

CAPÍTULO III

Herramientas y suites de inteligencia empresarial

El objetivo de este capítulo es mostrar las herramientas y el software existente para llevar a cabo el proceso de inteligencia empresarial. El capítulo se encuentra organizado de la siguiente manera: en la sección 3.1 se mencionan las herramientas que existen para las tecnologías de *datawarehousing*, OLAP y *data mining*, en la sección 3.2 se mencionan los sistemas que implementan todo el proceso de inteligencia empresarial y por último en la sección 3.3 las conclusiones del capítulo.

En la actualidad se han desarrollado diversos sistemas que ofrecen soluciones a los problemas de la inteligencia empresarial y que implementan todas o algunas de sus tecnologías. Para poder entender mejor el conjunto de aplicaciones para la inteligencia empresarial, en esta tesis se ha realizado una clasificación de dichos sistemas, la cual se muestra en la tabla 3.1 en donde los sistemas se dividen en 2 clases:

Clasificación	Definición
Herramientas simples	Herramientas especializadas en una sola técnica de inteligencia empresarial
Suites	Ambientes completos que ofrecen la implementación de las 3 técnicas de inteligencia empresarial: <i>datawarehousing</i> , OLAP y <i>data mining</i> .

Tabla 3.1 Clasificación de sistemas de inteligencia empresarial

A continuación veremos las herramientas y posteriormente las suites más populares que existen en el mercado de cada una de las clasificaciones propuestas, junto con sus características principales.

3.1 Herramientas simples

Las herramientas simples son aquellas que dan soporte a las tecnologías de la inteligencia empresarial individualmente. Estas herramientas han sido clasificadas a su vez de acuerdo a las tecnologías de inteligencia empresarial a la que dan soporte, en seguida mencionaremos las más importantes y populares para cada una.

3.1.1 Herramientas de integración de datos

En la figura 3.1 vemos algunas de las herramientas más populares que apoyan la integración de datos, las cuales se describirán con más detalle en seguida.



Figura 3.1 Herramientas de integración de datos (ETL)

3.1.1.1 Clover ETL

Es un ambiente de transformación de datos de código abierto basado en Java, para datos estructurados, capaz de funcionar como aplicación independiente (*standalone*) o estar incluida en otra aplicación [Clover, 2003].

Características principales:

- Maneja todas las bases de datos con conexión JDBC disponible para ellas.
- Realiza transformaciones gráficas basadas en XML para la descripción de los metadatos de los registros.
- Soporta valores nulos.
- Puede ejecutarse en múltiples CPU's utilizando una estrategia llamada "pipeline-parallelism".
- Se distribuye bajo la licencia LGPL, la última versión 3 (al momento de escribir este documento), se liberó el 29 de junio del 2007.
- CloverETL soporta 4 diferentes tipos de datos: string, numeric, date, bytes
- Arquitectónicamente, está conceptualizado en unidades lógicas separadas llamadas unidades de transformación que engloban funcionalidades de la transformación e inteligencia, cada una de las cuales puede ser utilizada como componentes de una aplicación independiente en otras aplicaciones y servicios. Cada componente corre como un hilo de ejecución separado, creado un ambiente más tolerante a fallas.

3.1.1.2 EnhydraOctopus

Conocida también como Octopus [Enhydra, 2007], es una herramienta de ETL para transformaciones de datos por conexión JDBC [JDBC, 2007]. Octopus solo soporta fuentes de datos que vengan con el manejador JDBC, incluye también drivers especiales que permiten la conectividad con archivos CSV, XML, MS-SQL y archivos de propietarios. Octopus utiliza

archivos XML para cargar los trabajos, así como para definir los parámetros de las transformaciones dadas.

Características principales:

- La característica principal de Octopus es el requerimiento de que para acceder a las fuentes de datos que lo utilicen, debe existir un manejador JDBC disponible.
- Herramienta poderosa capaz de: normalizar datos, crear llaves artificiales, tablas y llaves primarias.
- Todos los trabajos que corren en Octopus son independientes de los proveedores de bases de datos.

3.1.1.3 Kettle

Conocido actualmente como *Pentaho Data Integration*, es un proyecto belga de código abierto, ahora adoptado por Pentaho BI, (del cual hablaremos con más detalle en la sección de suites), que incluye un conjunto de herramientas para realizar ETL. Uno de sus objetivos es que el proceso de ETL sea fácil de generar, mantener y desplegar [Kettle, 2007].

Se compone de 4 herramientas:

- SPOON: permite diseñar de forma gráfica la transformación ETL.
- PAN: ejecuta un conjunto de transformaciones diseñadas con SPOON, conocidas como trabajos (*jobs*), creando dependencias entre dichas transformaciones.
- CHEF: permite, mediante una interfaz gráfica, diseñar la carga de datos incluyendo un control de estado de los trabajos.
- KITCHEN: permite ejecutar los trabajos batch diseñados con Chef.

