

4. Procesos de *business intelligence*

4.1 Metodología

En la presente unidad se analizarán cada una de las alternativas para la gestión y seguimiento de un proyecto de *business intelligence*.

4.1.1 Tipos de metodología

Un punto crucial para el éxito de todo proyecto, más allá de las tecnologías, es la utilización de una metodología para la gestión y seguimiento de este.

Un proyecto de *business intelligence* implica muchas etapas, cíclicas y en su gran mayoría ejecutándose en forma paralela. Es por ello que suele ser muy complejo poder elegir una única metodología a aplicar.

Según ATI (2012), los enfoques metodológicos utilizados en mayor o menor medida en proyectos de Business Intelligence BI son:

- ✓ *Plan-Driven approach o Requirement-Driven approach.*
- ✓ *Demand-Driven, User-Driven o Prototype-Driven Approach.*
- ✓ *Data-Driven Approach.*
- ✓ *Value-Chain Data Approach.*
- ✓ *Process-Driven Approach.*
- ✓ *Event-Driven Approach.*
- ✓ *Object-Process Driven Approach.*
- ✓ *Joint Approach.*
- ✓ *Goal-Driven Approach.*
- ✓ *Triple-Driven Approach.*
- ✓ *Model Driven Approach.*
- ✓ *Adaptive Business Approach.*
- ✓ *Agile Approach.*

A continuación se resumen sus principales objetivos y resultado de la aplicación en proyectos de Business Intelligence:

✓ ***Plan-Driven approach o Requirement-Driven approach:***

- **Objetivo:** Esta metodología se base en el seguimiento de un plan desde un principio a fin.
- **Comentarios:** El principal problema es que no tiene en cuenta los cambios o nuevas necesidades que surgen durante el desarrollo del proyecto. Por ende, puede darse que al final del mismo los entregables no sean realmente los que necesite el usuario.

✓ ***Demand-Driven, User-Driven o Prototype-Driven Approach:***

- **Objetivo:** Se basa en el uso de prototipos evolutivos que van mejorando la solución a medida que van pasando las etapas.
- **Comentarios:** En muchos casos esta metodología fracasa, porque el principal actor son los usuarios y muchas veces, en las etapas de relevamiento, no cuentan con los conocimientos de negocios necesarios para poder describir las necesidades.

✓ ***Data-Driven Approach:***

- **Objetivo:** La metodología pone en el centro de la escena a los datos, cómo están estructurados y cómo es la relación entre ellos.
- **Comentarios:** Pueden aparecer inconvenientes al momento de entregar la solución final, ya que en muchas ocasiones se deja de lado las necesidades de los usuarios en post de garantizar la coherencia de la información.

✓ ***Value-Chain Data Approach:***

- **Objetivo:** Esta metodología está basada en *Data-Driven* pero hace enfoque desde la recolección de requerimientos hasta la entrega de la solución.
- **Comentarios:** Si bien es una mejora de *Data-Driven*, no soluciona los problemas, ya que no se focaliza en los usuarios sino en cómo se construye la solución.

✓ ***Process-Driven Approach:***

- **Objetivo:** Esta metodología se base en el análisis de los procesos de negocio, cuáles son las entradas y salidas.
- **Comentarios:** Como la metodología se basa en los procesos, es muy posible que se pierda la visión integral del negocio, produciendo soluciones inconexas.

✓ **Event-Driven Approach:**

- **Objetivo:** En esta metodología, cada proceso de negocio está dividido bajo tres premisas: los datos, la función que tiene el proceso y como está organizado. Estas tres divisiones se interconectan entre sí mediante eventos.
- **Comentarios:** Esta metodología tiene por objetivo atacar la brecha que existe entre el área de negocio y el área de BI. Pero muchas veces fracasa porque las organizaciones no están del todo maduras.

✓ **Object-Process Driven Approach:**

- **Objetivo:** En esta metodología se intenta dar la misma importancia a los procesos de negocio, los datos, los usuarios y cualquier otro objeto que participe en el negocio.
- **Comentarios:** Esta metodología suele comenzar bien en sus primeras etapas, pero luego es tan grande el abanico que debe atacar que muchas veces se termina desorganizando el proceso.

✓ **Joint Approach:**

- **Objetivo:** En esta metodología se tiene en cuenta al proceso en su totalidad, es decir desde las entradas a las salidas, sin olvidar por todas las áreas de negocio por las que pasa en su duración. Es por ello que se divide a la organización en una matriz, para luego poder dibujar todos los procesos.
- **Comentarios:** Si bien conceptualmente es óptimo considerar todo el espectro del proceso, suele ser difícil de seguir, y los conflictos suelen surgir al momento de analizar los puntos de cruce.

✓ **Goal-Driven Approach:**

- **Objetivo:** Esta metodología se basa en cómo se obtienen los resultados de la organización. Se analizan los objetivos y cómo cada una de las partes de la organización trabaja en pos del cumplimiento de ese objetivo.
- **Comentarios:** Como el centro de la escena son los objetivos, las necesidades de información se basan en ese objetivo. Suelen surgir algunos problemas si cada uno de los objetivos a diferentes niveles no están alineados.

✓ **Triple-Driven Approach:**

- **Objetivos:** Esta metodología se basa en la premisa que ninguna de las metodologías anteriores por sí solas cumplen con las

necesidades. Es así como se toma lo mejor de las metodologías Goal, Data y User-Driven.

- **Comentarios:** Puede ser acertada en algunos casos y proyectos, pero los problemas que presentan las metodologías sobre las que se basan pueden sumarse.

✓ **Model Driven Approach:**

- **Objetivo:** En esta metodología se busca afianzar la relación entre las áreas de negocio y el área de business intelligence. Como resultado de esta unión, se busca obtener soluciones o entregables de rápida realización.
- **Comentarios:** Al trabajar sobre soluciones de rápida realización, los proyectos se dividen en etapas cortas, centrándose en la solución y la calidad de las mismas. Aun es difícil su implementación, porque las organizaciones no están maduras o porque las áreas intervinientes no logran entender el paradigma.

✓ **Adaptive Business Approach:**

- **Objetivo:** Esta metodología tiene como objetivo atacar aquellos aspectos en las que otras fallan. Es así como se centra en cómo las organizaciones van evolucionando para sobrevivir en un mundo tan cambiante. Es por eso que es tan relevante el conocimiento de los aspectos endógenos y exógenos a los que está afectada una organización.
- **Comentarios:** En esta metodología, el área de BI puede ser un actor importante, ya que ayuda en el conocimiento y a reducir la incertidumbre. Al centrarse en los aspectos más importantes, es posible que algunos de menor relevancia se dejen de lado, aunque puedan afectar a la organización.

