



UNIVERSIDAD CATÓLICA ANDRÉS BELLO
VICERRECTORADO ACADÉMICO
ESTUDIOS DE POSTGRADO
FACULTAD DE INGENIERÍA
POSTGRADO EN SISTEMAS DE INFORMACIÓN

TRABAJO DE GRADO DE MAESTRÍA

**MODELO DE IMPLEMENTACIÓN DE UN SISTEMA DE INTELIGENCIA DE
NEGOCIO CON BIG DATA PARA LA BANCA EN LINEA EN VENEZUELA**

Presentado por:

Bastidas Fuentes, William Eduardo

Para optar al título de
Magíster en Sistema de Información

Tutor

Bonillo Ramos, Pedro Nolasco

Caracas, 14 de noviembre de 2016

UNIVERSIDAD CATÓLICA ANDRÉS BELLO
VICERRECTORADO ACADÉMICO
ESTUDIOS DE POSTGRADO
FACULTAD DE INGENIERÍA
POSTGRADO EN SISTEMAS DE INFORMACIÓN

TRABAJO DE GRADO DE MAESTRÍA

**MODELO DE IMPLEMENTACION DE UN SISTEMA DE INTELIGENCIA DE
NEGOCIO CON BIG DATA PARA LA BANCA EN LINEA EN VENEZUELA**

Presentado por:

Bastidas Fuentes, William Eduardo

Para optar al título de
Magíster en Sistema de Información

Tutor

Bonillo Ramos, Pedro Nolasco

Caracas, 14 de noviembre de 2016

Caracas, 07 de Noviembre de 2016

UNIVERSIDAD CATÓLICA ANDRÉS BELLO

Postgrado en Sistemas de Información

Lic. María Esther Remedios

Presente.-

CARTA DE APROBACIÓN DEL TUTOR

Por la presente me permito comunicar que he sido el tutor del Trabajo de Grado de Maestría del estudiante William Eduardo Bastidas Fuentes, cédula de identidad nro. 19.671.995, quien opta por el título de Magíster en Sistemas de Información, titulado “Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela”.

Asimismo, hago constar que como tutor estoy conforme con el contenido presentado, por lo que cuenta con mi aprobación para ser inscrito como Trabajo de Grado de Maestría.

Sin otro particular al cual hacer referencia, se despide cordialmente,

Dr. Pedro Nolasco Bonillo Ramos

C.I. 10.868.538

DEDICATORIA

A Dios y la Virgen, por darme la fortaleza y la perseverancia.

A mi madre, padre y hermanas por darme todo el amor y apoyo necesario para seguir adelante y no desvanecer en el logro de mis aspiraciones.

A mi familia en general por brindarme su apoyo incondicional y creer en mí.

Dios nos mostrará el camino.

RECONOCIMIENTOS

El presente reconocimiento y agradecimiento al:

Dr. Pedro Bonillo, tutor de la presente investigación por su apoyo y dedicación.

A los Profesores María Esther Remedios, Profesora Lourdes Ortiz y Samuel Villegas, por el apoyo prestado.

A mis compañeros de trabajo por la colaboración prestada para el desarrollo de la presente investigación.

Especial reconocimiento a mi compañera de estudios y amiga, Belkis Carrasco, cuyo apoyo, dedicación y profesionalismo superaron todas mis expectativas.

Al Banco Venezolano de Crédito por su colaboración en el desarrollo de esta investigación.

A todos mi sentido de agradecimiento, por su colaboración y amistad, por brindarme su comprensión, tiempo y dedicación.

¡A todos gracias!



UNIVERSIDAD CATÓLICA ANDRÉS BELLO
VICERRECTORADO ACADÉMICO
ESTUDIOS DE POSTGRADO
FACULTAD DE INGENIERÍA
POSTGRADO EN SISTEMAS DE INFORMACIÓN

MODELO DE IMPLEMENTACIÓN DE UN SISTEMA DE INTELIGENCIA DE NEGOCIO CON BIG DATA PARA LA BANCA EN LINEA EN VENEZUELA

Autor: Bastidas Fuentes, William Eduardo

Tutor: Bonillo Ramos, Pedro Nolasco

Año: 2016

RESUMEN

Realizar inteligencia de negocios en toda organización es una de sus actividades fundamentales, donde la incorporación de nuevas tecnologías como Big data puede marcar la diferencia. Sin embargo, existen factores específicos en cada organización, y por otra parte, características particulares para aplicar este tipo de tecnología, lo cual no resulta fácil de identificar si no se dispone de un modelo que sirva como guía. El propósito de este trabajo de investigación es realizar un aporte al estudio y análisis del proceso de inteligencia de negocios y al tratamiento de grandes volúmenes de datos que este conlleva. Del mismo modo, el objetivo es desarrollar un modelo implementación de un sistema de inteligencia de negocio a través de la tecnología Big data para la banca en línea en Venezuela. Lo anterior implica analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero venezolano, identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data, y finalmente diseñar, desarrollar y evaluar un modelo de implementación mediante su aplicación a un caso de estudio. Esta investigación es del tipo aplicada y descriptiva, con un diseño de investigación de campo, apoyado en un diseño bibliográfico, en la que se utilizaron técnicas como entrevistas, encuestas y revisión de documentos para el levantamiento y análisis de la información. Finalmente se obtuvo como resultado un modelo que permite identificar la existencia de atributos que deben cumplirse para garantizar que se logre la implementación de un sistema de inteligencia de negocios usando Big Data con mayores posibilidades de éxito, permitiendo descubrir las necesidades y puntos de mejora en la organización.

Palabras Clave: Inteligencia de negocios, Banca en línea, Big data.

Línea de Investigación: Ingeniería del Software.

Área de Conocimiento: Gestión de Grandes Volúmenes de datos. (Big Data)

ÍNDICE GENERAL

DEDICATORIA	iv
RECONOCIMIENTOS.....	v
RESUMEN	vi
ÍNDICE DE FIGURAS	x
ÍNDICE DE TABLAS.....	xii
INTRODUCCIÓN	13
CAPITULO I: EL PROBLEMA.....	16
1.1 Planteamiento del Problema	16
1.1.1 Formulación del Problema.....	19
1.1.2 Sistemización del Problema	19
1.2. Objetivos	20
1.2.1 Objetivo General.....	20
1.2.2 Objetivos Específicos.....	20
1.3. Justificación de la Investigación.....	20
1.4. Alcance y Limitaciones de la Investigación	22
CAPÍTULO II: MARCO TEÓRICO	24
2.1 Antecedentes	24
2.1.1 Trabajos de Grado de Maestría (TGM).....	24
2.1.2 Artículos Técnicos.....	27
2.2. Fundamentos Teóricos.....	31
2.2.1 Inteligencia de negocio	32
2.2.2 Evolución de la Inteligencia de negocio	33
2.2.3 Arquitectura de una solución de inteligencia de negocio.	34
2.2.4 Marco de inteligencia de negocios.....	37
2.2.5 Método para implantar proyectos de BI	39
2.2.6 Metodología y modelo de ciclo de vida de una solución de BI.....	40
2.2.7 Big Data	42
2.2.8 Retos actuales	50
2.2.9 Arquitectura de Big Data.....	52

2.2.10 Apache Hadoop	56
2.2.11 Bases de datos NoSQL	59
2.2.12 Propiedades ACID vs BASE	63
2.2.13 Factores que intervienen en la creación de un modelo.....	66
2.2 Marco Referencial	71
2.4. Bases Legales.....	75
CAPITULO III: MARCO METODOLÓGICO.....	82
3.1 Tipo de Investigación	82
3.2 Diseño de la Investigación	83
3.3 Población y Muestra.....	83
3.4 Técnicas de Recolección de Datos	85
3.5 Fases de la Investigación.....	86
3.6 Procedimiento por Objetivos	87
3.7 Variables, Definición Conceptual y Operacional e Indicadores	88
3.8 Estructura Desagregada de Trabajo	92
3.9 Aspectos Éticos.....	94
3.10 Cronograma	95
3.11 Recursos	97
CAPÍTULO IV: ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	98
4.1 Fase Técnica.....	98
4.1.1 Recolección, procesamiento y análisis de los datos.....	98
4.1.2 Validación de los instrumentos	100
4.1.3 Análisis de los resultados del diagnóstico	100
CAPÍTULO V: DESARROLLO DE LA PROPUESTA	114
5.1 Características del modelo propuesto:.....	114
5.2 Métricas	137
CAPÍTULO VI: EVALUACIÓN DE LA PROPUESTA	145
6.1 Aplicación del modelo de implementación	145
CAPITULO VII: CONCLUSIONES Y RECOMENDACIONES	162
7.1 Conclusiones.....	162

7.2 Recomendaciones.....	165
REFERENCIAS BIBLIOGRAFICAS	167
ANEXOS	1675
ANEXO I Guía de Observación.....	171
ANEXO II Cuestionario	173
ANEXO III Formato de Validación de Instrumentos	181
ANEXO IV Resultados de la aplicación de la guía de observación.....	183
ANEXO V Resultados de la aplicación del cuestionario.....	187
ANEXO VI Matriz del Modelo diseñado	187
ANEXO VII Constancias de validación de los instrumentos.....	200

ÍNDICE DE FIGURAS

Figura	Pág.
1. Canales Transaccionales (% del volumen).	17
2. Evolución de las Transaccionales Web (Número de transacciones).....	17
3. Mapa del marco teórico de Inteligencia de negocio.	31
4. Mapa del marco teórico de Inteligencia de negocio desplegado.....	31
5. Evolución de la Inteligencia de negocio.	33
6. Principales productos de inteligencia de negocio.....	36
7. Fases de un proyecto de BI.....	39
8. Modelo de ciclo de vida de desarrollo de una solución de BI.....	41
9. Mapa del marco teórico de Big Data.	43
10. Mapa del marco teórico de Big Data desplegado en temas y subtemas.....	43
11. Organigrama del Sistema Bancario Nacional.....	72
12. Elementos que conforman el Sistema Financiero.	73
13. Estructura Desagregada de Trabajo.	93
14. Cronograma de trabajo1/2.....	95
15. Cronograma de trabajo 2/2.....	96
16. El estado del arte de las herramientas de inteligencia de negocios.	103
17. El estado del arte de las arquitecturas de Big Data.....	105
18. Análisis del entorno de inteligencia de negocio en la banca en línea.....	113
19. Características generales del modelo de implementación.	116
20. Características detalladas del modelo de implementación.....	117
21. Atributo Valor.....	120
22. Atributo Volumen.....	122
23. Atributo Variedad.....	123
24. Atributo Velocidad.	124
25. Atributo Veracidad.....	125
26. Atributo Gobierno.	126
27. Atributo Personas.....	127
28. Instrucciones para la aplicación del modelo.	146
29. Aplicación del Modelo atributo Valor.	147
30. Aplicación del Modelo atributo Volumen.....	148
31. Aplicación del Modelo atributo Variedad.	149
32. Aplicación del Modelo atributo Velocidad.....	150
33. Aplicación del Modelo atributo Veracidad.....	151
34. Aplicación del Modelo atributo Gobierno.....	152
35. Aplicación del Modelo atributo Personas.....	153

36.	Resultados Valor obtenido vs. Valor ideal del atributo Valor.....	154
37.	Resultados Valor obtenido vs. Valor ideal del atributo Volumen.	154
38.	Resultados Valor obtenido vs. Valor ideal del atributo Variedad.....	155
39.	Resultados Valor obtenido vs. Valor ideal del atributo Velocidad.....	155
40.	Resultados Valor obtenido vs. Valor ideal del atributo Veracidad.	156
41.	Resultados Valor obtenido vs. Valor ideal del atributo Gobierno.....	156
42.	Resultados Valor obtenido vs. Valor ideal del atributo Personas.	157
43.	Resultados Valor obtenido vs. Valor ideal total del Modelo.....	157

ÍNDICE DE TABLAS

Tabla	Pág.
1. Ciclo de vida de desarrollo de una solución de inteligencia de negocios.	42
2. Definiciones de Big Data.	44
3. Fases de la Gestión de Datos.	45
4. Unidades de Información.	45
5. Las 4Vs de Big Data.	48
6. Principales Definiciones del Término Modelo.	67
7. Procedimientos por objetivos planteados.	87
8. Definición de Variables.	89
9. Operacionalización de Variables.	90
10. Matriz de Recursos.	97
11. Métricas del atributo Valor del Modelo Propuesto.	138
12. Métricas del atributo Volumen del Modelo Propuesto.	139
13. Métricas del atributo Variedad del Modelo Propuesto.	140
14. Métricas del atributo Velocidad del Modelo Propuesto.	141
15. Métricas del atributo Veracidad del Modelo Propuesto.	142
16. Métricas del atributo Gobierno del Modelo Propuesto.	143
17. Métricas del atributo Personas del Modelo Propuesto.	144

INTRODUCCIÓN

Desde hace poco más de una década las operaciones de la banca en línea se desarrollan a través de la integración de sistemas de información y procedimientos de alta tecnología, cuyas innovaciones constituyen una de las principales causas de cambio y adaptación al nuevo entorno competitivo del sistema financiero, por lo que forman parte de un elemento de cambio estructural y un factor estratégico clave.

Dentro de este entorno competitivo la toma de decisiones juega un papel fundamental, y representa el proceso que distingue a las organizaciones exitosas de las que no lo son.

Diversos fenómenos y herramientas han influido en la humanidad en el proceso de toma de decisiones, desde la creación de las computadoras, las bases de datos, los sistemas para el soporte de decisiones, data warehouse, internet, la computación en la nube y finalmente la revolución de Big data, han dado origen al concepto de inteligencia de negocio, así como a los nuevos desarrollos que fortalecen aún más sus beneficios.

Big Data es la más reciente de las tecnologías creadas para el tratamiento e integración de grandes volúmenes de datos, resolviendo problemas asociados con el uso del modelo relacional de base de datos, que genera falta de velocidad y escalabilidad al manejar un conjunto de datos que supera la capacidad del software habitual para ser capturado, gestionado y procesado en un tiempo razonable. Su uso potencial dentro del ámbito de inteligencia de negocios está provocando profundos cambios en la sociedad y por extensión en los entornos empresariales. De acuerdo con una encuesta de Gartner Inc (2014), el 73% de las organizaciones están invirtiendo o tienen planificado invertir en Big Data en los próximos 24 meses.

Un buen manejo del proceso de toma de decisiones en las entidades financieras requiere de análisis y tecnología que permita manejar grandes cantidades de

información de manera rápida y efectiva en tiempo real, y a su vez gestione la creciente disponibilidad de datos y haga uso intensivo de los ya existentes.

El propósito de este trabajo de investigación es realizar un aporte al estudio y análisis del proceso de inteligencia de negocios, y del tratamiento de grandes volúmenes de datos que este conlleva, a través del desarrollo de un modelo implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.

El resultado esperado es obtener un modelo que permita evaluar la existencia de atributos que deben cumplirse para garantizar que se logre la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea con mayores posibilidades de éxito, permitiendo optimar la eficiencia en la provisión de servicios bancarios y aumentar la satisfacción de los clientes.

Lo anterior implica analizar el entorno de sistemas de Inteligencia de negocio en la banca en línea en Venezuela, identificar el estado del arte de los modelos de inteligencia de negocio y arquitecturas de Big Data asociadas al sistema financiero, y finalmente, diseñar, desarrollar y evaluar dicho modelo.

Con esta finalidad se ha estructurado el presente trabajo de grado de maestría de la siguiente manera:

En el primer capítulo se mostrará la idea de investigación transformada en el planteamiento del problema, abarcando los elementos de la investigación como lo son; planteamiento del problema, formulación del problema, objetivos, justificación y finalmente, alcance y limitaciones de la Investigación.

En el segundo capítulo se realiza una contextualización del problema de investigación planteado, mediante el desarrollo de una perspectiva teórica. Se detallan los antecedentes, los fundamentos teóricos, el marco referencial y las bases legales que apoyan la investigación.

En el tercer capítulo se desarrolla el marco metodológico de la investigación, en el que se define el tipo de investigación, el diseño de la investigación, población y

muestra, técnicas de recolección de datos, fases de la investigación, procedimientos por objetivos y se realiza la definición conceptual y operacional de variables e indicadores.

El cuarto capítulo consiste en el análisis e interpretación de los resultados obtenidos de la aplicación de los instrumentos de recolección con el objetivo de comprender y contextualizar la situación en la cual se realiza la investigación.

En el quinto capítulo se presenta el desarrollo del modelo propuesto el cual consiste en una adaptación del ciclo de vida del desarrollo de sistemas planteado por Kendall & Kendall (2005), y consta de ocho partes: siendo la primera; identificación de carencias, necesidades y expectativas, la segunda; selección del tipo de implementación, la tercera; determinación de requisitos de información, la cuarta; análisis de las necesidades del sistema, la quinta; diseño del sistema recomendado, la sexta; desarrollo y documentación del sistema, la séptima; prueba y mantenimiento, y la octava; implementación y evaluación.

Finalmente en el sexto capítulo se presenta la evaluación de la propuesta y en el séptimo capítulo las conclusiones y recomendaciones. Terminando el documento con las referencias bibliográficas consultadas para elaborar este proyecto y sus respectivos anexos.

CAPITULO I: EL PROBLEMA

En el presente capítulo se mostrará la idea de investigación transformada en el planteamiento del problema, abarcando los elementos de la investigación como lo son; planteamiento del problema, formulación del problema, objetivos, justificación y finalmente, alcance y limitaciones.

1.1 Planteamiento del Problema

Uno de los insumos básicos de las empresas financieras es la gran cantidad de datos que tienen de sus clientes, para muchas de sus operaciones requieren de un elevado componente informacional y esta es una de sus grandes ventajas competitivas.

Según McGuire (2013), los datos de servicios financieros son muy interesantes para mejorar los modelos de negocio, a diferencia de otro tipo de datos que permiten visualizar la intencionalidad de los consumidores, la información de los servicios financieros expresa las decisiones de compra de los mismos.

“La industria de los servicios financieros es uno de los sectores más sofisticados de la economía mundial en cuanto a manejo de datos, permitiendo la ejecución de complejas operaciones globales con solo pulsar un botón. Desde la banca móvil, la banca en línea y el comercio electrónico, este sector está generando montañas de información, de una gran diversidad de fuentes, enfrentando nuevos retos que les permitan estudiar e interpretar tal cantidad de datos, más allá de lo que hacen con sus herramientas tradicionales”. (Temsamani, 2014, pág. 1).

En Venezuela no existe una cantidad notable de investigaciones al alcance del público que exploren la situación actual de la banca en línea, y que evalué su importancia y potencialidades para el sistema Financiero Venezolano. Las investigaciones realizadas por los bancos nacionales se manejan en su mayoría de manera confidencial.

Según las últimas cifras estadísticas publicadas por el Banco Central de Venezuela (2013), se presenta la evolución de los canales bancarios para las operaciones de consultas de saldo, retiros y transferencias.

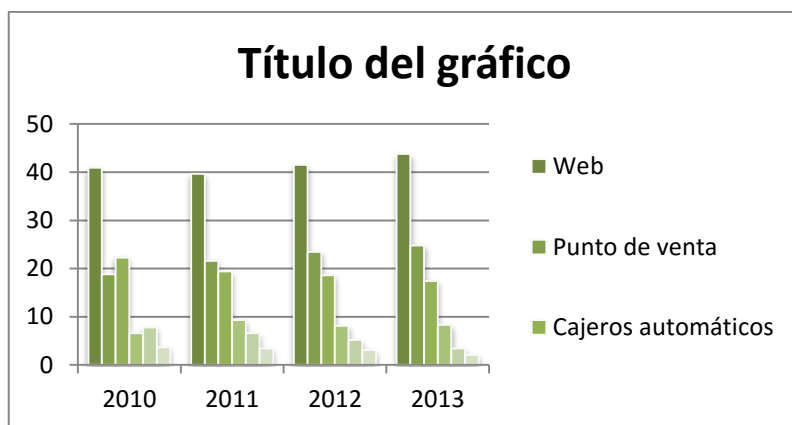


Figura 1. Canales Transaccionales (% del volumen).

Fuente: Banco Central de Venezuela (2013).

De acuerdo con la publicación de Serrano (2013), la figura 1, relativa a la evolución de los canales con mayor transaccionalidad de operaciones bancarias, resaltan la web, los puntos de venta y los cajeros automáticos, los cuales agrupan aproximadamente 84% del total, y se mantienen relativamente estables durante el periodo en estudio. Por su parte, los canales mensajería de texto, oficina y centros de llamadas telefónicas muestran una baja participación y un menor dinamismo en el tiempo.

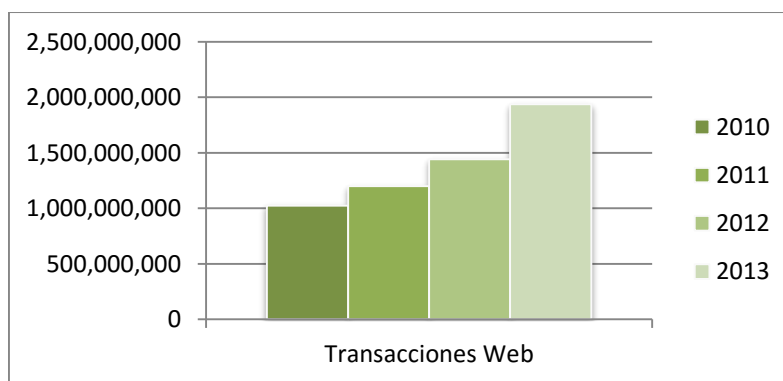


Figura 2. Evolución de las Transaccionales Web (Número de transacciones).

Fuente: Banco Central de Venezuela (2013).

La figura 2 indica que a partir del año 2010, según datos del Banco Central de Venezuela, el número de transacciones web ha tenido un incremento continuo de 15% en el año 2011 comparado con el año anterior, 17% en el 2012 y 26% en el 2013, lo que demuestra que cada vez es mayor el número de operaciones que migran hacia la

banca en línea (sin incluir datos de los años 2014 y 2015 por no contar con estadísticas oficiales).

El crecimiento y desarrollo de la banca en línea observado indica que la forma de realizar operaciones bancarias está cambiando desde hace algunos años, el mundo de las transacciones bancarias se dirige hacia una digitalización casi absoluta, por lo que se hace necesario explorar técnicas y herramientas tecnológicas que permitan elevar la competitividad y mejorar técnicas de negocio, junto a transacciones más rápidas, seguras y con mejor calidad de servicio en sector financiero Venezolano.

No obstante, a pesar del gran volumen de transacciones y datos que los bancos manejan, estos se encuentran dispersos en documentos, bases de datos, redes sociales, etc., esto ha generado la existencia de islas de datos independientes que en ocasiones limitan la visión de la empresa para la toma de decisiones, haciendo que casi toda la información quede ociosa y quienes la tienen en su poder no le sacan el provecho que podría brindarles, desperdiciando su potencial y la oportunidad de generar valor a través del análisis.

Según Bonillo (2014), una de las razones del problema planteado, es el uso del modelo relacional de base de datos, que por sí solo no logra ser una opción viable, por lo que nacen alternativas como las base de datos NoSQL, que junto a un ecosistema Hadoop, forman el corazón de Big Data, proporcionando ventajas en las carencias de las base de datos relacionales.

“Big Data ha dejado de ser solo una tendencia para convertirse en una realidad que está definiendo el escenario de las TIC en Venezuela y a nivel mundial, si los datos se tratan correctamente proporcionan un enorme valor a las organizaciones. (Flores, 2015, pág. 1).

En este contexto se pretende identificar y medir las características que la banca en línea de las instituciones financieras deben cumplir para iniciar la implementación de proyectos de inteligencia de negocios con tecnología Big Data, con el apoyo de investigaciones realizadas en este ámbito y contextualizadas al entorno nacional,

permitiendo determinar si están preparadas para este tipo de emprendimiento, y si el negocio requiere realmente de este tipo de soluciones.

“Big Data no es para todos, podría sonar como una propaganda que desestima el uso de esta solución, pero es una realidad que debe ser analizada antes de iniciar con proyectos de este tipo; entre los factores a analizar se encuentran ciertas características que deben tener las organizaciones que deseen aplicar Big Data” (Guama, 2015, pág. 52).

De acuerdo con lo anteriormente expuesto, se plantea la siguiente formulación del problema, con el fin de reflejar con claridad los aspectos desconocidos que marcan la investigación.

1.1.1 Formulación del Problema

De acuerdo a la descripción del problema, se plantea la siguiente pregunta de investigación:

¿Cuál es el modelo que permita a las organizaciones financieras determinar si están preparadas para implementar un sistema de inteligencia de negocio con Big data para la banca en línea?

1.1.2 Sistemización del Problema

A continuación se plantean las preguntas que constituyen el problema de esta investigación:

¿Cuál es el contexto actual de la Inteligencia de negocio en la banca en línea del sistema financiero Venezolano?

¿Cuál es el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero?

¿Cuáles son las dimensiones y características del modelo de implementación sustentadas en las investigaciones realizadas?

¿Cómo validar el modelo de implementación diseñado mediante un caso de estudio aplicado al Banco Venezolano de Crédito?

1.2. Objetivos

1.2.1 Objetivo General

Diseñar un modelo de implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.

1.2.2 Objetivos Específicos

1. Analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero Venezolano.
2. Identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero.
3. Desarrollar las dimensiones del modelo de implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.
4. Validar el modelo de implementación diseñado mediante su aplicación a un caso de estudio (Banco Venezolano de Crédito).

1.3. Justificación de la Investigación

Para obtener valor de los datos en las organizaciones que ofrecen servicios bancarios, estos deben proporcionar una oportunidad para actuar.

A través de la aplicación de Big Data las entidades financieras podrán enfocar sus propios negocios y los de sus clientes desde un nuevo punto de vista, que les permita desarrollar novedosos productos y servicios, que generen beneficios y oportunidades de negocio adicionales, al tiempo que mejoran sus servicios al cliente y a la propuesta de valor que les ofrecen.

Sin embargo la implementación de sistemas de inteligencia de negocios desde la perspectiva de Big Data dependerá de la visión que tenga la organización para determinar la necesidad de aplicar este tipo de soluciones.

Con la finalidad de presentar una visión uniforme y coherente de la inteligencia de negocio desde la perspectiva de Big Data, se propone un modelo de implementación de un sistema de inteligencia negocio con Big data para la banca en línea en

Venezuela, partiendo de la investigación de métodos existentes y aplicables al entorno nacional, tomando en cuenta características y atributos que estas organizaciones deben considerar para saber si están preparadas y si el negocio realmente requiere de este tipo de tecnología.

El desarrollo de este modelo es de gran relevancia para cualquier organización financiera que desee generar valor para el negocio a través de la adecuada gestión de datos, integrándolos de forma que pueda analizar la información y tener una visión más clara del desempeño de su organización y el descubrimiento de oportunidades.

El estudio de investigaciones existentes para aplicación de Big data, la realidad de la industria y su nivel de madurez permite identificar elementos, considerando las mismas variables mencionadas adaptadas a la realidad del sector financiero venezolano.

Poder identificar con claridad el propósito del uso de Big Data en una institución financiera, es un paso fundamental, establecer si se considera la aplicación de esta tecnología como una evolución del análisis de datos tradicional al análisis de datos no estructurados o como una estrategia para obtener valor de grandes conjuntos de datos que aún no han sido analizados, evidencia que, se tiene una idea sustentada para el desarrollo de un proyecto de inteligencia de negocios con Big Data.

Si por el contrario el propósito del uso de Big Data no está claro, y la organización no cuenta con una cultura de toma de decisiones fundamentada en el análisis de datos o con las habilidades, liderazgo e infraestructura adecuada, la implementación de este tipo de proyectos podría estar condenada al fracaso.

No obstante, tener solamente un propósito claro no garantiza el éxito de este tipo de proyectos, por su parte, características en volumen, velocidad, valor, veracidad y variabilidad de los datos son la base para determinar si realmente una organización requiere una aplicación de Big Data, o si por el contrario, simplemente innova en las soluciones de BI tradicionales o en el incremento de las características de la infraestructura.

Otros factores importantes a considerar para establecer si hacer inteligencia de negocios con Big data generará valor al negocio o se convertirá en un dolor de cabeza, están representados por el apoyo incondicional de la organización, el gobierno de los datos, los procesos existentes y el recurso humano necesario para hacer frente a este tipo de proyectos.

Condensar todas estas variables en un modelo, permite contar con una idea base para el inicio de un proyecto de inteligencia de negocios con Big Data, lo cual es crítico para establecer los objetivos y el alcance, de manera que se asegure la culminación del mismo determinando de forma adecuada los componentes necesarios para su ejecución y se reduzca el riesgo al fracaso.

Finalmente, en un entorno global en el que a diario se crean 2,5 millones de billones de bytes de información, las tecnologías de Big Data pueden ser determinantes, y es muy probable que esta realidad no sea distinta en el sistema financiero venezolano, pero el contraste de evaluar todos estos aspectos antes de aplicar este tipo de proyectos a través de un modelo, podrá marcar la diferencia, permitiendo en primer lugar descubrir las necesidades y puntos de mejora en la organización, y al mismo tiempo que se logre la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea con mayores posibilidades de éxito.

1.4. Alcance y Limitaciones de la Investigación

Este trabajo explora la tecnología de Big data para hacer inteligencia de negocio en la banca en línea de las instituciones financieras venezolanas. Para ello se analizó el entorno de Inteligencia de negocio asociado a la banca en línea, a través de la observación científica y la encuesta como técnicas de recolección de datos, utilizando una lista de chequeo y un cuestionario como instrumentos.

El trabajo no incluye la publicación de datos estadísticos asociados a información que podría llegar a considerarse confidencial por parte de las organizaciones financieras encuestadas.

Para identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero, se realizó una investigación bibliográfica, con el fin de recopilar, revisar y analizar fuentes de investigación especializadas, de diversos aspectos relacionados con inteligencia de negocio y Big Data, haciendo énfasis en los enfoques que puedan aportar valor a la investigación. Esta actividad no contempló el diseño o publicación de un artículo arbitrado.

En base a los resultados obtenidos se desarrolló el modelo, cuyo objetivo final es realizar un aporte, que sirva de guía para implementar un sistema de inteligencia de negocios con Big Data en la banca en línea con mayores posibilidades de éxito. La validación se efectuó a través del juicio de expertos y su aplicación a un caso de estudio en el Banco Venezolano de Crédito.

No se contempló la selección de un proveedor de soluciones tecnológicas de inteligencia de negocios con Big data. El desarrollo del modelo no pretende un paso a paso de cómo desarrollar este tipo de sistemas.

El tiempo para la elaboración del modelo de implementación fue de 10 meses.

Entre las limitaciones de la investigación se encuentran:

Poca información disponible de la experiencia de algunas instituciones financieras que ya están utilizando esta tecnología en diversas áreas de su negocio.

Difícil acceso a las instituciones financieras para obtener una muestra considerable.

Poca disposición y tiempo por parte de las instituciones financieras seleccionadas para participar en la investigación.

Falta de publicación de estadística de la evolución de los canales bancarios, por parte del Banco Central de Venezuela desde hace dos años.

CAPÍTULO II: MARCO TEORICO

En el capítulo se realiza una contextualización del problema de investigación planteado, mediante el desarrollo de una perspectiva teórica. Se detallan los antecedentes, los fundamentos teóricos, el marco referencial y las bases legales que apoyan la investigación.

2.1 Antecedentes

En la llamada "Era de la información", se producen cada vez más datos, en cantidades para muchos inimaginables, para lograr obtener y procesar estos grandes volúmenes de datos surge un nuevo paradigma, denominado Big data. En este punto se presentan aquellos estudios o investigaciones relacionados con el tema planteado.

2.1.1 Trabajos de Grado de Maestría (TGM)

Guama (2015), Estudio comparativo de métodos existentes para integrar la información estructurada y no estructurada de una industria enfocado en la generación de conocimiento, desde la perspectiva de una solución integral de Big data, para optar al título de Magister en Gerencia de Sistemas y Tecnologías de la Información en la Universidad de las Américas. Este estudio se basa en el análisis de las principales técnicas de análisis e integración de datos desde la perspectiva de Big data para generar conocimiento dentro de una organización a través de los resultados generados. En él se realiza una descripción de diferentes técnicas que pueden ser aplicadas en diferentes verticales de la organización, como la minería de datos, el análisis de Clúster, regresión, redes neuronales, Ensemble, Testing, redes, estadísticas, reconocimiento de patrones, entre otros. También se realiza un análisis de la evolución tecnológica de Big data, apoyado en estadísticas presentadas por Gartner, Inc, así como el crecimiento de este mercado en américa latina de IDC Big Data Report, Describe la perspectiva del negocio como un factor determinante para implementar este tipo de soluciones. Finalmente el autor indica que el análisis de los datos estructurados y no estructurados es una de las mejores opciones para la

generación de conocimiento y orientar la toma de decisiones, indicando que la participación del área de TI en la implementación de Big data debe ser una estrategia de negocio para apoyar el trabajo de los expertos en análisis de datos y en la administración de herramientas, y no se debe incluir al área de TI como el administrador de este tipo de proyectos pues es importante que su dirección este a cargo de un área de negocio.

Aportes: El estudio aborda un importante análisis en cuanto a la integración de datos estructurados y no estructurados, así como también, presenta algunos aspectos claves que ayudan a los expertos a definir el mejor momento para la aplicación de Big Data, lo cual es parte fundamental para desarrollo del modelo de inteligencia de negocio con tecnología Big data propuesto.

Palabras clave: Análisis de datos, integración de la información, Big Data, Data.

Espinoza (2015), Generar un marco de referencia para implementaciones de Big data en empresas de telecomunicaciones, caso de estudio corporación nacional de telecomunicaciones, para optar al título de Magister en Gerencia de Sistemas y Tecnologías de la Información en la Universidad de las Américas. En este trabajo se realiza una amplia conceptualización de los términos relacionados a Big data. Se refiere al término como una tecnología de exploración de la información caracterizada por el volumen, la variedad y la velocidad para generar valor y veracidad, combinando varias disciplinas como la computación, analítica avanzada, marketing virtualización entre otros. Concluye que la información existente en una empresa de telecomunicación procede de diferentes plataformas relevantes para el negocio, si esta información es analizada y procesada de la mejor manera será una oportunidad para la generación de nuevos ingresos, valor para la empresa y optimización de los servicios. Planificar Big data no solo se trata de entender lo nuevo o diferente, sino de cómo integrar lo nuevo a lo existente desde la infraestructura de base de datos, BI hasta herramientas informáticas y usuario final. Recomienda el uso de Hadoop y MapReduce como las tecnologías que más se adaptan y encajan con Big Data, ya que

proporcionan escalabilidad, tolerancia a fallos y soporta diferentes cambios en la arquitectura tanto a nivel de diseño o funcionalidades como a nivel de infraestructura.

Aportes: El marco de referencia presentado en este trabajo permitió contextualizar las propuestas, teorías y mejores prácticas relacionadas con la adopción e implementación de Big Data en la banca en línea del sistema financiero venezolano, además de aportar una importante base para el desarrollo del marco teórico de la investigación.

Palabras clave: Big Data, Data Analytics, Hadoop, MapReduce.

López (2013), Análisis de las posibilidades de uso de Big Data en las organizaciones, para optar al título de Magister en Empresas y Tecnologías de la Información y la Comunicación en la Universidad de Cantabria. El objetivo general de este trabajo es explicar en qué consiste el término Big Data y a qué hace referencia. En él se intenta aclarar en qué consiste, su alcance y su situación actual o estado del arte, abarcando términos relacionados con Big data, como la minería de datos, el Cloud Computing o Data Warehouse, analizando también distintas plataformas de software para el tratamiento de Big data. Igualmente se aclarara porqué surge, de donde procede y por que para muchos tecnólogos sugiere un cambio de etapa en el mundo de las TIC. Finalmente se realiza un breve análisis de casos de empresas que utilizan actualmente esta tecnología. El autor concluye que Big data no solo sirve para obtener grandes cantidades de datos, sino que también sirven para analizar esos enormes volúmenes de datos encontrando información y conocimiento, considera que a pesar de ser una tecnología emergente ha llegado para quedarse y marcará una nueva etapa en el mundo de las tecnologías de la información. Recomienda antes de instalar en una organización Big Data es necesario hacer un Business Case para ver si es viable o no el proyecto, y a las empresas que opten por implementarla tener una infraestructura muy completa para no sufrir robos de datos y sus indemnizaciones correspondientes a sus clientes.

Aportes: El estudio aborda la implementación de Big Data utilizando un Business case, como instrumento estratégico para valorar y tomar la mejor decisión respecto a un

proyecto de inversión, del cual fueron extraídos aspectos importantes a considerar para la implementación de sistemas de inteligencia de negocio con Big Data para la banca en línea en Venezuela.

Palabras clave: Big Data, Data Warehouse, Data Mining, Cloud Computing, Hadoop.

2.1.2 Artículos Técnicos

Montalvo (2015), El impacto del Big data en los servicios financieros. En este trabajo se analizan las posibilidades del “Big data” para mejorar los servicios financieros y aumentar la eficiencia en un contexto de presión, sobre la rentabilidad de los bancos y la necesidad de recuperar la reputación del sector. La utilización de técnicas de “Big data”, incluidos indicadores de reputación y capital social online, se ha extendido a la calificación crediticia de los solicitantes de crédito, la detección de fraude en tarjetas, la microsegmentación, los servicios de información a los clientes, el cumplimiento normativo, la prevención de blanqueo de capitales y operaciones de financiamiento al terrorismo, entre otras muchas actividades del sector. El trabajo también analiza algunos aspectos asociados a la utilización de “Big data” en cuestiones de privacidad de los datos y el cumplimiento de las regulaciones sobre utilización de la información.

Aportes: Este trabajo sirve para tener una visión más amplia de la aplicación de Big data en los servicios financieros, ya que muestra varios casos de estudio y también presenta algunos riesgos específicos para este sector, lo cual significa un importante aporte para el modelo planteado, por ser la industria en la que se desarrolla la presente tesis.

Palabras clave: “Big data”, “machine learning”, calificación crediticia, fraude, microsegmentación, privacidad.

Villalobos (2015), Relación entre las bases de datos NoSQL y Big Data. En el artículo publicado por la Escuela de Ingeniería de la Universidad Latinoamericana de Ciencia y Tecnología, se lleva a cabo un análisis de la situación actual de los SGBD, y su

evolución desde el surgimiento de los primeros sistemas de almacenamiento, hasta las nuevas tecnologías de bases de datos disponibles en el mercado, así como su implementación. El autor concluye que es necesario analizar detenidamente el requerimiento que se deberá satisfacer antes de tomar alguna decisión sobre las TIC por implementar, con la finalidad de disponer de los insumos que permitan seleccionar las herramientas necesarias, para solventar las necesidades que originan su adquisición de forma eficaz y eficiente, conforme a la capacidad de quien vaya a implementarlas.

Aportes: Este artículo nos da un importante aporte ya que establece una relación entre las bases de datos NoSQL y Big data, destacando términos y teoremas asociados a bases de datos, cuyas características son indispensables y deben ser evaluadas al momento de implementar sistemas de inteligencia de negocios con Big Data en la banca en línea.

