso se pretende dictaminar si una modalidad es más ventajosa que la otra de manera genérica, sino que ambas opciones serán valoradas, en función de las características del caso concreto sobre el que se realiza el proyecto. En este sentido, y con el propósito de tomar la decisión definitiva sobre cuál es la modalidad más conveniente para el presente Trabajo, se tendrá en cuenta un análisis posterior sobre aquellos factores que más repercusión tiene dentro del ámbito en el que se plantea llevar a cabo la solución.

A continuación, se explican algunas de las características, sin rigurosa exactitud, que definen cada tipo de software:

- Libre: Hace referencia a la libertad de los usuarios para ejecutar, copiar, distribuir, estudiar, modificar el software y distribuirlo modificado. La seguridad y fiabilidad depende en gran parte de los usuarios y desarrolladores que lo utilizan, por tanto, mientras mayor sea esta comunidad, mejores prestaciones presentará el producto.

Por lo general no se destinan muchos recursos al cuidado de la interfaz gráfica, y en consecuencia, las aplicaciones de este tipo pueden llegar a ser poco intuitivas y complicadas de manejar. Aunque este aspecto este cambiando favorablemente en el transcurso del tiempo. Posiblemente, la mayor ventaja que presenta para los usuarios finales, está relacionada con la inversión económica. En el caso del Software libre, el coste es muy bajo si se compara con programas propietarios destinados a cubrir la misma necesidad.

En muchas ocasiones, y especialmente cuando se trata de usuarios particulares, los productos ofrecidos son completamente gratuitos. También destacar que el usuario del programa no depende de la figura del creador del mismo.

- Propietario: Es todo programa informático en el que el usuario tiene limitaciones para usarlo, modificarlo o redistribuirlo (esto último con o sin modificaciones). Las compañías propietarias se encargan de la seguridad. Generalmente, se destinan muchos recursos a la interfaz gráfica, y se pretende que esta sea amigable para el usuario final. La documentación también va a cargo de la empresa, que intenta que esta sea lo más correcta y precisa posible. No todas las aplicaciones se desarrollan para las diferentes plataformas que coexisten en la actualidad. Además, se crea un vínculo de dependencia con la empresa desarrolladora del software, ya que es usual que los cursos de aprendizaje sobre el programa y el mantenimiento del mismo, sean llevados a cabo por parte del personal de la compañía propietaria. Destacar que en referencia al ámbito educativo, los planes de estudios de muchas universidades tienen un marcado enfoque al uso de herramientas propietarias, y son las propias organizaciones fabricantes quienes ofrecen a las universidades planes educativos de descuento muy atractivos. Con esta política, se aseguran disponer de clientes potenciales.

Una vez mencionadas algunas de las características más destacadas en ambas modalidades de software, se traslada al contexto de la implementación de soluciones *Business Intelligence*. Asimismo, también se tiene en cuenta el dominio sobre el que se desea aplicar el Trabajo.

Para el presente proyecto, se ha decidido enumerar algunos de los factores críticos, considerados clave en el éxito del proyecto.

- Costes de desarrollo: Es indiscutible que los costes en software de código libre son mucho menores. Este factor, es considerado crucial en la decisión, ya que no se trata de una implementación a gran escala, sino que se pretende abarcar todo tipo de compañías dedicadas a los servicios ciudadanos, pudiendo ser estas pymes con recursos limitados.

- Adaptación a los requisitos específicos: Existen estudios que afirman que los proyectos de BI requieren un elevado nivel de adaptación a las singularidades de las empresas en la que se desea desarrollar.

En este sentido, las soluciones de BI propietarias siguen en su mayor parte una configuración estándar y poco flexible debido a las licencias a las que van asociadas. El software libre, permite mayor margen de maniobra en este sentido. Dado que se trata de un proyecto destinado a cubrir unas necesidades muy particulares, se necesita que este pueda ser modelado al gusto del cliente.

- Imagen y aceptación: Debido a la situación económica actual, muchas entidades de carácter público han decidido aumentar el uso de programas libres, dando una imagen de aclimatación a las necesidades contemporáneas. Si se tiene en cuenta que a medida que el mercado de BI ha madurado, y que el incremento de soluciones libres disponibles es cada vez mayor, su uso puede aportar una percepción de compromiso con la realidad económica que rodea a la sociedad.

- TCO (Total Cost of Ownership) y ROI (Return Of Investment): Este aspecto se ve afectado primordialmente por factores como la posibilidad de desarrollar futuros proyectos dentro de las empresas, o la flexibilidad para la integración de las soluciones. Es necesario destacar que en caso de que las herramientas implementadas en el actual proyecto, den alcance a los propósitos ambicionados, la inexistencia de licencias va a permitir que esta se sume al resto de aplicaciones existentes en la organización a nivel global.

Una vez valorados los factores más trascendentes, se determina que el software libre es el más indicado para el proyecto que se expone en el actual Trabajo.

El siguiente paso, es determinar qué solución de inteligencia de negocio es la más acertada para generar las herramientas BI, y por tanto, a cubrir los objetivos del Trabajo.

## **4.2 Análisis de soluciones**

En el punto previo, y tras una valoración adaptada al contexto sobre el que se desarrolla el proyecto, se ha decidido que una aplicación de software libre es la opción más conveniente para desarrollar el Proyecto.

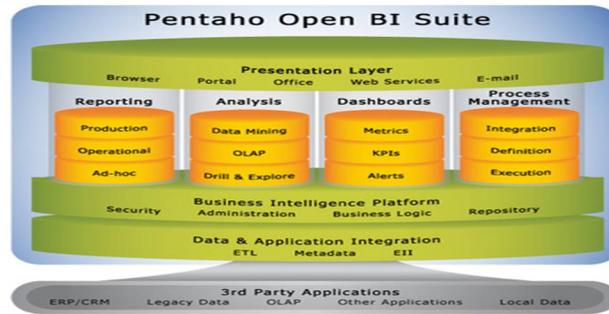
El apartado actual, se dedica a valorar diferentes alternativas de software libre. Se considera que dichas soluciones, pueden ser serias aspirantes a convertirse en la plataforma definitiva sobre la que se elabore el presente proyecto. Para ello se ha procedido haciendo una criba inicial, de la que han resultado 3 soluciones que el autor del presente Trabajo estima que tienen un mayor reconocimiento dentro de la comunidad OSBI (*Open Source Business Intelligence*). Se trata de Pentaho, Jasper y Palo.

Una vez revisadas individualmente, se ha creado un cuadro donde se expone de manera sencilla los puntos fuertes y débiles de cada una de ellas.

### **4.2.1 Pentaho**

Desde que fue creada en el año 2004, ha ido evolucionando hasta llegar a ser el actual líder en cuanto a soluciones de Business Intelligence Open Source. Ofrece, con soluciones propias, todo el espectro de recursos para desarrollar, mantener y explotar un proyecto de BI. Desde las ETL con *Data Integration* hasta los cuadros de mando con el *Dashboard Designer* o el *Community Dashboard Framework*.

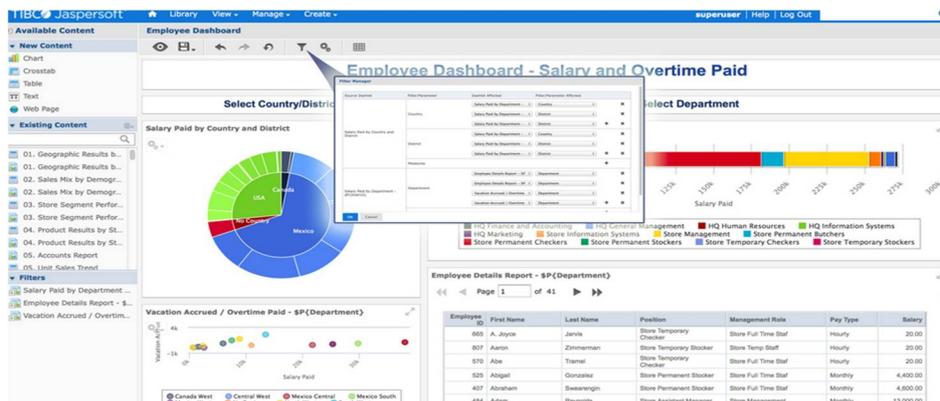
En la siguiente imagen se observa el diseño de la arquitectura de Pentaho:



#### 4.2.2 Jasper

Jaspersoft es la empresa que está detrás del famoso y extendido JasperReports. Esta solución *Open Source* de reporting es la preferida por la mayoría de desarrolladores para embeber en cualquier tipo de aplicación java que requiera un sistema de informes. La solución Jasper de *Business Intelligence*, funciona en torno a su motor de informes, y es que este motor de informes, es considerado superior a los que ofrece la competencia. En este sentido, se puede afirmar que El editor de informes *ad-hoc* es el mejor estructurado y con mejores funcionalidades de los que estamos analizando.

La siguiente captura, muestra la interfaz gráfica, y pretende dar una idea de la capacidad de esta herramienta en referencia a la creación de reports de calidad:

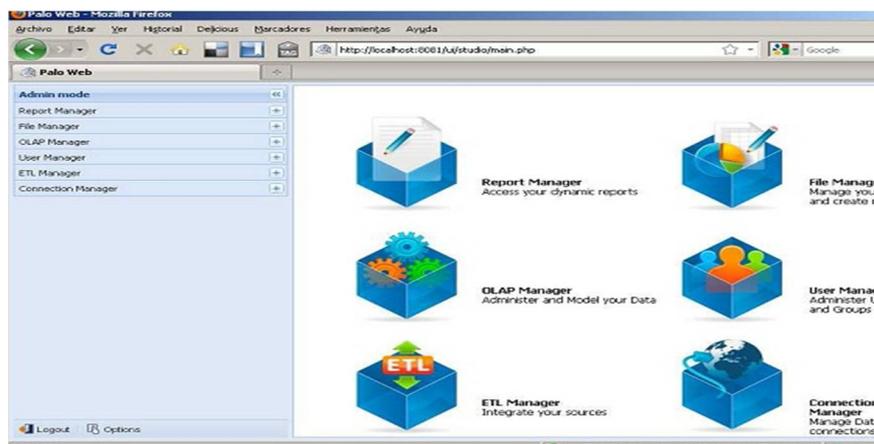


### 4.2.3 Palo

Es el motor MOLAP implementado por la empresa *Jedox*. Toda la propuesta *Business Intelligence* de *Jedox* gira en torno a su motor MOLAP que es su característica más diferenciadora. Si se quiere implementar un entorno de BI que permita realizar simulaciones, dibujar diferentes escenarios y trabajar con los datos manipulándolos para ver los diferentes posibles resultados Palo es una buena solución.

La funcionalidad más importante que aporta Palo, al ser MOLAP es que permite editar los valores y propagar los cambios en base a reglas de negocio.

Como sucede con el resto de soluciones, la interfaz es muy intuitiva:



### 4.2.4 Tabla comparativa: Selección de la plataforma

Se prosigue con una tabla donde se muestra la relación entre cada una de las herramientas comentadas anteriormente, y las características más relevantes que se esperan obtener de la aplicación de una solución BI.

Se denotarán como *completo*, *adecuado* o *insuficiente*, en función del nivel presentado por el producto en cada uno de los aspectos valorados. Es importante saber que de estas soluciones, existe versión *Community* y *Enterprise*, y se ha optado por valorarlas de manera global.

Característica	Pentaho	Jasper	Palo
ETL	Completo	Completo	Completo
Servidor BI / Servidor web	Completo	Completo	Completo
Reports	Completo	Completo	Completo
Reports Add-hoc	Suficiente	Insuficiente	Suficiente
Motor OLAP	Completo	Suficiente	Completo
Visor OLAP	Suficiente	Suficiente	Suficiente
Cuadros de mando	Completo	Suficiente	Insuficiente
Cuadros de mando Add-hoc	Insuficiente	Insuficiente	Insuficiente

Una vez evaluadas las soluciones, y sopesados los pros y los contras de cada una de ellas, se llega a la conclusión de que la más apropiada para llevar a cabo la implementación es **Pentaho Community Edition**.