✓ **Agile Approach:**

- **Objetivo:** Esta metodología tiene su base en la metodología Model Drive. Se busca que los proyectos estén divididos en porciones de corta duración, para así obtener soluciones de negocio lo más rápido posible. Estas soluciones tendrán una evolución en el tiempo, es por ellos que en un principio se busca resolver los temas más importantes e ir mejorando la solución entregable en entregable.
- **Comentarios:** Como el proceso se centra en el relevamiento y construcción en porciones cortas, se minimizan o reducen los aspectos de documentación y pruebas. En algunas situaciones esto puede generar ciertas dificultades. Sobre todo al momento de realizar el mantenimiento de la solución.

4.1.2 Metodología clásica

Mucho se ha hablado de las metodologías de gestión de proyectos y sobre todo de los éxitos y fracasos de cada una de ellas. En esta oportunidad analizaremos las metodologías clásicas de gestión de proyectos.

En principio, las etapas que conforman un proyecto son:

- ✓ Inicio.
- ✓ Planificación.
- ✓ Desarrollo.
- ✓ Monitoreo y Gestión.
- ✓ Cierre.

Históricamente esta ha sido la metodología más utilizada, con innumerables éxitos y fracasos. Si bien ha sido pensada para todo tipo de proyectos, los mayores resultados han sido sobre proyectos de corta duración, donde hay un estricto control y por lo tanto las variables que puedan surgir están bien contenidas.

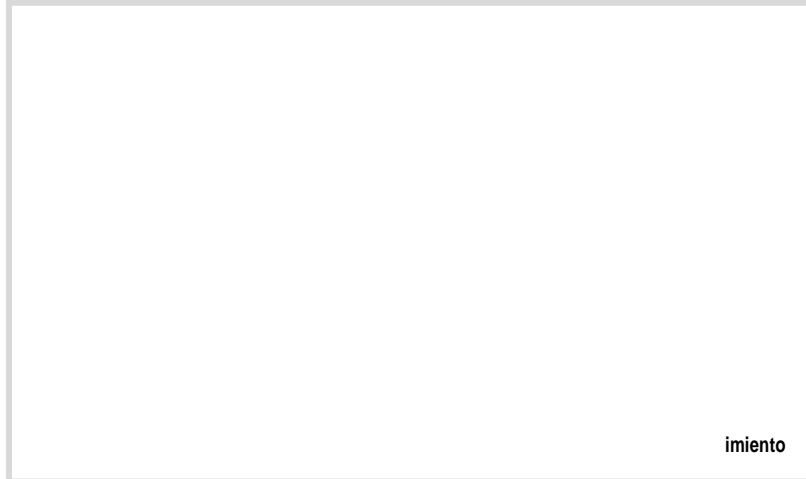
Otras metodologías tradicionales son:

- ✓ Cascada.
- ✓ Prince2.
- ✓ Proceso racional unificado (RUP).
- ✓ Desarrollo rápido de aplicaciones (RAD).

A continuación, el detalle de cada una de ellas:

- ✓ **Cascada:** Las etapas se encuentran organizadas en modo de cascada. El detalle se puede ver en la figura 1.

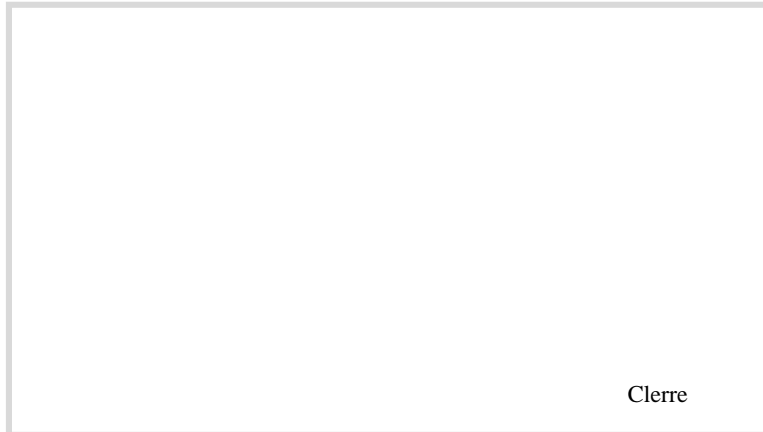
Figura 1: Cascada



Fuente: Elaboración propia.

- ✓ **Prince2:** Una versión mejorada para la administración de proyectos. El detalle se puede ver en la figura 2.

Figura 2: Cascada



Fuente: Elaboración propia.

- ✓ **Proceso racional unificado (RUP):** Las etapas se solapan y se organizan en forma iterativa. El detalle se puede ver en la figura 3.

Figura 3: RUP



Fuente: Elaboración propia.

- ✓ **Desarrollo rápido de aplicaciones (RAD):** Basada en prototipos pero con participación del usuario en todas las etapas. El detalle se puede ver en la figura 4.

Figura 4: RAD



Fuente: Elaboración propia.

4.1.3 Metodología ágil

Tomando las ventajas de cada uno de las metodologías antes mencionadas, podemos llegar a la conclusión de que la mejor manera de abordar un proyecto de *business intelligence* es dividirlo en etapas.

La metodología ágil consiste en una serie de prácticas que tienen como objetivo sacar el mejor provecho de la división en fases.

El proyecto generalmente se divide en etapas de entre una y cuatro semanas de duración. Se comienza inicialmente con el relevamiento con el usuario, el segundo paso es el desarrollo de este, por último se valida con el usuario los resultados y se ajustan de acuerdo a sus expectativas.

Esto proporciona las siguientes ventajas:

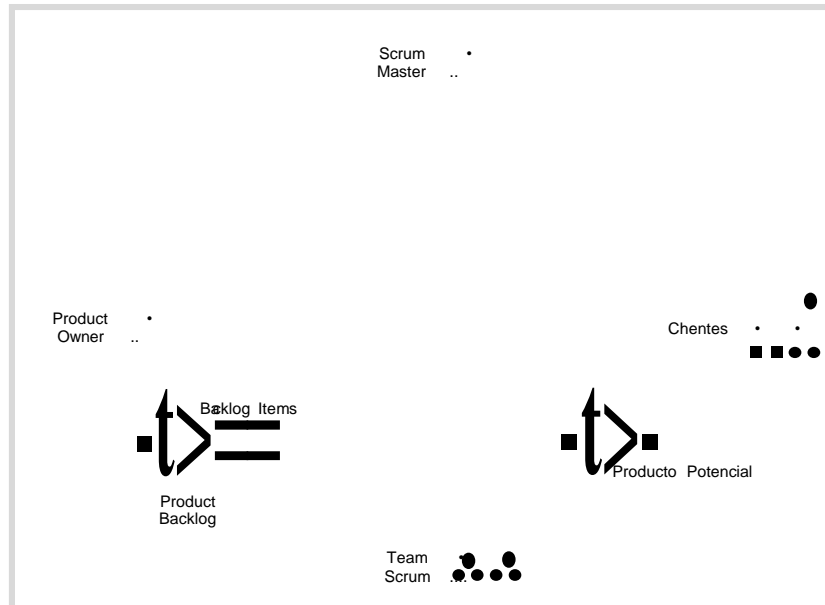
- ✓ El usuario puede validar lo solicitado en un corto período de tiempo.
- ✓ Al ser los entregables de corta realización, las modificaciones que puedan surgir terminan siendo menores que si se hicieran al final.
- ✓ Una tarea que suele demandar mucho tiempo es la validación de los datos. Al ser entregables pequeños, las pruebas también demandan menos esfuerzo.
- ✓ El usuario se siente partícipe del proyecto, colabora con este y, a medida que el proyecto avanza, vislumbra que los resultados van evolucionando.