Palabras clave: NoSQL, Big Data, Big Data Analytics, Bases de Datos

Ernst & Young (2014), Big Data en el sector financiero español. Contiene el estudio realizado en base a una encuesta sectorial sobre Big Data, realizada por EY, FrontQuery y Teradata para entender los principales retos que Big Data plantea al sector financiero, Banca y Seguros en el mercado español, así como conocer el estado y los planes de desarrollo de iniciativas de Big Data. Para ello, se remitió un cuestionario con ocho preguntas a directivos de las áreas comercial, marketing, sistemas / tecnología, operaciones, inteligencia de clientes y organización. El primer paso para derivar valor de Big Data no debería ser “hablar de cómo derivar valor de Big Data”, sino empezar preguntando “qué decisiones de negocio debo adoptar de modo más eficiente y eficaz”.

Aportes: El aporte de este estudio para la investigación está relacionado con la visión general que brinda en cuanto a oportunidades y aplicaciones de Big Data en el sector financiero, así como también, los retos que plantea, que extraídos directamente de directivos de distintas áreas de la industria financiera española, sirve como referencia para el desarrollo del modelo propuesto en el contexto venezolano.

Palabras clave: Big data, Analytics, financiero, Banca y Seguros.

Paredes (2015), Big Data: Estado de la cuestión. Es un artículo arbitrado en el que el autor se enfoca en el estado actual de Big Data, de dónde proviene la información, los tipos de datos se deben analizar, sus características principales y como ha venido desarrollándose. Por otra parte indica que aproximadamente la mitad de las empresas tratadas en una encuesta identificaron los objetivos centrados en el cliente como la máxima prioridad. Analiza las fuentes internas y externas de información, las fases de adopción de un modelo de Big data, los posibles obstáculos y finalmente algunos ejemplos de áreas de aplicación de esta tecnología.

Aportes: El estudio fue una importante base para el desarrollo del marco teórico de la investigación, además de dar a conocer algunas recomendaciones que deben considerar las empresas a medida que avanzan en la implementación de Big Data, las cuales fueron contextualizadas en el sistema financiero venezolano.

Palabras clave: Big Data, Cloud Computing, Web and Social Media, Machine-to-machine (M2M), Big Transaction Data, Hadoop.

Camargo (2014), Arquitectura Tecnológica Para Big Data. Es un artículo de investigación en el que se estudia, analiza y da a conocer de manera exhaustiva las diferentes arquitecturas de Big Data, con sus características, herramientas, tecnologías y estándares relacionadas con dicho termino, con el único objetivo de brindar una ayuda al sector empresarial para una posible implementación de Big Data.

Aportes: El aporte de este trabajo en cuanto a los componentes de las arquitecturas de Big data que plantea, permitió extraer aspectos y características importantes a tomar en cuenta para el desarrollo del modelo de implementación de un sistema de inteligencia de negocio con Big Data para la banca en línea en Venezuela.

Palabras Clave: Big Data, Hadoop, NoSQL, calidad de datos, modelo de datos, arquitectura.

Delrieux et al. (2014), Clasificación de Información en Big Data mediante la utilización de Técnicas de Inteligencia Artificial y Análisis de Redes Sociales. Se trata de un estudio que se enfoca a evaluar técnicas existentes e implementar desarrollos experimentales que permitan clasificar, ordenar, jerarquizar y analizar información sobre grandes volúmenes de datos heterogéneos (Big Data Analytics), mejorando y/o reformulando dichas técnicas. Dichas técnicas permitirán establecer las capacidades necesarias con las que debería contar una base de datos de información masiva, tanto desde la perspectiva de almacenamiento y técnicas de indexación, como de distribución de las consultas, escalabilidad y rendimiento en ambientes heterogéneos. Para ello se utilizarán Corpus científicos existentes de investigadores colaboradores del Centro Nacional Patagónico - CONICET, con la finalidad de contar con expertos que nos permitan realizar y estructurar los mecanismos de clasificación y evalúen los resultados de los desarrollos experimentales en estos Corpus. Adicionalmente a estos se analizará información pública digital en repositorios de portales y plataformas de información, tanto científica como de acceso general.

Aportes: El estudio presenta algunos aportes ya que nos habla de las diferentes técnicas para tratar grandes volúmenes de datos, del cual se extrajo información útil para desarrollar el modelo de implementación de un sistema de inteligencia de negocio con Big Data para la banca en línea en Venezuela.

Palabras clave: Big data analytics, extracción, clasificación, ontologías, escalabilidad.

2.2. Fundamentos Teóricos

Las figuras 3 y 4 muestran un mapa conceptual general y uno desplegado en temas y subtemas que formarán parte del desarrollo del marco teórico asociado a la inteligencia de negocio, el cual se profundizará a continuación.



Figura 3. Mapa del marco teórico de Inteligencia de negocio.

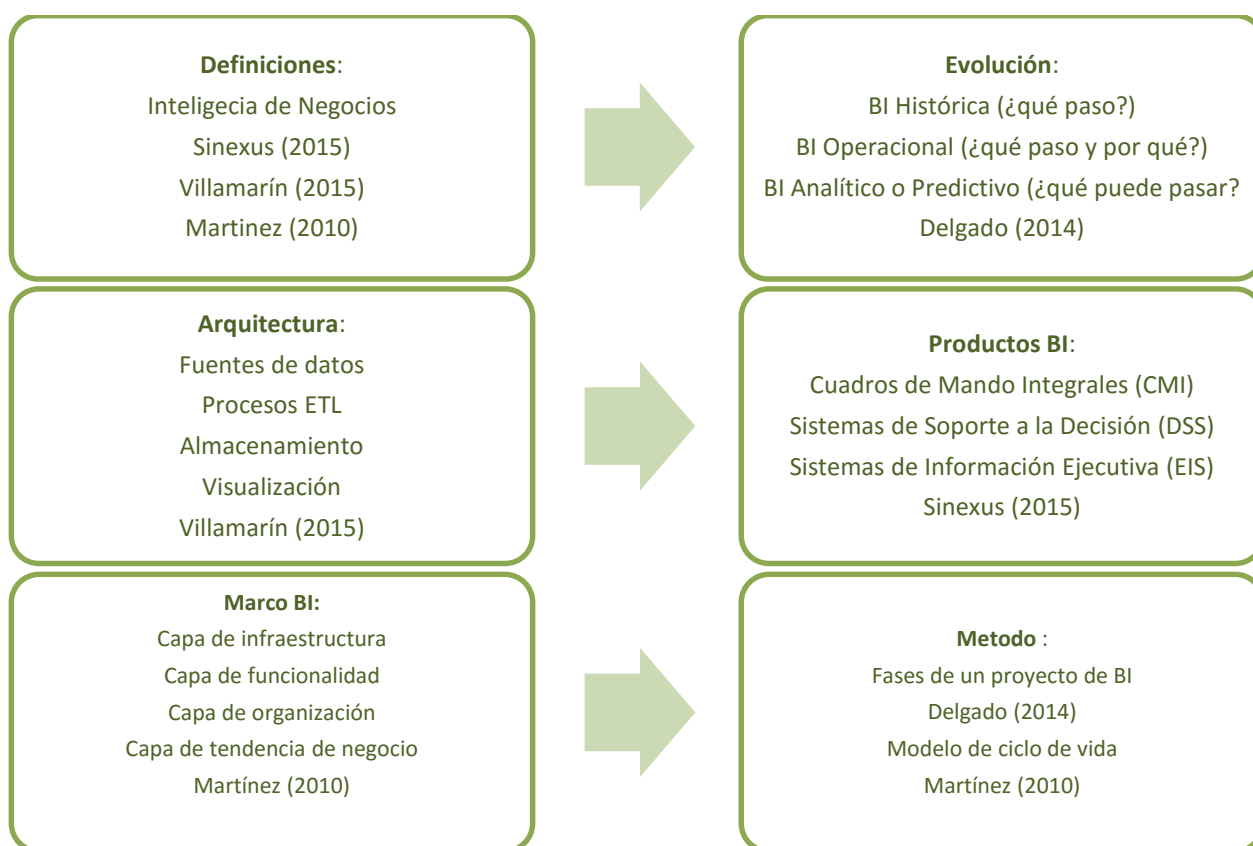


Figura 4. Mapa del marco teórico de Inteligencia de negocio desplegado en temas y subtemas.

2.2.1 Inteligencia de negocio

Roo & Boscán (2012), Afirman que la creciente demanda de información organizacional, ha originado la creación de nuevas herramientas y procesos acordes al ritmo cambiante y desmesurado que lo caracteriza, siendo la toma de decisiones un aspecto relevante dentro de la operación de toda empresa. Partiendo de esta premisa, surge el término inteligencia de negocios (business intelligence-BI), expresión que hace referencia al proceso de toma de decisiones en función del rendimiento organizacional, convirtiendo simples datos en información relevante para los diversos niveles, orientándose en primera instancia al área estratégica organizacional.

“Inteligencia de negocio es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios”. (Sinnexus , 2015).

Según Villamarín (2015), Es el proceso mediante el cual se hace uso de los recursos (humanos, tecnológicos y activos de una organización) para la generación de información oportuna, precisa y confiable; que permita una toma de decisiones ágil y efectiva, en la dirección de una organización, la comprensión del negocio y la ejecución de sus procesos. Con esto se busca, mantener el control sobre el cumplimiento de los objetivos, generando valor y ventaja competitiva en tiempo real, lo que se traduce en mayores ganancias para sus dueños o accionistas.

La inteligencia de negocio hace posible la gestión de información tanto estructurada como no estructurada, producto de la unión del procesamiento de datos y el uso de herramientas analíticas que agrupan la gestión de la información y del conocimiento, convirtiéndose en una ventaja competitiva en el horizonte organizacional.

“La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de la producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc...” (Sinnexus , 2015).

Martínez (2010), define inteligencia de negocio como un conjunto de técnicas y herramientas tanto de gestión empresarial como de aplicación tecnológica que permiten a partir de la formulación estratégica y teniendo como objetivo dar soporte a los procesos de planeación y control en las organizaciones, la extracción e integración de los datos que son generados como resultado de la operación de las diferentes áreas funcionales en una organización, su posterior procesamiento y distribución en forma de información.

“La inteligencia de negocios puede ser considerada en últimas como un término sombrilla bajo el cual se integran varias teorías, técnicas y herramientas para la gestión organizacional basados en tecnologías informáticas”. (Martínez, 2010).

2.2.2 Evolución de la Inteligencia de negocio

Delgado (2014), ilustra la evolución que ha seguido la inteligencia de negocio, desde los sistemas que respondían a la pregunta “¿qué pasó?”, pasando por “¿qué pasó y por qué?”, hasta llegar a las herramientas que hoy en día nos permite responder a “¿qué puede pasar?”.



Figura 5. Evolución de la Inteligencia de negocio.

Fuente: Elaboración propia a partir de Delgado (2014).

Considerando esta evolución y madurez que ha sufrido el BI, puede identificar tres categorías de sistemas de información empresarial:

BI Histórica (¿qué paso?): Recoge los datos de las aplicaciones que procesan las transacciones y su análisis posterior ayuda a quien toma decisiones a comprender qué pasó y cuáles situaciones fueron buenas o malas. Ejemplo: sistema de indicadores financieros.

BI Operacional (¿qué paso y por qué?): También denominada BI Operativa. Corresponde a un proceso iterativo y más complejo que el anterior. Recoge los datos de los sistemas que procesan transacciones pero se añaden datos de tiempo real. El proceso se hace por ensayo y error. Ejemplo: sistema de fijación de precios.

BI Analítico o Predictivo (¿qué puede pasar?): Revela sucesos que se vienen más adelante. Integra datos de fuentes internas y externas a la empresa, y con el uso de herramientas estadísticas y matemáticas, generar información de posibles eventos futuros. Ejemplo: un sistema para predecir el comportamiento de clientes ante el lanzamiento de un nuevo producto.

2.2.3 Arquitectura de una solución de inteligencia de negocio.

El trabajo de Villamarín (2015), basado en el artículo del Harvard Business School Press “The Architecture of Business Intelligence”, muestra la arquitectura de una solución de inteligencia de negocio, compuesto en primer lugar por las fuentes de datos, que contempla todos los procesos operacionales en los cuales se crean datos constantemente, que deben ser administrados (depurados, filtrados, adecuados) para que sean de utilidad e información referente a estos mismos datos del proceso (metadata). Una vez que se tiene establecido el insumo útil que permitirá generar la información relevante de un negocio, se aplican procesos conocidos como ETL con herramientas especializadas para este fin y, como todo proceso, entregar un producto para ser utilizado. Este producto es la información precisa y preparada para mostrar al usuario, para lo cual se emplean herramientas de análisis y visualización de datos, con una variedad de funcionalidades que permiten obtener fácilmente el estado de un

proceso en línea o consultar el comportamiento a través de un periodo de tiempo, según la necesidad o relevancia para tomar una decisión importante.

Extract, Transform, Load (ETL): Hace referencia a tres procesos que intervienen en la adecuación de la información para ser presentada (por sus siglas en inglés: Extract, Transform, Load). La extracción consiste en tomar de las fuentes de datos, aquellos que se requieran para combinar, calcular o mostrar al usuario; en el estado original en que se encuentran. La transformación consiste en aplicar las operaciones o tratamiento de los datos, que se requieran, para ser mostrados de tal manera que sea entendible y de valor para el usuario. La carga consiste en entregar o guardar los datos, en la herramienta que se vaya a utilizar para consulta por parte del usuario final.

Datamart y Datawarehouse: De acuerdo con Sinnexus (2015), son los principales componentes de orígenes de datos en la inteligencia de negocio actualmente.

“Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un datawarehouse, o integrar por sí mismo un compendio de distintas fuentes de información.” (Sinnexus , 2015).

Villamarín (2015), define DataWareHouse como un repositorio o almacén de datos, en donde se guardan todos aquellos datos tomados de diferentes fuentes, para ser posteriormente agrupados o procesados según las consultas que se requieran. Esta práctica permite obtener los datos relevantes para la toma de decisiones y a su vez ejecutar consultas sin que se afecte el rendimiento, ni la integridad de las bases de datos fuente.

Análisis y visualización de datos: En cuanto al análisis y visualización de datos Villamarín (2015), indica que es el resultado final de todo el proceso de depuración, transformación y carga de los datos, visible para un usuario. Las herramientas actuales ofrecen diversas opciones, para que el usuario pueda consultar los informes o indicadores del proceso; a través de un dispositivo móvil, una página web interna o en

Internet, un computador de escritorio según la necesidad u oportunidad de acceso en determinado momento.



Figura 6. Principales productos de inteligencia de negocio.

Fuente: Elaboración propia a partir de Sinnexus (2015).

En la publicación de Sinnexus (2015), muestra los cuadros de mando integrales (CMI), los sistemas de soporte a la decisión (DSS) y los sistemas de información ejecutiva (EIS) como los principales productos de inteligencia de negocio que existen hoy en día:

Cuadros de Mando Integrales (CMI): Según Sinnexus (2015), el Cuadro de Mando Integral (CMI), también conocido como Balanced Scorecard (BSC) o dashboard, es una herramienta de control empresarial que permite establecer y monitorizar los objetivos de una empresa y de sus diferentes áreas o unidades.

El Cuadro de Mando Integral se diferencia de otras herramientas de Business Intelligence, como los Sistemas de Soporte a la Decisión (DSS) o los Sistemas de Información Ejecutiva (EIS), en que está más orientados al seguimiento de indicadores que al análisis minucioso de información. Por otro lado, es muy común que un CMI sea controlado por la dirección general de una compañía, frente a otras herramientas de Business Intelligence más enfocadas a la dirección departamental. El CMI requiere, por tanto, que los directivos analicen el mercado y la estrategia para construir un

modelo de negocio que refleje las interrelaciones entre los diferentes componentes de la empresa (plan estratégico). Una vez que lo han construido, los responsables de la organización utilizan este modelo como mapa para seleccionar los indicadores del CMI. (Sinnexus , 2015).

Sistemas de Soporte a la Decisión (DSS): Un Sistema de Soporte a la Decisión (DSS) es una herramienta de Business Intelligence enfocada al análisis de los datos de una organización afirma Sinnexus (2015). El DSS es una de las herramientas más emblemáticas del Business Intelligence ya que, entre otras propiedades, permiten resolver gran parte de las limitaciones de los programas de gestión.

El principal objetivo de los Sistemas de Soporte a Decisiones es, a diferencia de otras herramientas como los Cuadros de Mando (CMI) o los Sistemas de Información Ejecutiva (EIS), explotar al máximo la información residente en una base de datos corporativa (datawarehouse o datamart), mostrando informes muy dinámicos y con gran potencial de navegación, pero siempre con una interfaz gráfica amigable, vistosa y sencilla.

Sistemas de Información Ejecutiva (EIS): Un Sistema de Información para Ejecutivos ó Sistema de Información Ejecutiva es una herramienta software, basada en un DSS, que provee a los gerentes de un acceso sencillo a información interna y externa de su compañía, y que es relevante para sus factores clave de éxito. Sinnexus (2015).

La finalidad principal es que el ejecutivo tenga a su disposición un panorama completo del estado de los indicadores de negocio que le afectan al instante, manteniendo también la posibilidad de analizar con detalle aquellos que no estén cumpliendo con las expectativas establecidas, para determinar el plan de acción más adecuado.

2.2.4 Marco de inteligencia de negocios

Martínez (2010), presenta un marco propuesto por Howard Dresner referido como “The BI framework”. Este marco sugiere que las empresas alineen sus iniciativas de

inteligencia de negocios en cuatro niveles, donde cada una de las capas afecta a las otras y el beneficio máximo de una iniciativa de BI sólo se consigue si la empresa comprende cómo las capas trabajan conjuntamente. Esto es sumamente importante a la luz de los resultados observados en este tipo de proyectos donde por ejemplo, es poco útil contar con una infraestructura si no hay funcionalidad de inteligencia de negocios que la utilice. Al igual no tiene sentido aplicar inteligencia de negocios de una forma en la que no se involucren los métodos de gestión y la cultura organizacional.

Capa de infraestructura: De forma general, es donde los datos son recogidos, integrados y puestos a disposición de las aplicaciones analíticas. En condiciones normales esto es hecho por una bodega de datos alimentada por procesos de extracción, transformación y carga. No obstante, hay escenarios en los cuales se busca una aproximación de tiempo real más cercana y por ende se puede incluir un almacén de datos operacionales (Operational Data Store ODS) vinculado a los flujos operativos. (Martínez, 2010).

Capa de funcionalidad: La capa de funcionalidad consiste de las aplicaciones y plataformas de Inteligencia de negocios al igual que las herramientas de consulta. Dado que generalmente ninguna herramienta es capaz de soportar el amplio espectro de requerimientos de negocio usualmente se hace necesario definir un portafolio de aplicaciones de BI. Cada una de estas aplicaciones sirve una o varias necesidades específicas dentro de un dominio de la empresa. (Martínez, 2010).

Capa de organización: Las características organizacionales conducen cuáles funciones de inteligencia de negocios son necesarias e igualmente importantes, la forma en que éstas se van implementando. Las empresas deben tener en cuenta varias inquietudes: ¿Qué tan madura es la gestión del desempeño? Es decir, la empresa se encuentra lista para aplicar alguna metodología como balanced scorecard, o en vez de ello necesitan mejorar la planeación y control de una manera más simple. ¿Cuál es el foco primario de la inteligencia de negocios? Es decir si se busca dar respuesta a inquietudes operacionales, tácticas o estratégicas. ¿Cómo se vive la cultura informacional en la empresa? ¿Qué habilidades analíticas tiene la empresa?

Es decir, ¿Para qué se usa la información? Se utiliza para tomar decisiones, planear la empresa, comunicar o es solo por llenar requisitos. ¿Cuáles tipos de usuario se identifican y como deben ser provistos de información? Información programada, por demanda, reportes estáticos, etc. (Martínez, 2010).

Capa de tendencia de negocio: En esta capa se mantiene actualizado el contexto de la organización así como el contexto tecnológico para mantener vigente la solución de inteligencia de negocios dentro de la empresa. (Martínez, 2010).

2.2.5 Método para implantar proyectos de BI

Delgado (2014), en su trabajo de grado de Mestria, hace referencia al método de desarrollo de proyectos de Inteligencia de negocios apoyado en el mapa de ruta propuesto por Larissa Moss en su libro “Business Intelligence Roadmap”. El método propone la realización del proyecto de BI en ocho fases, las cuales se pueden ver en a continuación:

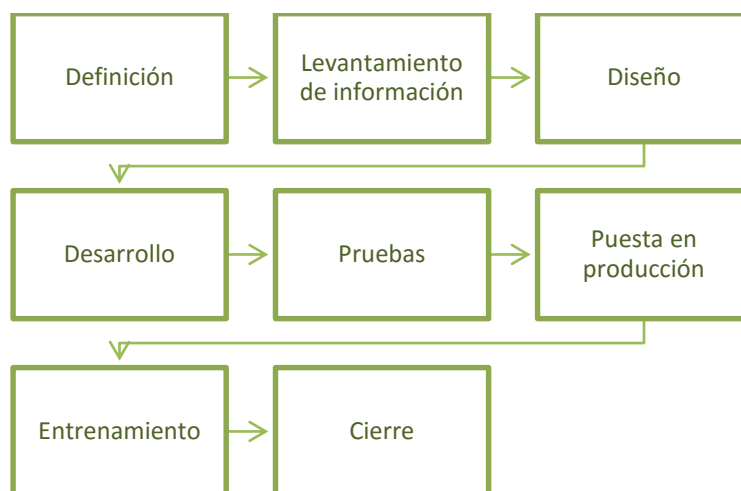


Figura 7. Fases de un proyecto de BI.

Fuente: Elaboración propia a partir de Delgado (2014).

La primera fase de definicion, involucra tareas de formalización del proyecto, formalización del equipo de trabajo y definición del plan. El levantamiento de información incluye; la identificación de las áreas y procesos de negocio, levantamiento

de la infraestructura tecnológica, identificación de aplicaciones operacionales y fuentes de datos, identificación de requerimientos de información, identificación de requerimientos de aplicación, requerimientos no funcionales y análisis de fuentes de datos. El diseño se refiere a; diseño del modelo del negocio, diseño del DWH, diseño ETL, diseño de reporte y otras interfaces de análisis, diseño de indicadores, diseño del modelo de análisis OLAP, diseño de la aplicación y el plan de capacidad de hardware y software. El desarrollo contempla; la instalación de la infraestructura de desarrollo, construcción del DWH, construcción de ETL, construcción de aplicación, construcción de esquemas OLAP, construcción de reportes y otras interfaces de visualización y el desarrollo de la documentación técnica.

La fase de pruebas contempla; definición del plan de pruebas, instalación del ambiente de pruebas, verificación del DWH, pruebas de aplicación, pruebas de ETL, pruebas de reportes e interfaces de visualización. Seguidamente la fase de producción incluye; preparación del ambiente de producción y la verificación del ambiente de producción. La fase de entrenamiento implica; la preparación del plan de entrenamiento, preparación del material y la ejecución del entrenamiento. Finalmente en la fase de cierre se llevan a cabo las tareas de entrega formal de los productos del proyecto.

2.2.6 Metodología y modelo de ciclo de vida de una solución de Inteligencia de negocios

Martínez (2010), describe brevemente una metodología y modelo de ciclo de vida para una solución de Inteligencia de negocios, propuesta por el grupo Gartner, el cual si bien es en sí mismo sólo una guía general, permite una mayor comprensión del proceso de implementación y así mismo le da rumbo a este proceso. Este modelo abarca ideas asociadas a ingeniería concurrente y metodologías de desarrollo iterativo de aplicaciones.

“A diferencia de otros modelos, este anima a todas las partes implicadas en el proyecto para romper las barreras tradicionales en la especificación, desarrollo y uso de las aplicaciones. Este factor es muy relevante puesto que como se verá en la investigación aplicada, es frecuente que la expectativa sea mayor que el resultado con lo cual se puede perder la motivación inicial en

este tipo de implementaciones. En consecuencia, romper estas barreras con un ciclo más fluido donde todas las partes participan a la vez en los diversos pasos y de esa forma ver resultados más rápidamente”. (Martínez, 2010, pág. 60).

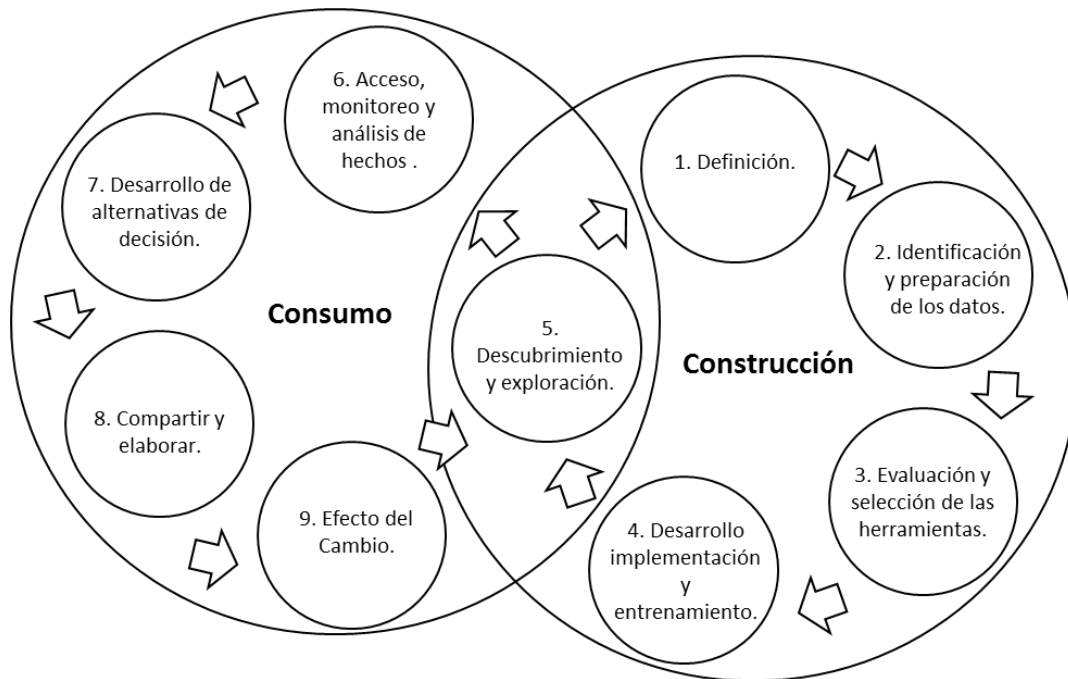


Figura 8. Modelo de ciclo de vida de desarrollo de una solución de inteligencia de negocios.

Fuente: Martínez (2010).

La metodología se divide en dos ciclos distintos: Por un lado el ciclo de construcción y por otro lado el ciclo de consumo, que sin embargo se engranan mutuamente permitiendo un proceso iterativo de mejoramiento de la solución de Inteligencia de negocios. Los pasos que se involucran en la metodología son:

En la tabla 2 se presenta una breve definición de los componentes del ciclo de vida de desarrollo de una solución de inteligencia de negocios.

Tabla 1. Ciclo de vida de desarrollo de una solución de inteligencia de negocios.

Definición:	En este paso se define la información que se requiere, y por lo tanto debe involucrar a los usuarios finales así como a los técnicos.
Identificación y preparación de datos:	Determinar qué fuentes de información pueden ser usadas para soportar las necesidades de negocio, así como las transformaciones que deben realizarse. En este paso debe identificarse la calidad de los datos en las fuentes.
Selección de las herramientas:	Las herramientas pueden variar desde tecnologías básicas de presentación de informes hasta sofisticadas plataformas de Inteligencia de negocios. La selección de la herramienta no sólo debe considerar el aspecto económico, sino principalmente las necesidades del usuario y la complejidad de la solución. Las herramientas deben ser evaluadas periódicamente puesto que los proveedores de software constantemente van implementando mejoras en sus sistemas.
Desarrollo, implementación y entrenamiento:	Los desarrollos dependen de la selección de la herramienta y la complejidad de la implementación. En algunos casos es necesario personalizar una herramienta, en otros casos el desarrollo debe hacerse a la medida con lo cual el tiempo de diseño y desarrollo se incrementa. Independientemente, el éxito en el uso de una tecnología radica en la forma como esta se involucra en las actividades de los usuarios finales, por esto es muy importante desarrollar sus habilidades mediante el entrenamiento. A su vez, los usuarios finales deben retroalimentar con el análisis de su experiencia al área técnica para inducir mejoras.
Descubrimiento y exploración:	Una vez dispuestas las herramientas de Inteligencia de Negocios debe comprenderse la forma como se va a usar. Esto debe hacerse puesto que cada organización es particular y así mismo es particular la forma como se usan las herramientas. En este paso se cuantifican y analizan los resultados de la utilización del sistema de Inteligencia de negocios y por ende es donde se engranan los ciclos de desarrollo y consumo.
Acceso, monitoreo y análisis de hechos:	Ocurre cuando los usuarios finales acceden, analiza e interpretan la información, llevándolo a una mejor perspectiva y comprensión del negocio y en últimas a la generación de conocimiento
Desarrollo de alternativas de decisión:	En este punto, el usuario final tiene una mejor comprensión de los fenómenos que ocurren en la organización y se encuentra en capacidad de proponer alternativas de solución a los interrogantes inicialmente planteados.
Compartir y colaborar:	La información generada así como las alternativas planteadas deben ser compartidas con otras personas de tal forma que estas se puedan contrastar con otras perspectivas. Este proceso pone de manifiesto situaciones en las que se pueden descartar alternativas gracias al conocimiento de otras personas sobre situaciones particulares.
Efectuar el cambio:	Una vez se ha tomado una decisión, esta debe llevarse a cabo y medirse sus resultados, con lo cual nuevas preguntas van surgiendo y por ende nuevamente comienza el ciclo de vida de la solución de inteligencia de negocios, enfocándose primordialmente en la re-evaluación, modificación, optimización y análisis de la información disponible.

2.2.7 Big Data

Las figuras 9 y 10 muestran un mapa conceptual general y uno desplegado en temas y subtemas que formarán parte del desarrollo del marco teórico asociado a Big data, el cual se profundizará a continuación.



Figura 9. Mapa del marco teórico de Big Data.

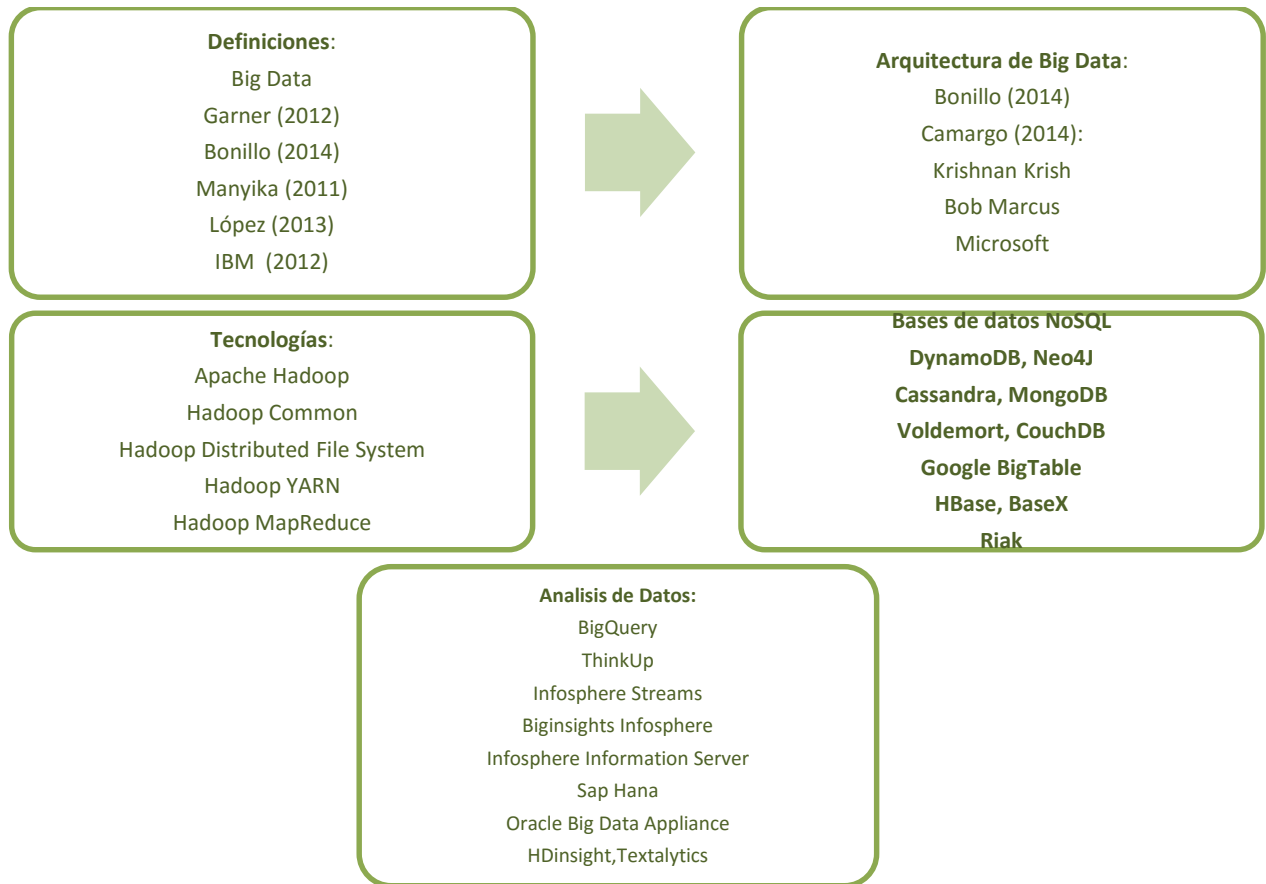


Figura 10. Mapa del marco teórico de Big Data desplegado en temas y subtemas.

Según Gartner (2014), desde la presentación del término por el MGI (McKinsey Global Institute) en Junio de 2011, han existido diversos intentos de acotación del concepto.

En la tabla 3 se presenta la definición de distintos autores para el concepto de Big Data.

Tabla 2. Definiciones de Big Data.

Autor:	Definición:
Gartner (2014) citado por López (2013)	Son activos de información caracterizados por su alto volumen, velocidad y variedad, que demandan soluciones innovadoras y eficientes de procesamiento para la mejora del conocimiento y toma de decisiones en las organizaciones.
Bonillo (2014)	Big Data (del idioma inglés “grandes datos”) es en el sector de tecnologías de la información y la comunicación una referencia a los sistemas que manipulan grandes conjuntos de datos (o data sets).
López (2013)	Definen Big Data como el conjunto de datos cuyo tamaño va más allá de la capacidad de captura, almacenado, gestión y análisis de las herramientas de base de datos.
López (2013)	Big Data es el término inglés que designa los conjuntos de datos de gran tamaño y generalmente desestructurados que resultan difíciles de manejar usando las aplicaciones de bases de datos convencionales
López (2013)	Big Data como “la enorme cantidad de datos que desde hace unos años se genera constantemente a partir de cualquier actividad.
López (2013)	El resultado de una encuesta a más de 1144 negocios y profesionales de TI de 95 países y docenas de expertos en la materia, las características que mejor describen el concepto de Big Data (en orden de importancia) son: Un mayor ámbito de información, nuevos tipos de datos y análisis, información en tiempo real, estrada de datos procedentes de nuevas tecnologías, formas no tradicionales de soporte, grandes volúmenes de datos, la última palabra de moda y datos de redes sociales.

De acuerdo con las definiciones presentadas en la tabla anterior, se puede concluir que el termino Big Data está asociado a la gestión de una cantidad de datos significativa, con una complejidad particular en cuanto a volumen, variedad y velocidad de los mismos, a través del uso de nuevas tecnologías.

Fases de la Gestión de Datos

Con el fin de lograr un mejor entendimiento de Big Data es necesario analizar la evolución de los datos y la forma en que se gestionan. Las fases nacieron de la necesidad de resolver un tipo de problema en la gestión de datos, la tabla 3 muestra las fases de gestión de datos. (Espinoza , 2015).

Tabla 3. Fases de la Gestión de Datos.

Fase	Problema a Resolver	Modo de Resolución
1. Creación de estructuras de manejo de datos	Manejo de archivos planos	Primeras estructuras RDBMS Modelo E-R
	Crecimiento de información	Aparecen DWH Refinamiento DataMarts
	Aparecen datos No/Semi-Estructurados	Tipo de datos BLOB ODBMS
2. Gestión de web y contenido	Nuevos Tipos de Datos Virtualización, Web, Cloud	Inclusión de metadata
3. Gestionando Big Data	3 Vs	Big Data

Fuente: Elaboración propia a partir de Espinoza (2015).

Big Data está construido en el tope de la evolución de la gestión de los datos y por otro lado ha alcanzado un punto común entre el costo del ciclo de computación y el almacenamiento, esto implicaba que las organizaciones se veían obligadas a almacenar únicamente cierta información por el costo que involucra pero esto limitaba a dejar a un lado datos que podrían ser útiles.

“Con Big Data es posible virtualizar los datos y almacenarlos de manera eficiente usando por ejemplo tecnología cloud y de esta forma optimizar el costo. Muchas tecnologías son el corazón de Big Data y la idea es que todas estas tecnologías trabajen juntas para obtener mejor conocimiento y tiempos basados en datos correctos los cuales son generados por las personas, máquinas o la web.” (Espinoza, 2015, pág. 6).

Cantidades de datos a los que hace referencia a Big Data

El término Big Data hace referencia al tratamiento y análisis de grandes repositorios de datos, originados de transacciones, interacciones y observaciones. Bonillo (2014), sugiere que el tamaño actual de big data esta en el orden de Petabytes de datos.

Tabla 4. Unidades de Información.

Nombre	Cantidad de bytes	Equivalente
Bit	Unidad Básica	
Byte	8 bit	1024
Kilobyte (KB)	1024	1024 bytes
Megabyte (MB)	1048576	1024 KB
Gigabyte (GB)	1073741824	1024 MB

Terabyte (TB)	1099511627776	1024 GB
Petabyte (PB)	1125899906842624	1024 TB
Exabyte (EB)	1152921504606846976	1024 PB
Zettabyte (ZB)	1180591620717411303424	1024 EB
Yottabyte (YB)	1208925819614629174706176	1024 ZB

Fuente: Elaboración propia a partir de Moreno (2015).

Datos Estructurados

Los datos estructurados son datos relacionales, ordenados, consistentes y fáciles de manejar y ordenar en una tabla de base de datos. (Guama, 2015).

La mayoría de organizaciones trabajan en la actualidad en base al análisis de datos generados a partir de sistemas transaccionales o capturados desde la web de acuerdo a ciertos niveles de ordenamiento y normalización.

Datos Semiestructurados

Los datos semiestructurados cuentan con una organización interna que facilita su administración. Como su nombre lo indica contiene características de los datos estructurados y no estructurados (Guama, 2015), algunos ejemplos incluyen: Archivos XML (Extensible Markup Language) y otros lenguajes de marcado, Correos electrónicos, Intercambio electrónico de datos por sus siglas en inglés EDI (Electronic Data Interchange), conjunto de estándares para el intercambios de información ente computadores.