Sin ánimo de extenderse sobre lo que ya se ha comentado sobre esta aplicación, decir que en este caso ha sido decisivo el nivel de madurez de esta herramienta, además de la gran comunidad de usuarios que posee. La documentación sobre la misma, alcanza niveles comparables a las de cualquier herramienta de licencia propietaria, lo que dota de gran ayuda a los usuarios noveles en la materia. Si además a ello le añadimos que según uno de los últimos informes de *Gartner*, Pentaho ofrece la plataforma más completa de BI *Open Source*, la decisión se hace más firme. En este caso, *Gartner* se basa tanto en las altas capacidades de integración de datos de diferentes fuentes que tiene, como en el amplio soporte para el análisis avanzado que ofrece. Asimismo, afirma que las principales áreas de enfoque que diferencian a Pentaho son: el análisis de negocio, el análisis *big data* y las herramientas de analítica incorporadas en la plataforma.

Como resultado de su apuesta por la innovación, Pentaho ha mejorado sustancialmente su posición en el *Cuadrante Mágico de Gartner*, moviéndose hacia el eje de los más visionarios en el mercado de la analítica y el *Business Intelligence*.

## 5. Instalación y configuración de Pentaho

Con la decisión final de utilizar Pentaho como plataforma sobre la que se desarrolla el Proyecto, se procede al paso de la instalación y su configuración.

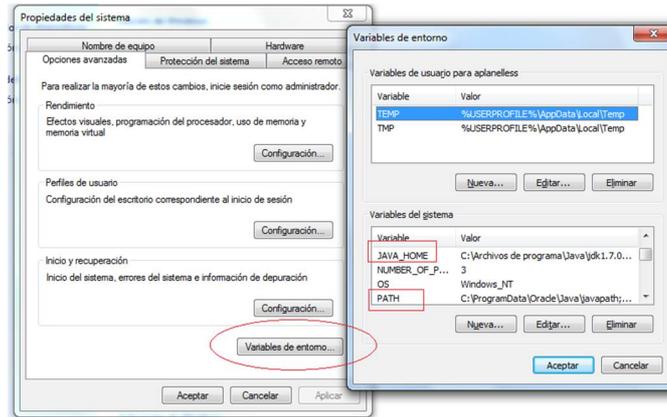
A partir de este punto, el Trabajo adopta básicamente un planteamiento práctico, en el que se incluyen las nociones técnicas adquiridas durante el estudio sobre Pentaho que ha sido llevado a cabo paralelamente al desarrollo del resto de puntos. Mencionar que de las versiones disponibles actualmente de Pentaho, se opta por la 5.4. En este sentido, ha prevalecido la madurez en detrimento de la novedad. Una versión estable y altamente utilizada, ayuda especialmente en aspectos como el soporte en cuanto a resolución de dudas por parte de la comunidad.

### 5.1 Descarga y configuración de componentes básicos

Cabe destacar que todos los componentes que componen Pentaho están desarrollados bajo la plataforma Java. Puesto que lo primero que debemos hacer es descargarnos la versión de java compatible con Pentaho Community. En este caso, se trata de “*Java Development Kit 7*”.

Una vez descargado, y antes de proceder con la instalación de Pentaho, es necesario configurar las variables de entorno de java. Dependiendo del sistema operativo donde se lleve a cabo dicha instalación, este proceso puede variar notablemente. En el caso del autor, la instalación se realiza tanto en *Ubuntu 14.04 LTS*, como en *Windows 7*. El resto de procesos, se ejecutan principalmente en *Windows 7*, ya que por motivos de rendimiento, es el entorno más favorable al que se tiene acceso por parte del autor para la realización de pruebas sobre la plataforma.

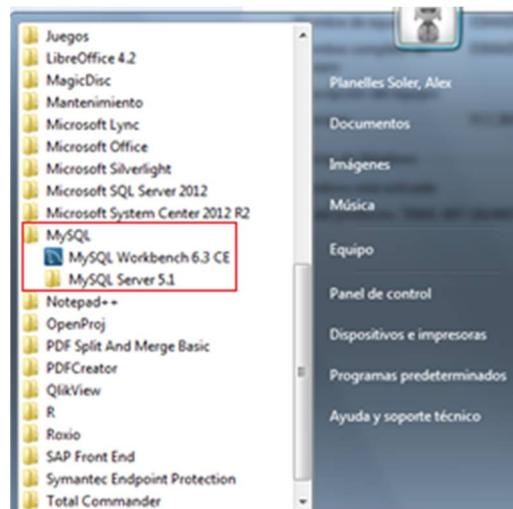
En la siguiente captura se muestran las variables de entorno necesarias (*JAVA\_HOME* y *PATH*) para el funcionamiento de la plataforma.



Para nutrir Pentaho BI server es necesaria la obtención de un origen de datos. En este sentido, se ha optado por el *SGBD*, sistema gestor de bases de datos, *MySQL*. Indicar que la gestión se realizará mediante “*MySQL Workbench*”, puesto que se trata de una herramienta que ofrece la sencillez requerida para el desarrollo del proyecto. Se genera una base de datos en la que posteriormente los datos a analizar podrán ser introducidos. Señalar que la *URL* es *localhost* y se mantiene el puerto por defecto *MySQL*, concretamente el 3306. En consecuencia, para acceder se debe introducir en el navegador: <http://localhost:3306/>

A continuación se muestran algunas capturas en relación a la creación de la base de datos:

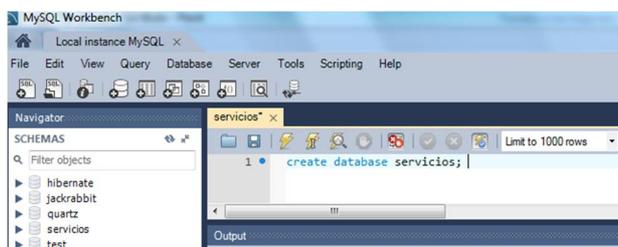
Ejecución del programa:



Introducir contraseña anteriormente indicada (durante la instalación):

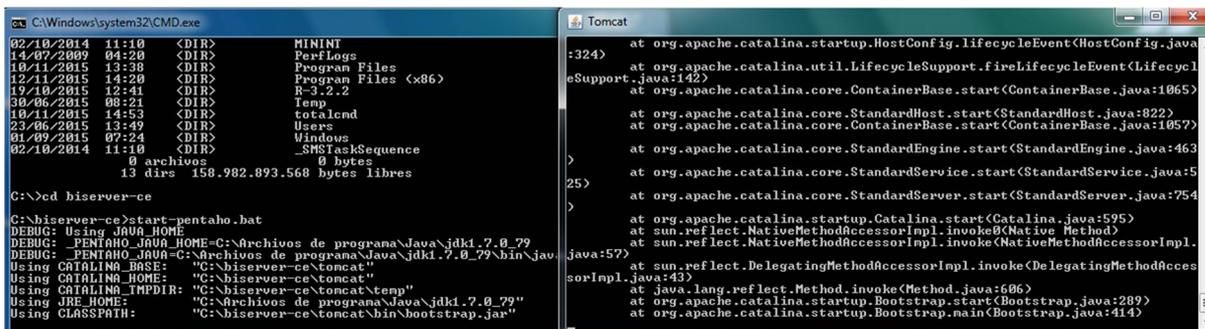


Se crea la BD "servicios":

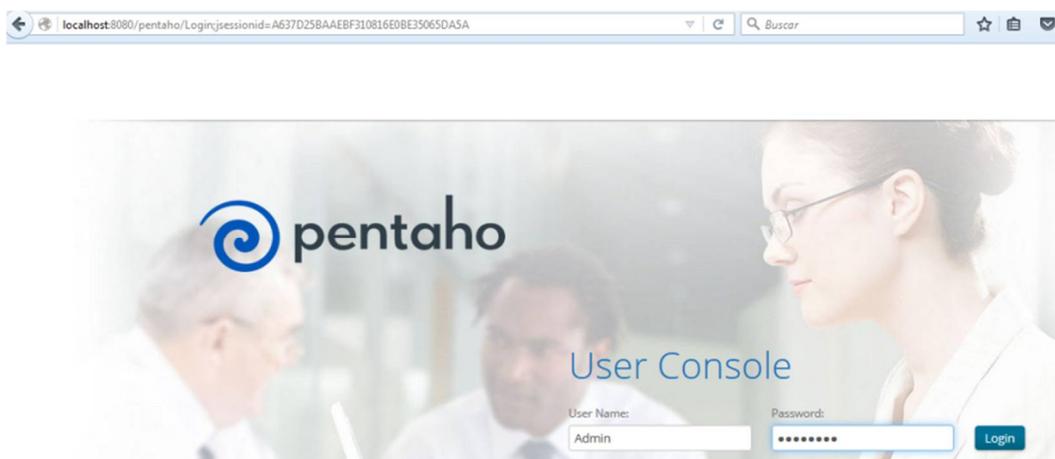


Posteriormente, se debe descargar la versión anteriormente comentada del servidor BI de Pentaho de la página de Pentaho Community (<http://community.pentaho.com/>).

Acto seguido, se ejecuta el archivo "start-pentaho.bat" residente en la carpeta "biserver-ce". En ese preciso instante, se abre la consola y se inicia el servidor Tomcat sobre el que se aloja Pentaho.



Haciendo uso de un navegador, ya es posible el acceso al servidor BI de Pentaho introduciendo en la barra de navegación “localhost:8080/pentaho”. En la consola se introduce el nombre de usuario: “Admin” y el password: “password”. En este punto, comentar que en otras versiones anteriores el nombre de usuario y contraseña varían.



Una vez se han introducido, ya se accede a la *Pentaho User Console*, tal y como se muestra a continuación:



## 5.2 El origen de los datos (Data Sources)

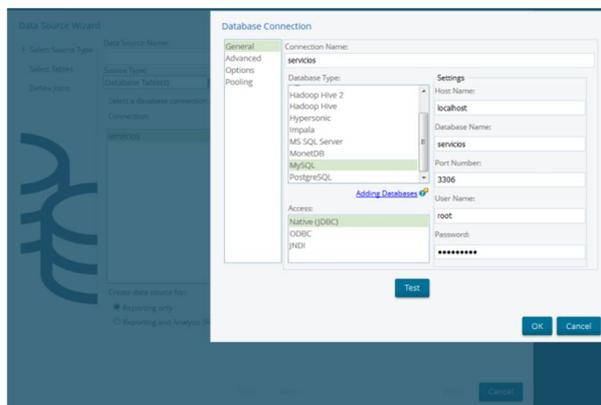
En este momento, y con *Pentaho BI server* y la base de datos *MySQL* ya instalados, es necesario añadir el origen de datos en Pentaho, para que sea posible el acceso a dichos datos. Estos, deben encontrarse disponibles en la base de datos. Para ello se accede a “Manage Data Sources”.

En la siguiente captura, se muestran los dos accesos posibles a la gestión de fuentes de datos:

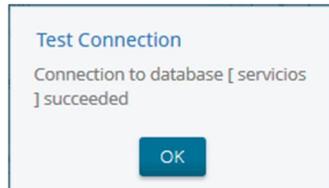


Aparecerá una ventana en la que se visualizan todos los orígenes de datos incluidos en Pentaho. Se debe acceder a “New Data Source”. A continuación, se escoge como Source Type la opción de Database Table(s), y posteriormente añade una nueva conexión pulsando “+”.

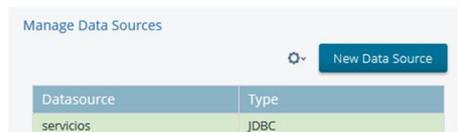
La siguiente captura muestra esta parte del proceso:



Una vez realizado el test de conexión con la base de datos “servicios” creada anteriormente en MySQL Workbench, se obtiene el siguiente mensaje:



A partir de este momento, ya se dispone de la fuente de datos “servicios”.