Características principales:

- Funciona en Windows, Unix and Linux.
- Tiene una interfaz visual con indicadores de las transformaciones.
- Es una aplicación escrita en Java con algunas características avanzadas escritas en JavaScript.
- Ofrece una licencia pública GPL.
- Basado en metadatos.
- Como soporte se encuentran los foros de Pentaho y la comunidad Pentaho.
- Soporta Oracle, DB2, SQL Server y Sybase así como MySQL, Postgres, Hypersonic, FireBird SQL e Ingres. También soporta la conectividad con SAP con una tarifa de licencia.
- Con respecto a escalabilidad, soporta la arquitectura de procesamiento en paralelo para distribuir las tareas de ETL a través de múltiples servidores.

3.1.1.4 Talend

Talend Data Integration [Talend, 2007], es una herramienta de ETL de código abierto, que efectúa operaciones tales como alimentación de un *datawarehouse*, sincronización de bases de datos, transformación y verificación de la integridad de datos. Su interfaz gráfica de usuario está basada completamente en Eclipse RCP (Rich Client Platform) e incluye numerosos componentes para procesos de modelado de negocios, así como implementaciones técnicas para extracción de información, transformación y mapeo del flujo de datos. Las funciones generales de Talend incluyen modelador de negocios, diseñador del trabajo y administrador de metadatos.

Características principales:

- Compatible con Microsoft SQL Server Integration Services (SSIS).
- Funciona en Windows, Unix y Linux.
- Interfaz de diseño *drag and drop* para componentes, conectores y relaciones.

- Genera componentes en Perl, Java o SQL para ETL o transformaciones ETL.
- Tiene una comunidad de soporte en Talend wiki, Talend Forum y un rastreador de errores.
- Conectividad con Oracle, DB2, MySQL, Sybase y Postgres y conectividad ODBC para otras bases de datos.

En la siguiente sección veremos las herramientas que existen para OLAP.

3.1.2 Herramientas para OLAP

En la figura 3.2 se muestran algunos de los proveedores de OLAP más populares que se describen más detalladamente a continuación.



Figura 3.2 Herramientas de OLAP

3.1.2.1 Mondrian

Mondrian [Mondrian, 2007], es un motor de ROLAP desarrollado en Java, que permite analizar grandes conjuntos de datos que se encuentran almacenados en el *datawarehouse*.

Mondrian se considera un motor, porque que se encarga de recibir consultas dimensionales en lenguaje MDX y devolver los datos del cubo que correspondan a la consulta. El cubo se representa como un conjunto de metadatos que definen cómo se han de mapear estas consultas dimensionales a sentencias SQL para obtener de la base de datos la información necesaria para satisfacer la consulta dimensional. Utiliza una memoria caché para almacenar los resultados de las consultas que se acceden múltiples veces.

Mondrian es usado para:

- Alto desempeño, análisis interactivo de grandes o pequeños volúmenes de información.
- Exploración dimensional de los datos, por ejemplo analizando ventas por marcas de productos, región o periodo de tiempo.
- Parseo de expresiones en lenguaje MDX a expresiones en SQL para recuperar respuestas a consultas dimensionales.
- Cálculos avanzados utilizando las expresiones de cálculo del lenguaje MDX.

Para definir el esquema que utilizará Mondrian existe una herramienta denominada CubeDesigner que se explica a continuación con más detalle.

3.1.2.1.1 *Cube Designer*

Cube Designer [Mondrian, 2007] es un ambiente gráfico que permite diseñar un documento XML que representa un cubo o hipercubo para analizar la información almacenada de forma multidimensional. Este esquema será interpretado por Mondrian para obtener la información de las consultas de MDX ejecutadas.

Un esquema define una base de datos multidimensional. Contiene un modelo lógico que consiste en cubos jerarquías, niveles, miembros y el mapeo de este modelo a un modelo físico, que es la fuente de datos que está representada a través del modelo lógico. Típicamente este modelo es un esquema estrella, que contiene un conjunto de tablas en una base de datos relacional.

Comentario [d1]: incluir diagrama bd → xml →mondrian

Para acceder a las funcionalidades que nos brinda Mondrian, se requiere de un cliente. Los clientes disponibles y más populares que existen se explican en seguida.

3.1.2.1.2 JPivot

JPivot [JPivot, 2003] es una librería de *Java Server Pages* (JSP) [JSP, 2007] personalizados que presentan tablas de OLAP y permite a los usuarios llevar a cabo operaciones típicas como, *slice & dice*, *drill down* y *roll up* y las consultas OLAP pueden realizarse por medio del lenguaje MDX.

Está diseñado para trabajar con diversos motores de OLAP, especialmente con XMLA y Mondrian. Por lo mismo, JPivot no utiliza las APIs de Mondrian directamente sino que implementa su propio modelo OLAP.