Datos No Estructurados

Los datos no estructurados son lo inverso a los datos estructurados. Son grandes cantidades de datos no relacionales, desordenados, cargados de texto y no son fáciles de representar en tablas tradicionales. Este tipo de datos ha sido recolectado por las organizaciones durante años, comprenden más del 80% de todos los datos en las

organizaciones y crecen 15 veces más rápido que los datos estructurados, pero, en su mayoría no están generando ningún tipo de valor al negocio. (Guama, 2015).

Según Guama (2015), Las principales características que se puede encontrar en este tipo de datos son: Volumen y crecimiento superior al de los datos estructurados, orígenes de datos diversos, arquitecturas de almacenamiento que emplean herramientas de Big Data, terminología e Idioma diferentes para la invocación y generación de datos, complejidad para otorgar niveles de seguridad a los diferentes tipos de datos. Respecto a los datos semiestructurados y no estructurados se debe señalar que son la base para el análisis empleando tecnología de Big Data y otro tipo de soluciones como bases de datos no relacionales e infraestructura de alta capacidad de almacenamiento.

Metadatos

Los metadatos en su definición más simple se refieren a datos sobre los datos, es decir, información estructurada acerca de los datos dentro de cualquier entorno (Guama, 2015). Los metadatos son un término que se viene utilizando desde hace varios años atrás, como ejemplo de sus inicios se tiene a las bibliotecas que lo utilizan aún para la gestión de libros y otros textos, y han facilitado la gestión el control, entendimiento y preservación de los datos al pasar del tiempo. Considerando el campo de aplicación de Big Data los metadatos constituyen una de las herramientas para la construcción de la información sobre los datos recopilados y su posterior procesamiento.

Tipos de Datos en Big Data

Smart Data: “Son todos los datos referentes al negocio (tanto online, como offline). En este grupo se puede encontrar desde las cifras de ventas, datos sobre los clientes, datos sobre nuestra actividad online, etc. Todo aquello que esté relacionado con la consecución de los objetivos de la empresa”. (Bonillo, 2014).

Identity Data: “Son todos los datos que nos permiten identificar a nuestros clientes actuales y potenciales: datos sobre sus gustos, historial de compras, perfil de internauta, tipo de interacción con nuestros contenidos (web, redes sociales, blog, mobile) entre otros. Toda esta información nos permite personalizar nuestra oferta de producto, contenido y canal de comunicación al máximo. En esta etapa es cuando entran en escena las herramientas de CRM y Business Intelligence.” (Bonillo, 2014).

Open data: “Agrupa al resto de datos externos a la empresa y que son accesibles por todo el mundo. Por su volumen y diversidad, nos será más difícil sacarles partido. Pero una vez encontrada la fuente de datos (e integrada), los beneficios que podemos sacar de esta información son enormes.” (Bonillo, 2014).

La tabla 5 a continuación muestra de manera resumida las 4Vs de Big Data.

Tabla 5. Las 4Vs de Big Data.

Las 4Vs		
1	Volumen	Cantidad de datos a ser ingerida, analizada y procesada
2	Velocidad	Rapidez con la cual se generan, procesan y entienden los datos
3	Variedad	Capacidad de procesar amplia variedad de tipos y fuentes
4	Valor	Extraer utilidad de todos los tipos de datos y fuentes

Fuente: Elaboración propia a partir de Espinoza (2015).

Según el análisis realizado por McKinsey Global Institute, “el 15% de la información es “información estructurada” o lo que lo mismo almacenada en base de datos el racionales (columnas y filas). Por otro lado el 85% de la información es “información no estructurada” tal como e-mails, video, blogs, conversaciones con el call-center y medios sociales, las cuales constituyen un desafío para las herramientas de inteligencia de negocios. Dispositivos que producen datos, tales como sensores, tabletas y teléfonos móviles continúan multiplicándose, las redes sociales también están creciendo aceleradamente conforme el mundo aumenta su capacidad de conexión. Estas nuevas maneras de compartir información representan un cambio fundamental en la forma como la gente, el gobierno y las empresas interactúan entre sí. (Espinoza, 2015).

Las características de Big Data dan forma a la manera en que las organizaciones ingieren, analizan, gestionan, almacenan, y distribuyen los datos en toda la empresa y en todo el ecosistema. (Espinoza, 2015, p. 12).

Volumen: El volumen de datos es el atributo principal de Big Data y está relacionado a la enorme cantidad de datos o la intensidad de datos que debe ser ingerida, analizada y procesada para tomar decisiones basadas en un análisis completo de los datos. Desde 1990 el aumento del volumen de datos ha sido un tema desafiante, la diferencia es que en la actualidad ha incrementado la cantidad de dispositivos inteligentes tanto en la red (como sondas) y conectados a la red (tales como smartphones, tablets, medidores y dispositivos M2M).

Velocidad: Se refiere a la rapidez en que los datos son generados o modificados, y de la misma manera son recibidos, entendidos y procesados. Para que las empresas mejoren la toma de decisiones, se requiere acceder y analizar los datos en tiempo real. Esta característica ha incrementado de manera proporcional al aumento de las fuentes de datos, la mejora en la conectividad (ancho de banda), y el aumento en la capacidad de computo en los equipos generadores de información (Espinoza, 2015).

Variedad: Se refiere a la capacidad de procesar los datos generados en una amplia variedad de tipos y fuentes. La variedad es lo que hace que Big Data sea de interés para muchas organizaciones donde el volumen es un reto menor, y está relacionado con la compleja gestión de los múltiples tipos de datos. “Las empresas requieren la integración y análisis de una compleja variedad de fuente de datos sea tradicional o no, y que sean generados en la organización o fuera de ella. (Espinoza, 2015).

Valor: Con el uso inteligente de Big Data, las empresas estarán preparadas para competir, crecer y crear nuevas fuentes de Valor. “Las empresas empiezan a invertir en Big Data porque esta nueva tecnología aumenta su Valor y las hace más competitivas, es así como surge la denominada cuarta V (Valor)”. (Espinoza, 2015).

Otros autores como Microsystem (2015), señalan que Big Data realiza un análisis de un gran Volumen de datos, a gran Velocidad, de una amplia Variedad de tipos de datos relacionados y sobre los cuales hay relativa certeza de su Veracidad.

Veracidad: “En este contexto, el concepto se refiere al sesgo, ruido y alteración de los datos. Son los responsables de la tecnología Big Data quienes tienen la responsabilidad de discriminar entre la información que es relevante para el problema en cuestión y la que no lo es”. (Microsystem, 2015).

2.2.8 Retos actuales

Bonillo (2014), plantea que la variedad, la escalabilidad, el modelo relacional y velocidad, son los principales retos actuales para tratar las grandes cantidades de datos que se generan.

Variedad: Han surgido nuevos tipos de datos que se desean almacenar: datos no estructurados. Las BD Relacionales no pueden almacenar este tipo de datos.

Escalabilidad: En búsqueda de la rapidez y rendimiento en consultas o procesamiento de datos se busca escalar siempre en horizontal. Es decir, al necesitar más rendimiento se añade una CPU al conjunto de trabajo para poder aumentar las prestaciones en conjunto y aumentar el rendimiento reduciendo el tiempo de búsqueda o almacenamiento.

“El problema es que actualmente las BD Relacionales no pueden estar distribuidas en nodos diferentes. Es impensable que una FK apunte a una tabla que está en otro nodo, por ello la única forma de conseguir estos dos objetivos en las BD Relacionales es añadiendo CPU y memoria, haciendo escalabilidad vertical. Pero esto no es suficiente, buscamos escalabilidad horizontal para tener todos los servidores que queramos trabajando en paralelo.” (Bonillo, 2014).

Modelo Relacional: “El modelo relacional no da soporte para todos los problemas. No se pueden atacar todos los problemas con el mismo enfoque, queremos optimizar al 100% nuestro sistema y no podemos ajustar nuestros sistemas a estas BD. Por ejemplo, en el modelo relacional no podemos tener herencia de objetos o no podemos tener columnas variables según las filas...” (Bonillo, 2014).

Velocidad: “La velocidad de generación de datos hoy en día es muy elevada, simplemente hay que verlo con las redes sociales actuales, aunque las empresas medias y muchas de las grandes no se ven afectadas por ello. Donde sí influye la velocidad es en el procesamiento de todo este conjunto ingente de datos, pues cuantos más datos tengamos más tiempo requieren. Por ello, se necesita un ecosistema que sea capaz de escalar en horizontal para trabajar en paralelo y ahorrar tiempo”. (Bonillo, 2014).

Soluciones a los retos con Big Data planteadas por Bonillo (2014)

Bases de datos Big Data: “Se ha observado que han surgido nuevos datos y nuevas formas de almacenar estos datos. Por ello, nacen nuevos gestores como los NoSQL que buscan dar solución a los problemas anteriores. Por ejemplo, las BD clave-valor consiguen almacenar de forma sencilla diferentes tipos de datos no estructurados, sería un símil a tener una fila con 2 columnas y después una fila con 20 columnas. Esto en las relacionales es impensable”. (Bonillo, 2014).

Sistema de archivos distribuido para garantizar escalabilidad: “El corazón del BigData, el concepto del algoritmo MapReduce y Hadoop, el primero es un algoritmo que permite procesar grandes volúmenes de información de forma sencilla y resumida, el segundo es una herramienta que garantiza ejecutar programas MapReduce hechos por usuarios en nodos distribuidos”. (Bonillo, 2014). Esta herramienta tiene un sistema de archivos HDFS el cual provee la distribución de trabajos a diferentes nodos que ejecutarán en paralelo este algoritmo de reducción. Una de las cosas más importantes es que el HDFS consigue hacer transparente o simplificar la creación de clústeres de nodos que trabajan en paralelo como un nodo solo.

Bonillo (2014), plantea estudiar big data desde el punto de vista de inteligencia de negocio desde 3 puntos de distintos: Fuente de datos, área intermedia y DataWarehouse.

Fuente de datos: Que salgan conectores a BD NoSQL desde herramientas ETLs o BI es vital en muchos proyectos BI. Actualmente existen muchas empresas que utilizando

MongoDB para guardar datos no estructurados, por ello, uno de los grandes retos para un consultor BI es poder obtener esa información. Al estar en JSON se ven afectados por la ralentización del "parseo" de todos los datos para dejarlos en una flatTable o en un StarSchema.

Área intermedia: El ecosistema Hadoop consiste en crear una base para poder escalar aplicaciones o aplicar MapReduce a nuestras fuentes de datos. Estas aplicaciones pueden ser BD las cuales funcionen de forma distribuida en el sistema de archivos HDFS.

Por ello, para los consultores es una opción el tener un stage donde se vuelquen todos los datos históricos y después se hagan agregaciones (MapReduce) de forma distribuida en nodos y obteniendo un resultado final el cual, como opción, podemos cruzar con otra fuente de datos estructurada y dejando el resultado final en un DW.

Datawarehouse: La finalidad de todos los proyectos BI es poder analizar los datos. Se puede tener un ecosistema Hadoop, el cual sirva de soporte para una BD distribuida que funcione como almacén de datos analítico al cual se aborde desde una herramienta BI con cubos, informes o dashboards. Con esto, se obtendría todos los datos de forma distribuida y a la hora de realizar consultas se puede tener mayor velocidad. Si se necesita aún más, se pueden añadir los nodos necesarios al clúster.

2.2.9 Arquitectura de Big Data

Camargo (2014), en su trabajo sobre "Arquitectura de Big data" presenta tres (3) de los modelos de arquitectura que existen actualmente en el mundo tecnológico.

Arquitectura de Procesamiento de Big Data propuesta por Krishnan Krish

"La arquitectura presentada por Krishnan Krish, es similar al tratamiento de la gestión de datos, que hace bastante tiempo se conoce. El cual consiste en cuatro (4) etapas: Recolección o recopilación, carga, transformación y extracción de datos, para este caso el autor le llama enfoque de procesamiento de Big Data". (Camargo,2014).

Etapas 1: Recolección o Recopilación o Fuentes de datos de Big Data: En la primera etapa los datos son recibidos de diferentes orígenes o fuentes, que pueden ser: páginas web, redes sociales, máquina a máquina (M2M), transacciones, biometría y generados por el ser humano.

Etapas 2: Carga de datos de Big Data: En esta etapa los datos se cargan aplicando el concepto de metadatos (datos que describen otros datos). Además, de la carga como tal, es la primera vez que los datos se estructuran.

Etapas 3: Transformación de datos de Big Data: En este punto los datos se transforman mediante la aplicación de las reglas del negocio y el procesamiento de los datos. Respecto al procesamiento en cada etapa producen resultados intermedios que se pueden almacenar para un posterior examen. El resultado de esta etapa es, unas cuantas claves de metadatos con modelo clave-valor.

Etapas 4: Extracción de datos de Big Data: El objetivo de la extracción de los datos, es obtener datos para su posterior análisis, para generar informes operativos, para la posible visualización de ellos y por último, y no el más importante, para su almacenamiento.

Arquitectura de Big Data propuesta por Bob Marcus

El modelo de arquitectura por niveles fue propuesto por Bob Marcus y está compuesto de la siguiente manera:

Fuentes de datos externas: Es parte de la arquitectura de datos que suministra las entradas de datos externas y la producción de los componentes internos de Big Data.

Secuencia y procesamiento ETL: La tarea que se desarrolla es la de filtrar y transformar los flujos de datos, provenientes de los recursos externos. El procesamiento de datos que se realiza es el llamado "en movimiento" entre los almacenes de datos.

Fundación altamente escalable: Respecto al componente existen tres (3) tipos de escalonamiento: El primero a nivel de la infraestructura, existe con el fin de poder atender el almacenamiento y procesamiento de grandes volúmenes de datos. El segundo escalonamiento se refiere a los almacenes de datos, tal como lo menciona Bob, “es la esencia de la arquitectura Big Data”, la cual sucede en forma de “escalabilidad horizontal usando componentes menos caros puede apoyar el crecimiento ilimitado de almacenamiento de datos. Sin embargo debe haber capacidades de tolerancia a fallas disponibles para manejar los fallos de sus componentes. Por último el autor propone escalar el procesamiento, buscando “aprovechar las ventajas de los almacenes de datos distribuidos escalables, es necesario contar con el procesamiento distribuido en paralelo escalable con tolerancia a fallos similares”.

Bases de datos operacionales y Analíticas: En el componente (D), se proponen tres (3) clases de bases de datos: Bases de datos Analíticas; el Análisis de bases de datos toma los datos procesados y escalonados de la sección anterior. Bases de datos operacionales; estas bases de datos mantienen una excelente operación en lectura y escritura en general de forma eficiente. En la memoria de datos Grids. Se refieren a los datos ubicados en memorias cachés, buscando minimizar escribir en disco los datos.

Analítica e interfaces de bases de datos: Se compone de tres (3) partes:

- Análisis de Interfaces de procesos en lotes. Se refiere al tipo de interfaz usada para el procesamiento de datos que provienen en lotes o Batch (p.e. Map-Reduce). También hace referencia a la interfaz de usuario para acceder a los datos en almacenes de datos escalables (por ejemplo, del sistema de archivos Hadoop).
- Análisis de interfaces interactivas. “Los almacenes de datos pueden ser bases de datos escalables horizontalmente sintonizados para las respuestas interactivas (p.e. HBase) o lenguajes de consulta en sintonía con los modelos de datos”.

- **Análisis de interfaces en tiempo real:** Se debe analizar con cuidado las interfaces que atienden o son parte de un sistema de tiempo real, pues, los eventos son complejos tanto en el procesamiento, como almacenamiento de los dato.

Aplicaciones e interfaz de usuario: Se refiere a las aplicaciones e interfaces de usuario, las cuales no deben ser algoritmos complejos, al usar grandes cantidades de datos distribuidos.

Arquitectura de Big Data propuesta por Microsoft

El modelo consta de cuatro (4) componentes: Datos fuentes (Data Sources), transformación de datos (Data Transformation), infraestructura de datos (Data Infrastructure) y el uso de datos (Data Usage).

Fuente de datos: Estos datos presentan tres (3) características que definen el Big Data: volumen, velocidad y variedad. Y como aspecto relevante este tipo de datos, son independientes de su contenido o del contexto. “Por lo general, los datos de "big data" se recoge para un propósito específico”. Además, “una vez que se recogen los datos, se pueden volver a utilizar para una variedad de propósitos, algunas potencialmente desconocida en el tiempo de recogida.

Transformación de datos: Se puede notar que esta etapa consta de cuatro (4) sub etapas. Cada una de ellas puede tener su etapa de: procesamiento específico, creación de metadatos, utiliza diferentes infraestructuras de datos especializados de acuerdo a sus tener su propia privacidad y por último, sus propias políticas

Colección: Es un proceso donde se recogen datos en diferentes tipos y formas, se pueden allegar datos de fuentes y estructuras similares o iguales, o combinadas. Además, se crean metadatos, para que sea más fácil hacer una búsqueda de los datos.

Agregación: consiste en adicionar datos a una colección más grande, cuando uno o varios metadatos tengan claves iguales. “Como resultado, la información acerca de cada objeto se enriquece o el número de objetos en la colección crece”.

Congruencia: En esta etapa se recogen datos con metadatos sin importar si son diferentes y se unen a una colección más grande. Obteniendo al final que cada objeto sea enriquecido.

Minería de datos: La minería o Data mining, es una extracción de datos, para luego poder hallar relaciones entre los datos. Existen “dos tipos de minería de datos: descriptivo, que proporcione información sobre los datos existentes, y predictivo, lo que hace que los pronósticos basados en los datos”.

Infraestructura de Big Data: Se considera la infraestructura como “un paquete de almacenamiento de datos o software de base de datos, servidores, almacenamiento y redes utilizados en apoyo de las funciones de transformación de datos y de almacenamiento de datos según sea necesario”.

Uso de los datos: Respecto al uso de los datos, depende del usuario y sus necesidades particulares, pero estos se pueden presentar en diferentes formatos y bajo ciertas consideraciones de seguridad.

Tecnologías de Big data

Camargo (2014), en su trabajo “Conociendo Big Data” presenta un conjunto de técnicas y tecnologías para el manejo de datos no estructurados o semiestructurados, como MapReduce o Hadoop, diseñado para el manejo de información estructurada, no estructurada o semiestructurada. A continuación un resumen de lo que se plantea.

2.2.10 Apache Hadoop

“Es una biblioteca de software que permite el procesamiento distribuido de grandes conjuntos de datos a través de grupos de ordenadores que utilizan modelos sencillos de programación. Está diseñado para pasar de los

servidores individuales a miles de máquinas, cada oferta local de computación y almacenamiento”. (Apache hadoop, 2015, pág. 1).

Camargo (2014), define Hadoop como un framework de código abierto, el cual permite escribir y ejecutar aplicaciones distribuidas que procesan grandes cantidades de datos. Tiene algunas características importantes:

- Fue diseñado para ejecutarse en grupos relativamente grandes de hardware, es decir, en clúster robustos.
- Es robusto, pues ante un mal funcionamiento del hardware puede superar tales situaciones sin mayor inconveniente.
- Tiene la ventaja de poder ser escalable, lo que indica que permite crecer o agregar nodos al clúster con relativa facilidad; por ejemplo, ante la forma vertiginosa como crecen las redes sociales, permite agregar más nodos con facilidad.
- Es simple, por lo que permite a los usuarios escribir código con eficiencia, para software distribuido.

De acuerdo con la información publicada por Apache hadoop (2015), Hadoop está compuesto por los siguientes módulos:

Hadoop Common: Las utilidades comunes que apoyan los otros módulos de Hadoop.

Hadoop Distributed File System (HDFS): Un sistema de archivos distribuido que permite el acceso de alto rendimiento a los datos de la aplicación. Un sistema de archivos altamente tolerante a fallos, escalable y con una arquitectura distribuida; puede llegar a almacenar 100 TB en un solo archivo, lo cual no es tan fácil en otros tipos de sistemas de archivos. Además, brinda la apariencia de estar trabajando en un solo archivo, pero realmente lo que se tiene es que están distribuidos en varias máquinas para su procesamiento.

Camargo (2014), menciona que HDFS se diseñó para el procesamiento por lotes, en lugar de uso interactivo por los usuarios; pero realmente se diseñó para resolver dos problemas importantes que se presentan en el procesamiento de datos a gran escala:

el primero es la capacidad de descomponer los archivos en varias partes y procesar cada una independientemente de las demás, y al final consolidar todas las divisiones del archivo en uno solo; el segundo problema era la tolerancia a fallos, tanto en el nivel de procesamiento de archivos como de forma general del software, al momento de realizar el procesamiento de datos distribuidos; lo que se busca es que el sistema pueda recuperarse de la falla que se pueda presentar sin afectar demasiado el proceso.

Según Camargo (2014), cuando se creó HDFS se propusieron tres objetivos:

- Permitir procesar archivos con tamaños de gigabytes (GB) hasta petabytes (PB).
- Poder leer datos a grandes velocidades.
- Capacidad para ser ejecutado en una máquina, sin solicitar hardware especial.

“La arquitectura de HDFS está compuesta por un nodo principal (NameNode) y varios nodos esclavos (DataNodes). El nodo principal es el servidor maestro, dedicado a gestionar el espacio del nombre de los archivos y controlar el acceso de los diferentes archivos de usuarios; además, el nodo maestro se encarga de gestionar las operaciones de abrir, cerrar, mover, nombrar y renombrar archivos y directorios.”. (Camargo, 2014).

“Los nodos esclavos (DataNodes), como su nombre lo indica, representan los esclavos de la arquitectura HDFS. En un HDFS pueden existir miles de nodos esclavos y decenas de miles de clientes HDFS por clúster; esto se debe a que cada nodo esclavo puede ejecutar múltiples tareas de aplicaciones de forma simultánea. La función del nodo esclavo es la de gestionar tanto la lectura como la escritura de los archivos de los usuarios, y realizar la replicación de acuerdo a como lo indique el nodo maestro.” (Camargo, 2014).

Hadoop YARN: Un marco para la planificación de tareas y gestión de recursos de clúster.

Hadoop MapReduce: Un sistema basado en el hilo para el procesamiento paralelo de grandes conjuntos de datos. Hadoop MapReduce es un marco de software creado con el fin de hacer aplicaciones que puedan procesar grandes cantidades de datos de forma paralela, en un mismo hardware. Cuando los datos entran para ser procesados

se dividen de manera independiente, para su procesamiento, es decir, de manera distribuida en diferente hardware que exista. MapReduce está compuesto de un maestro, llamado JobTracker, y un esclavo, TaskTracker, por cada nodo. El primero se encarga de programar las tareas, los componentes que manejan el esclavo, y éste ejecuta las tareas según las indicaciones del maestro. MapReduce es usado en soluciones donde se pueda procesar de forma paralela y, además, con grandes cantidades de información, es decir, con volúmenes de petabytes, de lo contrario no sería una solución adecuada. Otra condición es que se puede usar MapReduce en procesos que se puedan disgregar en operaciones map y reduce, las cuales se definen en función de datos estructurados. MapReduce se diseñó como un modelo de programación, para que se pudiera realizar procesamiento de datos de gran tamaño, y, de la misma forma, para que resolviera el problema existente de escalabilidad. MapReduce “es un modelo de programación para el procesamiento de datos”. Puede ser ejecutado en varios lenguajes de programación, como Java, Ruby, Python, and C++. (Camargo, 2014).

2.2.11 Bases de datos NoSQL

Camargo (2014), plantea que en 1998 aparece el término NoSQL, que significa no solo SQL. El nombre fue creado por Carlo Strozzi, para denominar su base de datos que no ofrecía SQL. Las NoSQL no presentan el modelo de las bases de datos relacionales; estas no tienen esquemas, no usan SQL, tampoco permiten joins (unión), no almacenan datos en tablas de filas y columnas de manera uniforme, presentan escalabilidad de forma horizontal, para su labor usan la memoria principal del computador; su objetivo es gestionar grandes volúmenes de información. Las bases de datos NoSQL tienen como característica principal que su estructura es distribuida, es decir, los datos se hallan distribuidos en varias máquinas. Las bases de datos NoSQL permiten obtener los datos con mayor velocidad que en otras con modelo relacional.

De acuerdo a la manera en que las bases de datos *NoSQL* almacenan sus datos es posible clasificarlas de la siguiente manera:

- Almacenamiento clave-valor
- Bases de datos orientadas a columnas
- Base de datos documentales
- Base de datos orientada a grafos

Existen varias clases de bases de datos NoSQL, dependiendo de su forma de almacenar los datos, tales como: almacenamiento Clave-Valor, orientadas a columnas y las orientadas a documentos. A continuación se presentan algunas Bases de Datos NoSQL:

DynamoDB: Fue desarrollada y probada de manera interna en Amazon; guarda muy fácil y económicamente cualquier cantidad de información. Los datos son almacenados en unidades de estado sólido SSD (Solid State Drive), las cuales permiten mayor velocidad a la hora de encontrar la información, pues estas unidades funcionan de manera diferente a como lo hace el disco duro de la computadora. Con el uso de SSD se tiene un excelente rendimiento, mayor fiabilidad y un alto grado de seguridad de los datos. Camargo (2014).

Cassandra: Proyecto iniciado por Facebook; es del tipo código abierto (Open Source). Se puede decir que después de la implementación de Cassandra, las redes sociales se dispararon en popularidad. Es una base de datos distribuida, y almacena los datos en forma de clave-valor; fue desarrollada en java, además, hoy en día es usada en la red social Twitter. Otras características importantes de Cassandra es que es descentralizada, lo que significa que cada nodo es idéntico, y, además, que no existe ningún punto único de fallo; que es escalable, es decir, que el software puede atender un número mayor de solicitudes de los usuarios sin que se note algún tipo de degradación en su rendimiento, y que es tolerante a fallos, es decir, que puede reemplazar nodos que fallen en el clúster sin perder tiempo (Camargo, 2014).

Voldemort: Voldemort fue creada por LinkedIn, con el fin de solucionar los problemas de escalabilidad que tenían las bases de datos relacionales; los datos los almacena

en forma de clave-valor; es de ambiente distribuido, los datos se replican automáticamente en los diferentes nodos o servidores, donde cada nodo es independiente de los demás; permite con cierta facilidad la expansión del clúster, sin necesidad de reequilibrar todos los datos. El código fuente está disponible bajo la licencia Apache 2.0. (The Apache Software Foundation, 2015).

Google BigTable: Fue creado por Google en el año 2004, con la idea inicial de que fuera distribuido para varias máquinas, por lo que necesitaban que fuese altamente eficiente. El sistema divide la información en columnas, y para almacenarla utiliza tablas multidimensionales compuestas por celdas. (Camargo, 2014).

HBase: HBase es una base de datos de tipo código abierto (Open Source); almacena los datos de forma clave-valor; también almacena y recupera los datos de forma aleatoria, es decir, que al momento de escribir los datos lo hace a su manera, y al leerlos funciona de igual forma. Trabaja con los tres tipos de datos: no estructurados, semiestructurados y estructurados, siempre y cuando no sean tan grandes. HBase no permite consultas SQL y, además, está diseñada para ejecutarse en un clúster de equipos, lo que indica que no puede trabajar en un solo servidor. En la medida que se aumenten más servidores, HBase no presenta inconvenientes en ese sentido, y, también, cuando uno de ellos presenta algún tipo de inconveniente se puede sustituir por otro sin mayor problema. (Camargo, 2014).

Riak: Es una base de datos que almacena la información en forma de clave-valor y es de ambiente distribuido, presenta la característica de que es tolerante a fallos, lo que indica que puede eliminar errores y sus efectos antes de que ocurra una falla, buscando de esta manera maximizar la fiabilidad del sistema. Utiliza JSON (JavaScript Object Notation - Notación de Objetos de JavaScript), que es un formato para el intercambio de datos. Además, Riak tiene mayor ventaja a la hora de trabajar en la Web, en la familia de bases de datos de su especie, medida en las peticiones de muchos usuarios simultáneamente. (Camargo, 2014).

CouchDB: Es el acrónimo en inglés de Cluster of Unreliable Commodity Hardware; fue creado en el año 2005, por Damien Katz. En el 2011 se hace el lanzamiento al público de la versión 1.1.1. Se considera que CouchDB es un servidor de base de datos documental, lo cual indica que los datos no los almacena en tablas, sino que la base de datos está compuesta por documentos, que a su vez trabajan como objetos. Hace uso de JSON, que es un formato para el intercambio de datos, usado cuando los datos son de gran volumen; por eso, para las consultas hace uso de JavaScript; debido a lo anterior, es muy usado por empresas como Yahoo y Google. (Camargo, 2014).

MongoDB: MongoDB es una base de datos con el perfil NoSQL orientada a documentos, bajo la filosofía de código abierto. La importancia de MongoDB radica en su versatilidad, su potencia y su facilidad de uso, al igual que en su capacidad para manejar tanto grandes como pequeños volúmenes de datos. Es una base de datos que no tiene concepto de tablas, esquemas, SQL, columnas o filas. No cumple con las características ACID, que es el acrónimo de Atomicity, Consistency, Isolation and Durability (Atomicidad, Consistencia, Aislamiento y Durabilidad, en español). MongoDB permite las operaciones CRUD, que es el acrónimo de Create, Read, Update and Delete (Crear, Obtener, Actualizar y Borrar); para almacenar y recuperar los datos hace uso de JSON, pero utiliza BSON, que es una forma binaria de JSON, el cual ocupa menos espacio al almacenar los datos. Además, BSON es más eficiente y rápida para convertir a un formato de datos de un lenguaje de programación. Otra característica de MongoDB es que realiza consultas dinámicas, es decir, puede realizar consultas sin demasiada planificación. MongoDB se desarrolló en C++. (Camargo, 2014).

Neo4J: Según Bonillo (2014), Es una base de datos orientada a grafos de código abierto desarrollada en Java, utiliza como modelo de datos los grafos, basados en propiedades. Es una de las pocas bases de datos orientadas a grafos que cuenta con un lenguaje de consulta propio denominado Cypher. Fue desarrollado por Neo Technology, una startup sueca con base en Malmö y San Francisco Bay Area en Estados Unidos.

BaseX: Es una base de datos de tipo documental, la cual permite almacenar, recuperar y gestionar datos de documentos; es de la clase de bases de datos NoSQL; tiene como característica importante que permite escalar y, además, que es de alto rendimiento. Su arquitectura es cliente/servidor, permitiendo realizar lecturas y escrituras de datos de manera simultánea. Cumple con el estándar, ACID (acrónimo de Atomicity, Consistency, Isolation and Durability-Atomicidad, Consistencia, Aislamiento y Durabilidad). Soporta grandes documentos en XML, JSON y formatos binarios. BaseX está desarrollado bajo Java y XQuery. (Camargo, 2014).

2.2.12 Propiedades ACID vs BASE

Teorema de CAP

Durante el simposio de “Principios de computación distribuida” de ACM en el año 2000, Eric Brewer, un profesor de la universidad Berkeley de California y cofundador de Inktomi, a través de una presentación titulada “Hacia sistemas distribuidos robustos”, estableció la conjetura que los servicios web no pueden asegurar en forma conjunta las siguientes propiedades: Consistencia (C), Disponibilidad (A) y Tolerancia a particiones (P), esto es lo que se conoce como el teorema de CAP. Posteriormente en el año 2002, Seth Gilbert y Nancy Lynch de MIT publicaron una demostración formal de la conjetura de Brewer, convirtiéndola en un teorema. El teorema de CAP establece que, en un sistema distribuido con datos compartidos, se debe optar por favorecer dos de las tres características: Consistencia, Disponibilidad y Tolerancia a particiones. Bajo estas restricciones, Brewer indica que se debe utilizar como criterio de selección, los requerimientos que se consideren más críticos para el negocio, optando entre propiedades *ACID* y *BASE*. (Brewer, 2000).

Propiedades ACID en sistemas distribuidos

Las propiedades ACID presentes en las transacciones de las bases de datos relacionales simplifican el trabajo de los desarrolladores de aplicaciones al ofrecer garantías en cuanto a:

- Atomicidad: Todas las operaciones en la transacción se completarán o ninguna lo hará.
- Consistencia: El estado de la base de datos se mantiene válido tanto al inicio como al final de la transacción. Cualquier operación puede ver los cambios en la base de datos.
- Aislamiento: Las transacciones se ejecutan de manera que una no puede afectar a la otra.
- Durabilidad: Los cambios realizados durante una operación serán persistentes y no se podrá deshacer, aunque falle el sistema.

Los proveedores de bases de datos se percataron de la necesidad de particionar, por lo que introdujeron el protocolo de *commit* a 2 fases para seguir manteniendo las propiedades ACID en las instancias de las bases de datos. El cual consiste en:

- Primera fase: el coordinador de la transacción solicita a cada base de datos involucradas que indiquen si es posible que realicen *commit*. Si es posible se procede a continuar con la segunda fase.
- Segunda fase: El coordinador de la transacción solicita a cada base de datos que realice el *commit*.

Sin embargo, si alguna base de datos no puede realizar el commit, se le solicita a todas las involucradas que realicen un *roll-back* de esa transacción. Sería análogo a que un avión no pudiese despegar hasta que todos los pasajeros estén en sus asientos, a pesar de que el vuelo haya sido programado para una hora en específico. En vista de esto podemos observar que se generan retrasos al realizar las transacciones, lo que afecta, de acuerdo al teorema de CAP, la disponibilidad del sistema. (Pritchett, 2008).

Propiedades *BASE*

Si las propiedades *ACID* se enfocan en la consistencia, las propiedades *BASE* se enfocan en la disponibilidad. La palabra *BASE* se refiere a básicamente disponible (BA), estado flexible (S) y eventualmente consistente (E).

Las propiedades *BASE* son diametralmente opuestas a las propiedades *ACID*. Donde *ACID* es pesimista y fuerza la consistencia al finalizar cada operación, *BASE* es optimista y acepta que la consistencia de la base de datos estará en un estado flexible. Aunque esto suene imposible de lidiar, en la realidad es manejable y puede llevar a niveles de escalabilidad que no se pueden obtener con *ACID*.

La disponibilidad en las propiedades *BASE* es alcanzada a través de mecanismos de soporte de fallas parciales, que permite mantenerse operativos y evitar una falla total del sistema.

Motores de Búsqueda

Los motores de búsqueda son programas que permiten hacer búsquedas por palabras claves y retornan una lista de documentos en donde se encontraron las palabras claves. Funcionan mediante la creación de índices con los cuales luego se realizan las búsquedas y a diferencia de las bases de datos, permiten integrar búsquedas de diversas fuentes de datos, permiten la escalabilidad, ranking por relevancia y hacer búsquedas por facetas.

Estas herramientas son necesarias para facilitar las búsquedas a los usuarios, ya que al hablar de grandes volúmenes de datos como es el caso de la web es muy probable que el usuario final encuentre información que no le sea de utilidad, por lo cual estas herramientas se manejan con análisis de contenido como es el caso del algoritmo de *Google*, *Page Rank*, en el cual se hace un ranking de las páginas más importantes dependiendo de la búsqueda realizada. Los principales motores de búsqueda en *BigData* son *Apache Lucene* y *Apache SolR*.

Herramientas de visualización y análisis

Una vez que se tienen todos estos datos en la arquitectura, se utilizan herramientas de visualización y análisis que permiten sacarle el valor a los datos, con los cuales es posible, mediante la aplicación de funciones estadísticas y de visualización, encontrar

patrones, tendencias, hacer modelos predictivos, e incluso sistemas de recomendación.

La importancia de la visualización de los datos radica en que permite al analista de datos, bien sea un gerente o un científico de datos, facilitarles la comprensión de los datos, de modo que puedan realizar sus funciones con mayor facilidad y puedan a su vez transmitir la información que encontraron en los datos.

Algunas herramientas de visualización más importantes podemos mencionar Apache Hue (Hadoop User Experience) y Banano.

Entre las herramientas de análisis más conocidas en el mundo del Big Data se encuentran:

- *RStudio* que es más conocida en el ámbito científico por su manejo de lenguaje de análisis R, el cual es muy utilizado por científicos de datos.
- Podemos mencionar *Apache Mahout*, el cual me permite utilizar algoritmos en su librería para hacer máquinas de aprendizaje
- *Apache Spark* el cual se integra con *Hadoop* y posee un módulo llamado *Spark MLIB* en el cual también se utiliza para hacer *machine learning*, además de poseer módulos para hacer consultas en SQL y análisis de grafos, con la peculiaridad de que *Spark* utiliza la memoria RAM de la máquina, lo que lo hace una herramienta más rápida que utilizar solo *Hadoop*.

2.2.13 Factores que intervienen en la creación de un modelo

Modelo

El término modelo deriva del latín *modellus* que significa medida, molde y es asociado a la representación proporcional de alguna cosa, según señala Hurtado (2010), de esta forma modelo hace referencia a lo deseado, a un ideal a lo perfecto o lo que debe ser. Así un modelo es la descripción sistemática, consciente y simplificada de una parte de

la realidad, realizada mediante signos, símbolos, formas geométricas o gráficas y palabras.

Ampliando en la definición de modelo Palella y Martins (2012), lo definen como una representación idealizada de una clase de objetos reales. Por otra parte Bonillo (2003), lo define como un esquema teórico de un sistema o realidad compleja que se elabora para facilitar su comprensión y estudio. Es una representación de la realidad, que ayuda a explicar una observación. Es una representación de un sistema o una teoría que permite predecir un comportamiento.

La siguiente tabla muestra las diferentes definiciones referentes término modelo presentadas por Bonillo (2003):

Tabla 6. Principales Definiciones del Término Modelo.

AUTOR	DEFINICIÓN
Booch, 1994	Es una relación entre una clase de objeto y un dato almacenado.
Fenton y Pfleeger, 1997	En general un modelo es una abstracción de la realidad, que permite eliminar los detalles y ver una entidad o concepto desde una perspectiva particular.
Hopkins, 1997	Un modelo es una relación entre variables, donde diferentes tipos variables, producen diferentes tipos de modelos.
Patrick, 2000	Un modelo es una representación visual controlada de una teoría o un sistema
Rumbaugh, 1998	Un modelo es una representación formal de un sistema en algún nivel de
Segovia, 1996	Los modelos son utilizados en los métodos de desarrollo de sistemas de información para describir, analizar, y para comunicar los relevantes de ellos en diferentes niveles de abstracción.

Fuente: Bonillo (2003).

Bonillo (2003), a partir de todas las definiciones presentadas concluye que un modelo es una representación de la realidad, es decir una abstracción de la misma que permite establecer relaciones entre diferentes variables para obtener diferentes tipos de configuraciones o formas de ver, a través de una representación visual o formal y que comunica los aspectos relevantes que están contenidos en la abstracción.

Así un modelo viene a ser una representación de la realidad con la que se pretende reducir la variedad y complejidad del mundo real mediante el lenguaje simbólico.

De esta forma los modelos permiten comprender fácilmente la realidad que rodea a un problema dado, como condición indispensable requiere de una abstracción de dicha realidad.