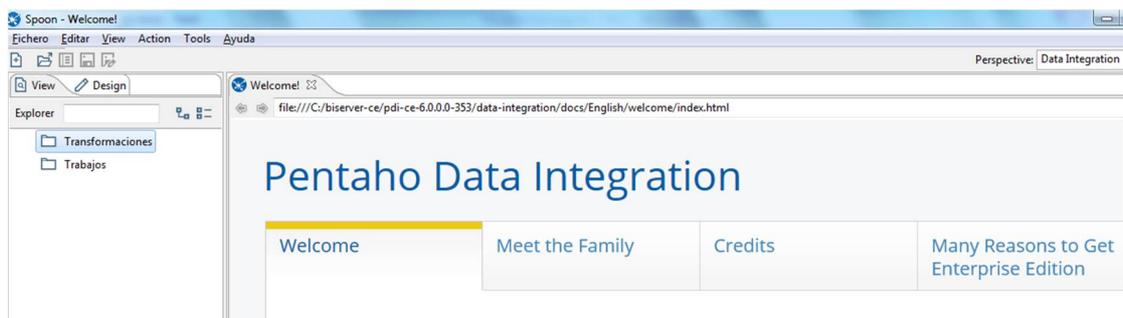
### 5.3 Resto de componentes y plugins

Para el resto de componentes simplemente se debe descargar el archivo comprimido desde la página web de Pentaho Community y descomprimirse. Dentro de la carpeta, una vez descomprimida, es posible ejecutar el archivo con el nombre de la aplicación para iniciar la aplicación.

Además de contar con Pentaho BI Server, convenientemente instalado y configurado, se requieren:

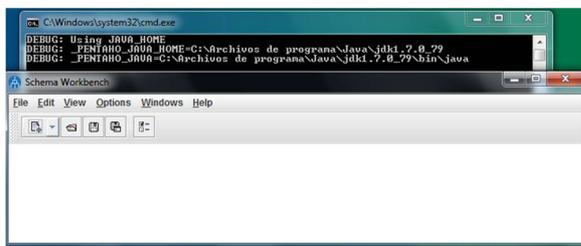
- ❖ Data Integration

Para iniciar “Data Integration”, se debe ejecutar el archivo “Spoon”:



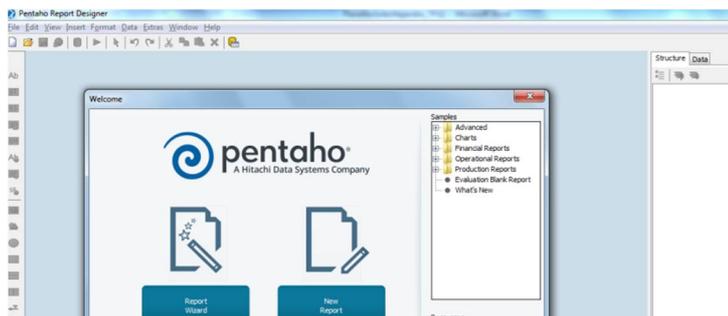
❖ Schema Workbench

En este caso, se debe ejecutar el archivo “workbench”:



❖ Report Designer

Para iniciar el diseñador de informes, se debe ejecutar el archivo “report-designer”:

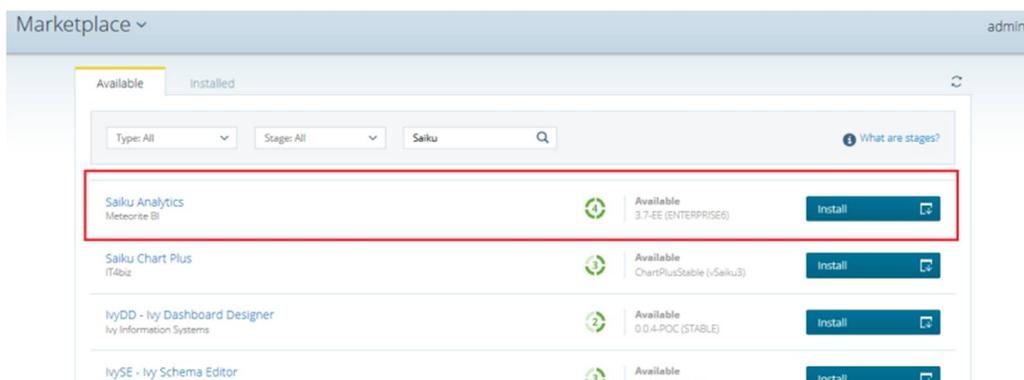


Además de los componentes anteriormente mencionados, Pentaho dispone del denominado *Pentaho Marketplace*. Se trata de un único lugar que permite el acceso, la gestión y la actualización de todos los *plugins* del servidor. En definitiva una única interfaz simple y sencilla que permite a los desarrolladores de *plugins* disponer de una vía de acceso rápida, y a los usuarios el fácil acceso a todos los *plugins* disponibles desde una única ubicación, sin tener que conocer todos los proyectos y sus correspondientes naturalezas. En la versión Pentaho sobre la que se trabaja en el presente proyecto, *Pentaho Marketplace* viene ya incorporado.

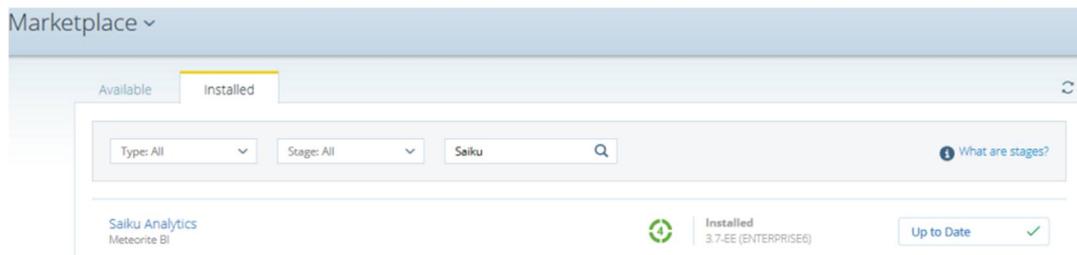
En este sentido, y tal como se muestra a continuación, la instalación de cualquier *plugin* que se precise, se convierte en un proceso relativamente simple.



El plugin que se decide instalar, debido a que permite efectuar consultas de una manera mucho más intuitiva y gráfica, es *Saiku Analytics* en su versión 3.7. Es interesante comprobar en qué estado de madurez se encuentra el plugin a instalar. Los valores, van desde el 1 hasta el 4, indicando el 1 que el plugin se encuentra en una fase de desarrollo prematura, hasta el 4 que es considerado el nivel más elevado en cuanto a madurez. Dicho esto, se puede observar que la versión de *Saiku Analytics* que se va a instalar, se encuentra en el nivel 4.



Una vez finalizada la instalación, en la pestaña “Installed” aparece *Saiku Analytics*.



## 6. Creación de herramientas BI

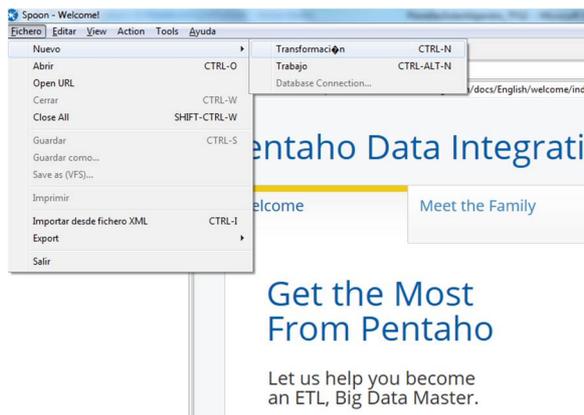
Con la plataforma Pentaho preparada, se puede proceder al inicio de la creación de las herramientas de inteligencia de negocio. Para alcanzar el propósito definido para el actual proyecto, se precisa del seguimiento de una serie de pasos, que se irán presentando en los próximos apartados.

### 6.1 ETL y Data Integration. Spoon (Kettle)

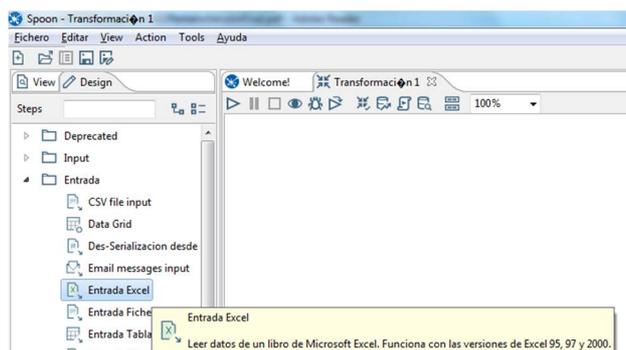
ETL hace referencia a las siglas en inglés de *Extract, Transform and Load*, es decir, Extraer, Transformar y Cargar. Este proceso, permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, para que finalmente puedan ser cargados en otra base de datos, *Data Mart*, o *Datawarehouse* para analizar, o en otro sistema operacional para apoyar un proceso de negocio.

La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

En el caso particular del presente proyecto, se obtendrán datos provenientes de archivos en formato Excel, que contienen la información que se ha estimado más interesante para ser objeto de análisis. La herramienta que se utilizará para llevar a cabo la integración de los datos (*Data Integration*), es *Spoon*, el entorno gráfico que permite trabajar con *Kettle*.

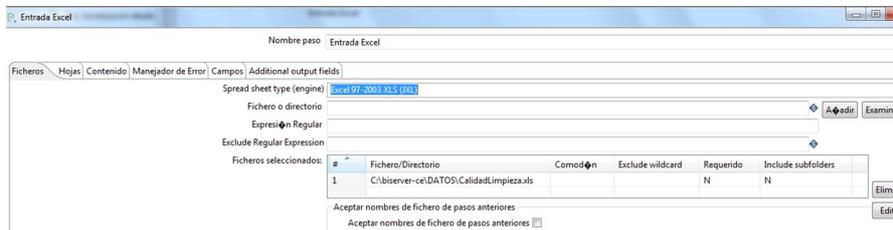


Para la entrada de datos de la nueva transformación, se selecciona la opción de "Entrada Excel", ya que es el formato en el que se encuentran los datos que se desean cargar.

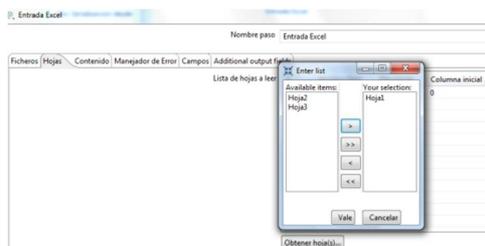


El siguiente paso es añadir el fichero. Para ello, se debe de examinar la ubicación de donde se encuentra, y posteriormente se añade a los ficheros seleccionados. En este punto, es interesante comentar que, tal y como se muestra en la captura anterior, existen restricciones sobre la versión de Excel.

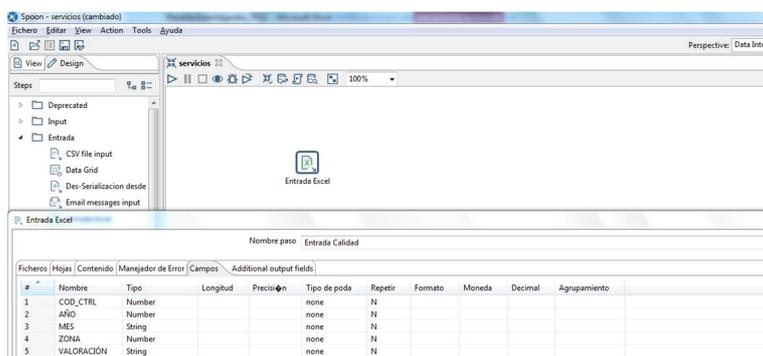
Por tanto, en caso de hacer uso de una versión más reciente, este debe ser guardado como “Libro de Excel 97-2003”.