Comentario [d2]: screenshot

3.1.2.1.3 JRubik

JRubik [JRubik, 2005] es un cliente OLAP realizado en Java/Swing y basado en los componentes de JPivot.

Este cliente es capaz de conectar fuentes OLAP basadas Mondrian, aún no se implementa la conexión con otros motores, sin embargo está considerada para realizarse en un futuro.

Las consultas OLAP también pueden realizarse por medio del lenguaje MDX. Sus componentes principales son: navegador OLAP, gestor de consultas MDX, visor de tablas, visualizador de gráficos, visor de mapas, menús y marcadores, datos estadísticos. Estos componentes están basados en componentes de JPivot, la diferencia radica en el cambio de la interfaz web basada en JSP por una hecha en Swing.

Además tiene otros componentes interesantes que no tiene JPivot como: los menús y los marcadores que almacenan el estado de los componentes para una consulta dada, así como el visor de mapas que está basado en otro proyecto abierto de visualizador GIS llamado *Geoclient* que permite generar mapas en SVG (*Scalable Vector Graphics*) que incorporan los resultados de una consulta OLAP.

Comentario [d3]: screenshot

3.1.2.2 Jedox PALO

Jedox Palo [Jedox, 2007] es un servidor de bases de datos multidimensional capaz de centralizar y administrar casi un número infinito de hojas de cálculo. El sistema opera en tiempo real, soporta la consolidación de jerarquías así como numerosas funciones de inteligencia empresarial y es un servidor de código abierto.

Palo es un servidor de datos multidimensional (MOLAP) orientado a celdas, específicamente desarrollado para almacenamiento y análisis de datos en hojas de cálculo.

Comentario [d4]: screenshot
<http://www.jpalo.com/Screens/PaloWebClient.jpg>

3.1.2.3 OLAP4J

OLAP4J (Online Analytical Processing for Java) [OLAP4J, 2007] es una interfaz de aplicación para el ambiente Java 2 Platform, Enterprise Edition, que soporta la creación, almacenamiento y administración de datos para una aplicación OLAP. Hyperion, IBM, y Oracle iniciaron el desarrollo de OLAP4J con la intención de que fuera un equivalente a la conexión JDBC pero específica para OLAP, que se liberará en diciembre 2007 y es compatible con Mondrian.

3.1.3 Herramientas para data mining

En la figura 3.3 vemos algunos de los proveedores más populares para data mining. Estos y otros más se describen con más detalle en seguida.



Figura 3.3 Herramientas de data mining

3.1.3.1 Clementine

La empresa SPSS, además de su conocido paquete estadístico, distribuye Clementine [Clementine, 2007], uno de los sistemas de *data mining* más populares de mercado. Se trata de una herramienta visual inicialmente desarrollada por ISL (*Integral Solutions Limited*). En la actualidad esta herramienta, comercializada por SPSS, posee una arquitectura distribuida (cliente/servidor).

Sus características principales son:

- **Acceso a datos:** fuentes de datos ODBC, tablas Excel, archivos planos ASCII y archivos SPSS.
- **Preprocesado de datos:** pick & mix, muestreo, particiones, reordenación de campos, nuevas estrategias para la fusión de tablas, etc.
- **Técnicas de aprendizaje:** árboles de decisión, redes neuronales, agrupamiento, reglas de asociación, regresión lineal y logística, combinación de modelos.
- **Técnicas para la evaluación de modelos** guiadas por las condiciones especificadas por el experto.
- **Visualización de resultados:** ofrece un potente soporte gráfico que permite al usuario tener una visión global de todo el proceso, que comprende desde el análisis del problema hasta la imagen final del modelo aprendido.
- **Exportación:** generación automática de informes (HTML y texto), volcado de los resultados del ejercicio de *data mining* en bases de datos, exportación de los modelos a distintos lenguajes (C, SPSS, HTML, estándar PMML, SQL para árboles de decisión y reglas).

3.1.3.2 Kepler

Kepler [Kepler, 2007] es un sistema desarrollado por un grupo de aprendizaje automático de *GMD (German General Research Center for Information Technologie)* y transformado en una herramienta comercial distribuida por Dialogis. Kepler está dotado de múltiples modelos de análisis y en su diseño se contemplan aspectos tales como la flexibilidad y la extensibilidad. Entre las herramientas de aprendizaje que proporciona Kepler se pueden encontrar: árboles de decisión, redes neuronales, regresión no lineal, vecinos más próximos, algoritmos multirrelacionales y utilidades estadísticas.