Palella y Martins (2012), plantean cinco aplicaciones o alcances diferentes de los modelos:

- Modelo como un tipo ideal (Weber)
- Modelo como paradigma (Merton)
- Modelo como mecanismo oculto fácilmente imaginado que explica los fenómenos observados (Pareto, Parsoms)
- Modelo como lo deseable
- Modelo como teoría matemática (p84)

También señala el referido autor que existen varias clasificaciones en cuanto a los tipos de modelos:

- Los modelos de tipo estático o de estructura
- Modelos de tipo dinámico o de funcionamiento
- Modelo de previsión o de decisión

¿Cómo elaborar un modelo teórico innovador?

Para la elaboración de un modelo teórico innovador se debe tomar en cuenta que en cada caso concreto es menester determinar qué tipo de modelo innovador se quiere y se puede desarrollar y además las teorías, hipótesis, supuestos y/o premisas en las cuales estaría sustentado. (Palella y Martins, 2012, pág. 26)

Según Palella y Martins (2012), se distinguen en cuatro fases en el proceso de creación de un modelo, estas son:

1. Definición de la situación problema o ámbito situacional
2. Formulación del modelo, que incluye:
 - Selección de variables a ser incluidas

- Selección y nivel apropiado de agregación clasificación y definición de los elementos epistemológicos (conocimiento), ontológicos (ser) y axiológicos (valores)
- Decisiones sobre el tratamiento del tiempo
- Especificaciones de los fines u objetivos innovadores para los que se va a utilizar el modelo.
- Valoración.

3. Simulación y/o validación del modelo

4. Evaluación del modelo

Se debe tener en cuenta que la cualidad innovadora del modelo, el cual es requisito indispensable para una genuina producción teórica, lo cual está asociado a la idea de producir una innovación que genera mejoras y acarrea cambios.

Finalmente y sintetizando todas la ideas anteriores se puede definir un modelo como una representación innovadora, racional, sistematizada, lógica, comprobable y original que permite construir los postulados y principios epistemológicos (conocimiento), axiológicos (valores) y ontológico (ser) de una innovación producto de la creatividad, en consecuencia crea una nueva manera de hacer, pensar o aprehender algo. Palella y Martins (2012).

Ventajas relacionadas a los modelos

Entre las ventajas señaladas por Hurtado (2010), relacionada con los modelos se tiene:

- Los modelos muestran los aspectos más relevantes de aquello que representan y facilitan su comprensión.
- Permiten establecer nexos entre diversas disciplinas.
- Permiten conectar los aspectos teóricos de una situación con sus consecuencias en la realidad.
- Hacen posible la confluencia de varias teorías, logrando establecer una visión compleja del aspecto estudiado.

- Tienen valor heurístico y permiten el surgimiento de nuevas investigaciones y aplicaciones prácticas del conocimiento.

Clasificación de los modelos.

Existen diversos tipos de modelos, según Hurtado (2010), se pueden clasificar en:

Modelos según su objetivo.

- **Modelo representativo:** Es una representación simplificada de una situación u objeto en base a sus características fundamentales, facilitando así la visualización del objeto en su conjunto.
- **Modelo explicativo o teórico:** Permite ilustrar la explicación de un proceso o fenómeno, integrando los diferentes eventos que forman parte del mismo y la manera como se relacionan e interactúan.
- **Modelo normativo:** Proporciona un conjunto de criterios y pautas para la acción, como lineamientos que permiten llevar a cabo cada una de las etapas de un proceso, a modo de plan de acción.
- **Modelo operativo:** Da cuenta del funcionamiento de ciertos objetos, artefactos o de algún fenómeno en general, visualizando sus procesos de funcionamiento internos.

Modelos según su nivel de abstracción.

- **Modelo formal:** Está compuesto por un conjunto de conceptos vinculados mediante ciertas relaciones de tipo lógica o analógica, se pueden derivar de las relaciones entre un conjunto de datos de forma inductiva o mediante procesos deductivos. El modelo formal puede estar basado en metáforas o analogías.
- **Modelo material:** Relaciona características concretas en lugar de conceptos o analogías, se basa en referentes empíricos, representan a un objeto o constructo, por lo general son representaciones a escala del objeto físico real.

Modelo según el tipo de representación utilizada.

- Verbales: Es aquel en el cual se utilizan las descripciones y explicaciones verbales como lenguaje para expresar el fenómeno que se desea representar.
- Gráfico o icónicos: Este tipo de modelos recurre a los gráficos tales como dibujos, imágenes y diagramas, como medio para expresar sus representaciones.
- Matemáticos: Son aquellos que utilizan símbolos numéricos y sus relaciones matemáticas como medio de representación, por lo general toman forma de ecuaciones.
- Modelo físico: Son los que utilizan materiales físicos para crear la representación.

Según las diversas clasificaciones anteriormente dadas, el modelo que propone la presente investigación según su objetivo es de tipo operativo, ya que propone la implementación de un sistema de información. Según su nivel de abstracción es un modelo formal, debido a que está compuesto por conceptos vinculados entre sí relacionados de manera lógica. Según su tipo de representación es un modelo tanto verbal como gráfico, ya que se utilizará descripciones y explicaciones tanto verbales como dibujos, imágenes y diagramas como medio de expresar su representación.

2.2 Marco Referencial

Según FinanzasDigital (2015), de acuerdo con el Informe del Sistema Financiero Venezolano al cierre del primer semestre de 2015, el sistema financiero Venezolano está compuesto por 23 bancos universales, 3 bancos comerciales y 4 bancos de desarrollo. A continuación el Organigrama del Sistema Bancario Nacional

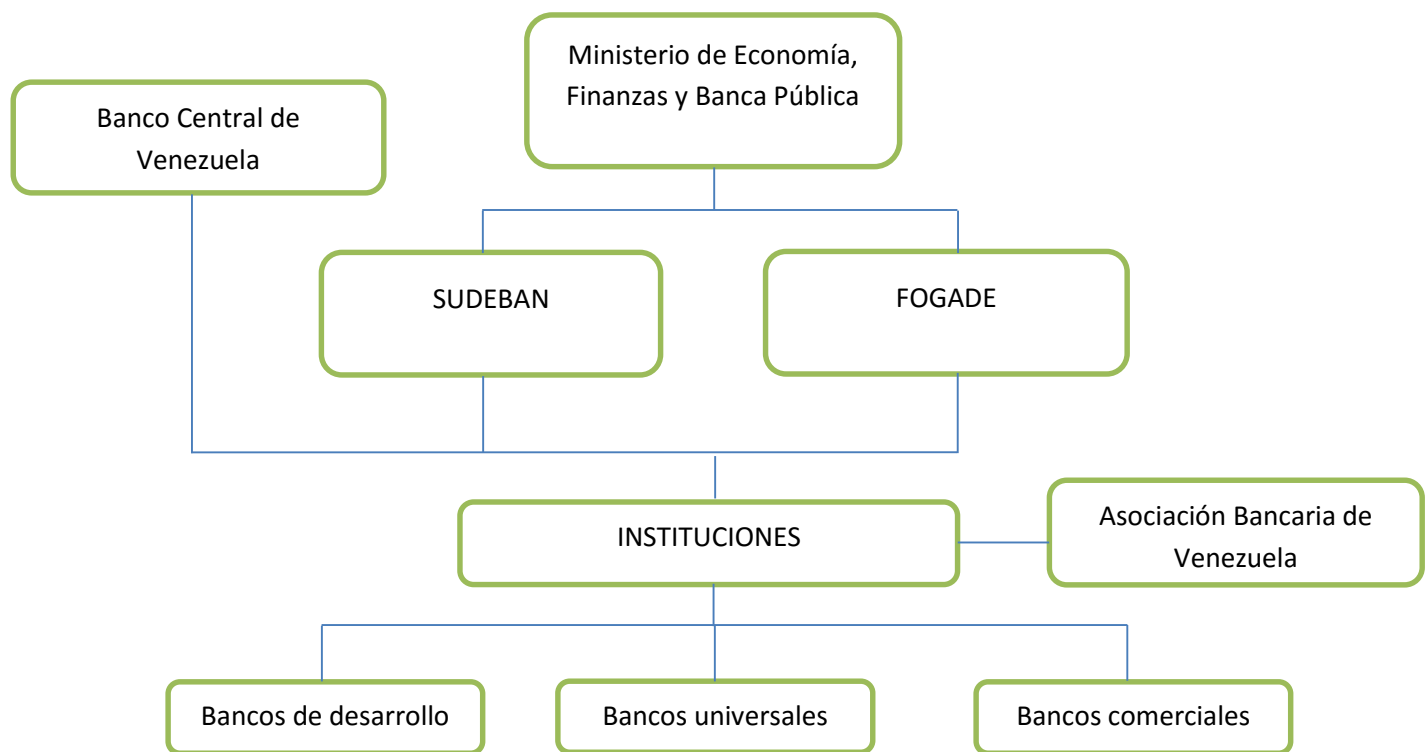


Figura 11. Organigrama del Sistema Bancario Nacional.

Fuente: Elaboración propia.

Sistema financiero

El sistema financiero es el conjunto de personas y organizaciones, tanto públicas como privadas, es decir, un conjunto de instituciones que previamente autorizadas por el Estado, captan, administran, regulan y dirigen los recursos financieros que se negocian entre los diversos agentes económicos, dentro del marco de la legislación correspondiente. El sistema financiero de un país está constituido por los intermediarios financieros, personas e instituciones que actúan en el mercado financiero, y comprende todas las operaciones financieras y comerciales entre los sujetos y los sectores económicos. (Dueñas, 2008)

Funciones Básicas del Sistema Financiero

- Fomento del ahorro

- Obtener estabilidad monetaria: el sistema financiero contribuye a salvaguardar el valor del dinero en los tres aspectos básicos de la moneda (unidad de cuenta, reserva de valor e instrumento de pago)
- Contribuir a las solvencias de las instituciones: tanto a nivel sistemático como en forma individual de los individuos que la componen
- Crear una variedad de activos financieros
- Conseguir una eficaz asignación de los recursos
- Conseguir bajo costo de intermediación



Figura 12. Elementos que conforman el Sistema Financiero.

Fuente: Dueñas (2008).

Activos Financieros

Los Activos Financieros son títulos librados por las Unidades de Gastos Deficitarias. Estos títulos son negociables y tienen tres características importantes: la liquidez, de fácil conversión en efectivo; el bajo riesgo, es decir la posibilidad mínima de pérdida; y la rentabilidad, que es la ganancia sobre la inversión. (Dueñas, 2008)

Mercados Financieros

Un mercado financiero, es un sitio físico o virtual al cual concurren compradores y vendedores, donde se negocian los activos financieros. (Dueñas, 2008).

Intermediación financiera

Se entiende por intermediación financiera a la actividad que realizan las instituciones bancarias y que consiste en la captación de fondos bajo cualquier modalidad y su colocación en créditos o en inversiones en títulos valores emitidos o avalados por la Nación o empresas del Estado mediante la realización de las operaciones permitidas por las leyes de la República. (Decreto1402, 2014)

Sector bancario público y privado

El sector bancario privado comprende el conjunto de las instituciones privadas, que previa autorización del ente regulador se dedican a realizar actividades de intermediación financiera. (Decreto1402, 2014).

Banco universal

Son las instituciones que realizan todas las operaciones de intermediación financiera y sus servicios conexos. Para operar requieren de un capital social mínimo suscrito y pagado de ciento setenta millones de bolívares, si tienen su asiento principal en el Área Metropolitana de Caracas, así como en las ciudades de Guarenas, Guatire, San Antonio de los Altos, Carrizal, Los Jeques, Los Valles del Tuy y en el Estado Vargas; y, de ochenta y cinco millones, si están situados en cualquier otra jurisdicción del país y han obtenido de la Superintendencia de las Instituciones del Sector Bancario la calificación de banco regional. (Decreto1402, 2014)

Instituciones bancarias especializadas

Son los bancos de desarrollo y a los bancos microfinancieros. Los primeros tienen por objeto fomentar, financiar y promover los proyectos de desarrollo industrial y social del país, así como actividades económicas y sociales para sectores específicos productivos del país. En cuanto a los bancos microfinancieros deben fomentar, financiar o promover las actividades de producción de bienes y servicios de las pequeñas y medianas empresas, de la economía popular y alternativa, de los microempresarios y las microempresas. (Decreto1402, 2014)

Activos del Sistema Financiero Venezolano

Según FinanzasDigital (2015), en junio de 2015, los activos del SFN llegaron a la suma total de Bs. 3.354.962 millones lo que se traduce en un incremento de Bs. 246.310 millones (7,92%) respecto al mes anterior, y un 79,3% adicional a lo alcanzado hasta el mismo mes del año 2014. La estructura del Activo Total se ha mantenido más o menos uniforme durante junio 2015 vs. Junio 2014; destaca la Cartera de Créditos Neta con una participación con el 48,43% y de 19,64% en Inversiones en Títulos Valores sobre el total de activos. Las Inversiones en Títulos Valores (sin las Colocaciones en el BCV) representan el 16,35%.

2.4. Bases Legales

Marco Legal del Sistema Financiero Nacional

Constitución de la República Bolivariana de Venezuela (Presidencia de la República Bolivariana de Venezuela, 1999)

El marco legal del Sistema Financiero Nacional se encuentra principalmente en la Constitución de la República Bolivariana de Venezuela en el Artículo 318, el cual establece: Las competencias monetarias del Poder Nacional serán ejercidas de manera exclusiva y obligatoria por el Banco Central de Venezuela. El objeto fundamental del Banco Central de Venezuela es lograr la estabilidad de precios y preservar el valor interno y externo de la unidad monetaria. La unidad monetaria de la República Bolivariana de Venezuela es el bolívar. En caso de que se instituya una moneda común en el marco de la integración latinoamericana y caribeña, podrá adoptarse la moneda que sea objeto de un tratado que suscriba la República. El Banco Central de Venezuela es persona jurídica de derecho público con autonomía para la formulación y el ejercicio de las políticas de su competencia. El Banco central de Venezuela ejercerá sus funciones en coordinación con la política económica general, para alcanzar los objetivos superiores del Estado y la Nación. Para el adecuado cumplimiento de su objetivo, el Banco Central de Venezuela tendrá entre sus funciones las de formular y ejecutar la política monetaria, participar en el

diseño y ejecutar la política cambiaria, regularla moneda, el crédito y las tasas de interés, administrar las reservas internacionales, y todas aquellas que establezca la ley.

Decreto con Rango, Valor y Fuerza de Ley de Instituciones del Sector Bancario

El decreto N° 1.402, Gaceta Oficial 6.154 Extraordinario de fecha 19/11/2014, dicta el Decreto con Rango, Valor y Fuerza de Ley de Instituciones del Sector Bancario.

“El objeto principal de este Decreto con Rango, Valor y Fuerza de Ley, consiste en garantizar el funcionamiento de un sector bancario sólido, transparente, confiable y sustentable, que contribuya al desarrollo económico-social nacional, que proteja el derecho a la población venezolana a disfrutar de los servicios bancarios, y que establezca los canales de participación ciudadana; en el marco de la cooperación de las instituciones bancarias y en observancia a los procesos de transformación socio económicos que promueve la República Bolivariana de Venezuela”. (Decreto1402, 2014)

Los institutos autónomos regidos por este decreto

La Superintendencia de las Instituciones del Sector Bancario y el Fondo de Protección Social de los Depósitos Bancarios.

Ministerio del poder popular de Economía y Finanzas

Es el Órgano Superior del Sistema Financiero Nacional, su objetivo es articular armónicamente las actividades del Ejecutivo Nacional en la regulación, formulación y seguimiento de políticas en materia de planificación estratégica, territorial, social, institucional, tanto en el mediano y en el largo plazo, así como la preparación de sus proyecciones y alternativas. De igual manera corresponde a este despacho la planificación y coordinación del desarrollo de los sistemas integrados de la administración de hacienda y finanzas del sector público, en el marco del modelo de desarrollo socialista. Todo ello en beneficio de la calidad de vida de nuestro pueblo soberano. (Ministerio del poder popular de Economía y Finanzas, 2015)

Superintendencia de las Instituciones del Sector Bancario

Según el Decreto 1402 (2014), la Superintendencia de las Instituciones del Sector Bancario es el ente de regulación del sector bancario. Se entiende por normativa prudencial emanada de la Superintendencia de las Instituciones del Sector Bancario, todas aquellas directrices e instrucciones de carácter técnico contable, legal, tecnológico y de riesgo de obligatoria observancia, dictada mediante resoluciones de carácter general y a través de las circulares enviadas a las personas naturales o jurídicas sometidas a su control.

Fondo de protección social de los depósitos bancarios (FOGADE)

El Fondo de Protección Social de los Depósitos Bancarios es un instituto autónomo con personalidad jurídica y patrimonio propio e independiente de los bienes de la República, adscrito al Ministerio del Poder Popular de Planificación y Finanzas a los solos efectos de la tutela administrativa.

Su misión es garantizar los depósitos del público en las instituciones del sector bancario venezolano mediante la ejecución de acciones orientadas a dar respuesta efectiva y oportuna a los depositantes, administrando eficientemente los recursos y otros bienes de su propiedad, así como ejercer la función de liquidador de aquellas instituciones que así lo requieran y de sus personas jurídicas vinculadas, conforme al ordenamiento jurídico vigente, a fin de contribuir con la confianza y estabilidad del Sistema Financiero Nacional, demostrando un alto sentido de justicia y compromiso social, en beneficio del país y de su pueblo en concordancia con las políticas públicas contenidas en el Plan de Desarrollo Económico y Social de la Nación. (FOGADE, 2015).

Ley Orgánica del Sistema Financiero Nacional

En el artículo 1, el cual establece: “La presente Ley tiene por objeto regular, supervisar, controlar y coordinar el Sistema Financiero Nacional, a fin de garantizar el uso e inversión de sus recursos hacia el interés público y el desarrollo económico y social, en el marco de la creación real de un Estado democrático y social de Derecho y de Justicia”

Marco Legal las Tecnologías de la Información y la comunicación en Venezuela

(Bonillo, González, Lobo, & Lugo, 2012)

Constitución de la República Bolivariana de Venezuela (1999) Artículo N° 108

Artículo N° 109 Artículo N° 110

Art 108 “Los medios de comunicación social, públicos y privados, deben contribuir a la formación ciudadana. El Estado garantizará servicios públicos de radio, televisión y redes de bibliotecas y de informática, con el fin de permitir el acceso universal a la información. Los centros educativos deben incorporar el conocimiento y aplicación de las nuevas tecnologías, de sus innovaciones, según los requisitos que establezca la ley”.

Art 109 “El Estado reconocerá la autonomía universitaria como principio y jerarquía que permite a los profesores, profesoras, estudiantes, egresados y egresadas de su comunidad dedicarse a la búsqueda del conocimiento a través de la investigación científica, humanística y tecnológica, para beneficio espiritual y material de la Nación...”

Art 110 “...El Estado destinará recursos suficientes y creará el sistema nacional de ciencia y tecnología de acuerdo con la ley. El sector privado deberá aportar recursos para los mismos...”

Ley Orgánica de Telecomunicaciones (2000) Artículo N° 2Capítulo II

Disposiciones Transitorias Artículo N° 208 Numeral 10 (Presidencia de la República Bolivariana de Venezuela, 2000)

Art 2 “...Promover el desarrollo y la utilización de nuevos servicios, redes y tecnologías cuando estén disponibles y el acceso a éstos, en condiciones de igualdad de personas e impulsar la integración del espacio geográfico y la cohesión económica y social...”

Art 208 numeral 10 “...Las disposiciones previstas en materia de contenido de transmisiones y comunicaciones cursadas a través de los distintos medios de telecomunicaciones...”

Ley Orgánica de Ciencias, Tecnologías e Innovación (2005) Decreto 1290: Administración y Funcionamiento de la Tecnología (Presidencia de la República Bolivariana de Venezuela, 2005)

Estipula la organización del Sistema Nacional de Ciencia, Tecnología la innovación y la definición de los lineamientos que orientarán las políticas y estrategias para la actividad científica, tecnológica y de innovación.

Ley Orgánica de Educación (2009) Artículo N° 9 (Presidencia de la República Bolivariana de Venezuela, 2009)

“...Los medios de comunicación social, como servicios públicos, son instrumentos esenciales para el desarrollo del proceso educativo y como tales deben cumplir funciones informativas, formativas y recreativas que contribuyan con el desarrollo de valores y principios establecidos en la Constitución dela República y la presente Ley, con conocimientos, desarrollo del pensamiento crítico y actitudes para fortalecer la convivencia ciudadana, la territorialidad y la nacionalidad...”

Ley Orgánica para la Protección de Niños, Niñas y Adolescentes (2007) Artículo N° 68 y 69 Artículo N° 72 y 73 Artículo N° 76 y 78 (Presidencia de la República Bolivariana de Venezuela, 2007)

Artículo 68. Derecho a la información. Todos los niños, niñas y adolescentes tienen derecho a recibir, buscar y utilizar todo tipo de información que sea acorde con su desarrollo y a seleccionar libremente el medio y la información a recibir, sin más límites que los establecidos en la ley los derivados de las facultades legales que corresponden a su padre, madre, representantes o responsables.

Artículo 69. Educación crítica para medios de comunicación. El Estado debe garantizar a todos los niños, niñas y adolescentes educación dirigida a prepararlos y formarlos para recibir, buscar, utilizar y seleccionar apropiadamente la información adecuada a su desarrollo.

Artículo 72. Programaciones dirigidas a niños, niñas y adolescentes. Las emisoras de radio y televisión tienen la obligación de presentar programaciones de la más alta calidad con finalidades informativa, educativa, artística, cultural y de entretenimiento, dirigidas exclusivamente al público de niños, niñas y adolescentes, en un mínimo de tres horas diarias, dentro de las cuales una hora debe corresponder a programaciones nacionales de la más alta calidad.

Artículo 73. El Estado debe fomentar la creación, producción y difusión de materiales informativos, libros, publicaciones, obras artísticas y producciones audiovisuales, radiofónicas y multimedia dirigidas a los niños, niñas y adolescentes, que sean de la más alta calidad, plurales y que promuevan los valores de paz, democracia, libertad, tolerancia, igualdad entre las personas y sexos, así como el respeto a su padre, madre, representantes o responsables y a su identidad nacional y cultural.

Ley Nacional de Juventud Artículo N° 28 y Artículo N° 38 (Presidencia de la República Bolivariana de Venezuela, 2002)

Art 28 “El Estado, a fin de preservar el acceso y la permanencia de los jóvenes y las jóvenes en el sistema educativo, fortalecerá la educación nocturna y la educación a distancia mediante el uso de la informática, y de cualquier otro instrumento que fortalezca los estudios no presenciales.”

Art 38 “Los jóvenes y las jóvenes tienen derecho a que les sean reconocidas como propias todas las invenciones, creaciones científicas, tecnológicas y culturales que realicen, de conformidad con la ley respectiva.”

Planes

Plan Nacional Simón Bolívar Plan Nacional de Tecnología de Información (Presidencia de la República, 2007)

Plan Simón Bolívar: Incorporar las tecnologías de la información y la comunicación en Líneas el proceso educativo (Suprema Felicidad Social) Enfoques Universalizar el acceso a los Objetivos diferentes tipos de comunicación. Fortalecer la práctica de la

Estrategias información veraz y oportuna. (Democracia protagónica Políticas revolucionaria)Infraestructura tecnológica

Plan Nacional De Tecnología de la Información: Objetivos Incorporar una Plataforma Nacional de Tecnologías de Información, Formar recursos humanos en cuanto a las tecnologías de la información, Propiciar una visión prospectiva, voluntad política y toma de decisiones.

Decretos

Decreto 3390 (Decreto 3390, 2000).

Artículo N° 1. La administración pública nacional empleara prioritariamente software libre desarrollado con estándares abiertos, en sus sistemas, proyectos y servicios informáticos.

Artículo N° 10.

El Ministerio de Educación La Administración Pública y Deportes, en coordinación con el Ministerio de Ciencia y tecnología, establecerá las políticas para incluir software libre desarrollado con estándares abiertos, en programas de educación básica y diversificada.

Decreto N° 825 (Decreto 3390, 2000).

N°1. Se declara el acceso y el uso de Internet como política para el desarrollo cultural, económico, social y político Artículo de la República Bolivariana de Venezuela.”

N° 5. “El ministerio de Educación, Cultura y Deportes dictará las directrices tendentes a instruir sobre el uso Artículo de Internet, el comercio electrónico, la interrelación y la sociedad del conocimiento.”

N° 8. “En un plazo no mayor de tres (3) años, el cincuenta por ciento (50%) de los programas educativos de Artículo educación básica y diversificada deberán estar disponibles en formatos de Internet...”

CAPITULO III: MARCO METODOLÓGICO

En este capítulo se desarrolla el marco metodológico de la investigación, en el que se define el tipo de investigación, el diseño de la investigación, población y muestra, técnicas de recolección de datos, fases de la investigación, procedimientos por objetivos y, se realiza la definición conceptual y operacional de variables e indicadores.

3.1 Tipo de Investigación

Según los objetivos extrínsecos o externos de la investigación (porque se refieren a la utilidad que, fuera del ámbito estrictamente científico, van a tener las conclusiones que saquemos), Sabino (1992), las clasifica en dos grandes tipos: puras y aplicadas.

“Son investigaciones puras aquellas en que los conocimientos no se obtienen con el objeto de utilizarlos de un modo inmediato, aunque ello no quiere decir, de ninguna manera, que estén totalmente desligadas de la práctica o que sus resultados, eventualmente, no vayan a ser empleados para fines concretos en un futuro más o menos próximo. La investigación aplicada persigue, en cambio, fines más directos e inmediatos. Tal es el caso de cualquier estudio que se proponga evaluar los recursos humanos o naturales con que cuenta una región para lograr su mejor aprovechamiento, o las investigaciones encaminadas a conocer las causas que provocan una enfermedad, con el fin de proteger la salud”. (Sabino, 1992, p. 52).

De acuerdo al problema y objetivos extrínsecos o externos, esta investigación está enmarcada en el tipo aplicada, pues, consiste en el diseño de un modelo viable para solucionar problemas, requerimientos o necesidades de las organizaciones financieras, en cuanto al tratamiento de grandes volúmenes de datos para hacer inteligencia de negocios.

Según los objetivos internos o intrínsecos de la investigación (en cuanto al tipo de conocimiento que el científico espera obtener al finalizar el trabajo), Sabino (1992), las clasifica como exploratorias, descriptivas y explicativas.

De acuerdo a los objetivos internos o intrínsecos esta investigación, está enmarcada en el tipo en el tipo de investigación descriptiva, ya que busca especificar propiedades, características y rasgos importantes del entorno de inteligencia de negocios del

sistema financiero, así como también, del estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data que fueron tomadas en cuenta para el diseño del modelo propuesto.

3.2 Diseño de la Investigación

En función del tipo de datos a ser recogidos para llevar a cabo una investigación, Sabino (1992) , categoriza a los diseños en dos grandes tipos básicos: diseños bibliográficos y diseños de campo.

Sabino (1992) , señala que en los diseños de campo los datos de interés se recogen en forma directa de la realidad, mediante el trabajo concreto del investigador y su equipo. Estos datos, obtenidos directamente de la experiencia empírica, son llamados primarios. Cuando, a diferencia de lo anterior, los datos a emplear han sido ya recolectados en otras investigaciones y son conocidos mediante los informes correspondientes nos referimos a datos secundarios, por consiguiente a diseños bibliográficos.

El diseño de la investigación que sustenta este proyecto esta enmarcado en el tipo de Campo, ya que se obtuvieron los datos directamente de la realidad, lo que permitió levantar la información necesaria para el modelo de implementación diseñado.

Como apoyo al diseño de Campo se utilizó la Investigación Bibliográfica, como base complementaria a la investigación central, con el fin de recopilar, revisar y analizar fuentes de investigación bibliográficas especializadas, de diversos aspectos relacionados con el estado del arte de las herramientas de inteligencia de negocio y arquitecturas de Big Data.

3.3 Población y Muestra

Población

Para Hernández, Fernández, & Baptista (2010), una población es un conjunto de todos los casos que concuerdan con determinadas especificaciones, este punto se centra en “qué o quiénes”, es decir, en los participantes, objetos, sucesos o comunidades de

estudio (las unidades de análisis), lo cual depende del planteamiento de la investigación y de los alcances del estudio.

En la presente investigación la población fue representada por todas las instituciones bancarias, ubicadas en Venezuela, que realizan transacciones financieras por canales digitales, específicamente a través de la banca en línea.

Muestra

Para Sabino (1992), una muestra, “en un sentido amplio, no es más que eso, una parte del todo que llamamos universo y que sirve para representarlo”.

Existen dos procedimientos para extracción de la muestra; muestreo probabilístico y no probabilístico.

En el primer tipo la característica fundamental es que todo elemento del universo tiene una determinada probabilidad de integrar la muestra, y que esa probabilidad puede ser calculada matemáticamente con precisión. En las muestras no probabilísticas ocurre lo contrario y el investigador procede en cierta forma a ciegas, pues no tiene una idea del error que puede estar introduciendo en sus apreciaciones. (Sabino, 1992, p. 100).

Las muestras no probabilísticas más utilizadas son las llamadas accidentales, por cuotas o intencionales.

En particular, esta investigación se utilizó el muestreo no probabilístico, de tipo intencional, debido a que se realizó un procedimiento de selección informal y un proceso arbitrario. La muestra comprende de tres entidades financieras de la banca privada venezolana, como lo son; Banco Venezolano de Crédito, Bancrecer y Banco Exterior.

El Banco Venezolano de Crédito se utilizó como caso de estudio, para la aplicación y validación del modelo de implementación diseñado.

3.4 Técnicas de Recolección de Datos

Las técnicas de recolección de información, según Palella y Martins (2012), son las distintas formas o maneras de obtener la información, para el acopio de los datos se utilizan técnicas como la observación, entrevista, encuesta y pruebas entre otras.

“Un instrumento de recolección de datos es, en principio, cualquier recurso de que se vale el investigador para acercarse a los fenómenos y extraer de ellos información”. (Sabino, 1992, p. 114).

Dentro de cada instrumento concreto pueden distinguirse dos aspectos diferentes: forma y contenido. La forma del instrumento se refiere al tipo de aproximación que establecemos con lo empírico. En cuanto al contenido éste queda expresado en la especificación de los datos necesarios a conseguir; se concreta, por lo tanto, en una serie de ítems que no son otra cosa que los mismos indicadores que permiten medir las variables, pero que asumen ahora la forma de preguntas, puntos a observar elementos a registrar, etc.

A continuación se mencionan las técnicas utilizadas durante el proceso de recolección de datos para la culminación de este estudio:

Datos secundarios (Revisión Bibliográfica): Según Hernández, Fernández, & Baptista (2010), implica la revisión de documentos, registros públicos y archivos físicos o electrónicos. En esta técnica se recurrió a la revisión bibliográfica de libros, documentos, revistas y seminarios, con el fin de recopilar y revisar información que permitiera identificar el estado del arte de las herramientas y arquitecturas de Big Data, y contrastar el aspecto teórico con la situación real o práctica dentro del modelo de implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea en Venezuela.

Observación científica: Sabino (1992), la define como el uso sistemático de nuestros sentidos en la búsqueda de los datos que se necesitan para resolver un problema de investigación. Dicho de otro modo, observar científicamente es percibir activamente la realidad exterior con el propósito de obtener los datos que, previamente, han sido definidos como de interés para la investigación. En esta etapa se tomaron en cuenta

todos los elementos, que de una u otra forman guardan relación con la investigación. Se efectuó una observación de tipo participante directa, ya que el autor forma parte de la institución donde se realizó la investigación y se utilizó como instrumento una guía de observación.

Encuesta: Entendiéndose ésta, “como un conjunto de preguntas recogidas en un cuestionario para conocer la opinión del público sobre un asunto determinado”. La encuesta, fue dirigida a la población seleccionada de manera intencional, a través de un cuestionario, que según Hernández, Fernández, & Baptista (2010), consiste en un conjunto de preguntas respecto de una o más variables a medir. Se utilizó el cuestionario como instrumento, el cual fue estructurado con una serie de preguntas cerradas.

En cuanto a la medición de los instrumentos se utilizó el juicio de expertos, que de acuerdo con Hernández, Fernández, & Baptista (2010), “se refiere al grado en que aparentemente un instrumento de medición mide la variable en cuestión, de acuerdo con expertos en el tema”, a fin de someterlos a la consideración de profesionales, tanto en sus aspectos de estructura como de contenido. De igual forma fueron tomadas sus observaciones, para hacer las correcciones necesarias, garantizando la calidad y efectividad modelo.

3.5 Fases de la Investigación

Las fases o etapas a seguir para lograr los objetivos específicos, se definieron mediante una adaptación del proceso de investigación y sus momentos, planteado por Sabino (1992), en el que asevera que la investigación científica se desarrolla de acuerdo a los lineamientos generales del proceso de conocimiento. En ella se asiste, a ese acercamiento del investigador hacia el objeto de estudio, por un lado, y a la verificación de las teorías que se elaboran al confrontarlas con los datos de la realidad, por el otro.

En este proceso existe la posibilidad de distinguir algunas grandes fases o momentos que, desde un punto de vista abstracto, muestran las sucesivas acciones que va desarrollando el investigador mientras trabaja.

A continuación las fases del proceso de investigación realizado:

Fase Proyectiva: Esta fase se realizó a través de la identificación y desarrollo la formulación del problema, junto a la contextualización de un marco teórico como punto de partida, siendo esta la fase inicial del proceso de investigación.

Fase Metodológica: En esta fase se estableció el marco metodológico para alcanzar los objetivos planteados, fijando una estrategia ante los hechos a estudiar. En esta fase se elaboraron y validaron los instrumentos de recolección de datos.

Fase Técnica: Luego de elegir la metodología o estrategia general para ejecutar el proyecto, se abordaron las formas y procedimientos concretos que permitieron recolectar y organizar la información requerida. En esta fase se realizó el trabajo de campo y bibliográfico para la obtención de los datos.

Fase de redacción final: Finalmente, al disponer de los datos proporcionados por los objetos en estudio, en esta fase se elaboraron los nuevos conocimientos, inferidos a través de los datos adquiridos, mediante la identificación de las dimensiones del modelo propuesto. En esta fase se realizó la validación del modelo de implementación diseñado mediante su aplicación a un caso de estudio.

3.6 Procedimiento por Objetivos

Tabla 7. Procedimientos por objetivos planteados

Objetivos	Procedimiento
Analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero Venezolano.	Para realizar el análisis del entorno de Inteligencia de negocio en la banca en línea en Venezuela, se aplicaron los instrumentos de recolección de datos diseñados (cuestionario y lista de chequeo), a través de las técnicas seleccionadas (observación científica y encuesta).
Identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero.	Para identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema

	financiero, a través de la técnica de datos secundarios o revisión Bibliográfica, recurriendo a la revisión de libros, documentos, revistas y seminarios, con el fin de contrastar el aspecto teórico con la situación real del modelo a proponer.
Desarrollar las dimensiones del modelo de implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.	De acuerdo al análisis de la información recolectada se elaboraron los nuevos conocimientos, inferidos a través de los datos adquiridos, mediante la identificación y caracterización de las dimensiones del modelo propuesto.
Validar el modelo de implementación diseñado mediante su aplicación a un caso de estudio (Banco Venezolano de Crédito)	Una vez elaborado el modelo de implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea, se procedió a validar el mismo mediante el juicio de expertos y su aplicación a un caso de estudio en el Banco Venezolano de Crédito.

Fuente: Elaboración propia.

3.7 Variables, Definición Conceptual y Operacional e Indicadores

Variables: Según Hernández, Fernández, y Baptista (2010), una variable es una propiedad que puede fluctuar y cuya variación es susceptible de medirse u observarse.

Definición conceptual: Según Hernández, Fernández y Baptista (2010), una definición conceptual trata a la variable con otros términos. Se tratan de definiciones de diccionarios o de libros especializados.

Definiciones operacionales: Según Hernández, Fernández y Baptista (2010), una definición operacional constituye el conjunto de procedimientos que describe las actividades que un observador debe realizar para recibir las impresiones sensoriales, las cuales indican la existencia de un concepto teórico en mayor o menor grado.

A continuación la tabla 8. Muestra la definición de variables:

Tabla 8. Definición de Variables.

Objetivos específicos	Variables	Definición
Analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero Venezolano.	Entorno de inteligencia de negocios.	Situación actual del entorno de inteligencia de negocio y factores claves asociados al tipo de organizaciones financieras.
Identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero.	Herramientas de inteligencia de negocios y Arquitecturas de Big Data.	Estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data que permitan implementar soluciones de BI.
Desarrollar las dimensiones del modelo de implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.	Dimensiones del Modelo	Fundamentado en el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data, y de acuerdo a la ssituación actual del entorno de inteligencia de negocio en las organizaciones financieras, , formular un modelo de implementación de un sistema de inteligencia de negocio con Big data para la banca en línea en Venezuela.
Validar el modelo de implementación diseñado mediante su aplicación a un caso de estudio (Banco Venezolano de Crédito)	Aplicación del Modelo	Aplicación del modelo de implementación diseñado al caso de estudio (Banco Venezolano de Crédito)

A continuación la tabla 9. Muestra la operacionalización de variables realizada:

Tabla 9. Operacionalización de Variables.

Variables	Dimensión	Indicadores
Entorno de inteligencia de negocios.	Soluciones de Inteligencia de negocios	Existencia de Soluciones de Inteligencia de negocios
	Inversión de tiempo	Tiempo invertido en la generación de reportes
	Importancia	Percepción de la Importancia de la aplicación de Soluciones de BI
	Costos	Costos de licenciamiento, mantenimiento y soporte.
	Satisfacción	Grado de satisfacción con respecto a la inteligencia de negocios realizada.
	Oportunidades	Oportunidades percibidas
	Beneficios	Beneficios esperados
	Apoyo	Nivel de apoyo organizacional
	Resistencia al cambio	Grado de resistencia al cambio
	Barreras	Principales barreras
	Recursos tecnológicos	Capacidad de los recursos tecnológicos disponibles
	Aspectos importantes	Aspectos importantes a considerar al implementar BI
	Big Data	Experiencia con Big Data
	Posición actual	Posición actual ante iniciativas de Big Data
	Capacidad de BI	Disponibilidad de datos básicos
		Calidad de los datos básicos requeridos
		Acceso, integración, análisis y presentación de datos básicos
		Facilidad de uso
	Organizacional	Existencia de cultura analítica
		Apoyo a los líderes
	Involucramiento de usuarios	Conocimiento de los usuarios respecto a las capacidades, funcionalidades y limitaciones del sistema
		Compromiso de los usuarios con el éxito del sistema
		Participación de los usuarios en la evaluación del sistema
		Contribución de los usuarios al continuo mejoramiento del sistema sugiriendo/revisando las mejoras
	Éxito de BI	Eficiencia y funcionamiento del proceso de BI según lo planificado

		Satisfacción con el retorno de inversiones en BI
Herramientas de inteligencia de negocios y Arquitecturas de Big Data.	Inteligencia de negocios	Definiciones
		Evolución
		Arquitecturas
		Productos de BI
		Método de BI
	Big Data	Definiciones
		Arquitecturas
		Tecnologías
		BD NoSQL
		Análisis de datos
Dimensiones del Modelo	Valor	Gestión de multicanalidad
		Segmentación avanzada de clientes
		Fidelización de clientes
		Análisis de riesgos
		Anonimizar los datos y compartir estadísticos
	Volumen	Volumen de la información almacenada
		Almacén de datos
		Datos frío o de bajo contacto
		Desperdicio de datos
		Retos técnicos por volumen de datos
	Variedad	Tipos de datos
		Formato del contenido
		Origen de los datos
		Predecir el tipo de datos
		Datos no incluidos
	Velocidad	Velocidad de generación de datos
		Incremento en la velocidad de generación de datos
		Velocidad de los sistemas de BI
		Cambios de requerimiento de BI
		Satisfacción en los tiempos de BI
	Veracidad	La autenticidad o exactitud de los datos
		Ambigüedad en los datos
		Compleitud de los datos
		Características ACID
		Características del teorema de CAP
	Gobierno	Seguridad y privacidad

		Normalización de los datos
		Tiempo en el cual los datos están disponibles
		Propiedad de los datos
		Usos permitidos
	Personas	Apoyo de la alta gerencia
		Científicos o especialistas de datos
		Personal con conocimientos técnicos en Big Data
		Consumidores de datos
		Tipo de procesamiento requerido
Aplicación del Modelo	Aplicación del Modelo	Aplicación del Modelo

3.8 Estructura Desagregada de Trabajo

Con el propósito de esquematizar el alcance del proyecto de investigación, la figura 13 muestra la Estructura Desagregada de Trabajo.