Cómo funciona el proceso

Los componentes de la organización del proceso en la Metodología Ágil se dividen en:

- ✓ **Product Backlog:** corresponde al listado de necesidades del negocio. Luego estas necesidades se organizan por prioridad según el peso que el Product Owner determine con el negocio. A partir de esta lista se tiene el detalle completo de la solución a construir.
- ✓ **Sprint Backlog:** Cada una de las necesidades planteadas en el Product Backlog, se dividen en unidades más pequeñas, con el objetivo de que los entregables sea rápidos de construir y puedan ser administrados durante el Sprint.
- ✓ **Sprint Planning Meeting:** Es una reunión inicial al comienzo de cada etapa o Sprint, se acuerdan las premisas a utilizar. Cada Sprint se administra por separado, con entregables bien atómicos.
- ✓ **Daily Scrum o Stand-up Meeting:** Se podría decir que es una reunión de seguimiento, pero centrada en lo que se hizo y lo próximo a realizar. El objetivo es plantear los problemas surgidos y/o futuros a fin de que todo el equipo pueda aportar su experiencia.
- ✓ **Sprint Review:** Al final de cada Sprint se realiza esta reunión, donde se analiza cada entregable y se acuerdan como presentarlo al negocio.
- ✓ **Sprint Retrospective:** Como parte del proceso de mejora, en esta reunión se revisa el proceso de desarrollo, se buscan lo bueno y lo no tan bueno para mejorar en la próxima etapa.

Figura 5: Metodología ágil



Fuente: Elaboración propia.

Según la figura 5, los participantes en la Metodología Ágil se son:

- ✓ **Product Owner:** Es el responsable por parte del Negocio. Conoce las reglas o tiene las facultades para poder por los usuarios.
- ✓ **Scrum Master:** Es el responsable de coordinar las reuniones y es el facilitador del equipo de desarrollo.
- ✓ **Scrum Team:** Es el equipo de desarrollo.
- ✓ **Cliente:** Son los usuarios de negocio y participan activamente de todo el proceso de desarrollo.

4.1.4 Implementación

Cada uno de los entregables se encuentra finalizado y aprobado por el usuario de negocio. Llegó el momento de realizar la instalación de estos.

Esta etapa, que parece ser simple, es clave para el éxito del proyecto y, aunque tiene pocas etapas, cada una de ellas es importante. Las detallamos a continuación:

- ✓ Plan de mantenimiento.
- ✓ Capacitación.
- ✓ Comunicación.
- ✓ Implementación.

✓ **Plan de mantenimiento**

- Esta etapa sirve para garantizar las tareas de soporte después de la implementación, para que la solución sea funcional y extensible en el tiempo.
- Se definen cada una de las tareas de mantenimiento, los puntos de contacto y responsabilidades.
- También se define un calendario, para las tareas de soporte preventivo y guardias.
- Se definen los tiempos de los acuerdos de servicio (SLA) para garantizar que las respuestas a incidentes durante la etapa productiva sean acordes a las necesidades de negocio.
- En el caso de que la instalación sea en múltiples sitios, se define para cada uno de los ellos el plan de personalización y plan de mantenimiento.

✓ **Capacitación**

- En esta etapa se define el plan de capacitación con el objetivo de sacar el mayor provecho a las funcionalidades construidas.
- Se define el plan de capacitación técnica para las áreas de soporte pos implementación. Garantizando que se conozcan los aspectos de arquitectura, componentes, aplicaciones, servidores, roles y estructuras de seguridad, redes y actualización de software entre otros.
- Se define, se coordina y se lleva a cabo el plan de capacitación a usuarios finales con el objetivo de que estos conozcan los alcances de las funcionalidades desarrolladas.

✓ **Comunicación**

- En esta etapa se coordina la estrategia comunicacional de cara a cada uno de los destinatarios del proyecto, que se realizará antes, durante y luego de la implementación; con el objetivo de disminuir la incertidumbre de los usuarios finales y garantizar el éxito del proyecto.
- Los aspectos comunicacionales incluyen las fechas de disponibilización de las funcionalidades, el plan de mejoras o Roadmap de la instalación y cómo se realizará el mantenimiento de lo implementado.

✓ **Implementación**

- En esta etapa se llevan a cabo las tareas para el despliegue y puesta en funcionamiento de cada uno de los componentes construidos.

- Si la instalación es en múltiples sitios, se coordina la instalación de los componentes básicos y cómo se despliegan cada una de las personalizaciones.
- En esta etapa también se realiza la configuración de los entornos productivos tanto a nivel hardware como de software, que fue definida durante la etapa de diseño.
- Se realizan las pruebas de puesta operativa, previa a la disponibilización a los usuarios finales.
- Se construye el plan de actualización y migración de los componentes tanto de hardware como de software.
- Luego de la instalación, se realiza la revisión de cada componente desplegado.
- Por último, se actualizan cada uno de los documentos de la aplicación.

4.1.5 Documentación

Cuando se habla del proceso de documentación en proyectos de desarrollo sobre cualquier tecnología, siempre se lo relaciona con procesos burocráticos sin sentido o documentos muy difíciles de confeccionar, que nadie lee posteriormente a la finalización del proyecto.

Es importante mencionar que como cualquier parte del proceso de desarrollo, la documentación es crucial para poder garantizar el cumplimiento de las mejores prácticas. Sin embargo, se debe realizar un balance entre las necesidades de documentación y el uso que se le va a dar a esta, considerando además sus tiempos de confección y actualización.

Basado en varios esquemas, se presenta a continuación los documentos más importantes en el proceso de desarrollo.

Figura 6: Documentos por etapa



Fuente: Elaboración propia.

- ✓ **Especificación de requerimientos (ERS):** Durante la etapa de relevamiento con los usuarios, se van recabando cada una de las necesidades y se escriben en la ERS. Básicamente se registra lo que se va a hacer y también lo que no va a formar parte del proyecto.