Cubre un amplio rango de problemas propios de *data mining*, que van desde la búsqueda de patrones en un conjunto de datos y agrupación de objetos, hasta la clasificación y predicción. Todo el sistema se maneja mediante una interfaz gráfica programada en Java, la cual permite:

- ***Preprocesado de datos.***
- ***Especificación de un experimento:*** elección de un modelo, instancias de sus parámetros, etc.
- ***Manipulación de la representación gráfica de los modelos obtenidos:*** aunque los resultados obtenidos pueden exportarse a otros entornos o herramientas de análisis o de visualización más específicas.

También dispone de un lenguaje de *script* que permite agilizar el diseño de los experimentos. Los datos de entrada pueden ser importados desde diversos formatos, entre los que se incluye el acceso a la información estructurada en bases de datos.

3.1.3.3 MLC ++

MLC++ (*Machine Learning library in C++*) [MLC, 2007] es un conjunto de librerías y utilidades para facilitar las tareas de probar y comparar la eficiencia y resultados proporcionados por diversos algoritmos sobre un mismo problema. Fueron desarrolladas originalmente, hasta la versión 1.3x y de dominio público, por la Universidad de Stanford. Las versiones posteriores se distribuyen por *Silicon Graphics* bajo dominio de investigación a excepción de las versiones 1.3x que aún son distribuidas bajo licencia de dominio público y por la misma empresa.

A nivel general se caracteriza por:

- **Acceso a datos:** archivos con formato plano siguiendo el estilo de los archivos del repositorio UCI.
- **Transformaciones de datos:** se dispone de un número de transformaciones sobre datos muy reducido.
- **Modelos de aprendizaje:** éstos se encuentran encapsulados mediante objetos. Las técnicas implementadas sobre las tablas de árboles de decisión perezosos, árboles de decisión con opciones, métodos bayesianos y el perceptrón.

3.1.3.4 WEKA

Weka es una herramienta de código abierto desarrollada por un equipo de investigadores de la universidad de Waikato en Nueva Zelanda [Weka, 2005] [Witten, et.al., 2000], implementada en Java. Sus características principales son:

- **Acceso a datos:** los datos son cargados desde un archivo en formato ARFF [Witten, et.al, 2000] (archivo plano organizado en filas y columnas). El usuario puede

observar en sus diferentes componentes gráficos, información de interés sobre el conjunto de muestras.

- **Preprocesado de datos:** existe una gran cantidad de filtros disponibles tales como: selección de atributos, discretización, tratamiento de valores desconocidos, transformación de atributos numéricos.
- **Métodos y técnicas:** árboles de decisión, tablas de decisión, vecinos más próximos, máquinas vectores de soporte, reglas de asociación, métodos de agrupamiento y modelos combinados.
- **Visualización:** la interfaz gráfica se compone de diversos entornos:
 - **Explorer:** que permite controlar todas las operaciones mencionadas anteriormente (filtrado, selección y especificación del método, diseño de experimentos, entre otros).
 - **El entorno de consola:** da la posibilidad de invocar textualmente las operaciones anteriores. También se puede acceder directamente a los métodos que implementan dichas tareas e incorporarlos en el código fuente de la aplicación que se esté programando.
 - **Experimenter:** facilita el diseño y la realización de experimentos complejos.
- Puede utilizarse como **aplicación independiente** o dentro de otra aplicación.

3.1.3.5 Xelopes

Xelopes (*eXtEnded Library fOr Prudsys Embedded Solutions*) [Xelopes, 2002] es una librería con licencia pública GNU para el desarrollo de aplicaciones de *data mining*, implementada por *Proudsys AG* en colaboración con *Russian MDA specialist Zsoft Ltd*. La librería implementa de una manera eficiente la mayoría de los algoritmos de aprendizaje, para desarrollar aplicaciones particulares de *data mining*. Además Xelopes es extensible, es decir, el

usuario puede incorporar sus propios métodos a la propia librería. Sus características principales son:

- **Acceso a datos:** existe una clase especial que permite dar uniformidad a todos los modos de accesos de datos permitidos. Así, el usuario puede acceder a archivos .log, archivos de bases de datos o implementar su propio formato de datos.
- **Modelos:** árboles de decisión lineales y no lineales, máquinas de vectores de soporte, redes neuronales, métodos de agrupamiento, métodos de reglas de asociación.
- **Exportación de datos:** existen métodos para exportar los modelos y sus resultados a otros entornos de *data mining*, soportando el estándar PMML.

La librería está desarrollada bajo el estándar MDA (*model drive architecture*), está disponible para C++ y Java, también existe interfaz para CORBA.