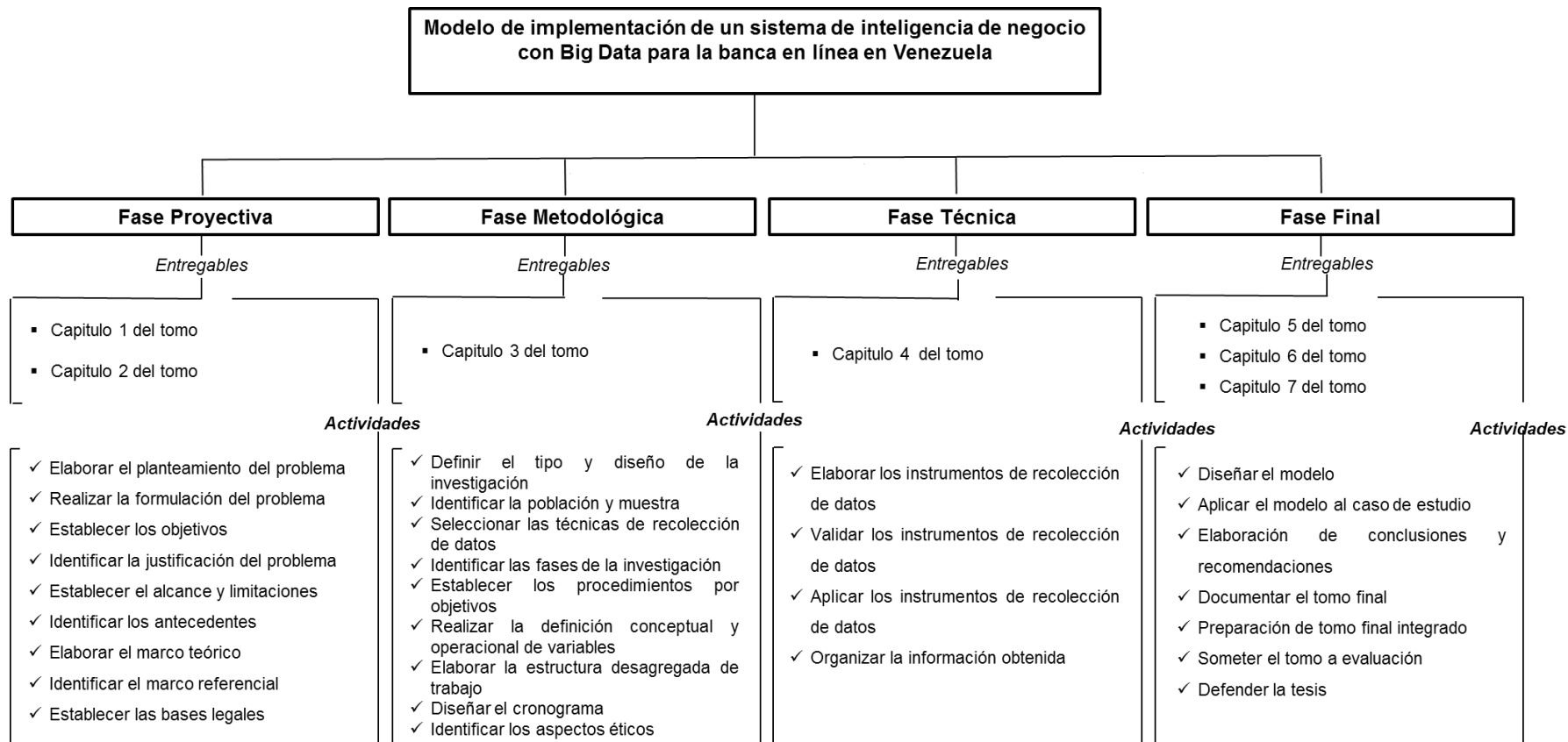


Figura 13. Estructura Desagregada de Trabajo.

3.9 Aspectos Éticos

En publicación de Gaceta Oficial N° 39.984, de fecha 13 de agosto de 2012, el Ministerio del Poder Popular de Planificación y Finanzas, por medio de la Superintendencia de las Instituciones del Sector Bancario, publicó la Resolución N° 099-12, de fecha 10 de junio de 2012, donde se dictan las Normas que permiten determinar el cumplimiento de los Requisitos de Calidad Moral y Ética exigidos para el Ejercicio de la Actividad Bancaria.

Norma: (SUDEBAN, 2011). Gaceta Oficial: # 39.714 de fecha 15 de julio de 2011. Ente Emisor: Superintendencia de las Instituciones del Sector Bancario (SUDEBAN). Norma: Resolución N° 199.11 del 14 de julio de 2011. Objeto/Título: Normas que permiten determinar el cumplimiento de los requisitos de calidad moral y ética exigidos para el ejercicio de la actividad bancaria.

Código de ética en los bancos en Venezuela

Las organizaciones bancarias que conforman el sistema financiero venezolano, tanto públicas como privadas, cuenta en su mayoría con un Código de Ética el cual constituye el conjunto de normas de carácter general que rigen la actividad bancaria, las normas internas de la Institución y los principios de actuación para los Directivos, Funcionarios y Empleados de cada Banco, tanto en sus relaciones interpersonales, como respecto a sus empleados y entorno social. Este es aprobado por la Junta Directiva.

Ética en este trabajo

Finalmente, el presente trabajo, respeta las normas éticas básicas de investigación, manteniendo la objetividad, revelando las fuentes de información y sus resultados, los cuales serán dispuestos para uso y beneficio de las organizaciones participantes en la muestra, ya que el modelo se desarrollará en función de aportar valor y mejorar las estrategias de negocio.

3.10 Cronograma

A continuación las figuras 14 y 15 muestran una representación gráfica de las actividades a desarrollar.

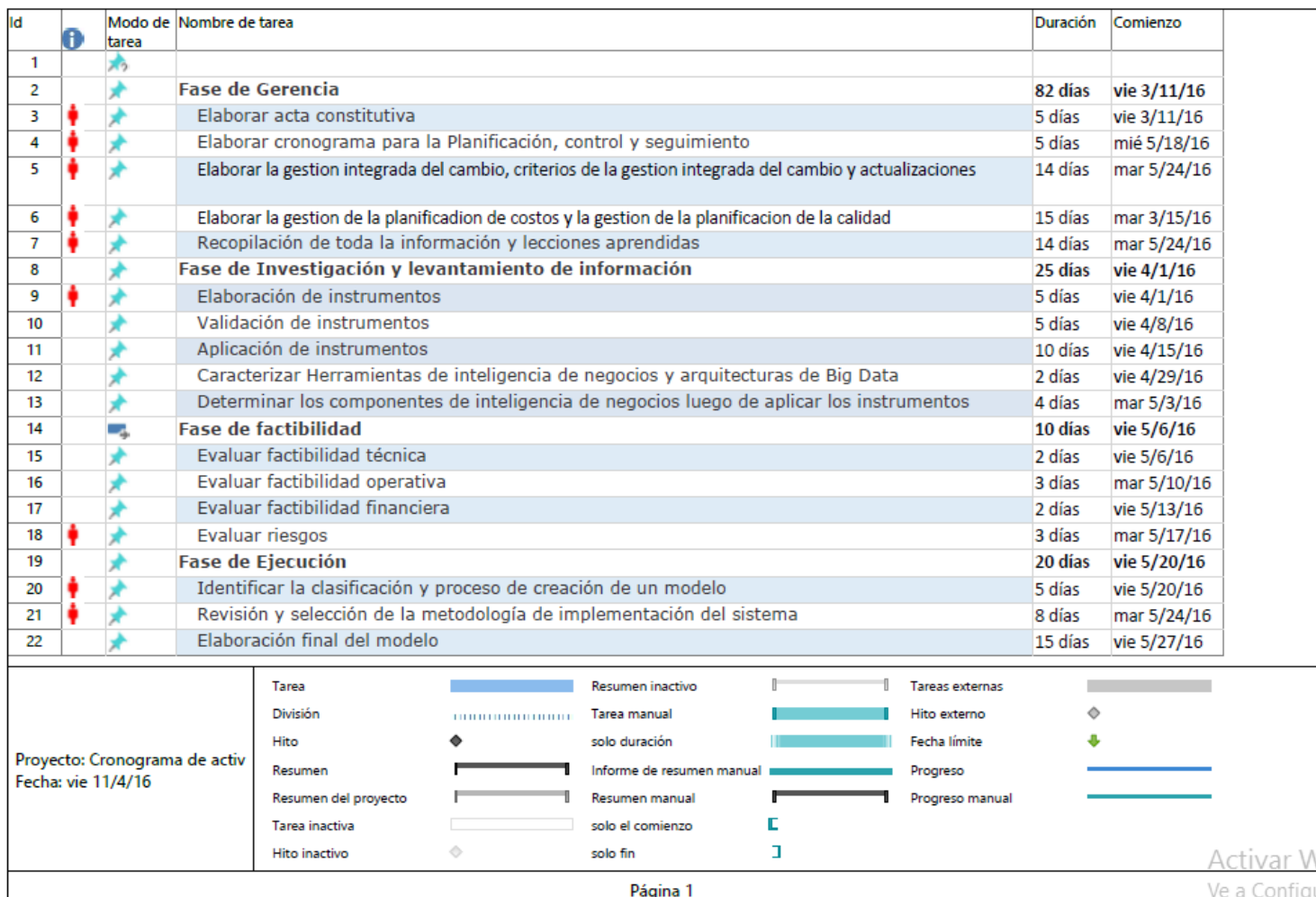


Figura 14. Cronograma de trabajo1/2.

Id		Modo de tarea	Nombre de tarea	Duración	Comienzo
23			Fase de evaluación y cierre	90 días	vie 6/10/16
24			Validar el modelo diseñado, a través de una encuesta	90 días	vie 6/10/16
25			Realizar retroalimentación de los resultados	85 días	vie 6/10/16
26			Coordinar asesoría final con el tutor	85 días	vie 6/10/16
27			Documentar el tomo final	85 días	vie 6/10/16
28			Preparación de tomo final integrado	85 días	vie 6/10/16
29			Someter el tomo a evaluación	85 días	vie 6/10/16
30			Defender la tesis	85 días	vie 6/10/16

Proyecto: Cronograma de activ

Fecha: vie 11/4/16

Tarea



División



Hito



Resumen



Resumen del proyecto



Tarea inactiva



Hito inactivo



Resumen inactivo

Tarea manual



solo duración



Informe de resumen manual



Resumen manual



solo el comienzo



solo fin



Tareas externas



Hito externo



Fecha límite



Progreso



Progreso manual



Página 2

Página 2

Figura 15. Cronograma de trabajo 2/2.

3.11 Recursos

La tabla 10 muestra los recursos utilizados para realizar la investigación:

Tabla 10. Matriz de Recursos

Nombre del recurso	Tipo	Iniciales	Costo/Uso	Capacidad máxima	Tasa estándar	Tasa horas extra	Acumular	Calendario base
Tesista	Trabajo	T	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Tutor	Trabajo	T	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Expertos	Trabajo	E	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Evaluable	Trabajo	E	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Jurado	Trabajo	J	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Usuarios	Trabajo	U	\$0.00	100%	\$0.00/hora	\$0.00/hora	Prorrateo	Estándar
Inscripción Seminario tópicos	Costo	I	5062.00				Prorrateo	
Inscripción Seminario I	Costo	I	6062.00				Prorrateo	
Inscripción Seminario II	Costo	I	10500.00				Prorrateo	
Inscripción TGM	Costo	I	31500.00				Prorrateo	
Impresión	Material	I	\$0.00		\$0.00		Prorrateo	
Computadora	Material	C	\$0.00		\$0.00		Prorrateo	
Internet	Costo	I					Prorrateo	

CAPÍTULO IV: ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

El presente capítulo consiste en el análisis e interpretación de los resultados obtenidos, mediante la aplicación de los instrumentos de recolección de datos, con el objetivo de comprender y contextualizar el entorno en el que se realiza la investigación.

4.1 Fase Técnica

En esta fase se realizó el trabajo de campo y bibliográfico para la obtención de los datos, permitiendo analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero Venezolano, así como también, identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data.

Para el cumplimiento de estos objetivos se utilizaron varias técnicas de recopilación de datos, tales como la revisión bibliográfica, la observación científica y la encuesta a través de la aplicación de instrumentos como la recolección de datos secundarios de fuentes bibliográficas, una guía de observación y un cuestionario.

La muestra comprende de tres entidades financieras de la banca privada venezolana, como lo son; Banco Venezolano de Crédito, Bancrecer y Banco Exterior.

4.1.1 Recolección, procesamiento y análisis de los datos

- a. Recolección de datos secundarios de fuentes bibliográficas:** La técnica de revisión bibliográfica a través de la recolección de datos secundarios permitió obtener las referencias teóricas y conceptuales que sustentan la investigación. A partir de estudios e investigaciones relacionados con el tema planteado, como las tesis de maestría realizadas por Guama (2015), Espinoza (2015) y López (2013), así como también, los artículos arbitrados elaborados por Montalvo (2015), Villalobos (2015), Ernst & Young (2014), Paredes (2015), Camargo (2014) y Delrieux et al. (2014), fueron determinantes para fundamentar teorías en las áreas de inteligencia de negocio y Big Data. Al mismo tiempo la

documentación mencionada, permitió identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero, lo cual apoyo finalmente al desarrollo del modelo. Para identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero estudiaron dichas fuentes, tomando en cuenta principalmente, definiciones, evolución, arquitecturas, producto y métodos de implementación de inteligencia de negocios, y en cuanto a Big Data, se consideraron principalmente definiciones, características, arquitecturas, tecnologías, bases de datos, ejemplos de aplicación. Finalmente la identificación del estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data fue incorporado al marco teórico de la investigación.

- b. Guía de observación:** En la aplicación de la técnica de observación científica se utilizó una guía de observación, la cual consta de cuatro aspectos a evaluar, incluyendo; la capacidad de BI actual, factores organizacionales, involucramiento de usuarios y el éxito de BI realizado, para los que se debió indicar con una X dentro de las casillas correspondientes el nivel de aplicación, donde 5 es *Siempre*, 4 es *Casi siempre*, 3 es *Algunas veces*, 2 es *Muy pocas veces* y 1 es *Nunca*, y en los casos necesarios se detalló en el campo de observaciones la información requerida. Una vez obtenidos estos datos, fueron codificados, transferidos a una matriz y almacenados, para finalmente proceder a analizarlos. La guía de observación puede ser consultada en el anexo I.
- c. Cuestionario:** La técnica de la encuesta se llevó a cabo a través del diseño de un cuestionario compuesto por 14 preguntas cerradas, entre las que se encuentran preguntas de elección única de tipo dicotómicas (con respuesta sí/no), politómicas (con varias alternativas), preguntas de elección múltiple, y preguntas de escala basadas Likert. Su desarrollo y aplicación se llevó a cabo por medio de la herramienta de google Forms. La encuesta fue aplicada a 7 personas con cargos de Vicepresidente, gerentes y coordinadores, de las áreas de inteligencia de negocios y administración de datos del Banco Venezolano de Crédito, Banco Exterior y Bancrecer. El cuestionario puede ser consultado en el

anexo II. Para el levantamiento de información se elaboraron preguntas relacionadas con las siguientes variables: la existencia de soluciones de inteligencia de negocios, el tiempo invertido en generación de reportes, la importancia que se le da a las soluciones de BI, sus costos en licenciamientos, mantenimiento y soporte, el grado de satisfacción con los procesos actuales, oportunidades percibidas, beneficios esperados, nivel de apoyo organizacional, el grado de resistencia al cambio, las posibles barreras, la capacidad del recurso tecnológico disponible, los aspectos que consideran importante estas organizaciones, la experiencia con proyectos relacionados con Big Data y finalmente la posición actual ante iniciativas de incorporar tecnología Big Data a proyectos de inteligencia de negocios. Una vez obtenidos estos datos, fueron codificados, transferidos a una matriz y almacenados, para finalmente proceder a analizarlos.

4.1.2 Validación de los instrumentos

La validación de los instrumentos se realizó a través del juicio de tres expertos, quienes ejercen actualmente como evaluadores en la Universidad Católica Andrés Bello, entre los que se encuentra MSc. Gloria Aponte, MSc. María Esther Remedios y finalmente el tutor de la tesis Dr. Pedro Nolasco Bonillo. Cada experto recibió un formato de validación vía email donde se recogió la información. Esta planilla contiene los siguientes aspectos de información por cada ítem congruencia, claridad, tendenciosidad, observación. Al recibir la revisión de los expertos se procedió a realizar las correcciones necesarias, para posteriormente aplicar los instrumentos. El formato de validación de los instrumentos puede ser consultado en el anexo III.

4.1.3 Análisis de los resultados del diagnóstico

El análisis de los resultados contiene la información hallada durante el transcurso de la investigación. El análisis de esta información fue realizado en función de la obtención de los objetivos específicos planteados. Inicialmente se presentan los resultados la

encuesta, posteriormente los resultados de la observación científica y finalmente la revisión bibliográfica realizada.

Revisión bibliográfica

La técnica de revisión bibliográfica a través de la recolección de datos secundarios, fue realizada tomando como referencia información obtenida de tres tesis de grado de maestría, seis artículos arbitrados y varios sitios web especializados en temas relacionados con este proyecto, lo cual permitió adquirir las referencias teóricas y conceptuales que sustentan la investigación, y al mismo tiempo, permitió identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero.

A continuación se analizan los resultados obtenidos para cada una de las características principales tomadas en cuenta para identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero:

Definiciones de inteligencia de negocios: De acuerdo con los resultados de la investigación, no existe una definición universal de lo que es “inteligencia de negocios”, sin embargo, los diversos autores estudiados coinciden en definir el término haciendo referencia directamente al proceso de toma de decisiones, sustentado en el uso de tecnologías, técnicas y herramientas, que permiten transformar datos en información, y esta información en conocimiento, con el fin de alcanzar objetivos, generar valor y ventajas competitivas en las organizaciones.

Evolución de la Inteligencia de negocio: La inteligencia de negocio a seguido una evolución desde los sistemas que respondían a la pregunta “¿qué pasó?”, pasando por “¿qué pasó y por qué?”, hasta llegar a las herramientas que hoy en día nos permite responder a “¿qué puede pasar?”, donde cada pregunta está asociada a una categoría de sistemas de información empresarial, entre las que se encuentra BI histórica

(indicadores financieros), BI operacional (sistemas de fijación de precios), y BI analítico o predictivo (predicción de comportamiento de clientes) respectivamente.

Arquitecturas de Inteligencia de negocio: Las arquitecturas de inteligencia de negocios están compuestas básicamente, en primer lugar por las fuentes de datos, que contemplan todos los procesos operacionales en los cuales se crean datos constantemente, una vez que se tiene establecido el insumo útil que permitirá generar la información relevante de un negocio, se aplican procesos de extracción, transformación y carga (conocidos como ETL), con herramientas especializadas para este fin y finalmente la entrega de un producto para ser utilizado. Este producto es la información precisa y preparada para mostrar al usuario, para lo cual se emplean herramientas de análisis y visualización de datos, con una variedad de funcionalidades.

Productos de Inteligencia de negocio: Los productos de Inteligencia de negocio son las herramientas que finalmente permiten el análisis y visualización de datos, actualmente ofrecen diversas opciones, para que el usuario pueda consultar los informes o indicadores del proceso; a través de un computador de escritorio o un dispositivo móvil, una página web interna o en Internet, según la necesidad u oportunidad de acceso en determinado momento. Los cuadros de mando integrales (CMI), los sistemas de soporte a la decisión (DSS) y los sistemas de información ejecutiva (EIS) son los principales productos de inteligencia de negocio que existen hoy en día.

Método para implantar proyectos de BI: Los métodos para implantar proyectos de inteligencia de negocios proponen las fases para la realización de este tipo de proyectos, los cuales finalmente solo sirven como un marco de referencia, ya que todo dependerá de las características particulares de cada organización. Generalmente se encuentran compuestos principalmente de fases como definición, levantamiento de información, diseño, desarrollo, pruebas, puesta en producción, entrenamiento, cierre, entre otras.

El estado del arte de las herramientas de inteligencia de negocios se encuentra resumido en la siguiente imagen.

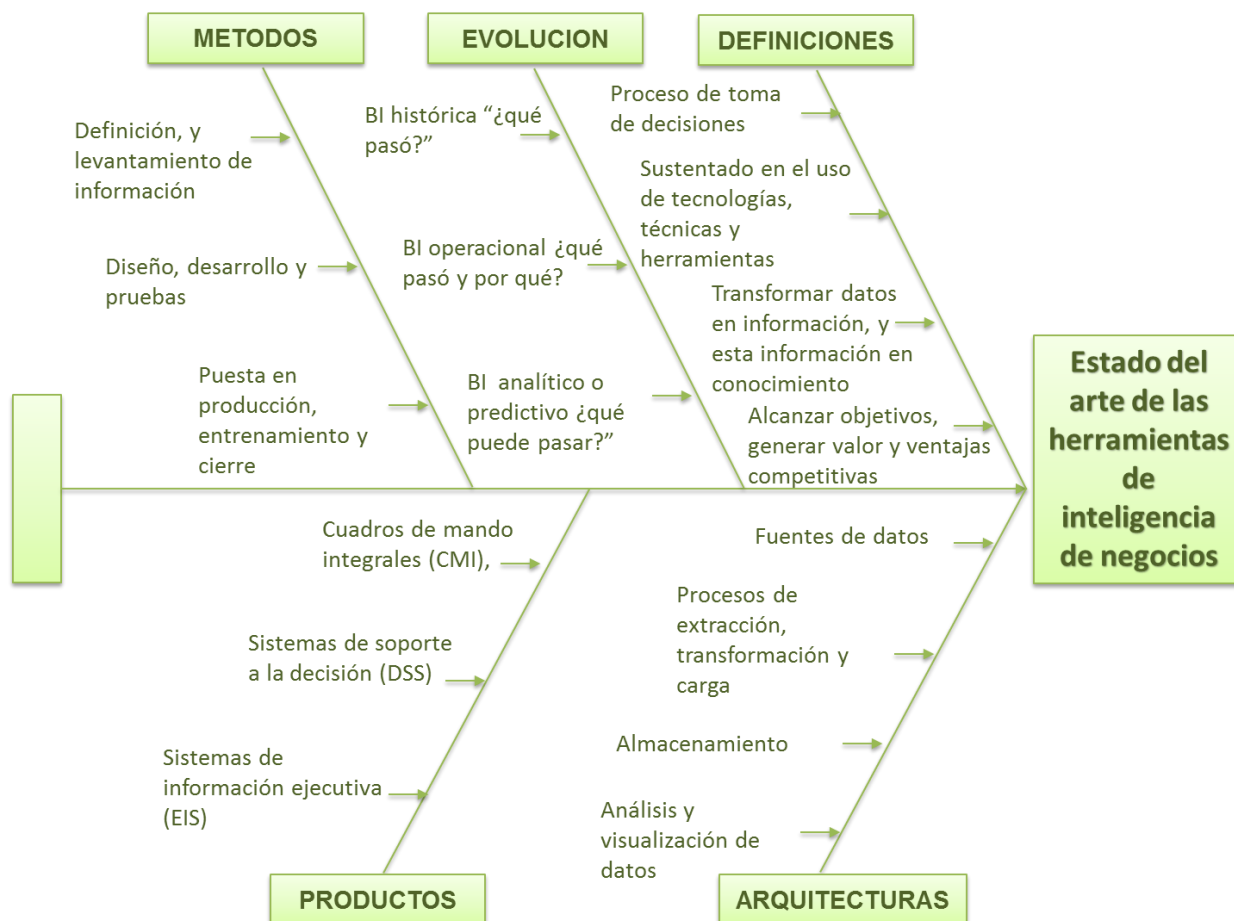


Figura 16. El estado del arte de las herramientas de inteligencia de negocios.

Definiciones de Big Data: El termino Big Data es relativamente nuevo, y de acuerdo con la investigación realizada no existe un consenso en cuanto a su definición, sin embargo los diversos autores estudiados coinciden en definir el término haciendo referencia directamente al proceso de gestión, administración y tratamiento de grandes volúmenes de datos, de distintos tipos, fuentes y formatos, a través del uso de la tecnología, cuyas características representan un nivel de complejidad para la gestión de datos convencional.

Características de Big Data: Big Data hace referencia al tratamiento y análisis de grandes repositorios de datos, originados de transacciones, interacciones y observaciones, cuyo tamaño actual esta en el orden de Petabytes de datos. Soporta el manejo de datos estructurados, no estructurados y semiestructurados. Tiene características particularidades en el tratamiento de los datos, a lo que diversos autores le han asignado el termino “las 3V de Big Data”, que con el transcurso del tiempo a pasado a ser 5 y hasta 7 “V” de Big data, pero que en general hacen referencia a las características principales las cuales son el volumen, variedad y velocidad de los datos.

Arquitecturas de Big Data: Las arquitecturas de Big Data, al igual que las arquitecturas de inteligencia de negocio, están compuestas básicamente por fuentes de datos, procesos de extracción, transformación y carga, almacenamiento y herramientas de visualización. Sin embargo tienen la particularidad de incluir nuevas herramientas tecnológicas en cada uno de estos componentes, permitiendo el tratamiento de grandes volúmenes de datos, más allá que el realizado con las herramientas de tecnológicas tradicionales.

Tecnologías Big data: La implementación de una solución Big Data se descompone en diferentes tecnologías que trabajan conjuntamente para lograr el objetivo final de almacenar y procesar grandes volúmenes de datos. Entre las tecnologías involucradas se encuentran el framework del proyecto Apache Hadoop, el uso de base de datos NoSQL entre las que podemos mencionar Cassandra, Hbase, MongoDB, Neo4j, entre otras. Sin embargo, también se hace necesario mencionar el Teorema de Brewer o Teorema de CAP, en el cual se enfocan las bases de datos NoSQL.

Ejemplos de aplicación: En este momento, existen varias prácticas de utilización de Big Data tanto en gigantes de la web como Google, Facebook o LinkedIn, así como también en compañías más tradicionales. En el sector público; con servicios de inteligencia, defensa y protección, en el sector salud; monitorización remota de pacientes, localización de emergencias y almacenamiento de historias clínicas, en el sector de telecomunicaciones; control de la red, venta de servicios de localización,

servicios de publicidad asociados al patrón de llamadas o las aplicaciones descargadas, en el sector financiero; prevención y detección de fraude, riesgo crediticio, segmentación de clientes, entre otros.

El estado del arte de las arquitecturas de Big Data se encuentra resumido en la siguiente imagen.

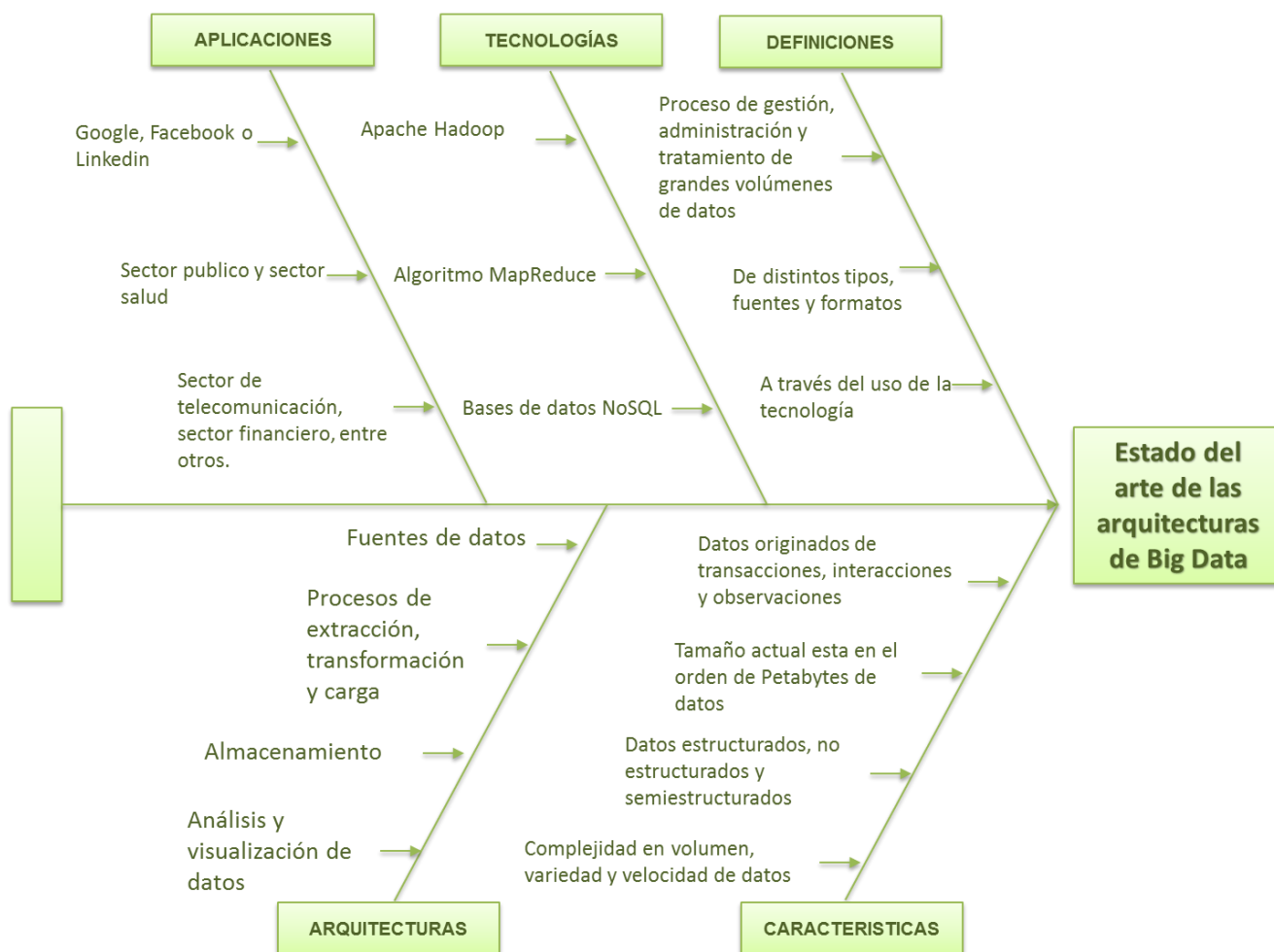


Figura 17. El estado del arte de las arquitecturas de Big Data.

Es un hecho que debido a los exponenciales avances de las tecnologías de la información y la comunicación, el uso de herramientas tecnológicas que apoyen el proceso de inteligencia de negocios con grandes volúmenes de datos en las organizaciones financieras, es una necesidad generalizada, pero cuando se trata de la implementación de proyectos de sistemas de inteligencia de negocios con Big data,

existen características deben ser identificadas claramente, como punto de partida, lo cual será determinante al momento de decidir si innovar con soluciones de inteligencia de negocios tradicionales o aplicar soluciones de inteligencia de negocios desde la perspectiva de Big Data. Por lo que es necesario apoyar este proceso desde su inicio, permitiendo en primer lugar descubrir las necesidades y puntos de mejora en la organización, y al mismo tiempo que se logre la implementación del proyecto con mayores posibilidades de éxito.

Observación científica

Para la observación científica, se observaron ciertos aspectos establecidos bajo una guía de observación compuesta por cuatro aspectos a evaluar, dentro de los cuales se encuentran un total de 7 ítems, con una puntuación de acuerdo al grado de aplicación del 1 al 5. El instrumento fue aplicado al Banco Venezolano de Crédito.

Se obtuvo como resultado que la institución financiera, en cuanto a capacidad de BI se encuentra por encima del punto medio, observándose una capacidad de BI aceptable sin llegar a ser excelente. Sin embargo se observa que cuando se trata de acceso, integración, análisis, presentación de datos y facilidad de uso, no se percibe que los sistemas de BI cumplan totalmente con las expectativas.

Por otra parte los factores organizacionales evaluados están compuestos por dos dimensiones: Cultura Analítica y Apoyo de los Líderes. En general se observa que la toma de decisiones se realiza de manera informada y respaldada en los datos, y la alta gerencia da un apoyo aceptable a las iniciativas de BI.

Así mismo, se observa que el involucramiento de los usuarios en general está por encima del punto medio, lo cual indica, que con un poco más de iniciativa se podría llegar al involucramiento ideal.

Finalmente se observa que la eficiencia y funcionamiento del proceso de BI según lo planificado, junto a la satisfacción con la capacidad de BI dentro de la organización financiera, se encuentra en un punto medio, demostrando que no están totalmente

satisfechos, ni insatisfechos con las capacidades de BI en sus organizaciones. Es importante destacar que estas hipótesis serán comprobadas directamente a través del cuestionario. La representación gráfica de los resultados del cuestionario puede ser consultada en el anexo IV.

Encuesta

El instrumento fue aplicado a los Bancos Venezolano de Crédito, Banco exterior y Banco Bancrecer. Se realizó el análisis cualitativo y se interpretaron los resultados obtenidos de un cuestionario compuesto por 14 preguntas cerradas, entre las que se encuentran preguntas de elección única de tipo dicotómicas, preguntas de elección múltiple y preguntas de escala basadas Likert. El uso de la herramienta de google forms permitió obtener los resultados de manera gráfica representados de forma general, a través de diagramas circulares basados en el cálculo porcentual para cada ítem.

La aplicación de este instrumento permitió analizar el entorno de Inteligencia de negocio en la banca en línea del Banco Venezolano de Crédito. La representación gráfica de los resultados de la encuesta puede ser consultada en el anexo V.

A continuación se analizan los resultados obtenidos para cada una de las variables asociadas al entorno de inteligencia de negocios y Big Data en las instituciones financieras encuestadas:

Existencia de soluciones tecnológicas de Inteligencia de negocios en las entidades financieras: 100 % de los encuestados coincide en afirmar que utilizan dentro de la organización financiera alguna herramienta tecnológica o programa para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en la institución..

Inversión de tiempo en la generación de reportes dentro de las entidades financieras: En cuanto al tiempo invertido en la generación de reportes 85.7% de los encuestados manifiesta estar insatisfecho, lo que hace evidente el descontento de los

participantes, quienes reflejan que la realidad es que, mientras más tiempo se requiera para generar reportes e indicadores, más tiempo se requiere para interpretar los mismos, lo que apunta finalmente a un proceso Inteligencia de negocios y toma de decisiones deficiente. Otro 14.3 % de los encuestados indica que se encuentran muy satisfecho con el tiempo invertido en la generación de reportes. Los reportes son el resultado final fruto de todo el proceso de inteligencia de negocios, y es lo que los directivos pueden obtener de forma “tangible” e inmediata como resultado de su inversión, por lo que además del tiempo invertido para generarlos, se deben tomar en cuenta otros factores como la cantidad de información a presentar, su precisión y efectividad.

Nivel de importancia que dan las entidades financieras a las soluciones tecnológicas de Inteligencia de negocios: El 42.9% de los encuestados concuerda en la importancia que tienen las soluciones tecnológicas de inteligencia de negocios en la organización financiera es muy alta y otro 42.9% indica que es alta, lo que ayuda a adquirir un mejor entendimiento de ella misma y de sus clientes, esto gracias a la capacidad de explotar su información, con la intención de poder manipularla de una manera más sencilla y entender su desempeño, o plantear escenarios a futuro, lo cual ayudará directamente a tomar mejores decisiones y generar valor. Finalmente un 14.3 % afirma que no es ni alta ni baja. La importancia de las herramientas tecnológicas de inteligencia de negocios destaca hoy en día en las instituciones financieras por el alto volumen de información que manejan de sus clientes y su aplicación se puede incorporar en cualquier ámbito en la medida que haya necesidad de desarrollar soluciones específicas a problemas específicos.

Costos de las herramientas de inteligencia de negocios actuales en las entidades financieras: A pesar de la importancia con la que se aprecia en las instituciones financieras a las herramientas de inteligencia de negocios, un elemento interesante es que el 71.4 % manifiesta estar insatisfecho con los costos de licenciamiento, mantenimiento y soporte de dichas herramientas, lo cual indica que se percibe el costo de software, hardware y servicios, asociados a la inteligencia de negocios, como

factores de costo elevado. Por otra parte 14.3 % de los encuestados indica estar muy insatisfecho y otro 14.3 % muy satisfecho.

Satisfacción de las entidades financieras con respecto a su inteligencia de negocios: El 71.4% de los encuestados manifiesta no estar ni satisfecho, ni insatisfecho con respecto a la inteligencia de negocios realizada en la institución financiera, aun cuando este grado de satisfacción no es bajo, se considera que es posible mejorar este conjunto de estrategias y herramientas que se tienen, enfocadas a la administración y creación de conocimiento mediante el análisis de datos de la organización. Por otra parte 14.3% de los encuestados indica estar muy insatisfecho y 14.3% está satisfecho.

Oportunidades percibidas por parte de las entidades financieras para implementar proyectos de inteligencia de negocios y Big Data: Desde el punto de vista de las oportunidades percibidas por parte de las organizaciones financieras para implementar proyectos de inteligencia de negocios y Big Data, las respuestas de los encuestados develan que la principal área de aplicación la constituye una mejor gestión de multicanalidad, tan importante en un contexto en el que las nuevas tecnologías ponen a disposición de los clientes un mayor número de canales de comunicación mejorados y alta capacidad de transacción con las empresas financieras.

La segmentación avanzada de clientes se perfila como la segunda aplicación de estas herramientas, la incorporación de nuevas fuentes de datos a las segmentaciones tradicionales posibilita la obtención de una visión y un entendimiento más completos de los clientes de las entidades financieras, lo que conlleva a la definición de propuestas de valor más adaptadas a los diferentes perfiles de clientes. Conocer mejor a los clientes permite realizar un targeting más ajustado, y esto se traduce en mayor eficiencia y rendimiento de las acciones comerciales.