Con el archivo añadido, en la pestaña “Hojas” se indican las hojas que contienen los datos. En este caso, simplemente se pretende obtener la “Hoja1”, ya que es la única que incorpora información.



El paso siguiente, es dirigirse a la pestaña “Campos”, y obtener los campos de la cabecera.

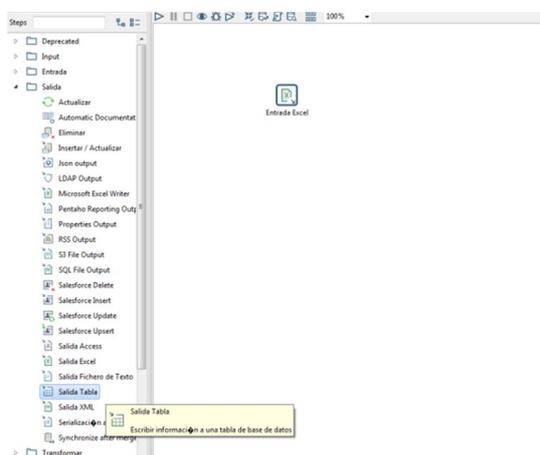


Como se puede apreciar, los datos ya están en un formato adecuado para ser tratados, y por tanto, simplemente hay que seleccionar la opción de “Salida tabla” y acto seguido, agregar una nueva conexión.

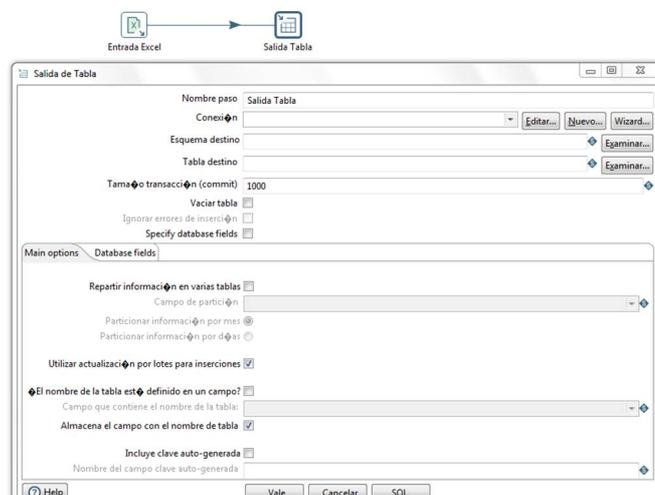
**Importante:** Previamente, se debe incluir la última librería de Java “*mysql-connector-java-5.1.37-bin*” en la carpeta a la carpeta \data-Integration\lib. Esta librería, se encuentra disponible en la web de *MySQL*.



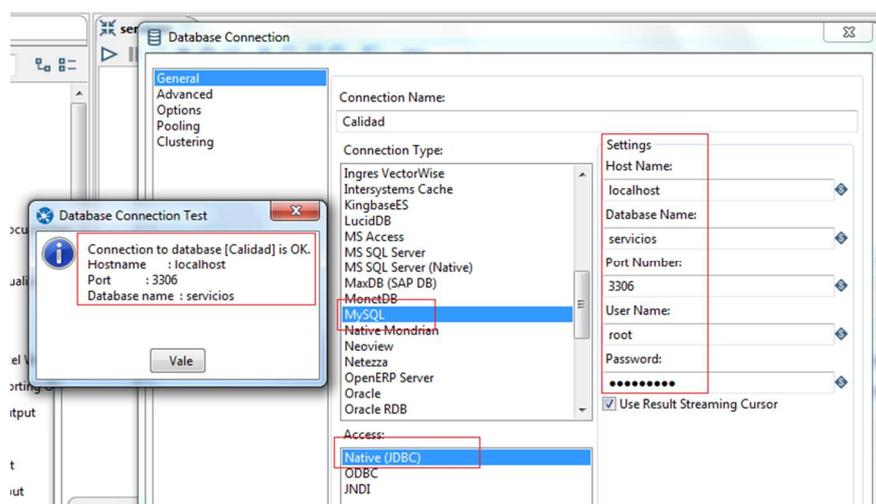
Una vez añadida la librería, ya se puede seleccionar la opción de “Salida tabla”:



A continuación, se procede a agregar la conexión (“Nuevo”):

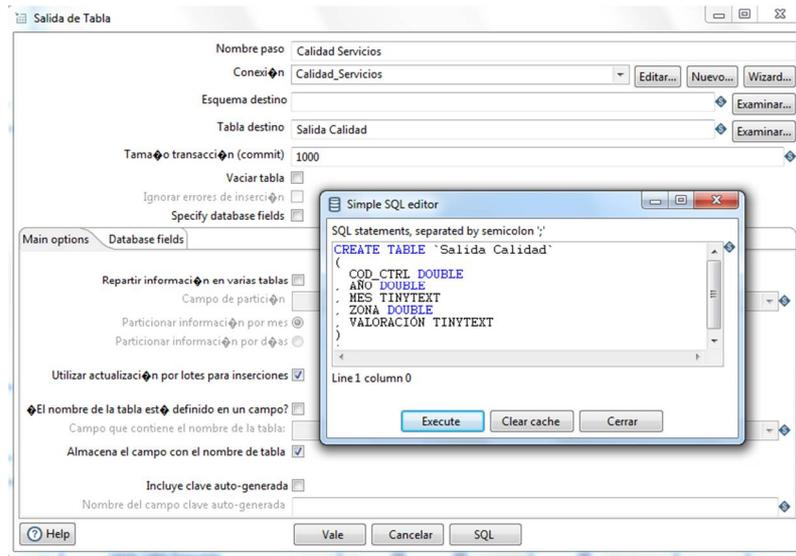


A continuación introduciremos los datos de conexión de la base de datos *MySQL*, y se hace clic en el botón “Probar”. Si todo ha funcionado correctamente, se obtiene un mensaje que indica que el proceso se ha realizado correctamente.

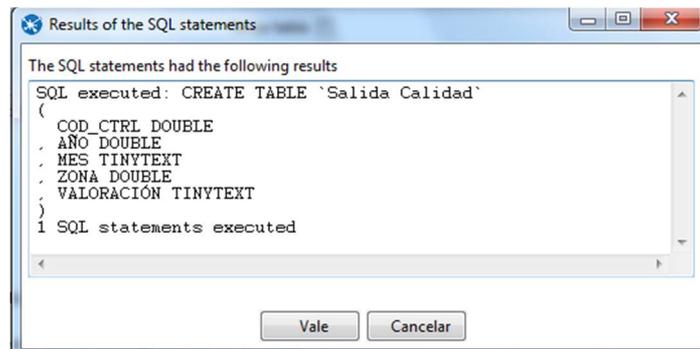


Como se puede observar, la conexión con la base de datos llamada “servicios” que se ha creado anteriormente en *MySQL*, se ha realizado correctamente.

Una vez queda configurada la conexión, se le da nombre a la tabla destino (“Calidad Servicios” en este caso) y se procede a pulsar sobre el botón “SQL”. Esta acción, provoca que de manera automática se genere un *script* con el nombre de la tabla a crear y los campos.



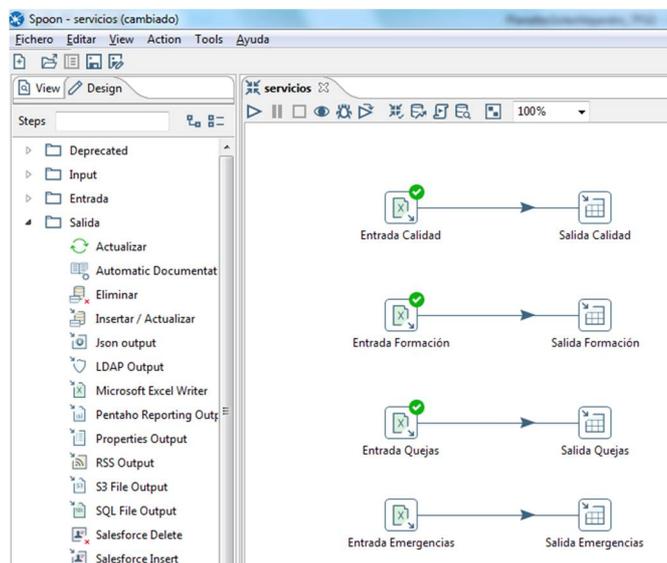
Una vez se ejecuta, haciendo clic en el botón “Execute”, se crean las tablas correspondientes a la BD.



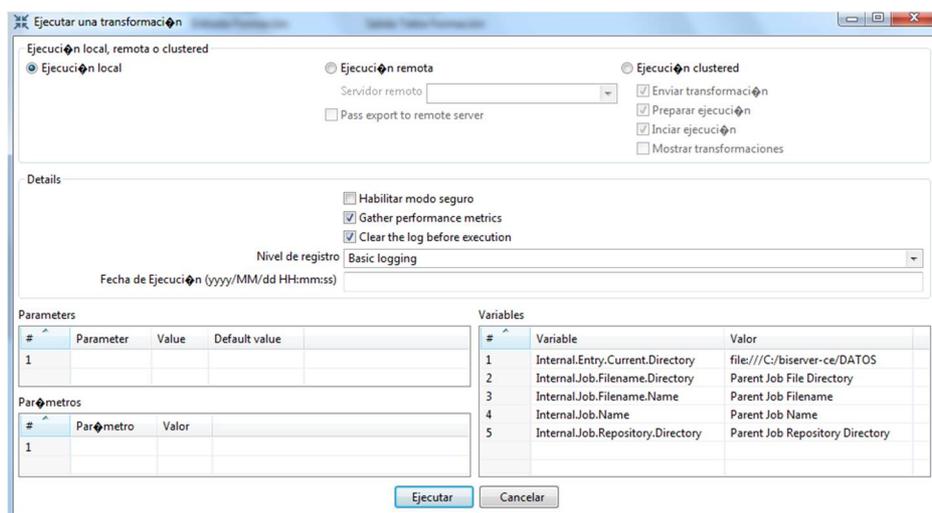
Este proceso se repite para los diferentes archivos Excel que serán la fuente inicial de datos que se encargan de nutrir al sistema. A continuación se muestran algunos de los utilizados:



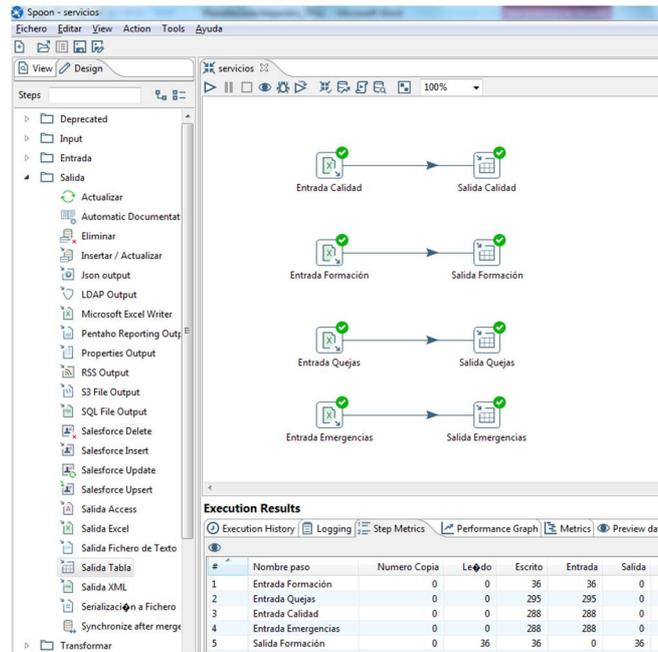
A continuación, se muestra la imagen del aspecto de la ventana “servicios” una vez realizadas todas las conexiones:



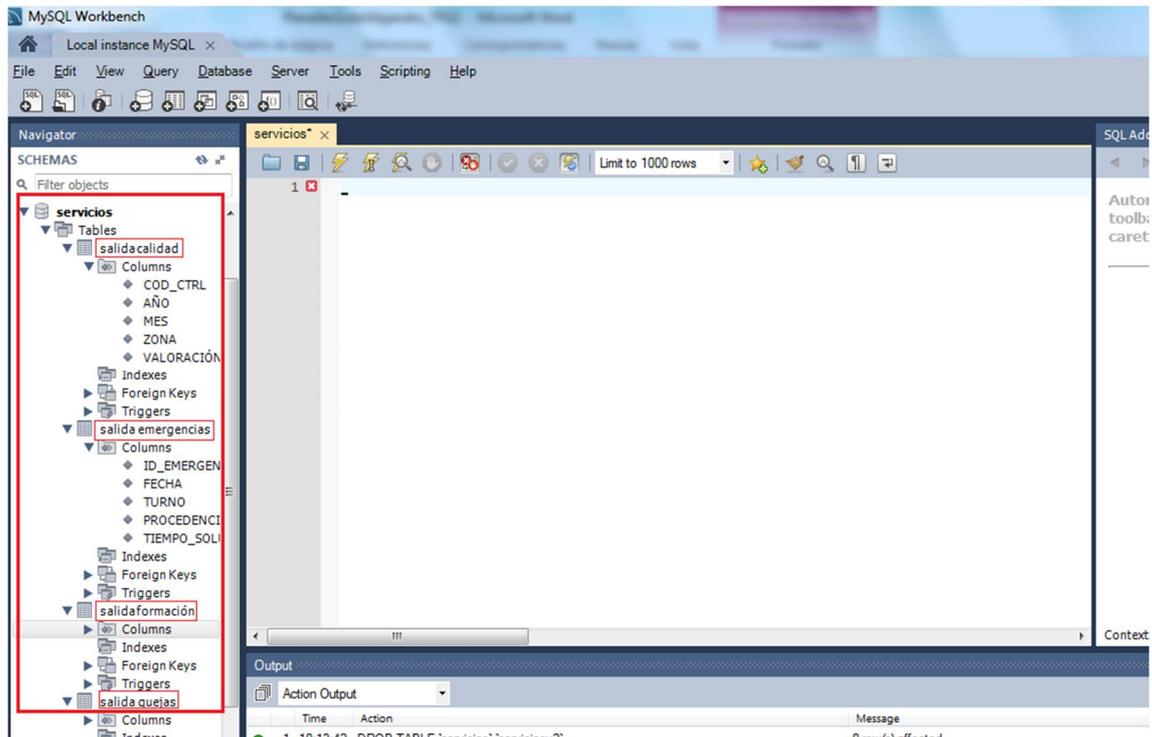
Después de ello se procede a ejecutar, y de esta manera llevar a cabo la carga de datos. Con dicha acción, los datos se añaden a la base de datos *MySQL*.



Cuando las conexiones se han realizado correctamente entre las tablas de entrada y salida, se muestra el símbolo de verificación y la información sobre las métricas de paso llevadas a cabo.

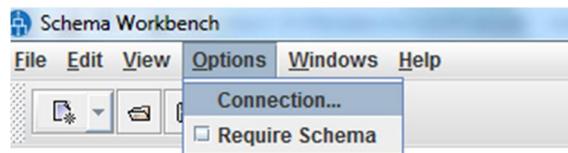


Como se ha comentado previamente, cuando se actualiza la base de datos “servicios” creada en MySQL, se puede apreciar como esta se ha cargado con las tablas (“salida calidad”, “salida emergencias”, “salida formación”, “salida quejas”) y columnas de los ficheros Excel sobre los que se trabaja. Posteriormente se realizan pruebas con otros archivos contenedores de datos.

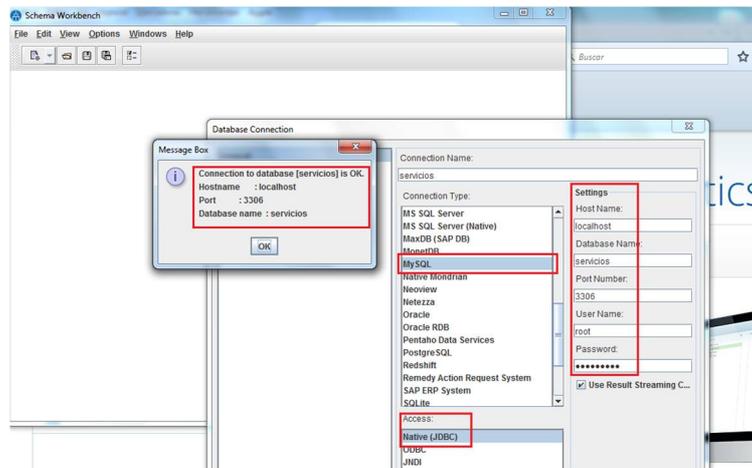


## 6.2 Cubo OLAP y análisis dimensional. Schema Workbench (Mondrian)

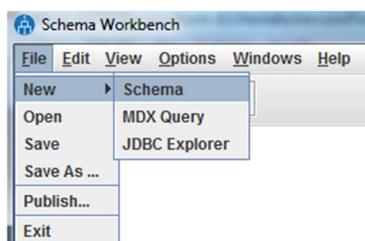
Respecto a la creación de cubos OLAP en Pentaho, la herramienta sobre la que se trabaja es *Schema Workbench*, la cual ha sido instalada previamente. Lo primero que se debe hacer es establecer la conexión con la base de datos *MySQL* desde el menú "Options".



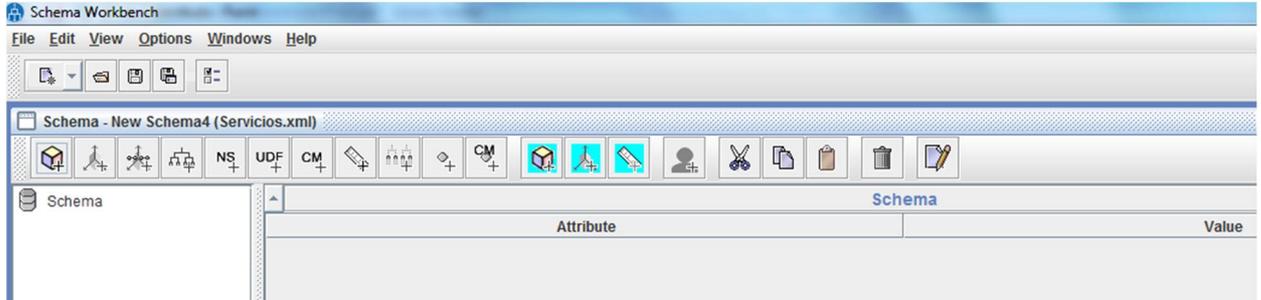
Se deben introducir los datos necesarios para establecer la conexión con *MySQL*. Una vez definida la conexión y los parámetros de la misma, se establece la comunicación:



Acto seguido se genera un nuevo esquema:



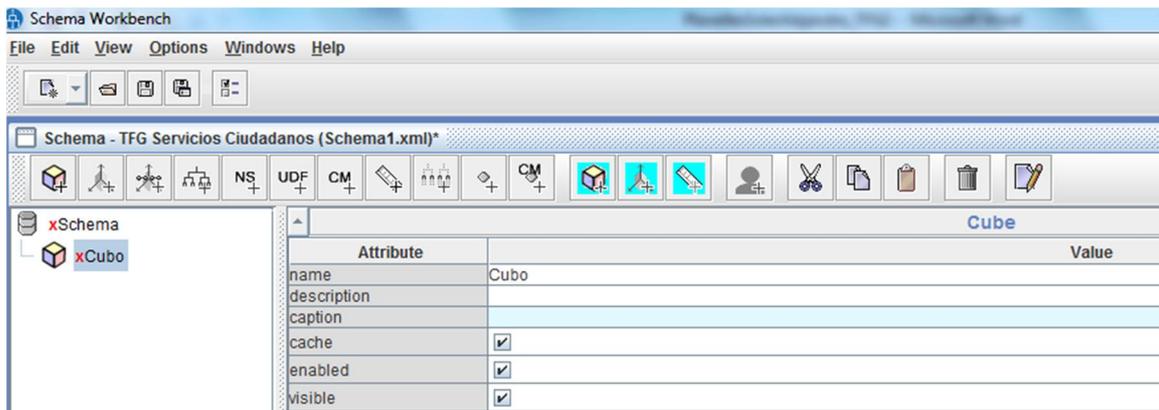
En la siguiente captura se observa el diseño de la interface sobre la que se trabaja:



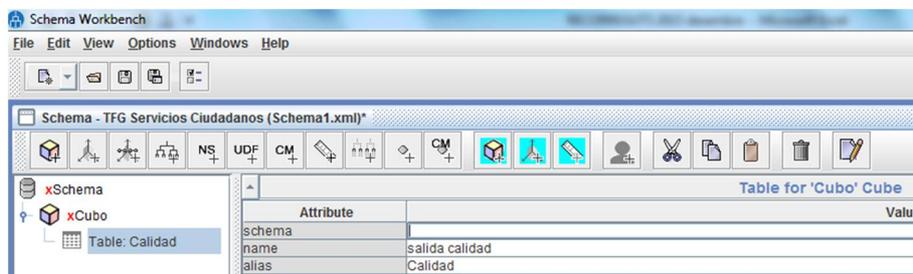
En primer lugar, se procede con la creación del Esquema (*Schema*).

Con la intención de aclarar el concepto, se puede decir que un esquema es un contenedor de cubos, que tendrá un único fichero XML, donde podremos crear tantos cubos como sea necesario. Las propiedades que se pueden indicar al crear un esquema son un nombre, la descripción, un nombre para la dimensión que agrupará a las medidas y un rol por defecto para utilizar en las conexiones de base de datos.

Al crear el Cubo, se indica un nombre y una descripción del mismo, pudiendo marcar además las opciones cache, para que el motor *Mondrian* trabaje con memoria cache en este cubo. Destacar también la opción "enabled", que se encarga de hacer visible al cubo. Si no se marca dicho *flag*, el cubo no aparecerá.



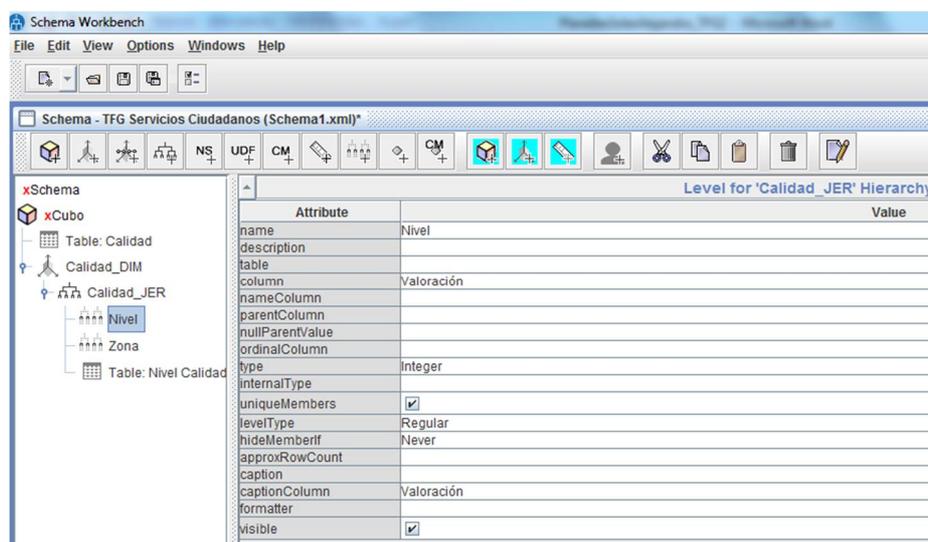
Acto siguiente, se debe seleccionar la Tabla de Hechos del cubo. A partir de esta tabla, se calculan tanto las medidas como los indicadores.