3.1.3.6 YALE (Yet Another Learning Environment) Rapid Miner

YALE, mejor conocido como Rapid Miner [RapidMiner, 2007], es un entorno para la realización de experimentos de aprendizaje automático implementado en Java por la Universidad de Dortmund. Las operaciones elementales se encapsulan en los llamados operadores, los cuales son configurables mediante archivos XML. Estos archivos a su vez, pueden ser especificados gráficamente. El sistema incluye operadores para:

- **Importación y pre-procesamiento de datos.**
- **Aprendizaje automático:** máquinas de vectores soporte, árboles de decisión, agrupamiento y algoritmos genéticos.
- **Validación de modelos.**

3.2 Suites

Las suites, son sistemas que ofrecen la implementación de todas las tecnologías de inteligencia empresarial, que hemos dividido en 2 conjuntos: código abierto y comerciales. En la figura 3.4 podemos ver algunas de las más populares que se describen a continuación.



Figura 3.4 Suites de inteligencia empresarial

3.2.1 Suites de código abierto

3.2.1.1 Pentaho BI

Pentaho BI [Pentaho, 2007] es la única suite que existe de código abierto, ya que ha adoptado las herramientas de Kettle, Mondrian y WEKA para conformar una sola plataforma de inteligencia empresarial.

Tiene las funcionalidades de reportes, análisis, *data mining* e integración de datos. Los proyectos involucrados en Pentaho BI son:

- Mondrian: servidor de OLAP.
- JFreeReport: reporteador.
- Kettle: integración de datos (ETL).
- Pentaho: plataforma de inteligencia empresarial.
- WEKA: *data mining*.

Permite a los desarrolladores de Java diseñar componentes que pueden ser rápidamente ensamblados en soluciones de inteligencia empresarial y a los usuarios finales desplegar rápidamente las soluciones existentes de inteligencia empresarial.

A pesar de que Pentaho BI se conforma ya como una suite que satisface las necesidades de inteligencia empresarial, sus componentes no se integran perfectamente ya que surgieron de manera individual. Sin embargo la compañía Pentaho ofrece soporte comercial para todas ellas como conjunto, por eso se considera una suite.

Las siguientes suites que se mencionarán en seguida, son comerciales y a diferencia de Pentaho sus componentes sí se integran perfectamente.

3.2.2 Suites comerciales

La figura 3.5 muestra una manera general de clasificación de las suites de acuerdo a la opinión de Gartner [Schlegel, et.al., 2007] de los principales proveedores de software comercial de inteligencia empresarial.

La figura muestra 4 cuadrantes: los *líderes*, *retadores*, *visionarios* y *aplicaciones de nichos de oportunidad*.



Figura 3.5 Clasificación de suites comerciales [Schlegel, et.al., 2007]

En seguida veremos las características principales de las suites que se encuentran en cada uno de los cuadrantes de la figura.

3.2.2.1 Líderes

Los líderes (*leaders*) son los proveedores razonablemente fuertes en las capacidades de sus plataformas de inteligencia empresarial, y aquellos que pueden ofrecer a las empresas, extensas

implementaciones que soportan amplias estrategias de negocios. Generan soluciones interesantes de negocio, basadas en la viabilidad y capacidad operacional de una plataforma global.

3.2.2.1.1 Business Objects

Bajo la dirección de John Schwarz, Business Object (BO) [BusinessObjects, 2007] expandió su línea de productos para complementar su plataforma de inteligencia empresarial añadiendo módulos de planeación, presupuesto, y la consolidación de aplicaciones en la integración de datos y calidad de sus productos.

Business Objects mejoró su plataforma con su lanzamiento Enterprise XI y la integración de Crystal Reports. Combina una vista enfocada hacia el usuario final, con una flexible administración del sistema, para facilitar la estandarización de todas las implementaciones de inteligencia empresarial de la empresa. BO ofrece:

- Escalabilidad y adaptabilidad en la arquitectura orientada a servicios.
- Vista enfocada hacia el usuario final.
- Fácil uso y soluciones manejables.
- Una amplia gama de servicios al cliente.

3.2.2.1.2 Cognos

Cognos 8 BI [Cognos, 2007] es uno de los productos que tiene un amplio rango de capacidades como: reportes, análisis, paneles, administración de eventos así como la integración de datos, todo provisto en una arquitectura individual.

De fácil integración, despliegue y uso, ofrece un ambiente simplificado que mejora la adopción del usuario, para una mejor toma de decisiones y tiene una escalabilidad de empresas para el desempeño en la administración. Cognos 8 BI ofrece:

- **Reportes:** un solo ambiente propietario con un amplio rango de reportes.
- **Análisis:** explora y analiza grandes volúmenes de datos con tiempos de respuesta rápidos y predecibles.
- **Marcadores:** enlaza iniciativas a estrategias, con métricas y mapa de estrategias.
- **Paneles:** entrega información de diferentes fuentes de datos en un sólo reporte visual.
- **Administración de eventos de negocios:** detección automática y administración de eventos de negocio importantes.
- **Integración de datos:** capacidad de ETL, y entrega un *datawarehouse* listo para el análisis y reporte.