Una tercera oportunidad de aplicación, no menos importante, viene representada por la fidelización de clientes, demostrando ser también una prioridad para las entidades

financieras, lo que se traduce en definir acciones para retención de clientes y en detectar abandono con suficiente antelación.

Beneficios esperados por las entidades financieras sobre los proyectos

Inteligencia de negocios y soluciones Big Data: En cuanto a los beneficios esperados por las entidades financieras sobre los proyectos de inteligencia de negocios y Big Data, la totalidad de los encuestados coincide en señalar que, el apoyo en procesos complejos de toma de decisiones y la extracción de “inteligencia” de la combinación de fuentes estructuradas y no estructuradas de información, como los primeros aspectos prometedores y diferenciales. Esto demuestra que el principal punto de partida al definir una estrategia de adopción de inteligencia de negocio y Big Data, debe enfocarse en determinar las cuestiones del negocio a las que se necesita dar una respuesta, siendo esta la mejor manera de asegurar el éxito en la generación de valor de estas iniciativas.

Ser capaz de trabajar con grandes volúmenes de información es otro aspecto relevante para esta organización, que combinado con los dos principales beneficios percibidos anteriormente, permitiría a la entidad aprovechar nuevas fuentes de datos para mejorar la calidad de la toma de decisiones, incrementando su eficiencia y aumento de sus ingresos.

Apoyo organizacional en las entidades financieras para implementar proyectos

de inteligencia de negocios y Big Data: 42,89 % de los encuestados considera que es el nivel de apoyo organizacional que existe para implementar proyectos de inteligencia de negocios y Big Data es alto, mientras que un 28.6% percibe que es muy alto y el 28.6% restante cree que no es alto ni bajo. Lo que representa un punto importante ya que una comprensión amplia por parte de la gerencia y apoyo a las iniciativas influirá directamente en el éxito de cualquier proyecto de inteligencia de negocios y Big Data.

Percepción de resistencia al cambio en las entidades financieras para

implementar proyectos de inteligencia de negocios y Big Data: 71.4 % de los

encuestados manifiesta que el porcentaje de resistencia al cambio no es ni alto ni bajo, otro 14.3 % indica que es muy alto y 14.3 % indica que es bajo, lo que representan una cifras importante para comenzar con la comprensión y el respeto a la cultura de la oficina. Por el contrario en entornos donde la cultura de resistencia al cambio es alta, planificar y ejecutar cualquier tipo de cambio es más difícil, y por lo general los gerentes de alto nivel suelen encontrar muchas excusas para no hacer un cambio. (El cambio será demasiado caro, lento o fuera de alcance.)

Barreras de entrada que enfrentan las entidades financieras para implementar proyectos de inteligencia de negocios y Big Data: Las principales barreras identificadas por los encuestados de las instituciones financieras para la ejecución de proyectos de inteligencia de negocios y Big Data se relacionan principalmente con la Falta de presupuesto económico, la falta de criterio sobre la rentabilidad que puede ofrecer este tipo de inversiones (ROI) y la falta de tiempo para abordar los mismos, lo que demuestra que existe una percepción de que es necesario realizar grandes inversiones en tecnología para poder implementar dichas iniciativas, algo difícil de asumir si se tienen en cuenta las restricciones presupuestarias actuales, la dificultad de estimar y cuantificar la rentabilidad de estas inversiones.

Finalmente en una disciplina tan novedosa como Big Data, la falta de conocimiento y talento es también un serio problema para muchas compañías, hecho que ha manifestado un poco más de la mitad de los encuestados, junto con las barreras relacionadas con la capacidad tecnológica de la entidad. Es necesario contar con científicos de datos, expertos en visualización, analistas de negocio, profesionales en la gestión de datos y otros especialistas como expertos en tratamiento de datos confidenciales, y el desarrollo de estas capacidades en el seno de las entidades es complejo en el corto plazo.

Capacidad de recursos tecnológicos disponibles en las entidades financieras para implementar proyectos de inteligencia de negocios y Big Data: El 57.1 % de los encuestados considera que la capacidad de los recursos tecnológicos que estarán disponibles al momento de implementar proyectos de inteligencia de negocios y Big

Data no es ni alta ni baja, 14.3 % indica que es baja, 14.3 % indica que es alta, mientras que 14.3 % indica que es muy alta, lo que se traduce en recursos tecnológicos mínimos disponibles para dar los primeros pasos e implementar proyectos pilotos.

Aspectos que consideran importantes las entidades financieras al momento de implementar proyectos de inteligencia de negocios y Big Data: 71.11 % de los encuestados afirma que el apoyo a la toma de decisiones es el aspecto más importante que consideraría la organización al momento de obtener una solución de inteligencia de negocios, para explotar los datos obtenidos como resultados de sus operaciones transaccionales. Otro 14.3% valora que permita manejar múltiples fuentes de información. Finalmente un 14.3 % de los encuestados indica que un aspecto importante es que debe proporcionar indicadores seguros.

Experiencia de las entidades financieras en proyectos de Big Data: El 100 % de los encuestados afirma no haber tenido experiencia directa con proyectos con Big Data.

Posición actual de las entidades financieras ante iniciativas de proyectos de inteligencia de negocios y Big Data: A pesar de que en la pregunta anterior donde 100 % de los encuestados afirma no haber desarrollado proyectos con Big Data, el 71.4 % manifiesta estar interesado en realizar proyectos piloto, como una primera aproximación a Big Data, que les permita probar los beneficios que pueden derivarse de estas nuevas metodologías y soluciones.

Un 14,3 % indica estar utilizando proyectos piloto, con resultados prometedores y otro 14.3 % se encuentra a la espera de que expertos demuestren respuestas concretas.

Se detecta una cierta cautela y escepticismo, con un alto porcentaje de personas ya conocedoras de estas nuevas tecnologías pero aún a la espera de ser capaces de identificar aplicaciones y utilidades concretas, este grupo está a la espera de que entidades con más capacidades analíticas y tecnológicas desarrollen las primeras

experiencias y sirvan como referencia para futuras implantaciones. Finalmente una minoría afirma estar realizando proyectos pilotos.

El análisis e interpretación de los resultados obtenidos en la aplicación del cuestionario y la guía de observación se encuentran resumidos en la siguiente imagen.

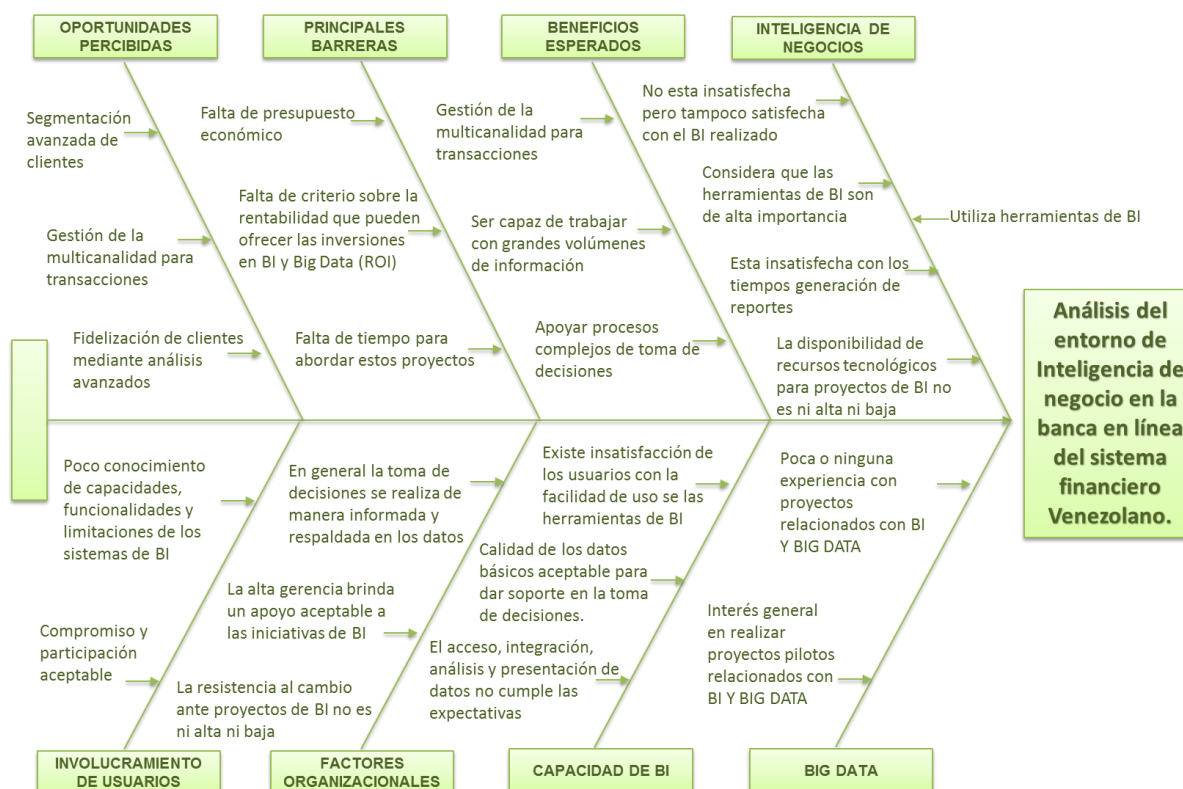


Figura 18. Análisis del entorno de inteligencia de negocio en la banca en línea del sistema financiero venezolano.

Los resultados obtenidos de la aplicación del cuestionario y la guía de observación permitieron analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero venezolano. Fueron identificadas un conjunto de características que pueden variar de acuerdo a las particularidades de cada organización y que deben ser tomadas en cuenta al momento de plantearse la implementación de proyectos de inteligencia de negocios con Big data, las cuales permitirán alcanzar mayores posibilidades de éxito.

CAPÍTULO V: DESARROLLO DE LA PROPUESTA

El presente capítulo tiene como finalidad presentar un modelo que permita implementar sistemas de inteligencia de negocio con Big Data en la banca en línea de las instituciones financieras en Venezuela. En función del diagnóstico realizado, fueron identificadas un conjunto de características que pueden variar de acuerdo a las particularidades de cada organización financiera y que deben ser tomadas en cuenta al momento de plantarse la implementación de proyectos de inteligencia de negocios con tecnología Big data.

Luego de analizado el entorno de inteligencia de negocios asociado la banca en línea del sector bancario, sus carencias, necesidades y expectativas, se opta por proponer un modelo de implementación de un sistema de inteligencia de negocios con Big Data, que sirva de guía a las organizaciones financieras, partiendo de la investigación bibliográfica realizada, tomando en cuenta características y atributos que estas organizaciones deben considerar para saber si están preparadas y si el negocio realmente requiere de este tipo de tecnología, permitiendo al mismo tiempo alcanzar mayores posibilidades de éxito.

5.1 Características del modelo propuesto:

Para el desarrollo del modelo se identificó el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data, vinculándolo posteriormente a las características, carencias, necesidades y expectativas obtenidas como resultado de la aplicación del cuestionario y la guía de observación.

El modelo propuesto se muestra en las figuras 23 y 24, el cual consiste en una adaptación del ciclo de vida del desarrollo de sistemas planteado por Kendall & Kendall (2005), y consta de ocho partes: siendo la primera; identificación de carencias, necesidades y expectativas, la segunda; selección del tipo de implementación, la tercera; determinación de requisitos de información, la cuarta; análisis de las necesidades del sistema, la quinta; diseño del sistema recomendado, la sexta;

desarrollo y documentación del sistema, la séptima; prueba y mantenimiento, y la octava; implementación y evaluación.

Cada fase se explica por separado pero nunca se realizan como pasos aislados, más bien es posible que algunas actividades se realicen de manera simultánea, y algunas de ellas podrían repetirse.

El modelo propuesto da un mayor énfasis a la primera parte de la implementación en la que se diseñaron las métricas para identificar de carencias, necesidades y expectativas de la organización tomada como caso de estudio.

Modelo de implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela

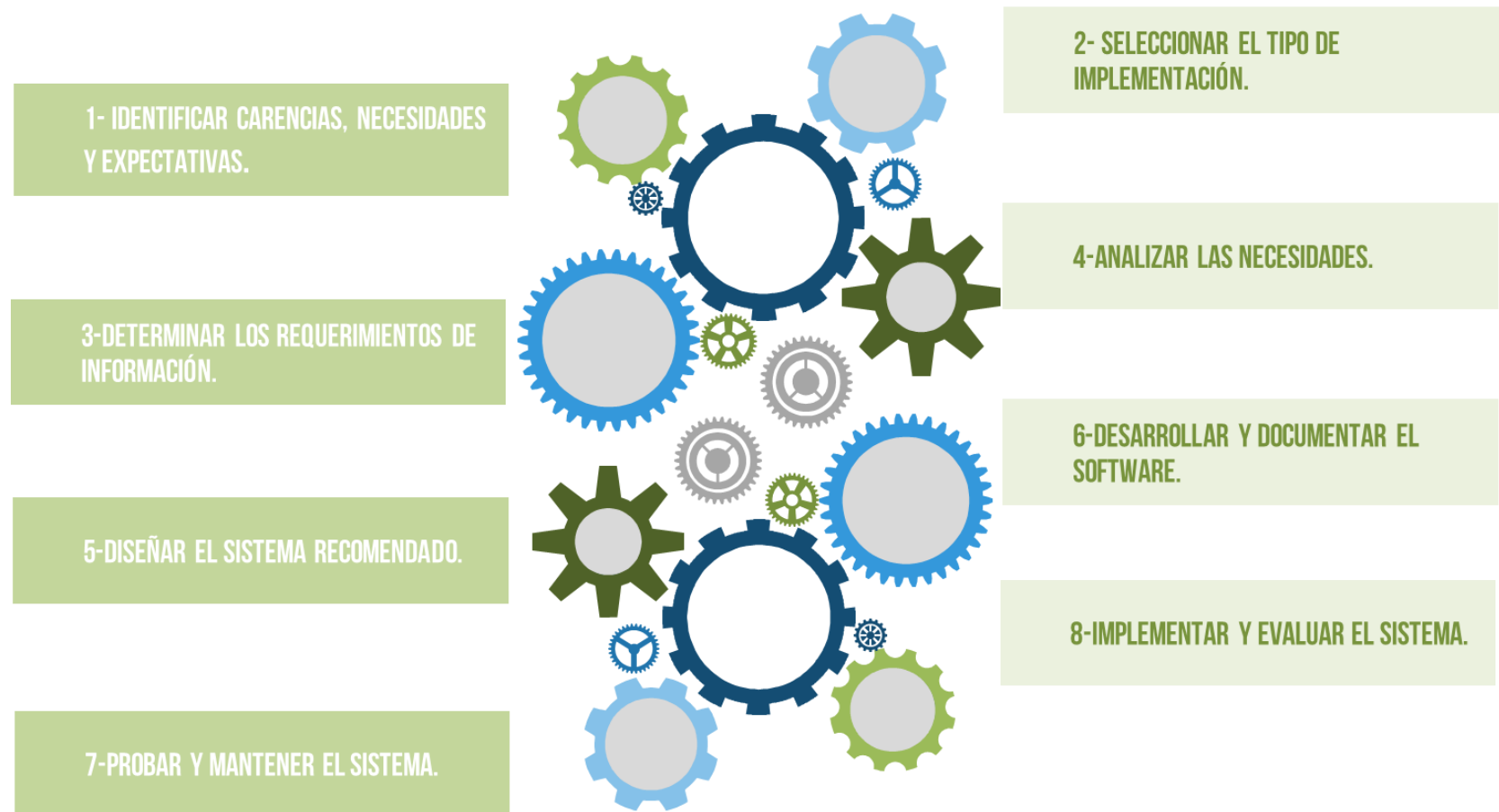


Figura 19.Características generales del modelo de implementación.

Modelo de implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela

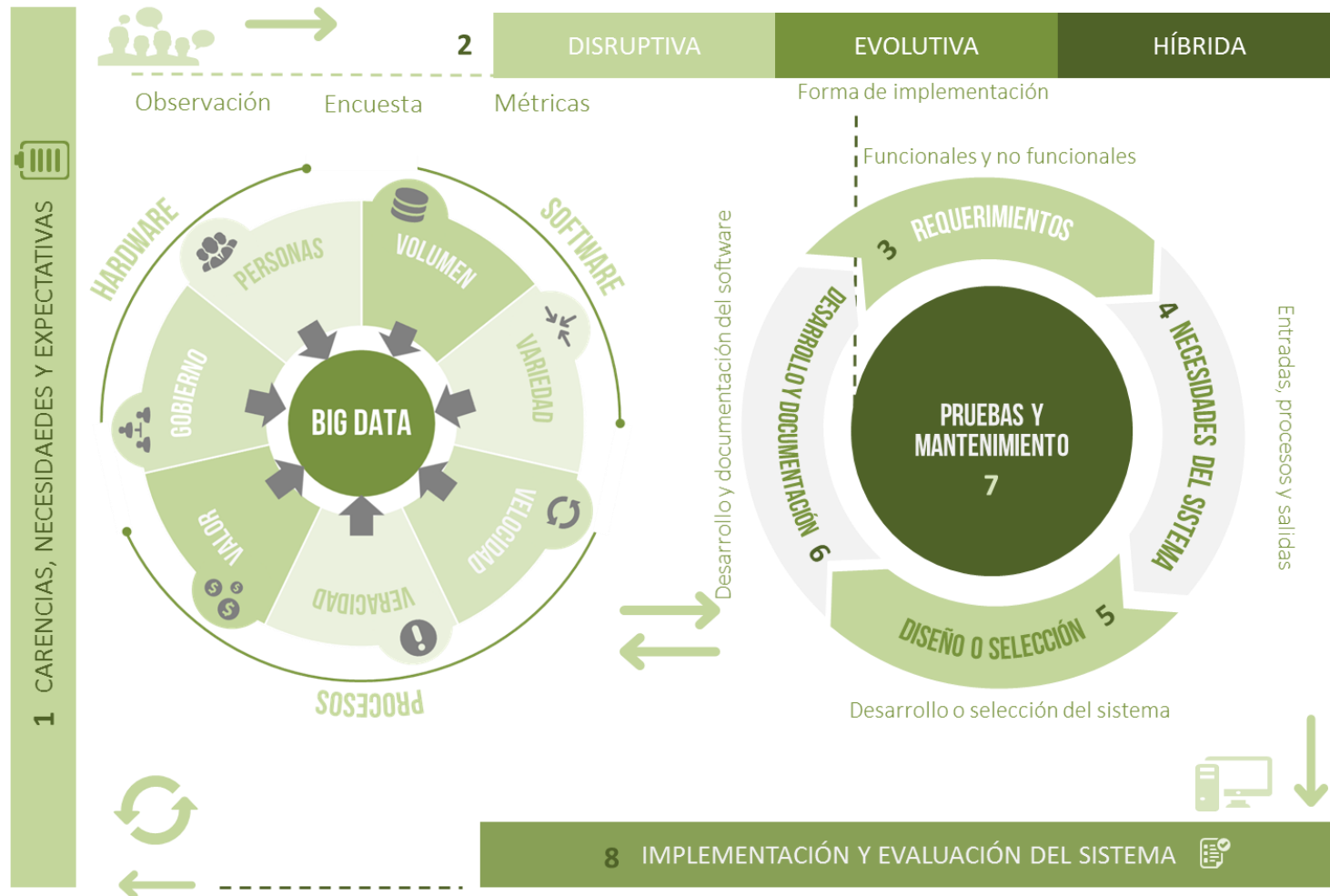


Figura 20. Características detalladas del modelo de implementación.

5.1.1 Identificar carencias, necesidades y expectativas.

En esta primera fase en la que se concentró el desarrollo de la propuesta, se deben identificar los problemas de la organización, detallarlos, examinar, evaluar las oportunidades y objetivos, esto incluye:

- Observación directa del entorno.
- Aplicación de entrevista para recolectar información.
- Sintetizar la información recolectada para construir objetivos.
- Estimar el alcance del proyecto.
- Identificar si existe una necesidad, problema u oportunidad argumentada.
- Documentar resultados.
- Estudiar los riesgos del proyecto.
- Presentar un informe de viabilidad.

Se debe identificar y evaluar los problemas existentes en la organización de manera crítica y precisa. Las oportunidades son situaciones consideradas susceptibles de mejorar utilizando sistemas de inteligencia de negocios con Big Data, lo cual le da mayor seguridad y eficacia a las organizaciones además de obtener una ventaja competitiva. Se debe identificar los objetivos, es decir, averiguar lo que la empresa trata de conseguir, se podrá determinar si algunas funciones de las aplicaciones de los sistemas de inteligencia de negocios con Big Data pueden contribuir a que el negocio alcance sus objetivos aplicándolas a problemas u oportunidades específicos. Los usuarios, los analistas y los administradores de sistemas que coordinan el proyecto son los involucrados en la primera fase. Las actividades de esta fase son las entrevistas a los encargados de coordinar a los usuarios, sintetizar el conocimiento obtenido, estimar el alcance del proyecto y documentar los resultados. El resultado de esta fase es un informe de viabilidad que incluye la definición del problema y un resumen de los objetivos. La administración debe decidir si se sigue adelante o si se cancela el proyecto propuesto.

Para alcanzar esta fase se identificaron dimensiones, atributos y subatributos que permiten caracterizar el entorno de inteligencia de negocios con Big Data asociado a la banca en línea del sistema financiero.

Los atributos a evaluar no hacen referencia a una dimensión en particular, ya que todas mantienen de alguna relación con el atributo y forman parte de un todo. El procedimiento a seguir durante esta fase, únicamente describe los objetivos que se deben cumplir para garantizar que se logre implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea. A continuación se define en que consiste cada una de las dimensiones identificadas:

Hardware, software y procesos

Hardware

Las entidades de hardware están relacionadas con los diversos puntos de partida a considerar para la implementación del sistema. Dentro de ella se encuentran los recursos de hardware requeridos para el procesamiento de los datos y la evaluación de las características tomadas en cuenta para la elaboración del modelo de implementación de un sistema de inteligencia de negocios con Big Data para la Banca en Línea.

Software

Las entidades de software están relacionadas con los sistemas informáticos y las características de software tomadas en cuenta para la elaboración del modelo de implementación de un sistema de inteligencia de negocios con Big Data para la Banca en Línea. La gestión de sistemas es crítica para realizar proyectos de Big Data, ya que implica muchos sistemas a través de clústers incluyendo gestión de los registros de los sistemas, máquinas virtuales, aplicaciones y otros dispositivos

Procesos

Las entidades de procesos incluyen actividades relacionadas con eventos de la organización que están asociadas a los tipos de atributos de hardware y software. Los

indicadores de procesos permiten a la organización tener una visión profunda de la eficacia de un proceso ya existente, en contraste con los requerimientos técnicos y humanos del sistema que se desea implantar.

Dentro de las entidades descritas (hardware, software y procesos) se distinguen un conjunto de atributos, los cuales contienen el conjunto de subatributos que pueden ser medidos examinando la cualidad del atributo en sí mismo o cómo este se relaciona con su entorno, permitiendo obtener las métricas para evaluar la implementación del sistema.

A continuación se define en que consiste cada atributo junto a sus subatributos:

Valor

Este atributo se refiere al valor que puede extraerse de los datos, teniendo como premisa que la gestión de datos debe estar acompañada y alineada con la estrategia de negocio. Para generar valor a través de la inteligencia de negocios mediante una solución de Big Data no existen directrices definitivas, los escenarios deben ser identificados por la organización, el objetivo final es generar valor de toda la información almacenada a través de distintos procesos de manera eficiente y con el coste más bajo posible. Este atributo sirve para identificar algunos de estos escenarios o casos de uso asociados al manejo de grandes volúmenes de datos de la banca en línea en la institución financiera.

Valor	Gestión de multicanalidad
	Segmentación avanzada de clientes
	Fidelización de clientes
	Análisis de riesgos
	Anonimizar los datos y compartir estadísticos

Figura 21. Atributo Valor.

Gestión de multicanalidad: Se refiere a la capacidad de poner a disposición de los clientes un mayor número de canales de comunicación mejorados y alta capacidad de transacción.

Segmentación avanzada de clientes: Se refiere a la capacidad de incorporar nuevas fuentes de datos a las segmentaciones tradicionales posibilitando la obtención de una visión y entendimiento más completo de los clientes, lo que conlleva a la definición de propuestas de valor adaptadas a los diferentes perfiles de clientes.

Fidelización de clientes: Se refiere a la capacidad de definir acciones para retención de clientes y en detectar abandono con suficiente antelación.

Análisis de riesgos: Se refiere a la capacidad de obtener una visión más amplia de los clientes, que permita no limitarse exclusivamente a su contabilidad e historia financiera.

Anonimizar los datos y compartir estadísticos: Es la capacidad de anonimizar los datos que se tienen de los clientes y compartir estadísticas con otros sectores, permitiendo generar nuevas oportunidades de negocios.

Volumen

Probablemente sea esta la característica principal con la que se asocia el término Big Data. Implica que se almacena y genera gran cantidad de datos, estos datos se producen en forma automática por los sistemas transaccionales de la banca en línea. Con este atributo se identifican las características en cuanto a volumen de datos del entorno de inteligencia de negocios de la banca en línea de la institución financiera

Volumen	Volumen de la información almacenada
	Almacén de datos
	Datos frío o de bajo contacto
	Desperdicio de datos
	Retos técnicos por volumen de datos

Figura 22. Atributo Volumen.

Volumen de la información almacenada: Se refiere al tamaño o volumen total de la información almacenada, descrita en unidades de almacenamiento de datos. (Bit, Byte, Kilobyte, Megabyte, Gigabyte, Terabyte, Petabyte, Exabyte, Zettabyte y Yottabyte.)

Almacén de datos: Se refiere a una colección de datos orientada a un determinado ámbito, integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la organización.

Datos frío o de bajo contacto: Se refiere a una cantidad de datos que son generalmente poco utilizados y no están siendo utilizados para derivar conocimiento del negocio.

Desperdicio de datos: Se refiere a una cantidad de datos que son eliminados debido a la incapacidad de almacenamiento o procesamiento.

Retos técnicos por volumen de datos: Se refiere a la existencia de retos técnicos y económicos para almacenar, buscar, compartir, analizar y visualizar datos utilizando los métodos tradicionales, como los motores de bases de datos relacionales.

Variedad

Este atributo se refiere a la diversidad de fuentes de las cuales proviene la información y el tipo de datos que se ingresa. Las nuevas fuentes de datos proporcionan nuevos y distintos tipos de datos, en diferentes formatos de información a los ya conocidos hasta ahora como datos estructurados, que un sistema de Inteligencia de Negocios con Big Data para la Banca en línea debe ser capaz de almacenar y procesar. Generar nuevos

productos y servicios en la banca en línea implica enfrentarse a estas nuevas fuentes de datos. Con este atributo se identifican las características en cuanto a variedad de datos del entorno de inteligencia de negocios de la banca en línea de la institución financiera.

Variedad	Tipos de datos
	Formato del contenido
	Origen de datos
	Predecir el tipo de datos
	Datos no incluidos en análisis de BI

Figura 23. Atributo Variedad.

Tipos de datos: Se refiere al tipo de datos que requiere ser procesado, como por ejemplo datos transaccionales, históricos, datos maestros y otros.

Formato del contenido: Se refiere al formato de los datos entrantes que pueden ser estructurados (RDMBS, por ejemplo), no estructurados (audio, video e imágenes, por ejemplo) o semiestructurados.

Origen de datos: Se refiere al origen de datos (donde se generan los datos) que pueden ser web, redes sociales, generados por máquina, generados por humanos, etc.

Predecir el tipo de datos: Se refiere a la propiedad de predecir o anticiparse al contenido y estructura de datos que debe ser procesado o analizado.

Datos no incluidos en análisis de BI: Se refiere a la existencia de una diversidad y volumen de datos que no han sido incluidos en los análisis de inteligencia de negocios por su complejidad y podrían generar valor.

Velocidad

El flujo de datos es masivo y continuo, por eso se requiere de sistemas que trabajan a toda máquina y sean capaces de procesar rápidamente los datos requeridos. El ritmo con que los datos de entrada fluyen desde las diversas fuentes en la banca en línea determinara la aplicabilidad de un sistema de inteligencia de negocios con Big Data. Con este atributo se identifican las características en cuanto a Velocidad de datos del entorno de inteligencia de negocios de la banca en línea de la institución financiera.

Velocidad	Velocidad de generación de datos
	Incremento en la velocidad de generación de datos
	Velocidad de los sistemas de BI
	Cambios de requerimiento de BI
	Satisfacción en los tiempos de BI

Figura 24. Atributo Velocidad.

Velocidad de generación de datos: Este atributo se refiere a la frecuencia con la cual los datos se ponen a disposición.

Incremento en la velocidad de generación de datos: Se refiere al incremento en la frecuencia con la cual los datos se ponen a disposición.

Velocidad de los sistemas de BI: Se refiere a la velocidad del sistema utilizado y es lo que determina la facilidad de acceso a la información, mientras más oportuno sea mayor es la posibilidad de entregar información valiosa, en el momento oportuno.

Cambios de requerimiento de BI: Se refiere a los cambios generados en los requerimientos de análisis de datos

Satisfacción en los tiempos de BI: Se refiere al grado de satisfacción de la organización con respecto a los tiempos de generación de reportes para hacer inteligencia de negocios.

Veracidad

Este atributo en este contexto, se refiere al sesgo, ruido y alteración de los datos. Son los responsables de la tecnología Big Data quienes tienen la responsabilidad de discriminar entre la información que es relevante para el problema en cuestión y la que no lo es. Una solución de inteligencia de negocios con Big Data para la Banca en línea podría ser adecuada si se identifica una complejidad razonable con la veracidad de los datos (junto con una complejidad en el volumen, variedad y velocidad de los mismos). Con este atributo se identifican las características en cuanto a complejidad en la Veracidad de datos del entorno de inteligencia de negocios de la banca en línea de la institución financiera.

Veracidad	La autenticidad o exactitud de los datos
	Ambigüedad en los datos
	Complejidad de los datos
	Características ACID
	Características del teorema de CAP

Figura 25. Atributo Veracidad.

La autenticidad o exactitud de los datos: Se refiere a la confiabilidad, autenticidad y exactitud del contenido de los datos que se almacenan y procesan.

Ambigüedad en los datos: Se refiere al carácter ambiguo (que pueda interpretarse de distintas maneras) de los datos que se almacenan y procesan.

Complejidad de los datos: Se refiere a la existencia de datos completos o por el contrario datos con información incompleta.

Características ACID: Se refiere a la necesidad de implementar un sistema de inteligencia de negocios cumpliendo características de atomicidad, consistencia, aislamiento y durabilidad

Características del teorema de CAP: Se refiere a la capacidad de implementar un sistema distribuido cumpliendo los requerimientos de BI que satisfaga solo dos de los siguientes atributos al mismo tiempo: Consistencia Fuerte, Alta Disponibilidad y Tolerancia a Particionamiento.

Gobierno

El gobierno de datos se trata de definir las directrices que ayuden a las empresas a tomar las decisiones correctas acerca de los datos. El gobierno de datos es de gran ayuda en el tratamiento de la complejidad, volumen y variedad de datos que están dentro de la banca en línea. Se requieren directrices y procesos fuertes para supervisar, estructura, almacenar y proteger los datos desde el momento en que entra en la empresa, se procesa, almacena, analizados y purgados o archivadas. Con este atributo se identifican las características en cuanto a Gobierno de datos del entorno de inteligencia de negocios de la banca en línea de la institución financiera.

Gobierno

Seguridad y privacidad

Normalización de los datos

Tiempo en el cual los datos están disponibles

Propiedad de los datos

Usos permitidos

Figura 26. Atributo Gobierno.

Seguridad y privacidad: Se refiere al cumplimiento de características de seguridad, privacidad y regulaciones de los datos que serán usados para hacer inteligencia de negocios.

Normalización de los datos: Se refiere a la existencia de normas que rigen los datos.

Tiempo en el cual los datos están disponibles: Se refiere a la disponibilidad de los datos para realizar los procesos de inteligencia de negocios.

Propiedad de los datos: Se refiere a quien es el dueño de los datos dentro de la organización, accesos y permisos requeridos para utilizarlos.

Usos permitidos: Se refiere al uso que puede darse a los datos disponibles.

Personas

Se requieren habilidades específicas para comprender y analizar los requisitos, y para mantener una solución de grandes volúmenes de datos. Estas habilidades incluyen conocimiento del sector, experiencia en el campo, y el conocimiento técnico sobre herramientas y tecnologías de Big Data. Científicos o especialistas en análisis de datos con experiencia en el modelaje, estadísticas, análisis y matemáticas son la clave para el éxito de cualquier iniciativa de Big Data dentro de la Banca en línea. Con este atributo se identifican las características en cuanto a Personas o recursos humanos del entorno de inteligencia de negocios de la banca en línea de la institución financiera.

Personas

Apoyo de la alta gerencia

Científicos o especialistas de datos

Personal con conocimientos técnicos en Big Data

Consumidores de datos

Tipo de procesamiento requerido

Figura 27. Atributo Personas.

Apoyo de la alta gerencia: Se refiere aceptación y apoyo por parte de los interesados y la alta gerencia para invertir en el proyecto

Científicos o especialistas de datos: Se refiere a la existencia de científicos o especialistas de datos disponibles para generar conocimientos y valor significativo a partir de los datos.

Personal con conocimientos técnicos en Big Data: Se refiere a la existencia de personal con conocimiento técnico del personal sobre herramientas y tecnologías de Big Data. (O personal que puede ser capacitado)

Consumidores de datos: Se refiere a todos los posibles consumidores de los datos procesados.

Tipo de procesamiento requerido: Se refiere al tipo de procesamiento requerido por parte del personal para realzar inteligencia de negocios. El procesamiento puede ser realizado en tiempo real, casi en tiempo real o en modalidad de procesamiento por lotes.

5.1.2 Selección del tipo de implementación de Big Data

Esta fase es útil para decidir el tipo de implementación de Big Data que desea la organización y con cual se siente más identificada, no existe una única manera para hacer una implementación de un sistema de inteligencia de negocios con Big Data, esto dependerá de:

- La madurez la inteligencia de negocio de la empresa.
- La infraestructura de que la organización disponga para la implementación.
- Su nivel de conocimiento del nuevo entorno, su expertise (data science).
- Su planteamiento, más innovador o más tradicional, que condicionará inequívocamente su elección.

En cualquier caso generalmente las empresas, atendiendo a su configuración y dependiendo de los factores precitados se agrupan atendiendo a tres modelos diferentes de implementación de Big Data:

- Disruptivo
- Evolutivo
- Híbrido

La implementación Disruptiva

Quienes optan por esta forma de implementar Big Data han decidido romper con todo y empezar de cero. Todas las empresas ya están funcionando, han generado datos y los han ido guardando y procesando a su manera. Hasta su primer contacto con Big Data han tomado decisiones en función de datos, pero siempre con un alto

componente de incertidumbre y una proporción de hechos que actualmente pueden parecer arriesgados.

Esta implementación requiere mover todos los datos disponibles al nuevo entorno. Y, junto con ellos, los informes, el modelado y la integración con los procesos de negocio. Desde el primer momento todo comenzará de cero en esta nueva plataforma.

Existen tres tecnologías que permiten llevarlo a cabo: hadoop, las bases de datos paralelas y las bases de datos in memory, para análisis en tiempo real. Mediante su táctica de "divide y vencerás" permiten trabajar con volúmenes inmensos de información en poco tiempo y a bajo coste y que se complementan con la posibilidad de moverlo todo al cloud.

Ventajas de este tipo de implementación:

- Agilidad del procesamiento de la información: ya que cada porción se procesa de forma mucho más rápida y lo que antes se procesaba en 6 horas, ahora se hace en minutos.
- Optimización de los recursos que implica una notable reducción de costes en comparación a otros métodos.
- Flexibilidad de uso gracias al sistema de nodos: que permite además poder calcular lo que se va a gastar y pagar sólo por lo que se usa.

Inconvenientes:

- Altos costos en lo referente a habilidades a lo que se sumaría la dificultad de encontrar candidatos con este perfil en el mercado, profesionales que conozcan y sepan manejar adecuadamente esta tecnología, sacándole todo el partido.
- Quizás también puede suponer un mayor tiempo de adaptación hasta que se pone en marcha, si se compara con otros sistemas, ya que se trata de un método totalmente nuevo.

La implementación Evolutiva

El método evolutivo es otra forma de acceder a este mundo que suele elegirse por empresas, quizás no tan pioneras, pero que ya contaban con un BI bastante maduro. Se habla de organizaciones que disponen de su data warehouse, su herramienta de visualización y reporting y que llevan años analizando datos, a la manera antigua pero con madurez.

Decantarse por el método evolutivo de implementación de Big Data supone que, manteniendo la estructura actual, simplemente se suman los datos que no tenían cabida en el sistema para pre-procesarlos desde la plataforma de Big Data. De esta forma, el sistema actual los puede ver y así se logran análisis, que siguen siendo del mismo tipo o parecidos, pero que ahora incorporan más datos.

La ventaja que tiene esta opción es que su umbral de entrada tiene un costo menor, ya que las empresas todavía pueden seguir empleando su herramienta, aunque de la extracción de datos y su estructuración se encargue Big Data. En concreto, Big Data se convierte en una entrada a la plataforma de BI existente. Los datos se acumulan y analizan, y los resultados se envían al data warehouse.

Ventajas de este tipo de implementación:

- Rendimiento. Los modelos ahora se nutren con muchos más datos y más diversificados.
- Volumen. Se multiplica el volumen de información que se obtiene, gracias a Big Data.
- Ahorro. Tanto en tiempo de implantación, como en costes, ya que se mantiene el BI existente.

Inconvenientes:

- La velocidad de extremo a extremo siempre estará limitada por el entorno actual de BI (que es mucho más lento).
- El nivel de percepción no es tan amplio, ya que fallaría la granularidad por el BI existente.

- Esta solución no es definitiva, porque llegará un momento en que el BI no tenga capacidad para hacer frente a los requerimientos informativos de la organización y se tenga que optar por hacer el traslado completo de los datos al exterior de la empresa.

La implementación Híbrida

Este sistema alterna el uso de una y otra tecnología en función del objetivo perseguido. Para determinado tipo de información, de análisis o de usuario se seguiría usando la BI existente, mientras que para análisis mucho más refinados, como los de tipo predictivo, simulaciones, etc. se usaría Big Data. Para lograrlo, bastaría con establecer un par de puntos de integración, consiguiendo que los datos de data warehouse se almacenen en el motor de análisis, lo que sería visto por el éste como un DataMart.

En esta opción Big Data no sólo es un anexo al BI tradicional que permite ahora ver datos que antes no podía, sino que además es una plataforma que sirve para hacer análisis avanzado mezclando datos del tradicional que siguen allí, con los nuevos como por ejemplo Redes Sociales o los datos No-Estructurados que los BI de antes no contemplaban a la hora de trabajar.