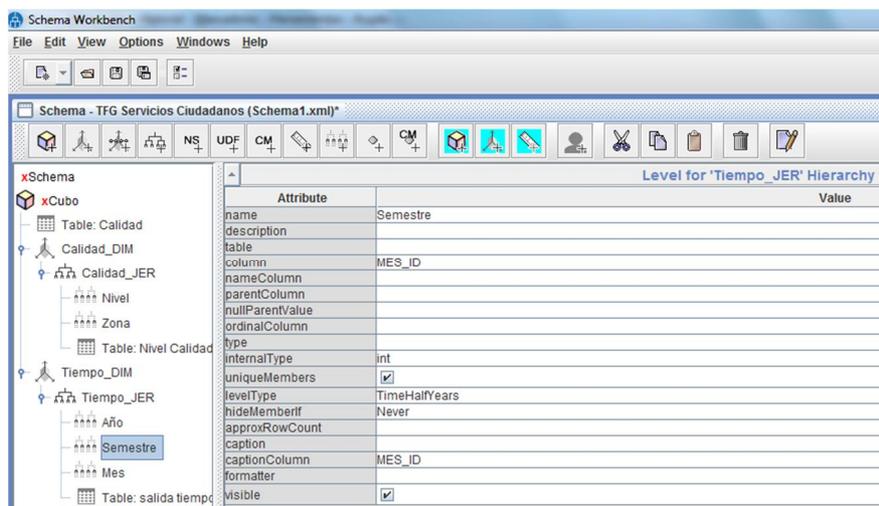
Antes de proceder a crear las medidas, se seleccionan las dimensiones que debe incluir el cubo, con la opción “Add Dimension Usage”. Se añaden todas las dimensiones necesarias. El cubo hereda todas las características que contenga la dimensión, abarcando todas las jerarquías y sus correspondientes atributos.

Se asigna un nombre significativo a la dimensión, y se indica el tipo. A continuación, se procede a crear las diferentes jerarquías que contenga la dimensión. Para cada jerarquía, se asigna una tabla de la dimensión, y a continuación se generan los *levels* o niveles que componen la jerarquía.

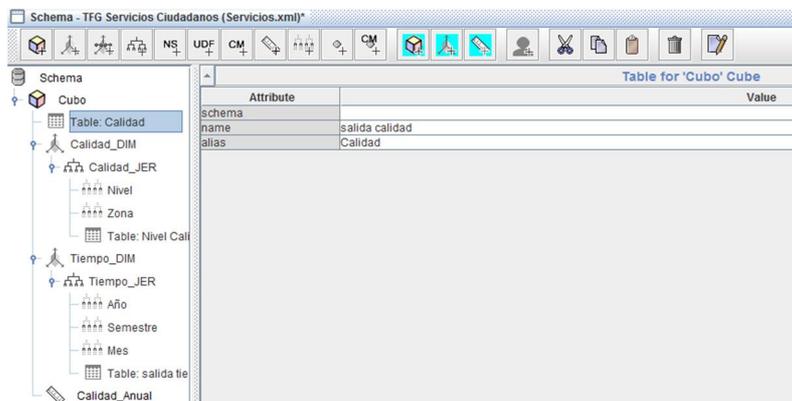
La siguiente imagen, muestra cómo se va desarrollando la creación del cubo, añadiendo las dimensiones necesarias, con las jerarquías y sus correspondientes niveles:



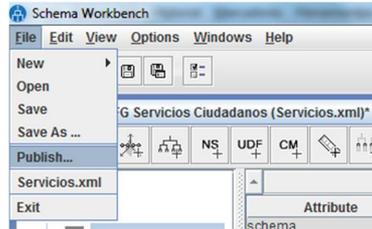
A continuación, se muestra mediante captura la dimensión tiempo. Esta dimensión es básica en la creación de cubos, ya que se considera indispensable para la realización de análisis, con independencia del área de aplicación.



Una vez creadas las dimensiones consideradas oportunas para llevar a cabo el posterior análisis, y como último paso en la creación del cubo, se definen las Métricas que van a ser los valores que forman parte del susodicho análisis. En el proyecto actual, y a modo de contacto con el funcionamiento de la herramienta, se elabora una Métrica para calcular la evolución de la calidad de la limpieza por zonas de una ciudad, teniendo en cuenta el factor tiempo entre los años 2012 y 2014 ambos incluidos.



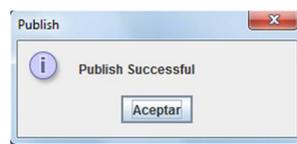
Llegados a este punto, se publica en el *Pentaho Server*, seleccionando la opción “Publish” del menú del *Schema Workbench*:



Simplemente se trata de introducir el enlace y las credenciales de Pentaho:



Mediante una ventana, se muestra que la publicación ha sido llevada a cabo con éxito.



### 6.2.1 Visualización del Cubo OLAP

Con la intención de verificar el correcto diseño del cubo y su posterior publicación, Pentaho dispone de herramientas como *JPivot*.

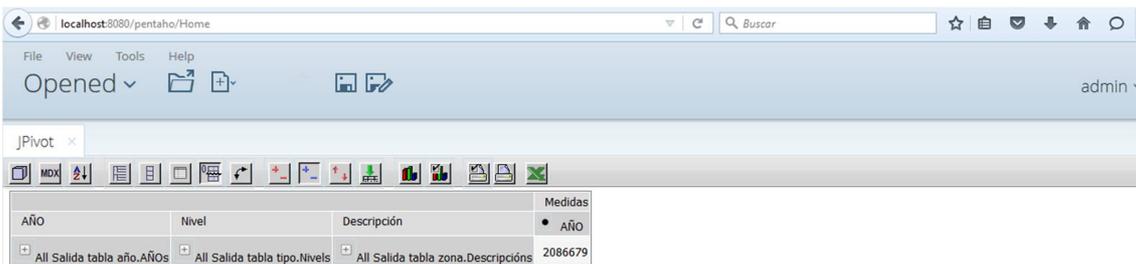
Para acceder a *JPivot*, dentro del servidor Pentaho, se debe hacer clic sobre el botón con el nombre “JPivot View” que hay incluido dentro del menú “Create New”.



A continuación se muestran algunos de los esquemas que han sido creados para el desarrollo del actual Proyecto durante las distintas fases de pruebas efectuadas.



Se selecciona el esquema y cubo que se corresponde a los datos que desean visualizar, y acto seguido se muestra que el proceso de la elaboración del cubo se ha llevado a cabo correctamente:

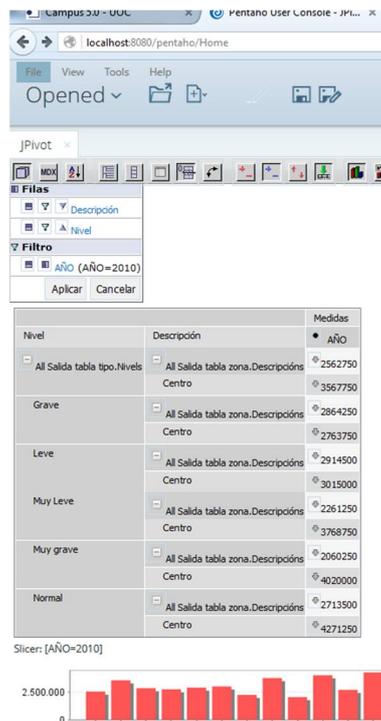


En la imagen anterior, se pueden observar las diferentes dimensiones que han sido empleadas durante la creación del Cubo. Partiendo de dichas dimensiones, el menú de *JPivot* permite, entre otras interesantes opciones, la posibilidad de ver las sentencias de las que se compone el cubo en formato MDX, ordenar la información, mostrar los datos que serán motivo de análisis en gráficas o importar a formato PDF/Excel.

En este sentido, *JPivot* también permite la presentación de la información desde el navegador OLAP.

Año	Nivel	Descripción	Medidas
2010	Grave	All Salida tabla zona.Descripcio	2086679
		Centro	2708315
		Districtos Centro	2615985
		Districtos afueras	2535490
	Leve	All Salida tabla zona.Descripcio	2384955
		Centro	2565425
		Districtos Centro	3420950
		Districtos afueras	3170925
	Muy Leve	All Salida tabla zona.Descripcio	2082625
		Centro	2565700
		Districtos Centro	2414750
		Districtos afueras	2666275
Muy Grave	All Salida tabla zona.Descripcio	2766400	
	Centro	2033400	
	Districtos Centro	2766850	
	Districtos afueras	2314300	
Normal	All Salida tabla zona.Descripcio	2515775	
	Centro		
	Districtos Centro		
	Districtos afueras		

A partir de los campos seleccionados que forman el Cubo OLAP, se puede representar la información en tablas que ayuden a su análisis.

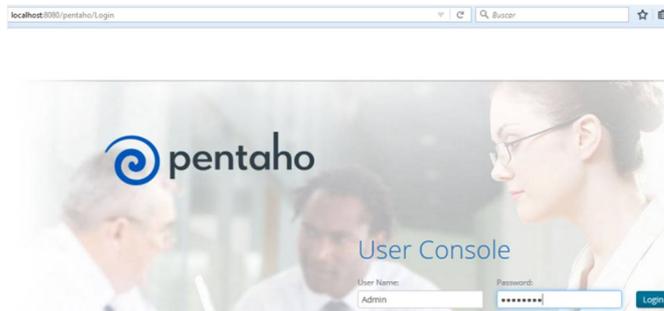


### 6.3 CDE Dashboard

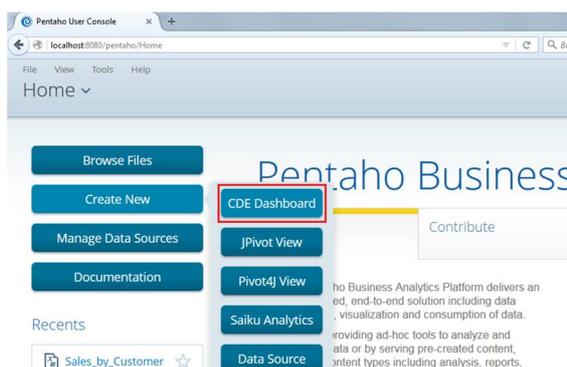
En el actual apartado, se prosigue con la fase de creación de herramientas de inteligencia de negocio. Concretamente, se trata de elaborar un *Dashboard*. Mediante la creación del tablero, se pretende manipular los datos de una manera intuitiva, y que ayude a las organizaciones a las que va orientado el presente Proyecto a la correcta toma de decisiones.

Siguiendo las pautas empleadas en los puntos anteriores, y que afectan a las diferentes fases de instalación, configuración o creación de herramientas BI, se tienen en cuenta los puntos clave en el desarrollo del tablero, proporcionando además una serie de imágenes que ayudan a su apropiada comprensión.

En primer lugar, se accede al portal al servidor BI de Pentaho. Para ello, se debe introducir en la barra de navegación “localhost:8080/pentaho” y luego se introduce el usuario (“Admin”) y la contraseña (“password”).



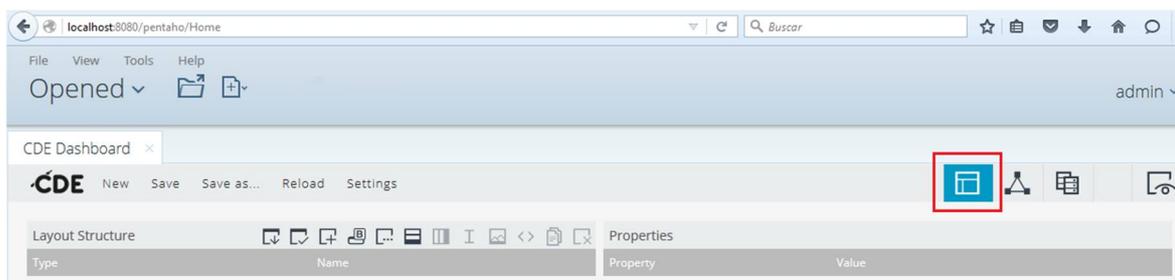
Una vez en la consola de usuario, se selecciona “CDE Dashboard” dentro del menú “Create New”.