3.2.2.1.3 *Hyperion Solutions*

El software de Hyperion System 9 BI+ [Hyperion, 2007] soporta la estandarización de estrategias de inteligencia empresarial, al enlazar las diferentes necesidades de reporte y análisis de cada usuario de cada departamento, con un sistema sencillo de usar y fácil de mantener. Los consumidores y los generadores de información tienen la flexibilidad de crear, acceder y compartir rápidamente la información vital que necesitan para poder tomar mejores decisiones en menos tiempo. Una interfaz sencilla pero potente permite a los usuarios de negocio con requerimientos mínimos de tecnologías de información para la generación de informes. Hyperion ofrece:

- **Reportes:** un rango completo de funcionalidades de gestión de reporte, combinando tanto la información financiera como la operacional.

- **Cuadros de mando:** todos los usuarios puede acceder a su propio cuadro de mando personalizado para monitorear fácilmente los indicadores clave de rendimiento.
- **Análisis:** incluye presupuestos, predicciones, modelos de rentabilidad y planificación de escenarios. Posee aplicaciones de análisis multidimensional y de *data mining*. Es posible tener acceso a la información y analizarla desde cualquier origen de datos, incluidos los sistemas transaccionales, *datawarehouses* y data marts.

3.2.2.1.4 Oracle

Oracle BI 10g [Oracle, 2007] es la plataforma más completa para la inteligencia empresarial, cubre un amplio rango de necesidades, incluidos los tableros interactivos, el análisis ad-hoc, alertas e inteligencia proactivas, publicación e informes avanzados, análisis predictivo en tiempo real, y análisis de tecnología móvil.

Existen varias versiones de las plataformas:

- **Oracle BI Suite Enterprise Edition (EE):** es una plataforma completa e innovadora que ofrece bases para crear soluciones empresariales desde fuentes de datos heterogéneas para la distribución de datos, con sistemas Oracle y que no son de Oracle.
- **Oracle BI Standard Edition (SE):** incluye Oracle Discoverer, está optimizado para trabajar con los datos y las aplicaciones Oracle, y ofrecer análisis e inteligencia avanzados al menor costo.
- **Oracle BI Publisher:** también denominado Oracle XML Publisher, ofrece la solución más eficiente y escalable para informes y publicaciones, disponible para entornos complejos y distribuidos. Disponible con Oracle BI Suite EE o como solución independiente.
- **Oracle Real-Time Decision:** combina los requerimientos comerciales y de información del cliente para hacer la mejor recomendación en cada interacción con el

cliente y en cada decisión operacional al adaptarse de manera inteligente, la información en constante cambio.

La versión Enterprise cuenta con las herramientas Oracle Data Integrator con las capacidades de ETL para la creación de *datawarehouses*, Oracle OLAP para el análisis de datos multidimensional y Oracle *Data mining* para el análisis y aplicación de algoritmos de *data mining*. Actualmente estas 3 herramientas se distribuyen también por separado.

3.2.2.1.5 SAS Institute

SAS Business Intelligence [SAS, 2007] integra los datos de la empresa y tiene un servicio de análisis y reportes, posee amplias capacidades de inteligencia empresarial y la aplicación de análisis e integración de datos para crear una solución completa y de fácil interacción. SAS BI ofrece:

- **Reportes:** wizards intuitivos, reportes ilimitados, reportes ad-hoc, integración de resultados y procesos analíticos.
- **Consultas y análisis:** consultas para todos los niveles de usuarios basados en sus necesidades y habilidades, y se pueden realizar sobre múltiples fuentes de datos utilizando datos relacionales o multidimensionales.
- **OLAP:** interfaces fáciles de utilizar para construir y mantener los cubos de OLAP. Los cubos pueden ser creados con base en datos almacenados en cualquier base de datos.
- **Visualización:** sumatorias y presentación de datos usando una gran variedad de gráficos.
- **Data mining** con Enterprise Miner.

3.2.2.2 Retadores

Los retadores (*challengers*) ofrecen una amplia funcionalidad en sus plataformas de inteligencia empresarial y están bien posicionados en el mercado. Sin embargo, pueden estar limitados a ambientes específicos o dominios de aplicación y su visión puede estar obstaculizada por la falta de estrategias coordinadas a través de varios productos de sus plataformas de inteligencia empresarial. O puede faltarles el canal de ventas, presencia geográfica y contenido específico de la industria ofrecido por los proveedores en el cuadrante de los líderes.