Ventajas de este tipo de implementación:

- Precisión: los sistemas ahora se nutren con una información estructurada, cuando antes sólo podían acceder a datos no estructurados.
- Big Analytics. posibilidad de hacer un análisis predictivo, creando un modelo estadístico con todos los datos e identificando las relaciones causales y las correlaciones; apoyándonos también en herramientas avanzadas de visualización.
- El internet de las cosas: que permite conocer en tiempo real todo lo que está sucediendo en cualquier parte y en relación a cualquier asunto.
- Rentabilidad. En definitiva, esta opción permite ofrecer mejores servicios que, claro está, también pueden cobrarse.

Inconvenientes:

- Resistencia: de las empresas menos evolucionadas, tecnológicamente hablando, a la implantación de un modelo de este tipo.
- Dificultad: de quienes están acostumbrados a limitarse a un tipo de análisis simplemente descriptivo, a la hora de lanzarse a las nuevas posibilidades.

5.1.3 Determinar de los requerimientos de información.

- Revisión de los objetivos.
- Identificar el dominio.
- Investigar la razón por la cual se implementa el sistema actual.
- Recolectar información sobre los procedimientos y operaciones que se desempeñan actualmente.
- Detallar específicamente quiénes son los involucrados, cuál es la actividad, regla y restricciones del negocio, entorno de desarrollo de las actividades, momentos oportunos de desarrollo de cada función, la manera en que se desempeñan los procedimientos actuales.
- Elaborar una lista detallada y organizada de todos los procedimientos.
- Separar requerimientos funcionales y no funcionales. Adicionar al informe de la primera fase, esta nueva información.

Esta fase es útil para confirmar la idea que tiene de la organización y sus objetivos.

Los implicados en esta fase son el analista y los usuarios, por lo general los trabajadores y gerentes del área de operaciones.

Es necesario conocer los detalles de las funciones del sistema actual:

El quién (la gente involucrada), el qué (la actividad del negocio), el dónde (el entorno donde se desarrollan las actividades), el cuándo (el momento oportuno) y el cómo (la manera en que se realizan los procedimientos actuales) del negocio que se estudia.

Al término de esta fase, se debe conocer el funcionamiento del negocio y poseer información muy completa acerca de la gente, los objetivos, los datos y los procedimientos implicados.

5.1.4 Analizar las necesidades.

- Evaluar las dos fases anteriores.
- Modelar las entradas, los procesos y las salidas de las funciones ya identificadas.
- Elaborar diccionario de datos y sus especificaciones.
- Elaborar diagramas de procesos de cada función.
- Elaborar propuesta del sistema con todos los diagramas de operaciones y de procesos.
- Realizar el análisis del riesgo sobre el realizado en las fases anteriores, tomando en cuenta el aspecto económico, técnico y operacional (estudio de factibilidad)
- Estimar en un diagrama de Gantt el tiempo que tomará desarrollar el sistema.

En esta fase se evalúan las dos fases anteriores, usa herramientas y técnicas como el uso de diagramas de flujo de datos para graficar las entradas, los procesos y las salidas de las funciones del negocio en una forma gráfica estructurada.

A partir de los diagramas de flujo de datos se desarrolla un diccionario de datos que enlista todos los datos utilizados en el sistema así como sus respectivas especificaciones.

El analista prepara en esta fase, una propuesta de sistemas que sintetiza sus hallazgos, proporciona un análisis de costo/beneficio de las alternativas y ofrece, en su caso, recomendaciones sobre lo que se debe hacer.

5.1.5 Diseñar el sistema recomendado.

- Evaluar las tres fases anteriores.
- Realizar el diseño lógico de todo el sistema.

- Elaborar procedimientos precisos para la captura de los datos que van a ingresar al sistema de información.
- Elaborar el diseño de la base de datos.
- Diseñar las diferentes interfaces de usuarios de cada operación, procedimiento y/o función.
- Diseñar controles y procedimientos de respaldos que protejan al sistema y a los datos.
- Producir los paquetes específicos de programas para los programadores.
- Elaborar una lista de las funciones genéricas y de las que será obligatorio crear.

En esta fase se utiliza la información recopilada en las primeras fases para realizar el diseño lógico del sistema.

Se diseñan lo procedimientos precisos para la captura de datos que aseguran que los datos que ingresen al sistema de información sean correctos.

Facilita la entrada eficiente de datos al sistema de información mediante técnicas adecuadas de diseño de formularios y pantallas.

La concepción de la interfaz de usuario forma parte del diseño lógico del sistema de información.

La interfaz conecta al usuario con el sistema y por tanto es sumamente importante.

También incluye el diseño de archivos o bases de datos que almacenarán gran parte de los datos indispensables para los encargados de tomar las decisiones en la organización.

En esta fase se interactúa con los usuarios para diseñar la salida (en pantalla o impresa) que satisfaga las necesidades de información de estos últimos.

Finalmente se debe diseñar controles y procedimientos de respaldo que protejan al sistema y a los datos y producir paquetes de especificaciones de programa para los programadores.

Cada paquete debe contener esquemas para la entrada y la salida, especificaciones de archivos y detalles del procesamiento.

5.1.6 Desarrollar y documentar software.

- Evaluar los procedimientos que va a ser desarrollados por el programador.
- Mostrar y explicar cada procedimiento, función y operación al programador.
- Elaborar manuales de procedimientos internos del sistema.
- Elaborar manuales externos de ayuda a los usuarios del sistema.
- Elaborar demostraciones para los usuarios y la interacción con distintas interfaces.
- Elaborar actualizaciones para los diferentes procedimientos. Elaborar un informe con el tiempo que se llevó construir cada procedimiento.

En la quinta fase del ciclo del desarrollo de sistemas, el analista trabaja de manera conjunta con los programadores para desarrollar cualquier software original necesario. Entre las técnicas estructuradas para diseñar y documentar software se encuentran los diagramas de estructuras, los diagramas de Nassi-Shneiderman y el pseudocódigo.

Durante esta fase el analista trabaja con los usuarios para desarrollar documentación efectiva para el software, como manuales de procedimientos, ayuda en línea y sitios web que incluyan respuestas a preguntas frecuentes en archivos “léame” que se integrarán al nuevo software.

La documentación indica a los usuarios cómo utilizar el sistema y qué hacer en caso de que surjan problemas derivados de este uso. Los programadores desempeñan un rol clave en esta fase porque diseñan, codifican y eliminan errores sintácticos de los programas de cómputo.

5.1.7 Probar y mantener el sistema.

- Realizar la programación de las pruebas del sistema.
- Realizar un instrumento para evaluar el sistema de información.
- El programador deberá elaborar un resumen de las pruebas del sistema.
- El analista deberá realizar un informe de sus pruebas y discutirlo con el programador.

- Elaborar la planificación de las horas del mantenimiento del sistema. Elaborar la lista de las operaciones que pudieran sufrir modificaciones de códigos.

Antes de poner en funcionamiento el sistema es necesario probarlo es mucho menos costoso encontrar los problemas antes que el sistema se entregue a los usuarios.

Una parte de la pruebas la realizan los programadores solos, y otra la llevan a cabo de manera conjunta con los analistas de sistemas. Primero se realizan las pruebas con datos de muestra para determinar con precisión cuáles son los problemas y posteriormente se realiza otra con datos reales del sistema actual.

El mantenimiento del sistema de información y su documentación empiezan en esta fase y se llevan de manera rutinaria durante toda su vida útil.

5.1.8 Implementar y evaluar el sistema.

- Planificar gradualmente la conversión del sistema anterior.
- Instalar los equipos de hardware necesarios para el funcionamiento del software creado.
- Capacitar por medio de talleres a los usuarios en el manejo de equipos y software creados.
- Evaluar la adaptabilidad de los usuarios al sistema.

En esta fase se capacita a los usuarios en el manejo del sistema. Parte de la capacitación la imparten los fabricantes, pero la supervisión de ésta es responsabilidad del analista de sistemas.

Se menciona la evaluación como la fase final del ciclo de vida del desarrollo de sistemas principalmente en áreas del debate. En realidad, la evaluación se lleva a cabo durante cada una de las fases.

El trabajo de sistemas es cíclico, cuando un analista termina una fase del desarrollo de sistemas y pasa a la siguiente, el surgimiento de un problema podría obligar a regresar a la fase previa y modificar el trabajo realizado.

5.2 Métricas

Una vez identificadas cada una de las dimensiones y atributos del modelo propuesto en la primera fase o etapa de la implementación, se plantea el escalamiento tipo Likert para obtener las puntuaciones, que según Hernández (2010), consiste en un conjunto de ítems presentados para en forma de afirmaciones o juicios, en este caso en forma de preguntas para cada subatributo, ante los cuales se pide o se mide la reacción de los participantes. Es decir, se presenta cada afirmación y se solicita al sujeto que extreme su reacción eligiendo uno de cinco puntos o categorías de la escala. A cada punto se le asigna un valor numérico. Así, el participante obtiene una puntuación para cada respuesta y al final su puntuación total, sumando las puntuaciones en relación a cada atributo.

Se plantean a continuación las métricas que definen la medida en la que una organización cumple con el modelo propuesto. La matriz del modelo de implementación integrada con los atributos, subatributos, métricas e indicadores puede ser consultada en el anexo VI. De igual forma en el siguiente capítulo será aplicada a un caso de estudio.

Las métricas del modelo de implementación de un sistema de inteligencia de negocios con Big data para la Banca en línea definen la medida en la que una organización cumple con el modelo propuesto, enfocado en la primera parte de la implementación para identificar carencias, necesidades y expectativas de la organización. Cada subatributo tendrá asignado una métrica o método de medición, a través del cual se indicará si la organización financiera cumple o no con el mismo.

Las métricas del modelo propuesto son las siguientes:

Tabla 11. Métricas del atributo Valor del Modelo Propuesto.

Valor				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Gestión de multicanalidad	¿Es la gestión de multicanalidad una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	Validar con gerentes del área de inteligencia de negocios la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2 = Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Segmentación avanzada de clientes	¿Es la segmentación avanzada de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	Validar con gerentes del área de inteligencia de negocios la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2 = Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Fidelización de clientes	¿Es la fidelización de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	Validar con gerentes del área de inteligencia de negocios la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2 = Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Análisis de riesgos	¿Es el análisis de riesgos una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	Validar con gerentes del área de inteligencia de negocios la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2 = Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Anonimizar los datos y compartir estadísticos	¿Anonimizar los datos que se tienen de los clientes y compartir estadísticas con otros sectores, permitiría generar nuevas oportunidades de negocios?	Validar con gerentes del área de inteligencia de negocios la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2 = Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 12. Métricas del atributo Volumen del Modelo Propuesto.

Volumen				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Volumen de la información almacenada	¿El volumen o tamaño de las bases de datos actuales se dimensiona de petabytes a exabytes, y en un futuro próximo, podría crecer hasta zettabytes?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Almacén de datos	¿Existe un entorno de almacén de datos que contiene un repositorio de todos los datos generados o adquiridos?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Datos frío o de bajo contacto	¿Hay una cantidad significativa de datos de frío o de bajo contacto que no está siendo analizada para derivar el conocimiento del negocio?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Desperdicio de datos	¿Se eliminan cantidades de datos debido a la incapacidad de procesamiento o procesamiento?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Retos técnicos por volumen de datos	¿Se plantean retos técnicos y económicos para almacenar, buscar, compartir, analizar y visualizar datos utilizando los métodos tradicionales debido al volumen de datos?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 13. Métricas del atributo Variedad del Modelo Propuesto.

Variedad				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Tipos de datos	¿El tipo de datos que requiere ser procesado es transaccional o histórico?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Formato del contenido	¿El formato de datos puede variar, incluyendo datos estructurados, semiestructurados y no estructurados?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Origen de datos	¿Los datos son generados por sistemas transaccionales?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Predecir el tipo de datos	¿El contenido y la estructura de datos a ser procesado no se pueden anticipar o predecir?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Datos no incluidos en análisis de BI	¿Existe una gran variedad de datos que nunca han sido incluidos para los análisis de inteligencia de negocios por su complejidad y que podrían generar valor?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 14. Métricas del atributo Velocidad del Modelo Propuesto.

Velocidad				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Velocidad de generación de datos	¿La velocidad de generación de datos es grande?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Incremento en la velocidad de generación de datos	¿La velocidad de generación de datos tiende a incrementar en el tiempo?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Velocidad de los sistemas de BI	¿Están las tecnologías y los métodos tradicionales de BI abrumados, ya no son adecuados para manejar datos que lleguen en tiempo requerido?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Cambios de requerimiento de BI	¿Está cambiando rápidamente y debe ser respondido de inmediato?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Satisfacción en los tiempos de BI	¿Está satisfecha la organización con los tiempos de generación de reportes para hacer inteligencia de negocios?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 15. Métricas del atributo Veracidad del Modelo Propuesto.

Veracidad				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
La autenticidad o exactitud de los datos	¿La autenticidad o exactitud de los datos es desconocida?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Ambigüedad en los datos	¿Los datos incluyen información ambigua?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Compleitud de los datos	¿No está claro si los datos están completos?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Características ACID	¿No es necesario que el sistema de BI cumpla con las características ACID?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Características del teorema de CAP	¿Cumpliría un sistema distribuido con los requerimientos de BI que satisfaga solo dos de los siguientes atributos al mismo tiempo: Consistencia Fuerte, Alta Disponibilidad y Tolerancia a Particionamiento?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 16. Métricas del atributo Gobierno del Modelo Propuesto.

Gobierno				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Seguridad y privacidad	¿Se tiene claro que datos se pueden almacenar y tener acceso? ¿Qué datos deben ser cifrados durante el movimiento o en reposo? ¿Quién está autorizado para ver los datos en bruto y los puntos de vista?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Normalización de los datos	¿Existen normas que rigen los datos?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Tiempo en el cual los datos están disponibles	¿Están los datos disponibles en un plazo que permita medidas que deben tomarse en el momento oportuno?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Propiedad de los datos	¿La solución tiene acceso y permiso apropiado utilizar los datos?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Usos permitidos	¿Está permitido el uso de los datos disponibles para la solución de inteligencia de negocios con Big Data planteada?	Validar con gerentes del área de Base de Datos la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

Tabla 17. Métricas del atributo Personas del Modelo Propuesto.

Personas				
Subatributo	Medición	Método de Aplicación	Indicador	Puntuación
Apoyo de la alta gerencia	¿Existe aceptación por parte de los interesados y la alta gerencia, y están dispuestos a invertir en el proyecto?	Validar con gerentes del área de Inteligencia de Negocio la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Científicos o especialistas de datos	¿Tiene científicos o especialistas de datos disponibles para generar conocimientos y valor significativo a partir de los datos?	Validar con gerentes del área de Inteligencia de Negocio la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Personal con conocimientos técnicos en Big Data	¿Existe personal con conocimiento técnico sobre herramientas y tecnologías de Big Data, o que puede ser capacitado?	Validar con gerentes del área de Inteligencia de Negocio la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Consumidores de datos	¿Se tiene identificado a todos los posibles consumidores de los datos procesados?	Validar con gerentes del área de Inteligencia de Negocio la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0
Tipo de procesamiento requerido	¿El procesamiento debe ser realizado en tiempo real, casi en tiempo real o en modalidad de procesamiento por lotes?	Validar con gerentes del área de Inteligencia de Negocio la respuesta	☐ 4 = Siempre ☐ 3 = Casi siempre ☐ 2= Algunas Veces ☐ 1 = Casi nunca ☐ 0 = Nunca	☐ 2.85 ☐ 2.14 ☐ 1.42 ☐ 0.71 ☐ 0

CAPÍTULO VI: EVALUACIÓN DE LA PROPUESTA

El presente capítulo se lleva a cabo con la finalidad de validar el modelo propuesto en el presente trabajo de investigación, concentrado en la primera fase del modelo propuesto, la cual consiste en identificar carencias necesidades y expectativas, a través de su aplicación a un caso de estudio, específicamente en el Banco Venezolano de Crédito, el cual fue seleccionado como muestra.

A continuación la matriz utilizada, así como el análisis, conclusión y recomendaciones para cada caso:

6.1 Aplicación del modelo de implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea

El enfoque utilizado consiste en responder preguntas dicotómicas que exploran siete atributos, (a través de cinco subatributos en cada uno), que asociados a lo que sabe la organización sobre su propio entorno, permitirá determinar porcentualmente el grado en que una solución de Big Data es apropiada o no.

Los atributos y subatributos evaluados con la aplicación del modelo, permiten apreciar la viabilidad de implementar un sistema de inteligencia de negocios con Big Data para la Banca en línea, específicamente en el Banco Venezolano de Crédito, entidad financiera tomada como muestra intencional.

A continuación los resultados de la aplicación del Modelo:

DATOS DE LA ENTIDAD FINANCIERA	
Nombre de la organización	Banco Venezolano de Crédito
DATOS DEL EVALUADOR	
Nombre del evaluador	William Bastidas
Cargo en la organización:	Analista de Sistemas
DATOS DE LA EVALUACION	
Fecha de evaluación:	09/06/2016
Áreas involucradas	Gerencia de Administradores de datos y Gerencia de Mercadeo
Responsable del área:	Lic. María Isabel Rodríguez y Lic. Berioska Rodríguez

Modelo de Inteligencia de Negocio con Big Data

[Ir a Métricas»](#)

Instrucciones:

1. Ingrese en la pestaña "Métricas" a través del siguiente enlace y coloque en cada ítem sus valores según la tabla a continuación
2. Observe los indicadores para cada atributo de acuerdo a las métricas introducidas.
3. Evalúe los resultados obtenidos en la pestaña "Resultados".



ítems				
Siempre	Casi siempre	Algunas veces	Muy pocas Veces	Nunca
4	3	2	1	0

Figura 28. Instrucciones para la aplicación del modelo.

Valor				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Gestión de multicanalidad	¿Es la gestión de multicanalidad una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	4	2.86%	2.86%
Segmentación avanzada de clientes	¿Es la segmentación avanzada de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	4	2.86%	2.86%
Fidelización de clientes	¿Es la fidelización de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	4	2.86%	2.86%
Análisis de riesgos	¿Es el análisis de riesgos una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?	3	2.14%	2.86%
Anonimizar los datos y compartir estadísticos	¿Anonimizar los datos que se tienen de los clientes y compartir estadísticas con otros sectores, permitiría generar nuevas oportunidades de negocios?	2	1.43%	2.86%
			12.14%	14.29%

Figura 29. Aplicación del Modelo atributo Valor.

Volumen				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Volumen de la información almacenada	¿El volumen o tamaño de las bases de datos actuales se dimensiona de petabytes a exabytes, y en un futuro próximo, podría crecer hasta zettabytes?	4	2.86%	2.86%
Almacén de datos	¿Existe un entorno de almacén de datos que contiene un repositorio de todos los datos generados o adquiridos?	4	2.86%	2.86%
Datos frío o de bajo contacto	¿Hay una cantidad significativa de datos de frío o de bajo contacto que no está siendo analizada para derivar el conocimiento del negocio?	4	2.86%	2.86%
Desperdicio de datos	¿Se eliminan cantidades de datos debido a la incapacidad de procesamiento o procesamiento?	2	1.43%	2.86%
Retos técnicos por volumen de datos	¿Se plantean retos técnicos y económicos para almacenar, buscar, compartir, analizar y visualizar datos utilizando los métodos tradicionales debido al volumen de datos?	3	2.14%	2.86%
			12.14%	14.29%

Figura 30. Aplicación del Modelo atributo Volumen.

Variedad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Tipos de datos	¿El tipo de datos que requiere ser procesado es transaccional o histórico?	4	2.86%	2.86%
Formato del contenido	¿El formato de datos puede variar, incluyendo datos estructurados, semiestructurados y no estructurados?	3	2.14%	2.86%
Origen de datos	¿Los datos son generados por sistemas transaccionales?	4	2.86%	2.86%
Predecir el tipo de datos	¿El contenido y la estructura de datos a ser procesado no se pueden anticipar o predecir?	2	1.43%	2.86%
Datos no incluidos en análisis de BI	¿Existe una gran variedad de datos que nunca han sido incluidos para los análisis de inteligencia de negocios por su complejidad y que podrían generar valor?	3	2.14%	2.86%
			11.43%	14.29%

Figura 31. Aplicación del Modelo atributo Variedad.

Velocidad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Velocidad de generación de datos	¿La velocidad de generación de datos es grande?	4	2.86%	2.86%
Incremento en la velocidad de genera	¿La velocidad de generación de datos tiende a incrementar en el tiempo?	4	2.86%	2.86%
Velocidad de los sistemas de BI	¿Están las tecnologías y los métodos tradicionales de BI abrumados, ya no son adecuados para manejar datos que lleguen en tiempo requerido?	3	2.14%	2.86%
Cambios de requerimiento de BI	¿BI Está cambiando rápidamente y debe ser respondido de inmediato?	4	2.86%	2.86%
Satisfacción en los tiempos de BI	¿No está satisfecha la organización con los tiempos de generación de reportes para hacer inteligencia de negocios?	3	2.14%	2.86%
			12.86%	14.29%

Figura 32. Aplicación del Modelo atributo Velocidad.

Veracidad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
La autenticidad o exactitud de los datos	¿La autenticidad o exactitud de los datos es desconocida?	2	1.43%	2.86%
Ambigüedad en los datos	¿Los datos incluyen información ambigua?	2	1.43%	2.86%
Compleitud de los datos	¿No está claro si los datos están completos?	2	1.43%	2.86%
Características ACID	¿No es necesario que el sistema de BI cumpla con las características ACID?	1	0.71%	2.86%
Características del teorema de CAP	¿Cumpliría un sistema distribuido con los requerimientos de BI que satisfaga solo dos de los siguientes atributos al mismo tiempo: Consistencia Fuerte, Alta Disponibilidad y Tolerancia a Particionamiento?	2	1.43%	2.86%
			6.43%	14.29%

Figura 33. Aplicación del Modelo atributo Veracidad.

Gobierno				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Seguridad y privacidad	¿Se tiene claro que datos se pueden almacenar y tener acceso? ¿Qué datos deben ser cifrados durante el movimiento o en reposo? ¿Quién está autorizado para ver los datos en bruto y los puntos de vista?	4	2.86%	2.86%
Normalización de los datos	¿Existen normas que rigen los datos?	4	2.86%	2.86%
Tiempo en el cual los datos están disponibles	¿Están los datos disponibles en un plazo que permita medidas que deben tomarse en el momento oportuno?	3	2.14%	2.86%
Propiedad de los datos	¿La solución tiene acceso y permiso apropiado utilizar los datos?	4	2.86%	2.86%
Usos permitidos	¿Está permitido el uso de los datos disponibles para la solución de inteligencia de negocios con Big Data planteada?	4	2.86%	2.86%
			13.57%	14.29%

Figura 34. Aplicación del Modelo atributo Gobierno.

Personas				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Apoyo de la alta gerencia	¿Existe aceptación por parte de los interesados y la alta gerencia, y están dispuestos a invertir en el proyecto?	3	2.14%	2.86%
Científicos o especialistas de datos	¿Tiene científicos o especialistas de datos disponibles para generar conocimientos y valor significativo a partir de los datos?	3	2.14%	2.86%
Personal con conocimientos técnicos en Big Data	¿Existe personal con conocimiento técnico sobre herramientas y tecnologías de Big Data, o que puede ser capacitado?	1	0.71%	2.86%
Consumidores de datos	¿Se tiene identificado a todos los posibles consumidores de los datos procesados?	3	2.14%	2.86%
Tipo de procesamiento requerido	¿El procesamiento debe ser realizado en tiempo real, casi en tiempo real? (Lo Contrario sería procesamiento por lotes)	2	1.43%	2.86%
			8.57%	14.29%

Figura 35. Aplicación del Modelo atributo Personas.

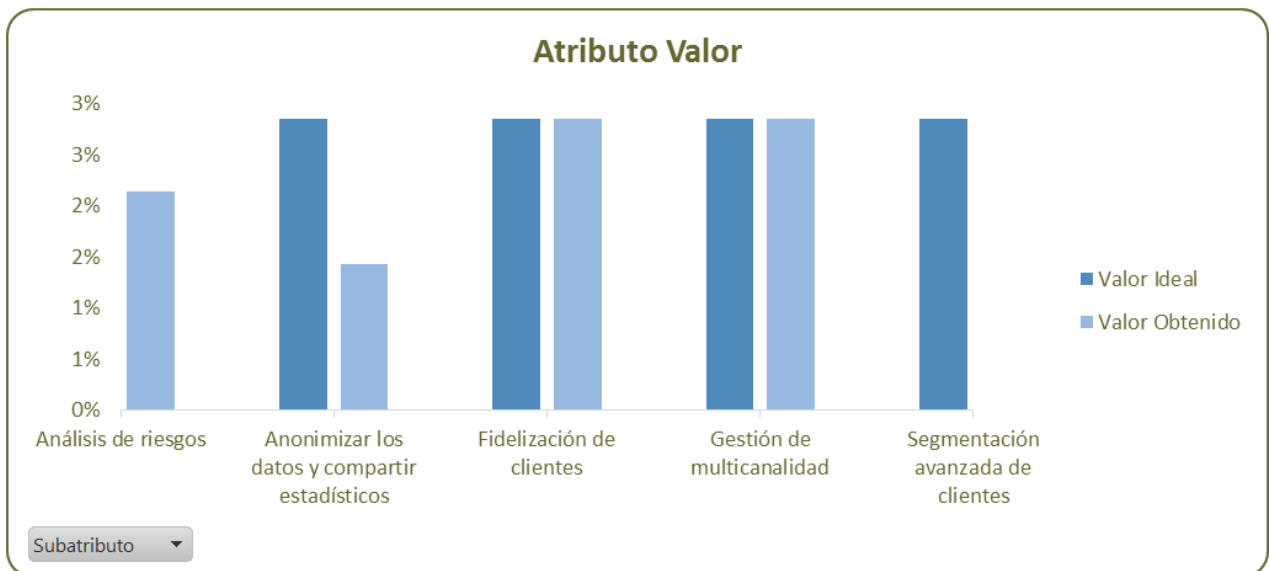


Figura 36. Resultados Valor obtenido vs. Valor ideal del atributo Valor.

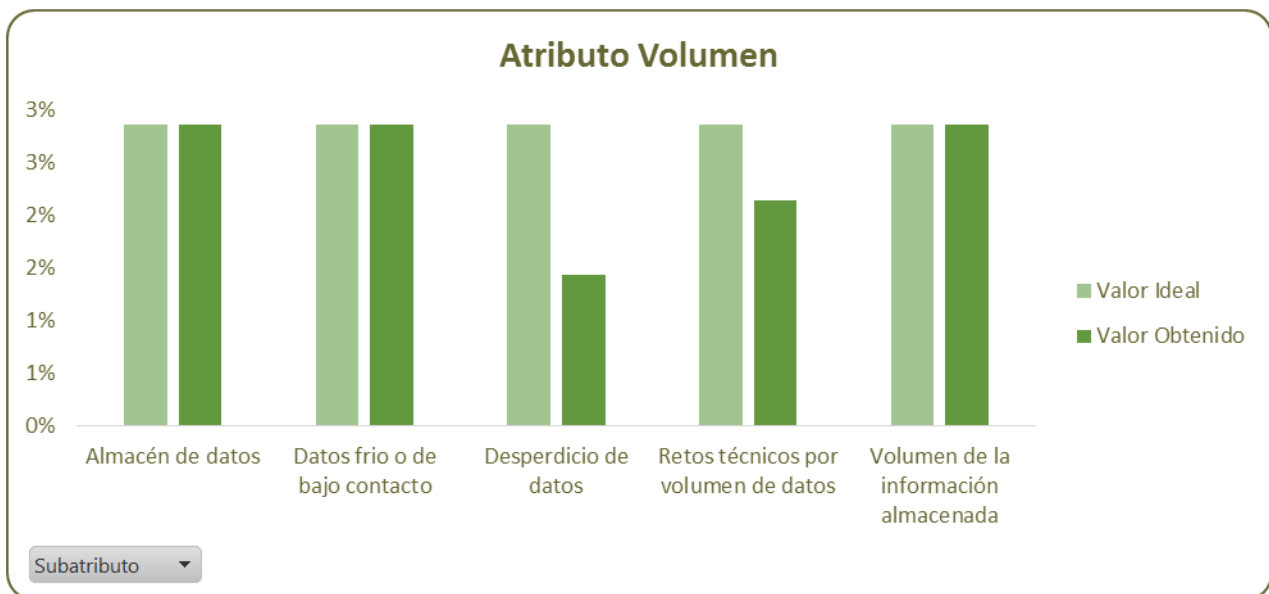


Figura 37. Resultados Valor obtenido vs. Valor ideal del atributo Volumen.

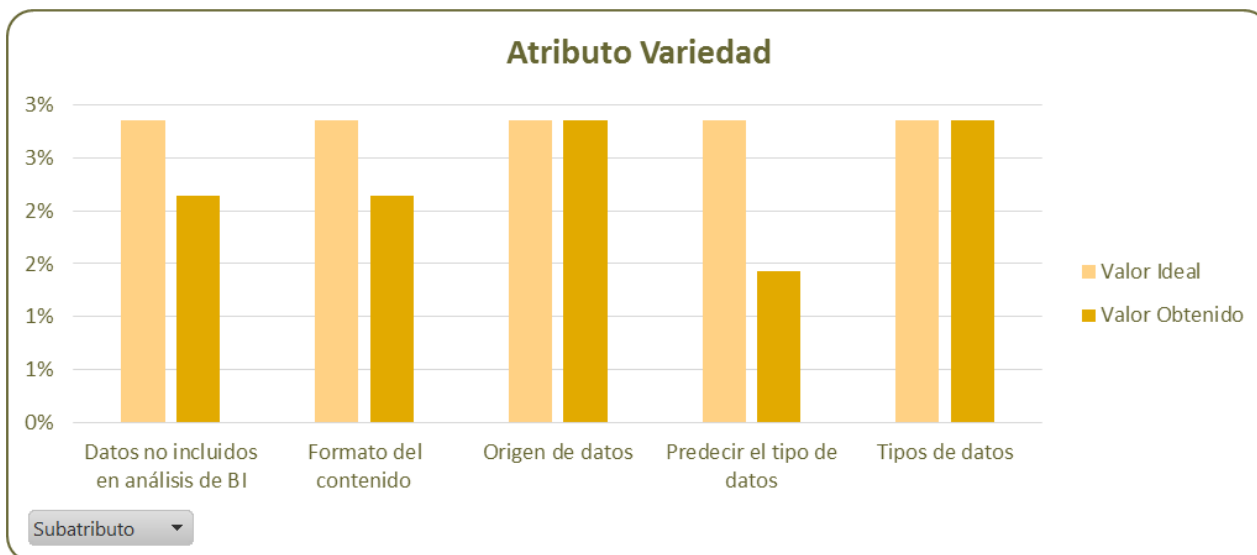


Figura 38. Resultados Valor obtenido vs. Valor ideal del atributo Variedad.

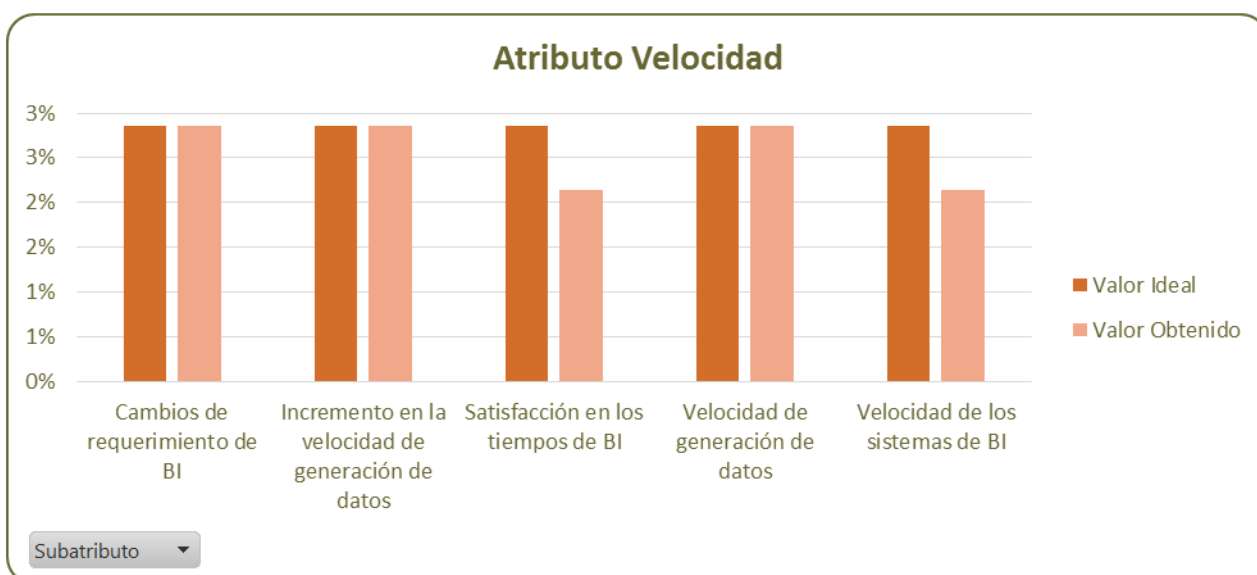


Figura 39. Resultados Valor obtenido vs. Valor ideal del atributo Velocidad.

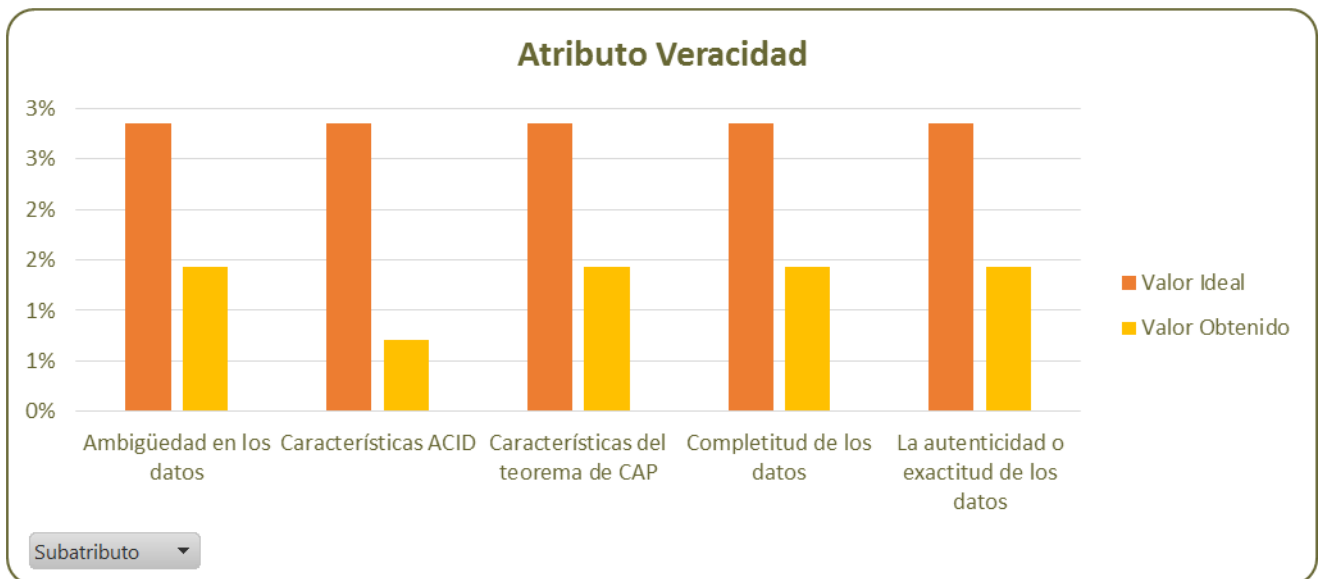


Figura 40. Resultados Valor obtenido vs. Valor ideal del atributo Veracidad.

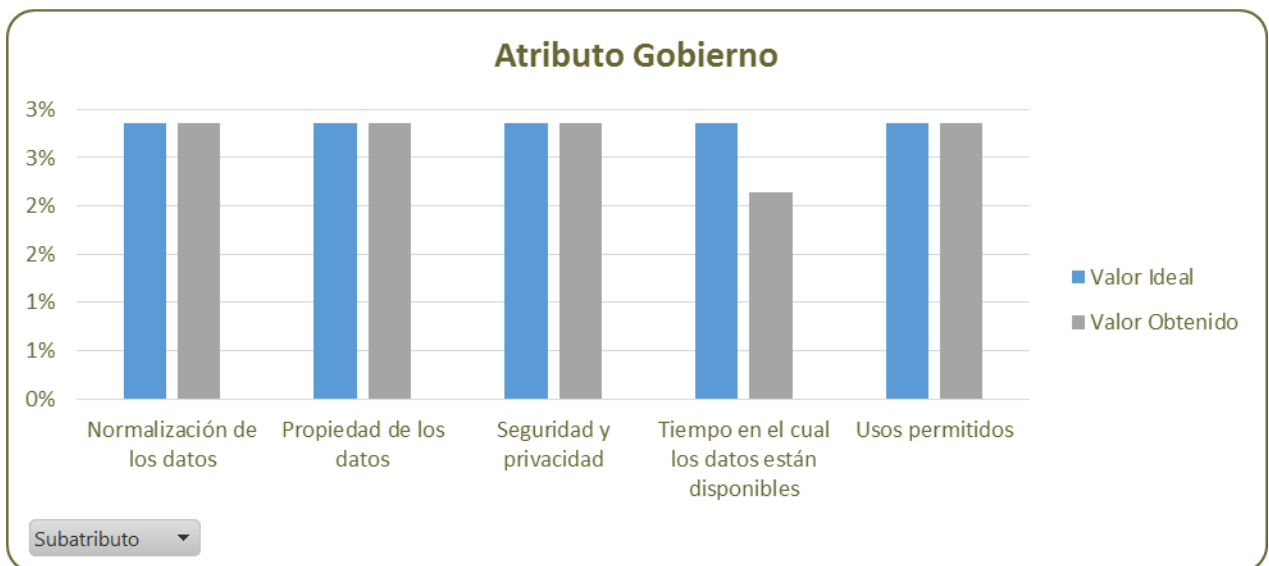


Figura 41. Resultados Valor obtenido vs. Valor ideal del atributo Gobierno.