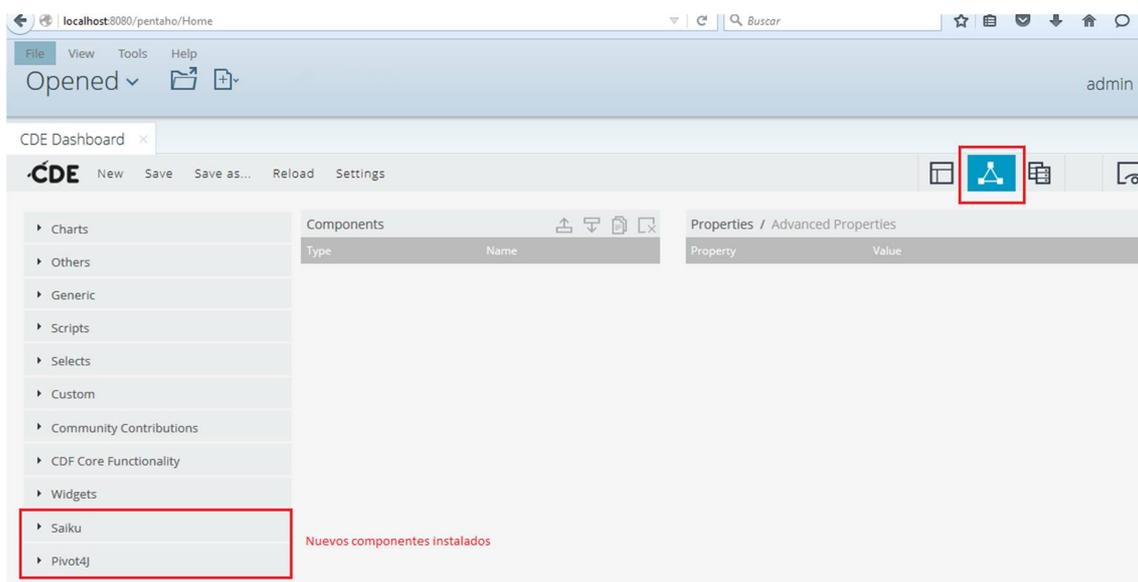
El entorno de trabajo sobre el que se irá desarrollando la aplicación, incluye tres tipos de paneles, los cuales que deben ser debidamente combinados para obtener una herramienta.

En este sentido, y con la intención de ofrecer una descripción de cada panel de los componen la interfaz de trabajo *del CDE Dashboard en Pentaho*, se propina una explicación de cada uno de ellos.

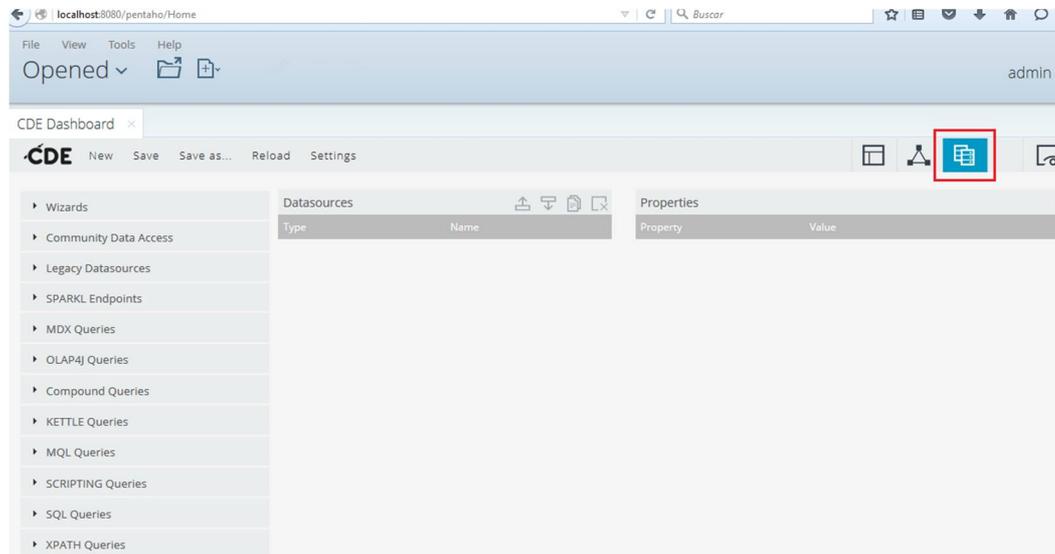
- *Layout Panel*: Se trata del panel dedicado al diseño. En dicho panel, se pueden realizar diferentes modificaciones sobre el aspecto final con el que será presentado el Dashboard.



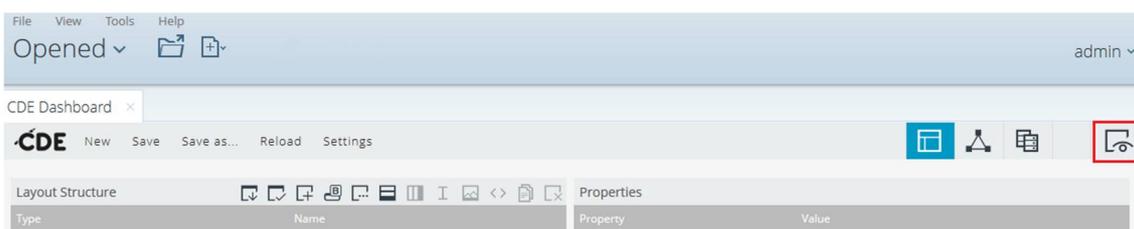
- *Components Panel*: Es el panel que contiene los componentes con los que se trabajará. En la parte izquierda del panel se aprecian dichos componentes. Además de los componentes que Pentaho incorpora, también aparecen los que se han instalado a posteriori.



- *Datasources Panel*: Como su nombre indica, incluye las fuentes de las que se obtienen los datos para la creación del tablero.



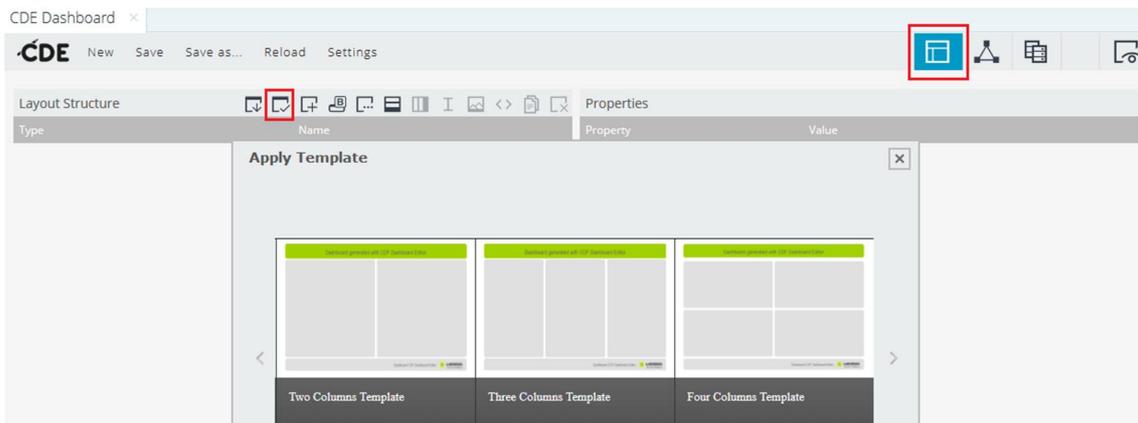
Además de los paneles explicados anteriormente, se debe comentar que pulsando sobre el icono de la parte superior derecha (“Preview your Dashboard”), es posible ver la evolución del proyecto, y por tanto, efectuar las modificaciones que se consideren oportunas para obtener el resultado deseado.



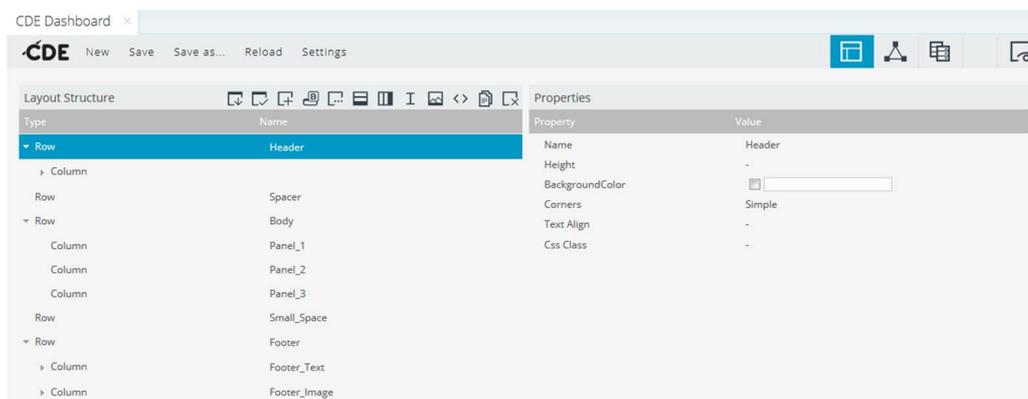
Una vez se está familiarizado con el entorno gráfico sobre el que se trabajará, se procede al desarrollo de cada uno de los paneles que forman parte en la construcción del *Dashboard*.

### 6.3.1 Layout Panel

Existen diferentes combinaciones para mostrar la salida de la información en forma de plantillas prediseñadas. Estas pueden verse al hacer clic sobre el botón “Apply Template”. Se considera la manera más cómoda y sencilla de empezar a desarrollar el *Dashboard*. Obviamente, también es posible plantear el diseño desde cero, y hacerlo totalmente personalizado.

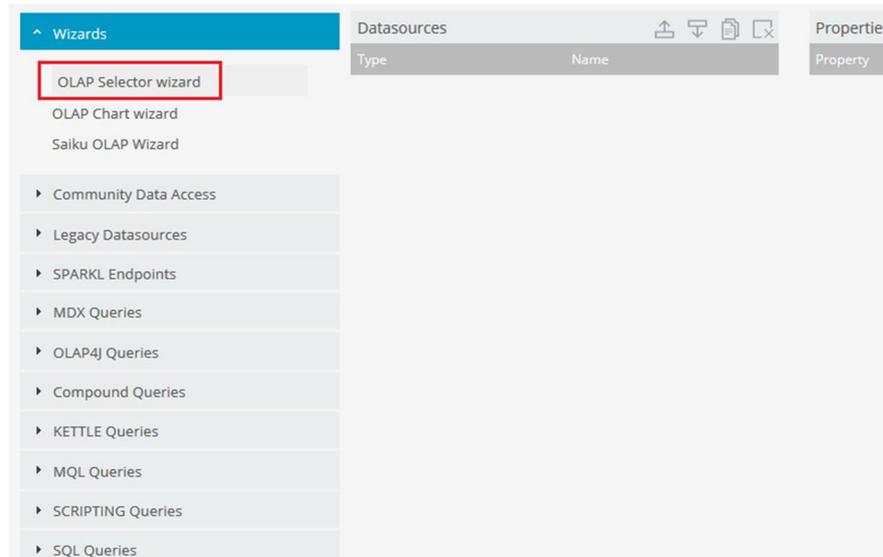


Se escoge el *Template* que mejor se adapte en función de los requisitos del proyecto sobre el que se proyecta el diseño. En el actual Trabajo, el autor considera que una plantilla de tres columnas es lo más recomendable. Basada en esta selección, se genera automáticamente una estructura basada en filas y columnas, que posteriormente se puede personalizar para darle una apariencia en consonancia con el tipo de corporación que vaya a utilizar la herramienta BI.