3.2.2.2.1 *Information Builders*

Information Builders WebFOCUS [InformationBuilders, 2007] es una suite de herramientas, que reduce el costo y tiempo para desarrollar y desplegar las aplicaciones de inteligencia empresarial, mientras provee análisis y reportes escalables y poderosos. WebFOCUS ofrece: integración de datos, consultas y análisis, reportes y visualización de resultados

3.2.2.2.2 *Microsoft*

Microsoft BI Platform [Microsoft BI, 2007], es una plataforma completa para el análisis de datos e inteligencia empresarial, está basada en Microsoft SQL Server y proporciona capacidades de reportes, análisis e integración de datos. Microsoft BI ofrece principalmente:

- ***Microsoft SSIS (SQL Server Integration Services)***: integra datos de diferentes orígenes, construye administra y despliega soluciones de integración escalables y construye o alimenta un *datawarehouse*.

- **Microsoft SSAS (SQL Server Analysis Services):** provee, una vista unificada e integrada para todos los datos de la empresa, reportes, análisis OLAP, marcadores de indicadores de desempeño y *data mining*.
- **Microsoft SQL Server 2005 Data mining:** es un conjunto de capacidades de *data mining* fáciles de usar que permiten análisis predictivos en los datos almacenados.

3.2.2.2.3 SAP

SAP NetWeaver BI [SAP, 2007], contiene una poderosa infraestructura de inteligencia empresarial y comprende un conjunto de herramientas entre las que destacan: planeación, simulación y funcionalidades de *datawarehousing* dirigidos hacia un portal empresarial.

Se tiene la posibilidad de integrar datos de toda la empresa y transformarlos en información práctica y manejable que permita la toma de decisiones y entregue resultados sólidos. SAP BI ofrece:

- **Datawarehousing:** Administración de un *datawarehouse*, modelado del negocio, proceso de ETL para la implementación del *datawarehouse*, modelado de la información de acuerdo a la estructura de la empresa y administración de datos de múltiples fuentes.
- **Análisis:** OLAP, *data mining*, y avisos para proveer el acceso y presentación de datos, búsqueda de patrones e identificación de excepciones.
- **Planeación:** un ambiente de planeación que asegura un flujo de datos, planeación y presupuestos consolidados en los datos de la empresa.
- **Vistas de negocio:** diseño de consultas, análisis, reportes, y una aplicación Web que permite al usuario crear reportes de análisis para la toma de decisiones a cualquier nivel.

3.2.2.3 Visionarios

Los visionarios (*visionaries*) son proveedores que tienen una fuerte visión para distribuir sus plataformas de inteligencia empresarial. Son distinguidos por su apertura y flexibilidad en las arquitecturas de sus aplicaciones, y ofrecen profundidad en la funcionalidad de las áreas que dirigen, pero puede faltarles una amplia funcionalidad de requerimientos.

3.2.2.3.1 *MicroStrategy*

MicroStrategy BI Developer Kit [Microstrategy, 2007] es un paquete de software especialmente diseñado para el rápido desarrollo de aplicaciones de inteligencia empresarial.

Consiste en:

- ***MicroStrategy Desktop***: herramienta de desarrollo y consulta avanzada
- ***MicroStrategy Architect***: herramienta de diseño
- ***MicroStrategy módulos analíticos***

Los módulos analíticos de MicroStrategy son modelos portables de soluciones analíticas para diseñar soluciones a los problemas del negocio. Cada uno de los módulos viene con un modelo lógico, un modelo físico y numerosos reportes e indicadores. Los módulos analíticos de MicroStrategy están bajo el esquema de la portabilidad, que permite trabajar sobre soluciones existentes y escalar sobre inversiones ya realizadas.

3.2.2.3.2 QlikTech

Las características de QlikView 8 [QlikView, 2007] se basan en 4 aspectos fundamentales:

- **Análisis para la acción:** lleva a cabo análisis inteligentes gracias a la creación de campos parametrizados, lo que permite tomar los datos en memoria y que puedan ser tratados como cualquier otro dato, que sean introducidos/editados y distribuidos a lo largo de las agregaciones si es necesario.
- **Análisis colaborativo:** proporciona la inteligencia empresarial de forma colaborativa mediante: objetos de documentos privados y compartidos, marcadores, envío por e-mail de marcadores, análisis desarrollado de forma local y compartido en servidores locales integrando QlikView en aplicaciones cliente como Microsoft Office.
- **Despliegue para grandes grupos de usuarios:** realiza el despliegue a los grupos de usuarios a través de: requerimientos simplificados para QlikView Server y QlikView Publisher, el módulo que automatiza la actualización de los datos y asegura el acceso apropiado para usuarios específicos.
- **Despliegue para grandes grupos de documentos:** sincroniza las versiones de QlikView Server y QlikView Publisher.

QlikView 8 también incluye la visualización de los datos. Los usuarios pueden ver los resultados en 24 estilos distintos de formatos de gráficos, desde gráficos de tipo embudo y pareto hasta mapas térmicos y con múltiples indicadores para mejorar la administración de la información.