Dentro de las necesidades, se plantean las funcionalidades que debe tener cada elemento a construir. También suele expresarse, si así fuera necesario, cualquier necesidad técnica que sea necesario cumplir. La ERS es un documento que finaliza con la etapa de relevamiento, y este debe ser aprobado por la persona que el negocio haya nombrado como responsable del proyecto. Esto constituye un compromiso para las dos partes, para el usuario es un compromiso de lo que solicitó, y para el equipo de desarrollo es un compromiso de lo que debe entregar al finalizar cada entregable.

- ✓ **Especificación de diseño (EDS):** una vez aprobada la ERS, comienza el proceso de análisis y diseño de la solución. A partir de las necesidades planteadas y acordadas en la ERS, se especifica cómo se va a construir cada uno de los entregables. La EDS contendrá elementos gráficos, tablas, descripciones y cualquier objeto que ayude a comprender de qué manera se realizara la producción de cada entregable. La EDS es un documento vivo durante todo el proyecto, primero como guía de cómo se va a llevar la tarea y luego si durante el desarrollo, pruebas o implementación algo cambia, también se deberá modificar en la EDS. Incluso la EDS continua vigente durante la etapa de mantenimiento. Es uno de los documentos más importantes. También requiere su aprobación al final de la etapa de diseño, y el aprobador es alguien que el área de negocio designa para tal fin.
- ✓ **Plan de testing:** el documento de plan de *testing*, consiste en una planificación de las pruebas que se llevaran a cabo para garantizar el correcto funcionamiento de cada entregable. Este plan incluye los tipos de prueba a utilizar, quiénes serán los responsables de la construcción de los casos de pruebas y quiénes la llevarán a cabo. Este documento requiere su aprobación al final de la etapa de diseño, y el aprobador es alguien que el área de negocio designa para tal fin. Este documento será la guía durante la etapa de pruebas.
- ✓ **Seguridad:** el documento de *Checklist* de seguridad contiene la planificación de los pasos a seguir para la implementación de cada entregable. Contiene la lista de los componentes a instalar, los prerequisites (si fueran necesarios) y las tareas posteriores a la instalación. El documento es aprobado por alguien del equipo de Implementación de la organización y debe respetar las normas que esta ha acordado.
- ✓ **Roadmap:** durante la etapa de relevamiento, desarrollo y pruebas de usuario, pueden surgir nuevas necesidades o pedidos de mejorar. Generalmente estos pedidos pueden ser incluidos en el proyecto, pero en el caso de que no sea posible por distintos factores, estos se incluyen en el documento de *Roadmap*. Básicamente es un plan de

mejora, en donde se especifican las próximas versiones del proyecto y qué mejoras o nuevas funcionalidades serán sumadas al proyecto.

- ✓ Otros documentos a considerar son:
 - Documento de detalle de pruebas.
 - Documento de registro de las pruebas.
 - Documento de administración de cambios.
 - Documento de administración de riesgos.
 - Diagrama Gantt del proyecto.
 - Minutas de reuniones.

4.2 Otros conceptos

En la presente unidad se analizarán cada una de las alternativas para la gestión y seguimiento de un proyecto de *business intelligence*.

4.2.1 Ciclo de vida de BI

Cuando se habla de *business intelligence* en una compañía, generalmente se la mide por la percepción que tienen los usuarios de negocio sobre el desempeño del área en brindar soluciones.

Una forma de analizar la percepción de los usuarios es utilizando los modelos de madurez. Pero ¿a qué se refieren estos modelos?

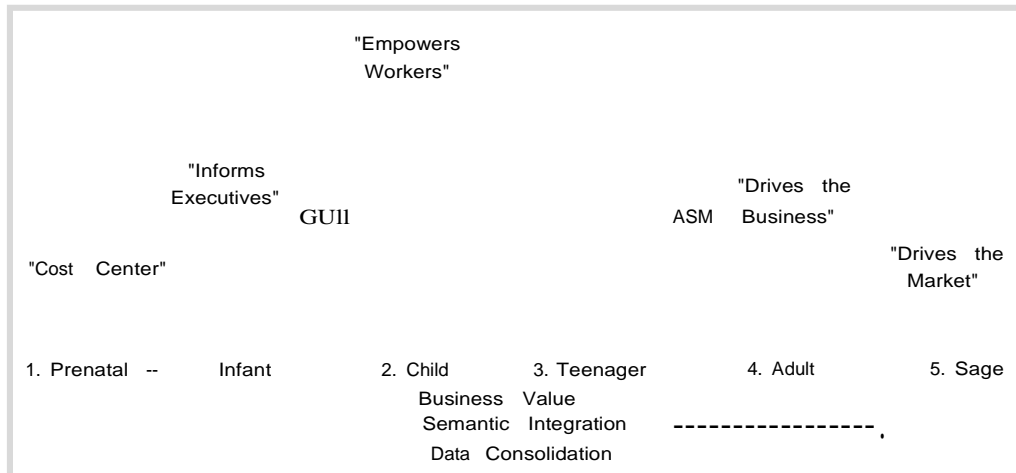
Los modelos de madurez de *business intelligence* son la manera de identificar qué tan alineadas están las iniciativas que promueve el área de *business intelligence* a los objetivos estratégicos de la compañía, midiendo a su vez el aporte que brinda BI con sus soluciones de información para la toma de decisiones. Existen varios modelos de madurez que se aplican a *business intelligence*.

- TDWI BI Maturity Model.
- Garner BI Maturity Mode.
- HP BI Maturity Model.

En este caso, centralizaremos dos de los más relevantes, el de TWDI y el de Garner.

TDWI BI Maturity Model

Figura 7: TDWI BI Maturity Model



Fuente: Eckerson, 2009. Recuperado de <https://goo.gl/yLmihK>.

Este modelo se divide en seis etapas:

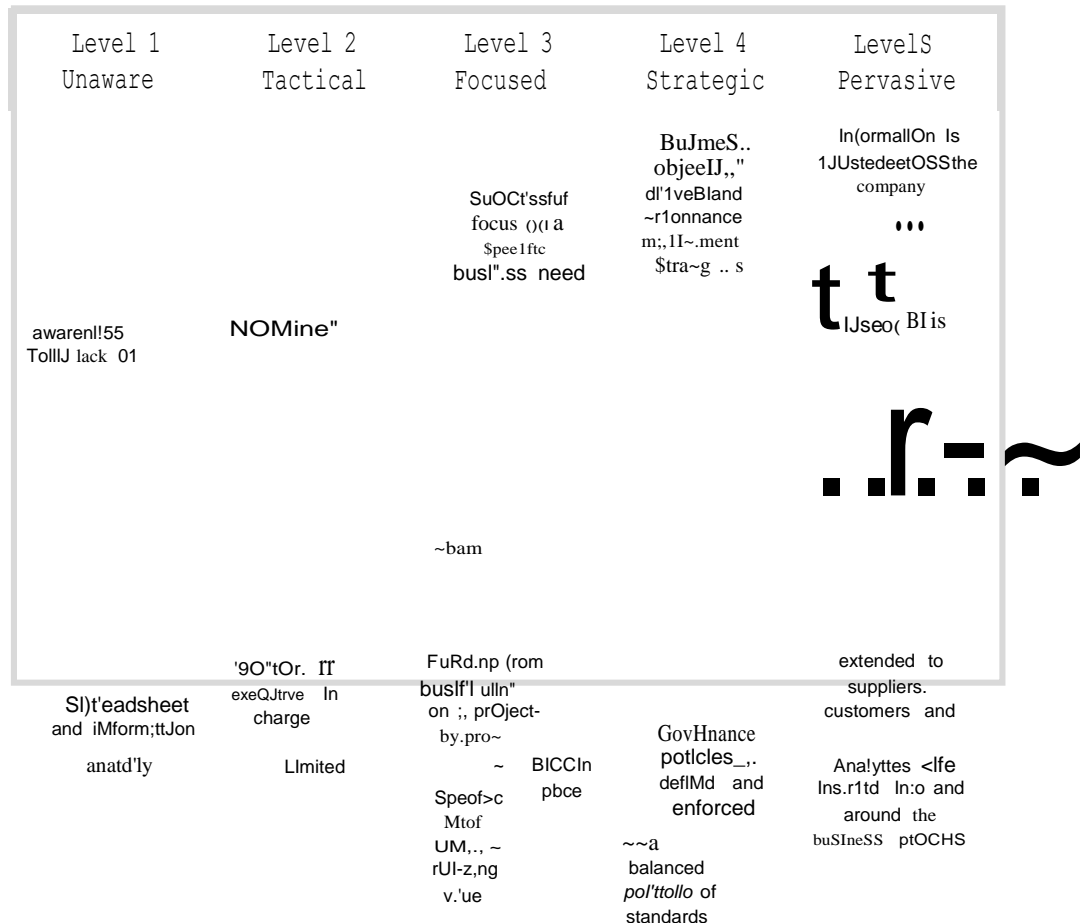
- ✓ **Etapa prenatal:** el negocio o la organización no promueven el análisis de la información para la toma de decisiones. No hay una visión integrada de la información. Solo se analizan los datos recientes y se dejan de lado los datos históricos.
- ✓ **Etapa de infancia:** existe información, pero se encuentra dispersa por toda la organización en planillas y archivos administrados por los usuarios. Existen varias versiones de la verdad. Cada sector analiza la información con su propio criterio, sin tener una visión corporativa y estratégica. La gestión se realiza consumiendo informes directamente de los sistemas transaccionales.
- ✓ **Etapa de la niñez:** existen soluciones de business intelligence para procesar y explotar la información. La información está incompleta, ya que se encuentra a cierto nivel de agregación. La información va evolucionando modelo a modelo y según las necesidades puntuales. Existen limitaciones para extender el análisis a los usuarios, sobre todo al momento de intentar realizar cruces entre modelos e indicadores. Existen mucha dependencia del área de IT.
- ✓ **Etapa de adolescencia:** existe un Datawarehouse corporativo. Existencia del concepto de única verdad. Hay acceso generalizado para las distintas capas de usuarios. Los requerimientos y las necesidades de información se encuentran alineados con los objetivos estratégicos de la organización. Los tableros de comandos o Dashboard, tienen la información adecuada o son efectivos. Existe

automatización de indicadores basados en el monitoreo de ciertas medidas de negocio.

- Etapa de madurez:** existe proactividad de parte de los usuarios y del área de BI. Los usuarios hacen uso intensivo de consulta Ad-hoc. El uso de Dashboard se extienden, con visualizaciones avanzadas que ayudan a comprender la información. Toda la información se encuentra en el DW, es por ellos que se extiende la toma de datos a fuentes externas. Existencia del uso de herramientas de Self Service. Existe metodología para el gobierno de los datos. Se utilizan modelos de desarrollo adaptativos y ágiles.
- Etapa de sabiduría:** existen herramientas analíticas dentro de los sistemas transaccionales. Se extiende el uso de BI a clientes y proveedores (portales, servicios, aplicaciones, etc.). Se hace un aprovechamiento al máximo de la información. Se reducen las ineficiencias en la entrega de información.

Garner BI Maturity Model

Figura 8: Garner BI Maturity Model



Fuente: Garner (2008). Recuperado de <http://goo.gl/FTXD8S>

Según la figura 8, las etapas del modelo son:

- ✓ **Inconsciente (unware):**

- No hay concepción sobre la organización de los datos.
- La información se encuentra desperdigada por toda la organización.
- Las planillas de cálculo son las herramientas por excelencia.
- La responsabilidad por proveer la información recae sobre el área de IT.

✓ **Táctico (Tactical):**

- Existe un área formal de business intelligence o gestión de información.
- La medición del negocio se realiza por silo y no existe una concepción estratégica de la gestión de competencias.
- La responsabilidad sigue siendo del área de IT, ahora en la figura de business intelligence, y ésta área es la encargada de “evangelizar” a la organización.
- Las herramientas son estándares con poca funcionalidad y los usuarios a menudo no están capacitados para aprovechar toda la potencialidad.
- Se desprecia de la calidad de la información que las áreas de negocio suben a la dirección.

✓ **Centrado (Focussed):**

- El área de business intelligence toma el centro de la escena.
- Se confía en la información y se comienza a pensar en cómo obtener beneficios del tratamiento de la misma.
- Se hace extensivo el uso de tableros a nivel directivo.
- Aún se analiza a la organización en forma de silos, pero comienza a descubrir la importancia de obtener indicadores centralizados.
- Las herramientas cubren las necesidades de la organización, pero muchas veces existe más de una y no están interconectadas entre sí.
- Se hace necesaria la implementación de una BICC.

✓ **Estratégico (Strategy):**

- El área de Business Intelligence toma completamente el control de la información, con estrategias bien definidas. La dirección confía plenamente en el área y pone a disposición los presupuestos para su normal desempeño.
- La información fluye por la organización. Los usuarios ya no tienen que solicitarla sino que está disponible en todos los ámbitos, incluso a proveedores y clientes.

- Existe una clara estrategia para indicadores a todo nivel de la organización, interconectados entre sí y con una política de seguimiento, medición y mejoramiento de cada uno.
- Los usuarios están altamente capacitados para ir a buscar la información, utilizar las herramientas puestas a disposición y sumar nuevas fuentes de análisis.