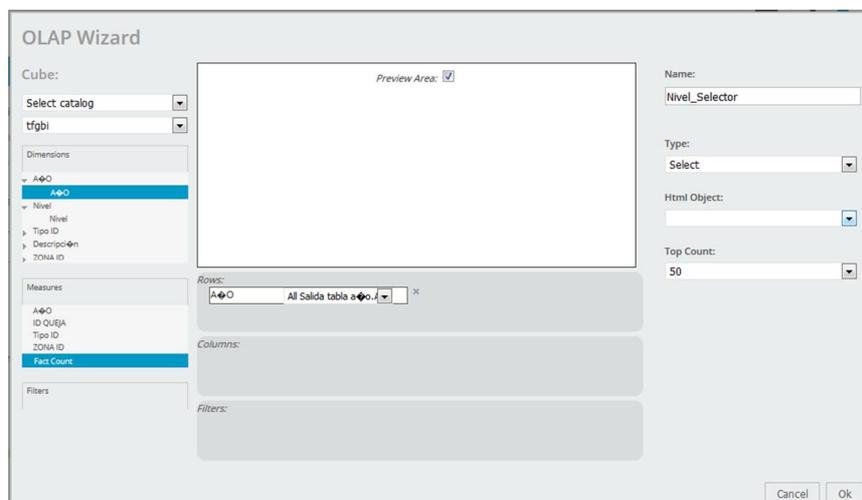


### 6.3.2 Datasources Panel

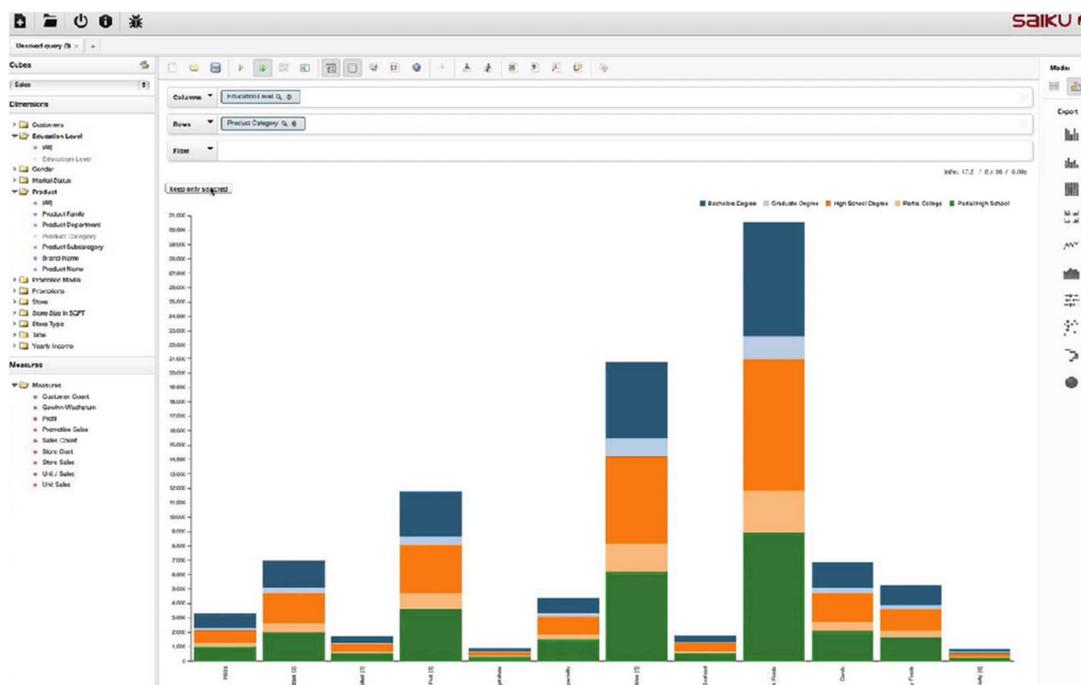
En este panel, se tratan de definir los selectores y las consultas que forman los cimientos sobre los que se construyen los gráficos. Para crear un selector pulsaremos en la parte izquierda sobre OLAP Selector Wizard.



Posteriormente aparece una pantalla que incorpora las opciones para la creación de selectores. En dicha pantalla se debe seleccionar el cubo y la dimensión sobre la que se pretenda trabajar.



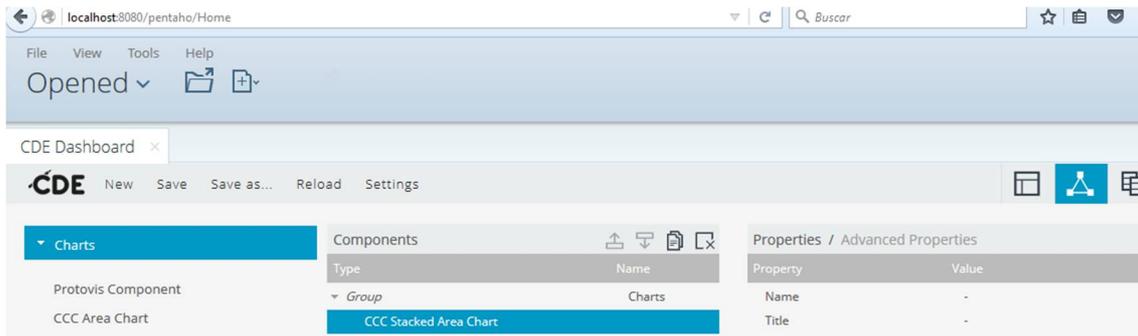
En la pestaña Datasources, se crean las consultas que se consideren necesarias, y que constituyen la base de los gráficos del *Dashboard*. Para llevar a cabo dicho proceso, se hace uso de la herramienta *Saiku Analytics* que ha sido instalada en una fase anterior de este mismo Proyecto. Como se ha comentado, ofrece un entorno intuitivo y atractivo para la realización de consultas y la posterior generación de gráficos asociados a dichas consultas. La siguiente imagen, muestra una salida de gráficos obtenidos mediante Saiku.



### 6.3.3 Components Panel

En el panel de componentes, lo que se pretende es añadir los gráficos del tipo que se deseen, utilizando como bases datos obtenidos como resultado de las consultas realizadas.

Para construir el gráfico de línea del presente Trabajo, se debe agregar un nuevo "CCC Stacked Area Chart". Para ello, se debe seleccionar del conjunto de *Charts* disponibles en Pentaho para la creación de un *Dashboard*.



Para rellenar los campos del *Chart*, hay que prestar especial interés a los escuchadores o *listeners* del gráfico, encargados de detectar cambios en los parámetros de interés para automáticamente modificar las gráficas que estén asociados a estos.

Acto seguido, se observa el resultado de las pruebas que se han llevado a cabo durante el desarrollo del Proyecto. Para ello, primero se debe seleccionar el “Preview your Dashboard”.



Posteriormente, se observan los resultados de las herramientas BI creadas.



## 7. Conclusiones y líneas de futuro

En líneas generales, estoy muy satisfecho con el Trabajo realizado. Desde un punto de vista teórico, he podido comprender muchos de los aspectos que envuelven al universo de la inteligencia de negocio y a los que anteriormente les asignaba definiciones poco claras y carentes de precisión. No obstante, y sin lugar a duda, la parte práctica del Proyecto ha sido la que me ha dotado de mayor entendimiento sobre la materia. Elaborar desde cero herramientas de *Business Intelligence* sobre una plataforma OSBI como Pentaho, me ha generado muchos obstáculos que se han convertido en conocimiento cuando eran superados.

Respecto a los objetivos planteados, se han cubierto prácticamente en su totalidad, ya que la evolución de las herramientas, y su planteamiento en distintos escenarios dentro del mismo sector, han aportado mayor conocimiento del esperado en el momento del inicio del Proyecto. Hay que tener en cuenta que al tratarse de un Proyecto focalizado a todo un sector, no se disponían de los datos concretos de una compañía en particular, y por tanto, los resultados mostrados están basados en supuestos y tablas con datos ficticios. El resto de pruebas se ha decidido que queden al margen de la memoria del Trabajo.

La intención es demostrar el alcance de estas herramientas, y por tanto, una vez valorada la capacidad que poseen, poder aplicarse sobre cualquier organización dedicada a la prestación de Servicios Ciudadanos. Obviamente, cambiando los indicadores y las fuentes de datos, se puede aplicar a cualquier tipo de negocio, independientemente de su naturaleza. Reconocer que si se hubiese trabajado con datos fidedignos, y los análisis hubieran sido más precisos.

En lo referente a la metodología, el haber iniciado el Trabajo en base a una metodología como el *PMBOK*, ha proporcionado rigurosidad sobre el seguimiento de las pautas establecidas inicialmente. No obstante, y debido a la condición de esta tipología de proyectos, se pueden considerar las metodologías ágiles como una alternativa apropiada.

Se han introducido cambios respecto a la planificación inicial, pero que no han sido lo suficientemente determinantes como para variar el rumbo definido del Proyecto.

Es evidente que iniciarse en el *Business Intelligence*, y tratar de crear herramientas sobre un software del que no se tiene conocimiento previo, es una tarea compleja. A su vez, el reto de explorar un campo nuevo, ha supuesto una motivación extra por parte del autor del presente Proyecto. Por tanto, soy consciente de la gran cantidad de aspectos que han quedado pendientes de ser explorados en un futuro próximo. El presente Proyecto es la entrada a un nuevo mundo del que espero seguir disfrutando.

Simplemente deseo que llegue el día en que relea el presente Trabajo, y pueda detectar sin demasiado esfuerzo las carencias existentes y mejoras aplicables. Sera una señal de que mi avance en esta materia prospera favorablemente.

## **7. Glosario**

A continuación se listan algunos de los términos y acrónimos considerados más relevantes por parte del autor, que han sido utilizados durante la elaboración del Trabajo.

BI: *Business Intelligence*. Abarca todo el conjunto de técnicas y herramientas necesarias para poder transformar los datos e informaciones de una empresa en conocimiento con el cual poder tomar decisiones estratégicas.

Conocimiento: Es posible transformar la información en conocimiento cuando se dispone de unos datos previos sobre el mismo suceso, experiencia.

Dato: Es una representación simbólica de una variable cuantitativa. Un dato en sí mismo no ofrece ningún tipo de información.

ETL: *Extract, Transform and Load*. Proceso mediante el cual se extraen datos desde cualquier fuente de datos de la organización, se transforman como sea requerido y posteriormente se cargan con la intención de ser analizados.

Información: Cuando es posible obtener un conjunto organizado de los datos y así obtener algún tipo de conocimiento acerca de ellos se puede hablar de información.

OLAP: *Online Analytical Processing*. Se puede definir como una base de datos multidimensional en la cual el almacenamiento físico de los datos se realiza también en un vector multidimensional.

## 8. Referencias

- «Sistemas de Información BI: Estado Actual y Herramientas de Software Libre» [En línea]
- [http://openaccess.uoc.edu/webapps/o2/bitstream/10609/8175/1/Sduras\\_TFM\\_0611.pdf](http://openaccess.uoc.edu/webapps/o2/bitstream/10609/8175/1/Sduras_TFM_0611.pdf) [Último acceso: Noviembre 2015].
- «Palo» [En línea]. Available: <http://sourceforge.net/projects/palo/>. [Último acceso: Diciembre 2015].
- «MySQL» [En línea]. Available: <https://www.mysql.com/>. [Último acceso: Diciembre 2015].
- «MySQL connector Java» [En línea]. Available: <http://dev.mysql.com/downloads/connector/j/>. [Último acceso: Diciembre 2015].
- «Pentaho Community Edition» [En línea]. Available: <http://community.pentaho.com/>. [Último acceso: Enero 2016].
- «Pentaho [En línea]» <http://www.pentaho.com> [Último acceso: Enero 2016].

- «Documentación Pentaho Community» [En línea]. Available: <http://community.pentaho.com/>. [Último acceso: Enero 2016].