3.2.2.4 Nichos de oportunidad

Las aplicaciones que buscan nichos de oportunidad (*niche players*) son aquellas que están enfocadas en dominios específicos o ciertos puntos de la inteligencia empresarial, sin embargo

les falta profundizar en la funcionalidad de otros aspectos y tienen capacidades limitadas para innovar y superar a otros proveedores en el mercado.

3.2.2.4.1 *Arcplan*

Arcplan [Arcplan, 2007] se enfoca principalmente en las soluciones de administración que requieren análisis simultáneo, y múltiples fuentes de datos operacionales y analíticos, con las cuales trabaja muy bien. Destaca su facilidad de uso, y sus capacidades visuales, que ofrecen un enfoque diferente para la conexión a fuentes de datos y para la creación de aplicaciones analíticas, aunque carece de modelos de predicción y de integración. Permite monitorear en tiempo real los procesos de negocio y analizar los impactos de los resultados operacionales y metas estratégicas. Sin embargo, carece de un gran número de referencias para determinadas capacidades clave como metadatos y consultas ad-hoc.

3.2.2.4.2 *Actuate*

Las capacidades principales de Actuate [Actuate, 2007] de escalabilidad en hojas de cálculo y presentación de informes han permitido a la compañía posicionarse con un nicho único en el mercado, especialmente en la industria de servicios financieros. Ofrece una amplia gama de productos de inteligencia empresarial, provee una variedad de formatos intuitivos y personalizables que presentan datos integrados de todas las fuentes relevantes. Sus opciones de presentación son muy flexibles y escalables para otros proyectos grandes o pequeños.

3.2.2.4.3 *Applix*

Applix, actualmente adquirido por Cognos, [Applix, 2007] tiene la plataforma más robusta basada en web para la presentación de informes, consultas ad-hoc y capacidades de tableros de inteligencia empresarial. Ofrece las funcionalidades de presupuesto, predicción, análisis de ventas, administración de inventarios, producción y planeación. Contiene 4 herramientas las cuales son:

- ***Turbo Integrator***: permite manejar la extracción de datos de otros sistemas.
- ***Applix Integra***: permite administrar y monitorear el rendimiento de los procesos de negocio.
- ***Applix Interactive Planning***: administra los procesos de planeación, presupuestos, y predicción de toda la organización.
- ***TMI Web***: fortalece el acceso a los datos desde cualquier parte en la organización.

3.2.2.4.4 *Panorama Software*

Panorama NovaView 5 [Panorama, 2007] es un producto de inteligencia empresarial que incluye funcionalidades analíticas, reportes, paneles, marcadores, visualización avanzada y modelado. También ofrece una solución innovadora que ayuda a los usuarios a enfocarse en los datos relevantes, toma acciones estructuradas y automatiza rutinas integrando la inteligencia empresarial con los procesos del negocio. Panorama ofrece:

- ***NovaView Analytics***: permite realizar consultas rápidas y análisis a través de una interfaz gráfica de fácil de uso.
- ***NovaView Smart Reporting***: ayuda a diseñar y dar formato a los reportes.

- ***NovaView Dashboards***: muestra una vista general del estado de la empresa basada en los indicadores de desempeño, que pueden personalizarse.
- ***NovaView Visuals***: incluye múltiples tipos de gráficas y presentaciones visuales, que permiten mostrar imágenes interesantes en vez de hojas de cálculo.
- ***NovaView Business Modeler***: permite construir un ambiente que crea automáticamente cubos y tablas, necesarios para monitorear la información.
- ***NovaView Server***: es escalable y soporta miles de usuarios y terabytes de información, provee un motor de análisis para consultas MDX que administra las consultas y la información de toda la solución de NovaView.

3.2.2.4.5 *Spotfire*

Spotfire [Spotfire, 2007] es una suite interactiva y visual, que elimina la necesidad de los reportes estáticos mediante una visualización interactiva de la información a los usuarios en tiempo real de cualquier conjunto de datos. Utiliza información de inteligencia empresarial de cualquier fuente como Oracle, IBM o Microsoft y da a los usuarios una representación visual que puede ser fácilmente manipulable por ellos.

Los usuarios pueden explorar patrones en las transacciones, identificar excepciones, descubrir relaciones escondidas entre los datos que pueden ser explotadas para generar las ventajas competitivas. A diferencia de otras suites de inteligencia empresarial, Spotfire da una completa flexibilidad y métodos intuitivos en la visualización de datos que permiten un análisis claro y rápido de la información.

3.3 Discusión final

A lo largo del capítulo se mostró una diversidad de herramientas que ayudan a la implementación del proceso de inteligencia empresarial. Sin embargo, debido a esta gran diversidad existe la complejidad de selección de una(s) herramienta(s) que sean útiles ¿Cómo saber cuál es la mejor para la problemática definida en el capítulo I? Para esto, es importante conocer los requerimientos de las PyMEs y la naturaleza de los datos que vamos a manejar, lo cual veremos en el siguiente capítulo.