✓ **Omnipresente (Pervasive):**

- El área de Business Intelligence o gestión de la información se hace imprescindible. Participando en todas las actividades de la organización como un actor más que importante.
- EL BICC se encuentra afianzado y forma parte de la estrategia de la información.
- Los objetivos de la organización, desde su concepción estratégica, cuentan con mecanismos para su medición y seguimiento.
- El usuario no tiene dependencia directa del área de BI. Están capacitados para poder consumir la información a cualquier nivel y son capaces de analizar cualquier tipo de información.
- Socios de negocio como proveedores, clientes o competencia forman parte del circuito de información.

4.2.2 BICC

Un Centro de Competencia de *Business Intelligence* (Business Intelligence Competency Center, BICC) es un equipo multidisciplinario de profesionales, organizado en una estructura organizativa permanente y formal. El cual es administrado con tareas definidas, roles, responsabilidades y procesos, con el objetivo de dar soporte y promover el uso eficaz de los recursos de *business intelligence* en toda la organización.

Los objetivos principales son:

- ✓ Modelar, estandarizar y mejorar tanto los procesos de BI como sus iniciativas.
- ✓ Garantizar que el conocimiento y beneficio del uso de BI esté disponible y sea aprovechado por toda la organización.
- ✓ Trabajar de manera organizada de manera tal de reaccionar eficazmente a los cambios de los componentes internos y externos de la organización.
- ✓ Aprovechar al máximo las inversiones realizadas en BI.

Así como toda área de una organización, el BICC tiene sus funciones bien definidas:

- ✓ **Programa de BI:** es responsable de definir, supervisar y coordinar todas las actividades del BICC, siendo la cara visible hacia el resto de la organización.

- Define la estrategia de BI, los objetivos y factores críticos.
 - Debe garantizar alinear la estrategia de BI con la estrategia de la organización.
 - En la PMO (Oficina de Proyectos) de todos los proyectos asociados a BI.
 - Debe velar por la adopción de BI, poniendo al corriente a toda la organización tanto de las buenas prácticas como de las nuevas tecnologías y tendencia que surgen en el mercado de BI, y cómo ellas pueden ser aprovechadas para el beneficio de la organización.
- ✓ **Soporte técnico:** generalmente en una estructura de soporte a usuarios, es el segundo nivel, resolviendo lo que esté a su alcance y acudiendo a otros profesionales (internos y externos) para encontrar la solución.
 - ✓ **Formación y capacitación:** es el responsable de la formación de todos los usuarios del negocio para que estos puedan obtener el mejor rendimiento de las aplicaciones de BI. También debe ser el trasmisor de cada uno de los conceptos asociados a cada uno de los modelos.
 - ✓ **Gestión de datos:** debe administrar la Metadata física de la arquitectura de BI y garantizar que esté alineada con la Metadata del negocio. Es responsable de data *Quality*, *data Cleansing* y el gobierno de los datos.
 - ✓ **Advanced Analytics:** debe proveer a la organización de las herramientas de data Mining para el uso de modelos de segmentación, modelos predictivos y estadísticos que permitan sacar el mayor provecho de la Información.
 - ✓ **Gestión de contratos:** debe gestionar todo lo referente a los contratos con proveedores:
 - Licencias de usuario.
 - Licencias de software.
 - Niveles de servicio en servicios de soporte.
 - Evaluación de productos.
 - Presupuestos.
 - ✓ **Adquisición de datos:** participa en las actividades relacionadas con la extracción, transformación y carga de la información (ETL). Esta actividad abarca las tareas de diseño, desarrollo, pruebas e implementación de las soluciones integración de datos y el Datawarehouse.

- ✓ **Entrega BI:** debe definir, desarrollar y gestionar cada uno de los componentes de explotación disponibles de cara al usuario, garantizando la calidad de estos y su correcto aprovechamiento.

El proceso de implementación y crecimiento de un BICC tiene un proceso definido.

Figura 9: Proceso de implementación BICC



Fuente: Elaboración propia.

- ✓ **Evaluación:** los líderes de negocios visualizan una oportunidad o una amenaza potencial comercial en el horizonte. Esta etapa se basa en la exploración y la experimentación.
- ✓ **Requerimientos:** se comienza a definir los objetivos del BICC, se detectan clientes y objetivos potenciales del servicio y se definen los problemas que el BICC deberá resolver.
- ✓ **Lanzamiento:** el BICC comienza a funcionar, y como todo producto o servicio nuevo, debe de ser promovido. Aparecen pilotos y pequeños proyectos (*quick wins*).
- ✓ **Desarrollo:** el BICC se encuentra formalmente funcionando y debe obtener la confianza del negocio. Su objetivo es perfeccionar la experiencia, identificar iniciativas, candidatos con mayor riesgo y mayor beneficio, y desarrollar canales potenciales.
- ✓ **Monitoreo:** esta etapa puede durar de seis a 12 meses después del lanzamiento. Dentro del seguimiento se añade la cuantificación para medir el desempeño y beneficio del BICC.
- ✓ **Transición:** tanto las unidades de negocio y los líderes de TI deben decidir sobre el futuro del BICC, el cual tiene tres posibles rumbos. El primero es que ha madurado y el centro de competencia se ha adoptado ampliamente y suficiente para ser institucionalizado, el segundo que se dé de baja porque la demanda fue satisfecha o bien deba de ser transformado.

4.2.3 Big data

Desde hace algunos años, para nadie es un secreto que la información crece segundo a segundo. Tecnologías como internet, dispositivos inteligentes y las redes sociales, generan un cúmulo de información que muchos de nosotros ni siquiera puede imaginar.

Esta información más allá de ser un gran cúmulo de datos, presenta interrogantes sobre su concepción, uso y tratamiento. Las respuestas que se

pueden encontrar en esta gran maraña de información pueden ayudar a las organizaciones a seguir creciendo y acrecentando su patrimonio.

El concepto de “Big Data” viene a tratar de conquistar un lugar que las tecnologías tradicionales no han podido ocupar para poder almacenar la inimaginable información disponible.

Inicialmente se piensa a “Big Data” como una tecnología que provee de herramientas para el almacenamiento de grandes volúmenes de información, con el objetivo que puedan ser consumidos por los usuarios con un buen nivel de performance.

El gran desafío que tiene “Big Data”, es convertir la información disponible (Internet, redes sociales, texto, imágenes, voz, video, sonido y cualquier otra tecnología) en conocimiento, para que la organización pueda aprovecharla para la toma de decisiones.

Las 4 v

¿Cuál es entonces la diferencia entre las aplicaciones analíticas y de gestión y los nuevos conceptos de *big data*?

La respuesta a esta pregunta se asocia a 4 palabras:

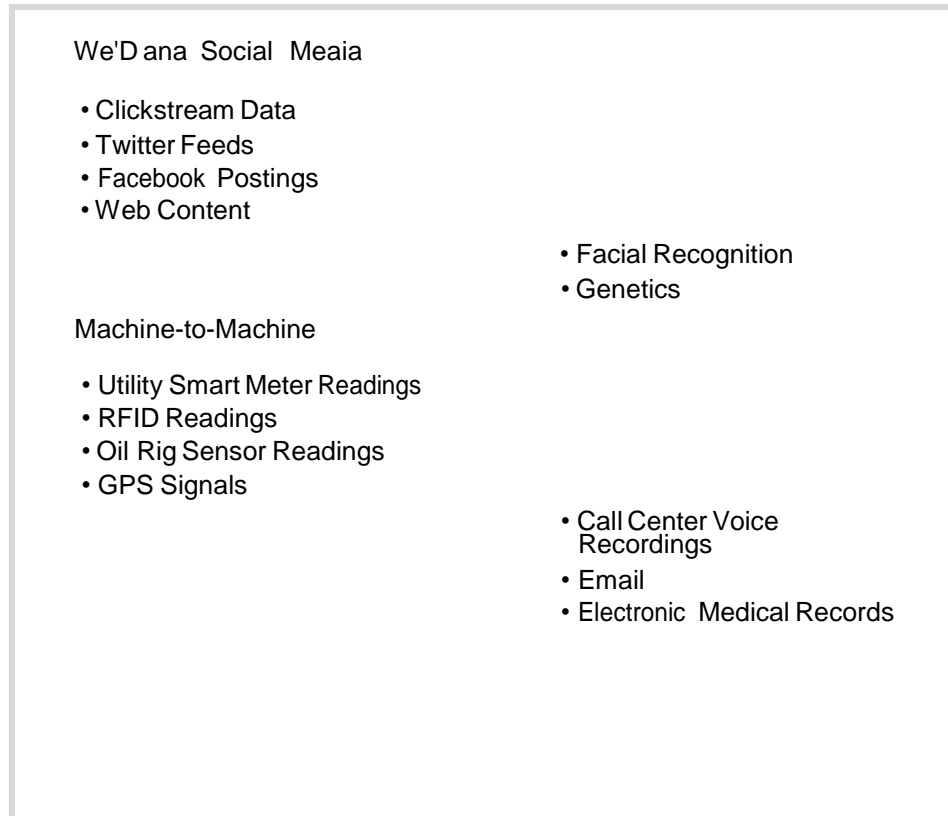
- ✓ **Volumen:** la cantidad de datos a procesar o analizar es muy alta.
- ✓ **Velocidad:** los datos se generan a una frecuencia muy elevada.
- ✓ **Variedad:** formado tanto por datos estructurados como semiestructurados y no estructurados, provenientes de las más diversas fuentes.
- ✓ **Valor:** la calidad, validez e importancia de los datos.

¿Qué datos debo analizar?

Muchas organizaciones se enfrentan a la pregunta ¿qué información es la que se debe analizar?, sin embargo, el cuestionamiento debería estar enfocado hacia: ¿qué problema es el que se está tratando de resolver?

Si bien sabemos que existe una amplia variedad de tipos de datos a analizar, una buena clasificación nos ayudaría a entender mejor su representación, aunque es muy probable que estas categorías puedan extenderse con el avance tecnológico.

Figura 10: Tipos de datos de big data



Fuente: Fragoso. (2012). Recuperado de <https://goo.gl/MSXCjm>.

¿Cuáles son los componentes de plataforma de big data?

Gran parte de la habilidad que tiene *big data* para procesar grandes volúmenes de Información se debe a Hadoop.

Según Fragoso (2012), *Hadoop* está basado en un proyecto de Google denominado “*Google File System (GFS)*” y en el paradigma de programación *MapReduce*. *MapReduce* es una técnica o algoritmo que busca dividir en dos tareas (*mapper*, *reducer*), con el objetivo de obtener un sistema altamente performante.

Estas técnicas tienen por objetivo que al momento de cargar y consumir la información, la tarea no recaiga sobre un único o algunos pocos servidores. En este caso las tareas se distribuyen en varios puntos denominados nodos que juntas forman un clúster.

De acuerdo a este paradigma siempre es mejor dividir la tarea y no concentrarla en un único punto.

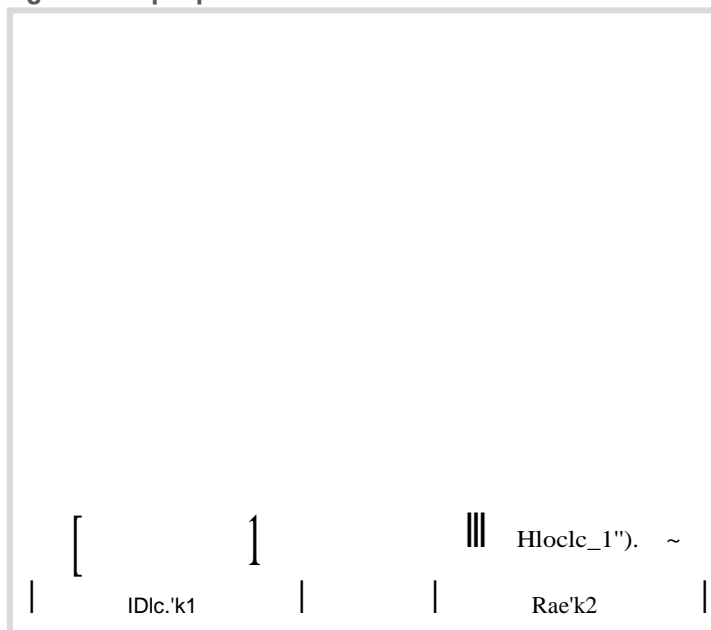
Es así que Hadoop se divide en tres componentes:

- ✓ Hadoop Distributed File System (HDFS)
- ✓ Hadoop MapReduce
- ✓ Hadoop Common

El detalle de cada uno de estos componentes se detalla a continuación

- ✓ **Hadoop Distributed File System (HDFS):**
 - Como se indicó con anterioridad, se dividen las tareas en porciones pequeñas denominadas bloques.
 - Tanto la función de Map y la de Reduce se dividen y se ejecutan en pequeñas unidades de procesamiento.
 - Al dividir a carga de trabajo en pequeñas unidades, la tarea es ejecutada en múltiples unidades de procesamiento y por ende la performance y el tamaño de almacenamiento es mayor que en otras tecnologías. También esta arquitectura provee de escalabilidad ya que se puede sumar más unidades de procesamiento al clúster.

Figura 11: Ejemplo de HDFS

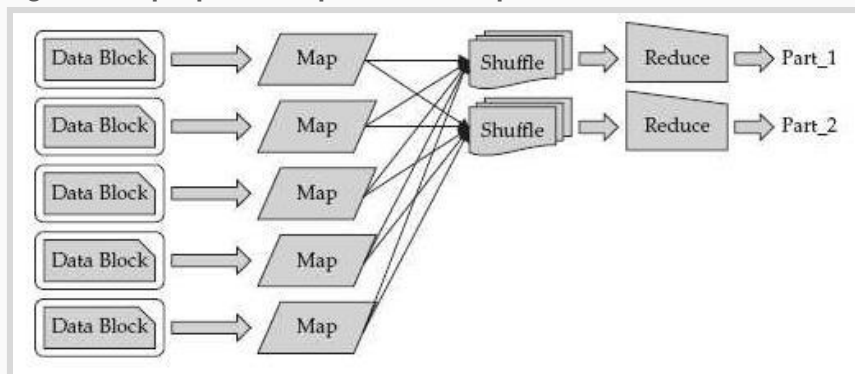


Fuente: Fragoso. (2012). Recuperado de <https://goo.gl/MSXCjm>.

- ✓ **Hadoop MapReduce:**
 - Es la parte central del Clúster de Hadoop.
 - Formado por dos partes Map y Reduce.

- Map, a partir del conjunto de información principal, la divide en conjuntos más pequeños y que siguen una regla determinada. Estos conjuntos más pequeños se denomina Tuplas y tienen dos entidades: Código y Valor.
- Reduce: Toma los conjuntos generados por Map, y a partir de ellos los combina y los ubica en conjuntos aún más pequeños. Estos bloques serán procesados en la etapa denominada Shuffle. En dicha etapa, se determinará que nodo procesará cada bloque para así obtener el resultado final del procesamiento.

Figura 12: Ejemplo de un proceso de MapReduce



Fuente: Fragoso. (2012). Recuperado de <https://goo.gl/MSXCjm>.

✓ Hadoop Common:

- Son las librerías o paquetes de componentes que permitirán las ejecuciones de otras implementaciones de proyectos o subproyectos de Hadoop.

Además de estos tres componentes principales de Hadoop, existen otros proyectos relacionados (Avro, Cassandra, Chukwa, Flume, Hbase, Hive, Jaq, Lucene, Oozie, Pig, Zookeeper).

4.2.4 Técnicas de performance

Una parte crucial de la arquitectura de *business intelligence* es que la información siga llegando a los usuarios de negocio como lo hacía desde el primer momento.

Esto muchas veces no es posible, sobre todo porque la información crece a diario, incrementándose el tamaño de almacenamiento y por lo tanto acrecentando los tiempos de acceso.

Revisaremos varias técnicas que tienen por objetivo optimizar distintos puntos de la arquitectura para seguir cumpliendo con el objetivo de encontrar la información en el tiempo justo.

- ✓ **Ejecuciones en paralelo:** si bien suele ser una propiedad de la base de datos, es crucial que cada proceso de extracción y carga de información sea diseñado para utilizar las capacidades de paralelismo. Básicamente, esta técnica divide cada tarea en hilos de ejecución que se ejecutan en forma independiente y luego se unen al final para dar respuesta a la necesidad.
- ✓ **Particiones:** consiste en dividir las tablas, índices y otros objetos de almacenamiento en partes más pequeñas. El objetivo es almacenar una porción de información a la que luego se podrá acceder en forma paralela. Existen, a su vez, varias técnicas: *Range Partitions*, *Hash Partitioning*, *List Partitioning*, *Index Partitioning*.
- ✓ **Utilización de índices:** los índices son objetos de base de datos que ayudan a optimizar los tiempos de acceso a la información. Existen algunos tipos de índices que suelen ser usados en ambientes de *Datawarehouse*: índices *Bitmap* e índices *B-Tree*.
- ✓ **Uso de estadísticas:** se suele llamar: los datos de los datos. Es información de *Metadata* que guarda la historia de las ejecuciones, buscando optimizar cada consulta, de manera que la próxima vez se utilice la misma solución.
- ✓ **Vistas materializadas:** las vistas son objetos de base de datos compuesto por consultas en *Transac SQL*. Estas vistas pueden estar formadas por *queries* muy complejas y pueden tardar demasiado tiempo en responder. La forma de resolverlo es almacenar físicamente esas vistas para obtener mejora en los accesos de información.
- ✓ **Agregaciones o Facts agregadas:** como vimos con anterioridad son tablas resumen de cierta información con el objetivo de mejorar determinado acceso a los datos.

El principal objetivo al consultar una base de datos es leer poco. Las operaciones de I/O son las operaciones más costosas de una base de datos. Algunas de las técnicas que se pueden utilizar son las siguientes:

- ✓ **Sistemas balanceados:** consiste en una arquitectura física y distribuida, con el objetivo de poder dividir la carga de trabajo. Aquí

aparece un objetivo muy importante, el balanceador, que corresponde a un nodo encargado de distribuir la tarea lo más democráticamente, de manera de no castigar y no beneficiar a determinado nodo.

- ✓ **Layout de discos:** el *layout* de discos de un sistema de almacenamiento consiste en ubicar de manera estratégica cada porción de la información en unidades de discos diferentes, con el objetivo de aprovechar las propiedades de paralelismo.
- ✓ **Almacenamiento Columnar:** es una técnica que tienen determinadas base de datos. Consiste en almacenar los datos en columnas, de esta manera se genera un vector con los valores que puede tomar esa columna. Esto posibilita reducir los tiempos de acceso y paralelizar la lectura de distintas columnas.
- ✓ **Base de datos con discos de estado sólido (SSD):** la utilización de discos de estado sólido mejora los tiempos de acceso, ya que la tecnología de estos discos tiene mejores tasas de respuesta al no contar con elementos mecánicos.
- ✓ **Sistemas en memoria:** esta tecnología permite llevar toda o una parte de la información a la memoria de los servidores. El acceso a los datos ubicados en memoria tiene mejores tasas de respuesta que cualquier elemento físico.

Referencias

ATI. (2012). *Enfoques metodológicos de Business Intelligence. Data Prix.* Recuperado de: <http://www.dataprix.com/enfoques-metodologicos-business-intelligence>.

Gutiérrez, C. (2014). *¿Para qué sirve el Scrum en la Metodología Ágil?* Recuperado de: <http://www.i2btech.com/blog-i2b/tech-deployment/para-que-sirve-el-scrum-en-la-metogologia-agil/>.

Eckerson, W. (2009). *Crossing The Chasm Part 1: Delivering Strategic Value.* (Traducción Propia) TDWI. Recuperado de: <https://tdwi.org/blogs/tdwi-blog/2009/10/chasm-1.aspx>.

Garner (2008). Recuperado de <http://www.jamesserra.com/archive/2013/06/business-intelligence-maturity-assessment/>

Fragoso, R. (2012). *¿Qué es Big Data?* Recuperado de: <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>