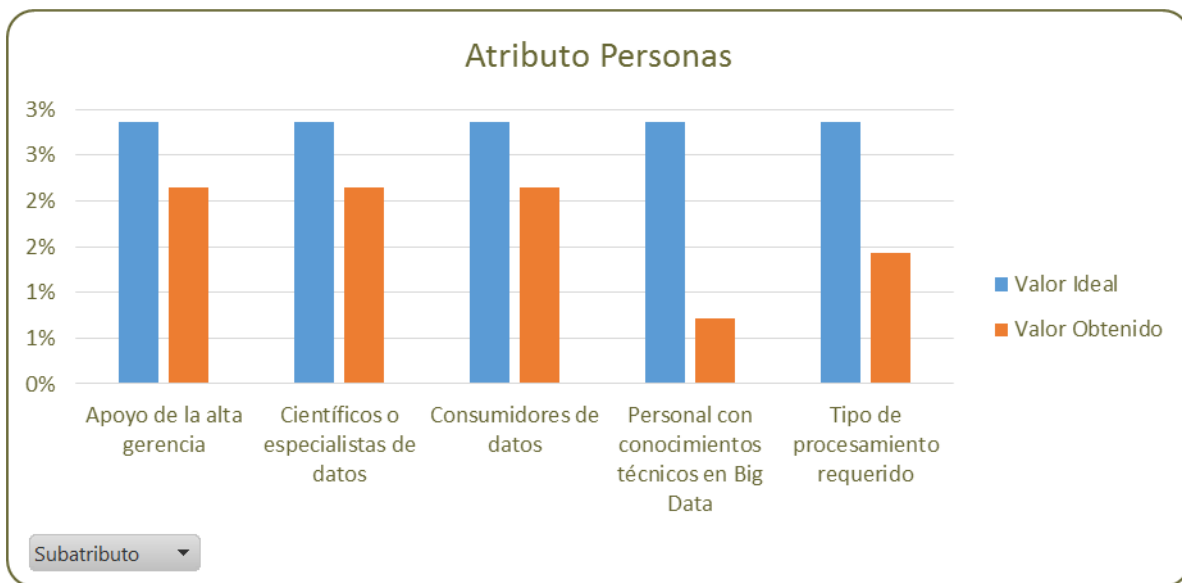
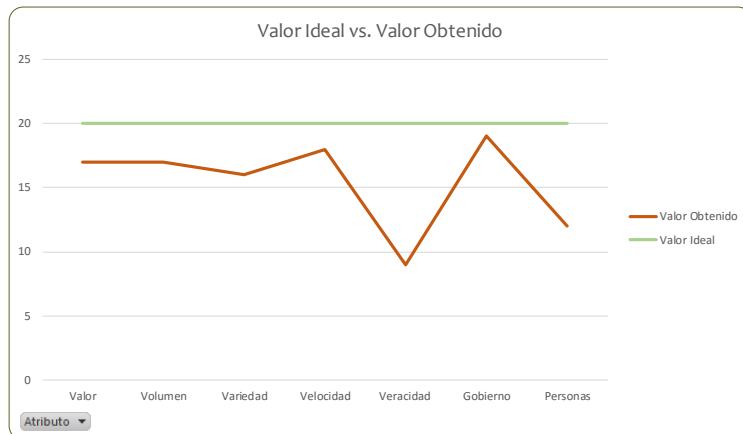


Figura 42. Resultados Valor obtenido vs. Valor ideal del atributo Personas.



Resultados

[« Volver a Instrucciones](#)

El valor ideal representa una estimación deseable y alcanzable que debe tener cualquier organización financiera al momento de implementar proyectos de Inteligencia de Negocios y Big Data. El valor obtenido es el resultado de las métricas introducidas de acuerdo a las características particulares de la organización a la que se aplica el modelo las cuales permitirán en primer lugar descubrir las necesidades y puntos de mejora de la organización, y al mismo tiempo que se logre la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea con mayores posibilidades de éxito.

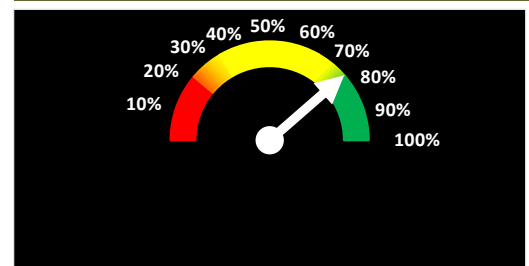


Figura 43. Resultados Valor obtenido vs. Valor ideal total del Modelo.

Conclusiones y Recomendaciones de la aplicación del modelo

En La última figura se presenta un resumen de los resultados obtenidos para cada dimensión o atributo evaluado vs. El valor ideal de acuerdo a este modelo, los cuales se interpretan de la siguiente manera:

Valor: Para generar valor a través de la inteligencia de negocios mediante una solución de Big Data no existen directrices definitivas, los escenarios deben ser identificados por la organización, el objetivo final es generar valor de toda la información almacenada a través de distintos procesos de manera eficiente y con el coste más bajo posible. La puntuación obtenida en esta dimensión demuestra que la organización se identificó con los casos de uso que, de aplicarse, aportarán un valor significativo para el negocio, Esto se traduce en cuanto a “Valor” generado por la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea, la solución es aplicable. La recomendación en este punto para la organización, en primer lugar está en estudiar el mercado para ver que está haciendo la competencia y sobretodo que están buscando los clientes. En segundo lugar, es el momento de evaluar el valor generado por cada uno de los casos de uso que se plantee, en contraste con el costo de inversión de una solución de este tipo, lo que finalmente será trascendental para determinar la viabilidad de la solución.

Volumen: Probablemente sea el “Volumen” la característica principal con la que se asocia el término Big Data. Implica que se almacena y genera gran cantidad de datos, lo cual debe ser una propiedad fundamental a evaluar al momento de plantearse la idea de implementar una solución de inteligencia de negocios con Big Data, que podría ser adecuada si hubiera complejidad razonable en el volumen. El porcentaje sobre la puntuación máxima obtenida (12.14 % / 14.29%), representa 85/100 % de esta dimensión particular, el cual demuestra la aplicabilidad en cuanto a “Volumen”, de la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea en la entidad financiera.

La recomendación para la organización en esta dimensión es evaluar y categorizar los datos que existan internamente, para determinar de cuales podría extraerse valor. Indiscutiblemente los datos de las entidades financieras representan ventajas y son muy interesantes para mejorar los modelos de negocio, a diferencia de otro tipo de datos que permiten visualizar la intencionalidad de los consumidores, la información de los servicios financieros expresa las decisiones de compra de los mismos. Otra recomendación importante es evaluar los recursos existentes para desarrollar nuevas soluciones de Big data, antes de contratar externamente dichos recursos, debe tener en cuenta el costo de la infraestructura, hardware, software, y el mantenimiento requerido por las herramientas y tecnologías de grandes volúmenes de datos.

Variedad: La variedad de datos se asocia a diversidad de fuentes de las cuales proviene la información y el tipo de datos que se ingresa. Una solución de grandes volúmenes de datos podría ser adecuada si hubiera complejidad razonable en la variedad de los datos a procesar. El porcentaje sobre la puntuación máxima obtenida (11.43 % / 14.29%), representa 80/100 % de esta dimensión particular, lo que representa una complejidad media-alta en el formato y tipo de datos que requiere ser procesado para hacer inteligencia de negocios en la banca en línea en la entidad financiera. Para el caso de la banca en línea esta puede ser una característica difícil de cumplir en un 100% ya que, generalmente la información que almacenan es el resultado de operaciones transaccionales y en su mayoría se encuentra de forma estructurada, en un tipo de datos y formato poco variable, casi siempre predecible. La recomendación para la organización es que tome en cuenta que los nuevos tipos y formatos de datos pueden surgir de fuentes que no están siendo consideradas para el proceso de Inteligencia de negocios, o de la creación de nuevos productos y servicios identificados a partir de la implementación de este tipo de sistemas.

Velocidad: El flujo de datos es masivo y continuo, por eso se requiere de sistemas que trabajan a toda máquina y sean capaces de procesar rápidamente los datos requeridos. El porcentaje sobre la puntuación máxima obtenida (12.86 % / 14.29%), representa 90/100 % de esta dimensión particular, esto indica que el ritmo con el que se generan los datos en la banca en línea y la necesidad de velocidad de

procesamiento, comprueban la aplicabilidad en cuanto a “Velocidad”, de la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea en la entidad financiera. La recomendación para la organización en este aspecto es que tome en cuenta que a pesar de tener una aplicabilidad de 86% en cuanto a “Volumen”, es útil definir claramente el alcance del proyecto. Que el volumen de datos sea bastante grande no quiere decir que requiera necesariamente un gran proyecto o gran alcance, esto podría traer dificultades para conseguir apoyo y completarlo en un lapso de tiempo razonable, si por el contrario, se plantea un alcance demasiado pequeño, podrían no verse los beneficios para la organización y perder el apoyo obtenido. Es importante encontrar un punto medio, implementando una solución de Big Data de forma incremental para ganar la confianza de las partes interesadas.

Veracidad: Este atributo en este contexto, se refiere al sesgo, ruido o alteración que pudiesen tener los datos. El porcentaje sobre la puntuación máxima obtenida (6.43% / 14.29%), representa 45/100 % de esta dimensión particular, lo que representa una complejidad media-baja en cuanto a “Veracidad” de datos. Una solución de inteligencia de negocios con Big Data para la Banca en línea podría ser adecuada si se identifica una complejidad razonable con la veracidad de los datos (junto con una complejidad en el volumen, variedad y velocidad de los mismos). La recomendación para la organización es que considere que la complejidad del volumen, variedad y velocidad de datos actuales, unida a la creación de nuevos productos y servicios, a mediano plazo van a repercutir en la veracidad de los mismos, y a mayor complejidad en la veracidad, volumen, variedad y velocidad de datos, mayor aplicabilidad tienen las soluciones de Big Data, por el contrario para datos menos complejos, deben evaluarse soluciones tradicionales.

Gobierno: Al decidir si se debe implementar una solución de Big Data, una organización podría estar buscando nuevas fuentes de datos, donde la propiedad no está claramente definida, esto unido a los problemas de gobierno de TI y ciertas regulaciones, pueden complicar el panorama. El porcentaje sobre la puntuación máxima obtenida (13.57% / 14.29%), representa 95/100 % de esta dimensión particular, esto indica la aplicabilidad en cuanto a “Gobierno”, de la implementación de

un sistema de inteligencia de negocios con Big Data para la banca en línea en la entidad financiera. La recomendación para la organización en esta dimensión es que tenga cuenta al momento de implementar este tipo de proyectos, temas relacionadas con la gobernabilidad como la seguridad y privacidad de los datos, la normalización, la disponibilidad, propiedad y usos permitidos de los mismos.

Personas: Para comprender, analizar y mantener un sistema de inteligencia de negocios con Big Data para la banca en línea se requieren habilidades específicas, que incluyen conocimiento del sector, experiencia, y conocimiento técnico sobre herramientas y tecnologías de grandes volúmenes de datos. El porcentaje sobre la puntuación máxima obtenida (8.57% / 14.29%), representa 60/100 % de esta dimensión particular. La recomendación para la organización en este punto es que considere al momento de implementar este tipo de proyectos que contar con científicos o especialistas en análisis de datos con experiencia en el modelaje, estadísticas, análisis y matemáticas son la clave para el éxito de cualquier iniciativa de Big Data.

Finalmente se observa que la aplicabilidad total de la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea en la entidad financiera, obtuvo 78/100 % sobre la puntuación máxima. Lo que representa una aplicabilidad bastante alta de este tipo de soluciones tecnológicas.

Finalmente, dada la puntuación obtenida por la organización, es recomendable avanzar a la siguiente fase del modelo propuesto en la que debe identificarse con alguno de los tres tipos de implementación de inteligencia de negocios con Big Data planteados y de esta forma orientar el proceso de implementación, donde seguidamente se debe confirmar la idea que tiene de la organización y sus objetivos, evaluar las fases anteriores, utilizar la información recopilada en las primeras fases para realizar el diseño lógico del sistema, documentar, probar y mantener el mismo.

CAPITULO VII: CONCLUSIONES Y RECOMENDACIONES

El presente capítulo se lleva a cabo con la finalidad presentar las principales conclusiones y recomendaciones obtenidas luego del desarrollo de los objetivos del trabajo de grado de Maestría realizado, en el cual se propone un Modelo de implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en línea en Venezuela, con el fin de realizar un aporte al estudio del proceso de inteligencia de negocios y al tratamiento de grandes volúmenes de datos en las organizaciones financieras.

7.1 Conclusiones

La adopción de herramientas y tecnologías de Big Data para hacer Inteligencia de negocios es una tarea ardua, no todos los requerimientos de análisis y presentación de informes en este ámbito, requieren de una solución de grandes volúmenes de datos o Big Data.

Antes de invertir en una solución de Inteligencia de Negocios con Big Data las organizaciones financieras deben plantearse una idea de que beneficios podrían obtenerse al analizar las enormes cantidades de datos que tienen disponible, evaluar estos datos y los recursos necesarios para definir, diseñar, desarrollar e implementar una plataforma de grandes volúmenes de datos es trascendental.

En esta investigación se plantearon cuatro objetivos específicos que han sido cumplidos a lo largo del cronograma de trabajo, por medio de un conjunto de actividades, de las que se concluye lo siguiente:

El primer objetivo propone analizar el entorno de Inteligencia de Negocios asociado a la banca en línea del sistema financiero venezolano, para ello se aplicó un cuestionario a tres entidades financieras (Banco Venezolano de Crédito, Banco Exterior y Banco Bancrecer), complementado por una guía de observación aplicada en el Banco Venezolano de Crédito, como resultado de un muestreo intencional, involucrando específicamente las gerencias asociadas a la inteligencia de negocios y la administración de datos. Esto permitió recabar información del estado actual en el

que se encuentra el entorno de inteligencia de negocios dentro de las organizaciones financieras, identificando sus características, carencias, necesidades y expectativas.

Se observó la existencia de soluciones tecnológicas de inteligencia de negocios en todas las entidades financieras que conforman la muestra. Hay un alto grado de insatisfacción con los tiempos de espera en la generación de reportes y los costos de dichas soluciones. A nivel general las entidades financieras dan un nivel de importancia alto a las soluciones tecnológicas de Inteligencia de negocios y tienen un alto apoyo organizacional cuando se habla de este tipo de proyectos. La gestión de multicanalidad, la segmentación avanzada de clientes y la incorporación de nuevas fuentes de datos a las segmentaciones tradicionales son las principales oportunidades percibidas por parte de las organizaciones financieras para implementar proyectos de inteligencia de negocios y Big Data. El 100 % de los encuestados afirma no haber desarrollado proyectos con Big Data, y aunque los recursos tecnológicos y el capital humano disponible para implementar este tipo de proyectos es limitado, 71.4 % manifiesta estar interesado en realizar proyectos piloto

En el segundo objetivo se identificó el estado del arte de la inteligencia de negocio y las arquitecturas de Big Data que permitirán implementar soluciones de grandes volúmenes de datos asociadas a la banca en línea del sistema financiero venezolano, donde fueron seleccionados elementos considerados vigentes y aplicables al entorno financiero, a partir de estudios e investigaciones relacionados con el tema.

Las características principales tomadas en cuenta para identificar el estado del arte de las herramientas de inteligencia de negocios y arquitecturas de Big Data asociadas al sistema financiero estuvieron compuestas por definiciones de inteligencia de negocios, su evolución, arquitecturas de Inteligencia de negocio, principales productos de Inteligencia de negocio, método para implantar proyectos de Inteligencia de negocio, definiciones de Big Data, características de Big Data, arquitecturas de Big Data, tecnologías Big data y ejemplos de aplicación. Finalmente se obtuvieron dos diagramas que resumen el estado del arte identificado.

El tercer objetivo fue el centro de la investigación planteada donde se desarrollaron las dimensiones del modelo de implementación de un sistema de inteligencia de negocios para la banca en línea en Venezuela, partiendo del análisis del entorno de inteligencia de negocio y tomando como referencia la investigación bibliográfica realizada. Como resultado se obtuvo un modelo que consta de ocho partes: siendo la primera; identificación de carencias, necesidades y expectativas, la segunda; selección del tipo de implementación, la tercera; determinación de requisitos de información, la cuarta; análisis de las necesidades del sistema, la quinta; diseño del sistema recomendado, la sexta; desarrollo y documentación del sistema, la séptima; prueba y mantenimiento, y la octava; implementación y evaluación, donde cada fase se explica por separado pero nunca se realizan como pasos aislados, más bien es posible que algunas actividades se realicen de manera simultánea, y algunas de ellas podrían repetirse.

El modelo propuesto da un mayor énfasis a la primera parte de la implementación en la que se diseñaron las métricas que permiten identificar de carencias, necesidades y expectativas asociadas a la inteligencia de negocio de la banca en línea del sistema financiero, seguidamente le permite a la organización decidir el tipo de implementación de Big Data que desea y con cual se siente más identificada, los cuales fueron agrupados en tres modelos diferentes; Disruptivo, Evolutivo e Híbrido, en los que se plantean las ventajas e inconvenientes de cada uno de ellos, las siguientes fases permiten orientar el proceso de implementación, donde se debe confirmar la idea que tiene de la organización y sus objetivos, evaluar las fases anteriores, utilizar la información recopilada en las primeras fases para realizar el diseño lógico del sistema, documentar, probar y mantener el mismo.

En el cuarto y último objetivo se validó el modelo de implementación diseñado mediante su aplicación a un caso de estudio, siendo este el Banco Venezolano de Crédito, en el que se analizaron los resultados por cada dimensión o atributo evaluado y se establecieron finalmente las recomendaciones para cada uno de ellos. En el resultado final se observa que la aplicabilidad total de la implementación de un sistema de inteligencia de negocios con Big Data para la banca en línea en la entidad

financiera, obtuvo 78/100 % sobre la puntuación máxima. Lo que representa una aplicabilidad bastante alta de este tipo de soluciones tecnológicas.

7.2 Recomendaciones

Las organizaciones financieras que deseen implementar sistemas de inteligencia de negocios con Big Data para la banca en línea deben considerar:

- ✓ En primer lugar se debe estudiar el mercado para ver que está haciendo la competencia y sobretodo que están buscando los clientes, y en segundo lugar evaluar el valor generado por cada uno de los casos de uso que se plantee, en contraste con el costo de inversión de una solución de este tipo, lo será trascendental para determinar la viabilidad de la solución.
- ✓ Es recomendable evaluar y categorizar los datos que existan internamente, para determinar de cuales podría extraerse valor, así como también analizar los recursos existentes para desarrollar nuevas soluciones de Big data, antes de contratar externamente dichos recursos, debe tener en cuenta el costo de la infraestructura, hardware, software, y el mantenimiento requerido
- ✓ Se debe tomar en cuenta que nuevos tipos y formatos de datos pueden surgir de fuentes que no están siendo consideradas para el proceso de Inteligencia de negocios, o de la creación de nuevos productos y servicios identificados a partir de la implementación de este tipo de sistemas.
- ✓ Se debe considerar que a mayor complejidad en la veracidad, volumen, variedad y velocidad de datos, mayor aplicabilidad tienen las soluciones de Big Data, por el contrario para datos menos complejos, deben evaluarse soluciones tradicionales.
- ✓ Se deben tener en cuenta temas relacionadas con la gobernabilidad de TI como la seguridad y privacidad de los datos, la normalización, la disponibilidad, propiedad y usos permitidos de los mismos, al momento de implementar este tipo de proyectos.

- ✓ Contar con científicos o especialistas en análisis de datos con experiencia en el modelaje, estadísticas, análisis y matemáticas son la clave para el éxito de cualquier iniciativa de Big Data.
- ✓ A las organizaciones que obtengan un porcentaje de aplicabilidad alta en la primera fase del modelo se recomienda avanzar a la siguiente fase en la que deben identificarse con alguno de los tres tipos de implementación de inteligencia de negocios con Big Data planteados y de esta forma orientar el proceso de implementación, donde seguidamente se debe confirmar la idea que tiene de la organización y sus objetivos, evaluar las fases anteriores, utilizar la información recopilada en las primeras fases para realizar el diseño lógico del sistema, documentar, probar y mantener el mismo.
- ✓ Finalmente se recomienda continuar con el diseño de este modelo, donde la siguiente etapa debería estar enfocada en profundizar los temas relacionados con el desarrollo del software como tal o de la selección de una solución tecnológica de inteligencia de negocios con Big Data comercial.

REFERENCIAS BIBLIOGRAFICAS

- Apache hadoop. (2015). *Apache hadoop*. Recuperado el 05 de 12 de 2015, de <http://hadoop.apache.org/>
- Banco Central de Venezuela. (2013). *Banco Central de Venezuela*. Recuperado el 11 de 15 de 2015, de <http://www.bcv.org.ve/snp/estadisticas.htm>
- Bonillo, P. (2003). *Modelo Sistémico de Evaluación Financiera de Proyectos de Inversión de Tecnología de la Información en Venezuela*. Obtenido de oocities: <http://www.oocities.org/es/pedrobonillotesis/proyecto/index.html>. [Consultado Abril 2016].
- Bonillo, P. (2014). *Phd consultores CA*. Recuperado el 05 de 11 de 2015, de <http://educacion.phd2014consultores.org/>
- Bonillo, P., González, C., Lobo, N., & Lugo, M. (2012). *Basamento legal de las tic en Venezuela*. Recuperado el 01 de 12 de 2015, de <http://es.slideshare.net/MarolLugo/basamento-legal-de-las-tic-en-venezuela-13162285>
- Brewer, E. (2000). *Toward Robust Distributed Systems*. Obtenido de <http://www.cs.berkeley.edu/~brewer/cs262b-2004/PODC-keynote.pdf>
- Camargo V., C. O. (2014). *Arquitectura Tecnológica Para Big Data*. *Revista Científica*, 7-18.
- Decreto 3390. (2000). *Gaceta Oficial de la República de Venezuela N° 36.955*. Caracas.
- Decreto1402. (2014). *Decreto con Rango, Valor y Fuerza de Ley de Instituciones del Sector Bancario*. Recuperado el 28 de 11 de 2015, de http://sudeban.gob.ve/?wpfb_dl=481
- Delgado, P. (2014). *Inteligencia de Negocios para Empresas de Construcción y la Gestión de Proyectos con enfoque en las mejores prácticas*. Cuenca, Ecuador.
- Delrieux, C. (2014). *Clasificación de Información en BigData mediante la utilización de Técnicas de Inteligencia Artificial y Análisis de Redes Sociales*. Universidad Nacional de la Patagonia San Juan Bosco.

- Dueñas, P. (2008). *Introducción al sistema financiero y Bancario*. Bogotá: Politécnico Grancolombiano.
- Ernst & Young. (2014). *Big Data en el sector financiero español*. Ernst & Young, Teradata FrontQuery.
- Espinoza, S. (2015). *Generar un marco de referencia para implementaciones de Big data en empresas de telecomunicaciones, caso de estudio corporación nacional de telecomunicaciones*. . Universidad de Las Américas.
- FinanzasDigital. (Julio de 2015). *FinanzasDigital.com*. Recuperado el 28 de 11 de 2015, de <http://www.finanzasdigital.com/wp-content/uploads/2015/07/BCOM201506.pdf>
- Flores, J. (2015). *Hombre al día*. Recuperado el 06 de 11 de 2015, de <http://www.hombreadia.com/tecnologia/analizar-datos-para-generar-valor-iv-foro-big-data-venezuela-2015/>
- FOGADE. (2015). *Fondo de protección social de los depósitos bancarios*. Recuperado el 30 de 11 de 2015, de <http://fogade.gob.ve/>
- Gartner Inc. (2014). *Gartner Survey Reveals That 73 Percent of Organizations Have Invested or Plan to Invest in Big Data in*. Obtenido de Website Gartner: <http://www.gartner.com/newsroom/id/2848718>
- Guama, S. (2015). *Estudio comparativo de métodos existentes para integrar la información estructurada y no estructurada de una industria enfocado en la generación de conocimiento, desde la perspectiva de una solución integral de big data*. Universidad de Las Américas.
- Hernández, R., Fernández, C., & Baptista, P. (2010). *Metodología de la investigación*. México D.F.: McGraw-Hill.
- Hurtado, J. (2010). *El Proyecto de Investigación, Comprensión holística de la metodología y la investigación*. (6ª Ed.). Ediciones Quiron.
- Kendall K. & Kendall J. (2005). *Análisis y Diseño de Sistemas* (6ª Ed.). Ediciones Pearson.Mexico.
- López, D. (2013). *Análisis de las posibilidades de uso de Big Data en las organizaciones*. Universidad de Las Américas.
- Martínez, J. (2010). *La inteligencia de negocios como herramienta para la toma de decisiones estratégicas en las empresas*. Bogotá.

- McGuire, T. (06 de 03 de 2013). *centrodeinnovacionbbva*. Recuperado el 03 de 11 de 2015, de <http://www.centrodeinnovacionbbva.com/innovation-edge/big-data/big-data-en-servicios-financieros>
- Microsystem. (2015). *Microsystem*. Recuperado el 01 de 07 de 2015, de <http://www.microsystem.cl/2015/02/conozca-las-4-v-del-big-data/>
- Ministerio del poder popular de Economía y Finanzas. (2015). Recuperado el 30 de 11 de 2015, de <http://www.mefbp.gob.ve/>
- Montalvo. (2015). *El impacto del "Big data" en los servicios financieros*. Universitat Pompeu Fabra.
- Moreno, P. (2015). Big Data: *Estado de la cuestión*. *International Journal of Information Systems and Software Engineering for Big Companies (IJISEBC)*, Vol. 2, Num. 1, pp. 38-59.
- Palella, S., & Martins, P. (2012). *Metodología de la investigación cuantitativa*. Fedupel.
- Paredes, A. (2015). *Big Data Estado de la cuestión*. *International Journal of Information Systems and Software Engineering for Big Companies*.
- Presidencia de la República. (2007). *Proyecto Nacional Simón Bolívar*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (1999). *Constitución de la República Bolivariana de Venezuela*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (2000). *Ley Orgánica de Telecomunicaciones*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (2002). *Ley Nacional de Juventud*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (2005). *Ley Orgánica de Ciencias, Tecnologías e Innovación*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (2007). *Ley Orgánica para la Protección de Niños, Niñas y Adolescentes*. Caracas.
- Presidencia de la República Bolivariana de Venezuela. (2009). *Ley Orgánica de Educación*. Caracas.
- Pritchett, D. (2008). *BASE: An Acid Alternative*. Obtenido de <http://queue.acm.org/detail.cfm?id=1394128>

- Roo, A., & Boscán, N. (09 de 2012). *Inteligencia de negocios en la banca nacional: Un enfoque basado en herramientas analíticas*. *Revista Venezolana de Gerencia*, 548-563. Recuperado el 2015
- Sabino, C. (1992,). *El proceso de investigación*. Caracas: Panapo.
- Serrano, M. (2013). *Publicación de estadísticas sobre canales transaccionales en la página electrónica del Banco Central de Venezuela*. *Boletín Económico (bvcvozeconomico)*, 1-4.
- Sinnexus . (2015). *sinnexus.com*. Recuperado el 2015 de 11 de 28, de http://www.sinnexus.com/business_intelligence/
- SUDEBAN. (2011). *Normas que permiten determinar el cumplimiento de los requisitos de calidad moral y ética exigidos para el ejercicio de la actividad bancaria*. Caracas.
- Temsamani, D. (2014). *Estrategia y negocios*. Recuperado el 03 de 11 de 2015, de <http://www.estrategiaynegocios.net/tecnologia/tendencias/773434-330/el-poder-del-big-data>
- The Apache Software Foundation. (2015). Recuperado el 02 de 12 de 2015, de <http://cassandra.apache.org/>
- Villalobos, A. (2015). *Relacion entre las bases de datos NoSQL y Big*. Universidad Latinoamericana de Ciencia y Tecnologia, ULACIT.
- Villamarín, W. (2015). *Conceptualización de una unidad de inteligencia de negocios en una empresa de tercerizacion de servicios*. Bogotá.

ANEXO I: Guía de Observación

Universidad Católica Andrés Bello.

Vicerrectorado Académico.

Estudios de Postgrado

Área de Ingeniería.

Postgrado en Sistemas de Información.

Proyecto de trabajo de grado de maestría: Modelo de implementación de un sistema de inteligencia de negocios con Big Data para la Banca en línea en Venezuela.

Guía de Observación

Consta de cuatro aspectos a evaluar, para las cuales se debe indicar con una X dentro de las casillas correspondientes al nivel de aplicación, donde 5 es *Siempre*, 4 es *Casi siempre*, 3 es *Algunas veces*, 2 es *Muy pocas veces* y 0 es *Nunca*, y en los casos que aplique escribir en el campo de observaciones la información requerida.

#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
1	Disponibilidad de los datos básicos requeridos para dar soporte en la toma de decisiones.						
2	Calidad de los datos básicos requeridos para dar soporte en la toma de decisiones.						
3	Acceso, integración, análisis y presentación de datos requeridos para la de toma de decisiones gerenciales						
4	Facilidad de uso o habilidad de la tecnología para proporcionar el soporte necesario en la toma de decisiones del usuario del sistema						

#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
5	Existencia de una cultura analítica respecto a la medición, las pruebas y la evaluación cuantitativa de evidencias para la toma de decisiones.						
6	Apoyo a los líderes a través del entendimiento del rol de BI en la toma de decisiones de negocios						

#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
---	--------------------------	---	---	---	---	---	---------------

7	Conocimiento de los usuarios respecto a las capacidades, funcionalidades y limitaciones del sistema						
8	Compromiso de los usuarios con el éxito del sistema						
9	Participación de los usuarios en la evaluación del sistema						
10	Contribución de los usuarios al continuo mejoramiento del sistema sugiriendo/revisando las mejoras						

#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
11	Eficiencia y funcionamiento del proceso de BI según lo planificado						
12	Satisfacción con el retorno de inversiones en BI						

ANEXO II: Cuestionario

Estimado participante

El presente cuestionario tiene por objeto recabar información para el desarrollo de la investigación titulada “Modelo de Implementación de un Sistema de Inteligencia de Negocio con Big Data para la Banca en Línea en Venezuela”.

Quisiera solicitar su colaboración para responder las preguntas que se presentan a continuación. Sus respuestas serán de carácter confidencial y anónimo, por lo que puede sentirse libre de expresar su opinión e ideas. La información que Ud. proporcione será considerada en la investigación en términos globales y no individualmente, los resultados serán compartidos con los participantes garantizando la confidencialidad de la información. La información será utilizada solo con fines confidenciales.

Las personas seleccionadas para responder este cuestionario, fueron elegidas por el tipo de actividad laboral dentro de las organizaciones del sector financiero, debido a que se considera que pueden aportar información valiosa al estudio.

Se le solicita conteste cada pregunta con la mayor objetividad posible, no existen respuestas correctas o incorrectas.

A continuación el Cuestionario:

1. ¿Según su criterio, cuál es su grado de satisfacción con el tiempo invertido en la generación de reportes? *

- ☐ Muy satisfecho
- ☐ Satisfecho
- ☐ Ni satisfecho, ni insatisfecho
- ☐ Insatisfecho
- ☐ Muy insatisfecho

2. ¿Según su criterio, cuál es el grado de importancia que tiene para su organización el uso de soluciones tecnológicas de Inteligencia de negocios? *

- ☐ Muy alta
- ☐ Alta
- ☐ Ni alta, ni baja
- ☐ Baja
- ☐ Muy baja

3. ¿Utiliza su organización actualmente alguna solución tecnológica de Inteligencia de negocios? Si su respuesta es "NO" ir a la pregunta 6. *

- ☐ Si
- ☐ No

4. ¿Según su criterio, cuál es su grado de satisfacción con respecto a costo de licenciamiento, mantenimiento y soporte de las herramientas de inteligencia de negocios en su organización? *

- ☐ Muy satisfecho
- ☐ Satisfecho
- ☐ Ni satisfecho, ni insatisfecho
- ☐ Insatisfecho
- ☐ Muy insatisfecho

5. ¿Según su criterio, cuál es el grado de satisfacción con respecto a la inteligencia de negocios realizada por su empresa? *

- ☐ Muy satisfecho
- ☐ Satisfecho
- ☐ Ni satisfecho, ni insatisfecho
- ☐ Insatisfecho
- ☐ Muy insatisfecho

6. ¿Cuáles serían los tres aspectos que en su opinión representan mayores oportunidades estratégicas, basadas en inteligencia de negocios y soluciones Big Data? *

- ☐ Segmentación avanzada de clientes
- ☐ Gestión de la multicanalidad para transacciones
- ☐ Marketing en detección de alertas y otro tipo de eventos complejos
- ☐ Fidelización de clientes mediante análisis avanzados
- ☐ Seguimiento avanzado de la competencia
- ☐ Otra...

7. ¿Cuáles serían los tres aspectos que considera Usted más prometedores * y diferenciales en su negocio, basados en inteligencia de negocios y soluciones Big Data?

- ☐ Apoyar procesos complejos de toma de decisiones
- ☐ Ser capaz de integrar fuentes externas e internas de datos
- ☐ Ser capaz de extraer "inteligencia" de la combinación de fuentes estructuradas y no estructurada
- ☐ Ser capaz de conjugar datos heterogéneos
- ☐ Ser capaz de trabajar con enormes cantidades de información
- ☐ Otra...

8. ¿Según su criterio, cuál considera que es el nivel de apoyo organizacional * que existe para implementar proyectos de inteligencia de negocios y Big Data?

- ☐ Muy alto
- ☐ Alto
- ☐ Ni alto, ni bajo
- ☐ Bajo
- ☐ Muy bajo

9. ¿Según su criterio, cuál considera que es el porcentaje de resistencia al cambio por parte de los funcionarios para implementar proyectos de inteligencia de negocios y Big Data? *

- ☐ Muy alto
- ☐ Alto
- ☐ Ni alto, ni bajo
- ☐ Bajo
- ☐ Muy bajo

10. ¿Cuáles considera que serían las tres principales barreras para implementar proyectos de inteligencia de negocios y Big Data?

*

- ☐ Falta de presupuesto económico.
- ☐ Barreras relacionadas con la capacidad tecnológica de la entidad
- ☐ Falta de criterio sobre la rentabilidad que pueden ofrecer las inversiones en Big Data (ROI)
- ☐ Falta de tiempo para abordar estos proyectos
- ☐ Falta de conocimiento especializado dentro de la Organización
- ☐ Otra...

11. ¿Según su criterio, cuál considera usted que es la capacidad de los recursos tecnológicos que estarán disponibles para implementar proyectos de inteligencia de negocios y Big Data?

*

- ☐ Muy alta
- ☐ Alta
- ☐ Ni alta, ni baja
- ☐ Baja
- ☐ Muy baja

12. ¿Según su criterio, al momento de obtener una solución de inteligencia de negocios, cuál sería el aspecto más importante a considerar? *

- ☐ Económicamente accesible
- ☐ Proporcione indicadores seguros
- ☐ Apoye a la toma de decisiones
- ☐ Que permita múltiples fuentes de información. (Base de datos)
- ☐ Entrega de reportes inmediatos
- ☐ Otra...

13. ¿Ha tenido su organización experiencia directa en el desarrollo de proyectos de Big Data? *

- ☐ Si
- ☐ No

14. De las siguientes posibles posiciones ante los proyectos de inteligencia de negocios y Big Data, ¿Cuál reflejaría mejor su actual situación? *

- ☐ Por momentos soy escéptico
- ☐ A la espera de que expertos demuestren utilidades concretas
- ☐ Interesado en realizar proyectos piloto
- ☐ Estoy realizando proyectos piloto, sin resultados evidentes
- ☐ Estoy realizando proyectos piloto, con resultados prometedores
- ☐ Otra...

ANEXO III: Formato de Validación de Instrumentos

Caracas, 24 de Febrero de 2016

Formatos de Validación

Estimado profesor _____, el presente formato de validación del cuestionario y la guía de observación, hace parte de la investigación titulada “Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela”. La evaluación de los instrumentos es de gran relevancia para que los resultados obtenidos a partir de éstos sean utilizados eficientemente, contribuyendo al logro de los objetivos de la investigación. Agradezco su valiosa colaboración.

De acuerdo con los siguientes indicadores califique cada uno de los ítems según corresponda:

Congruencia:	<i>Correspondencia del ítem con el de la variable a medir.</i>
Claridad:	<i>Se refiere a la redacción precisa y sencilla, fácil de comprender.</i>
Tendenciosidad:	<i>Tendencia o inclinación del ítem hacia determinada respuesta.</i>

Cuestionario

Ítem	Congruencia		Claridad		Tendenciosidad		Observaciones
	Si	No	Si	No	Si	No	
1							
2							
3							
4							
5							
6							
7							
8							
9							

10							
11							
12							
13							
14							

Guía de Observación

Ítem	Congruencia		Claridad		Tendenciosidad		Observaciones
	Si	No	Si	No	Si	No	
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							

Para dar cumplimiento a lo anteriormente expuesto, se le hace entrega formal del cuestionario y la guía de observación, con lo que se pretende medir las variables que permiten analizar el entorno de Inteligencia de negocio en la banca en línea del sistema financiero venezolano.

Sus observaciones y recomendaciones en esta validación, serán de gran ayuda para la elaboración de la versión final de los instrumentos, por lo que se le agradece su gran colaboración.

Agradeciendo de antemano su atención,

Atentamente,

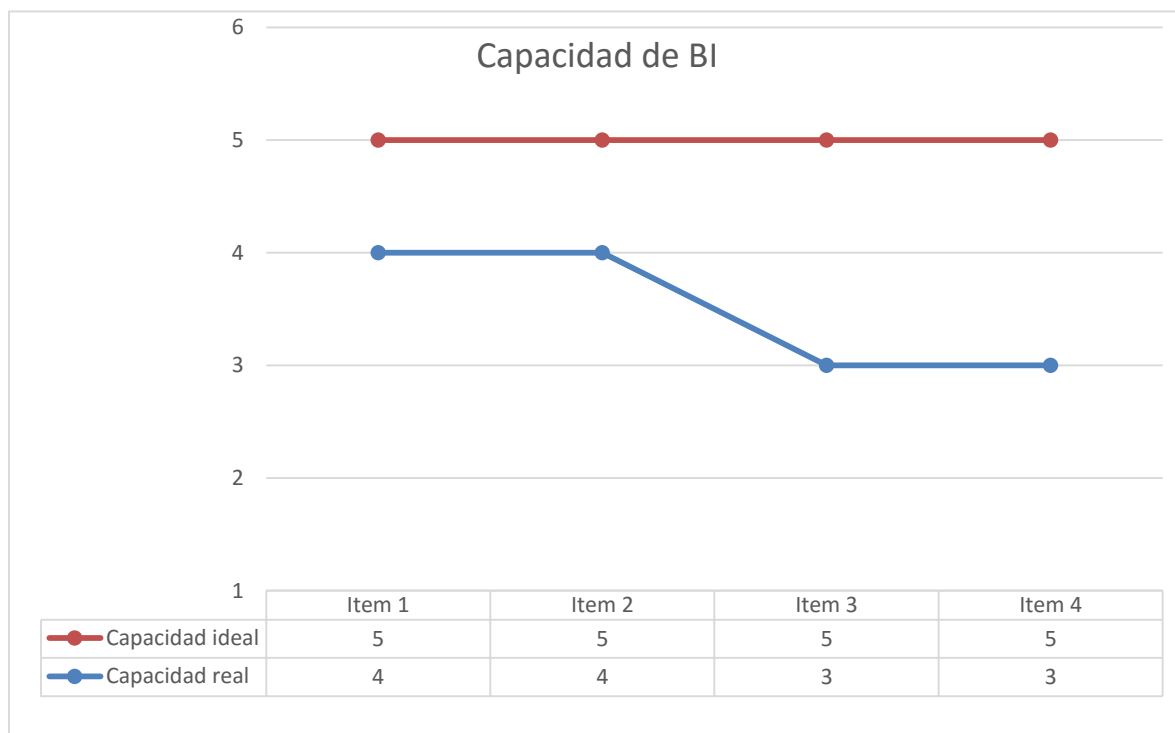
Ing. William Eduardo Bastidas Fuentes

C.I.: 19.671.995

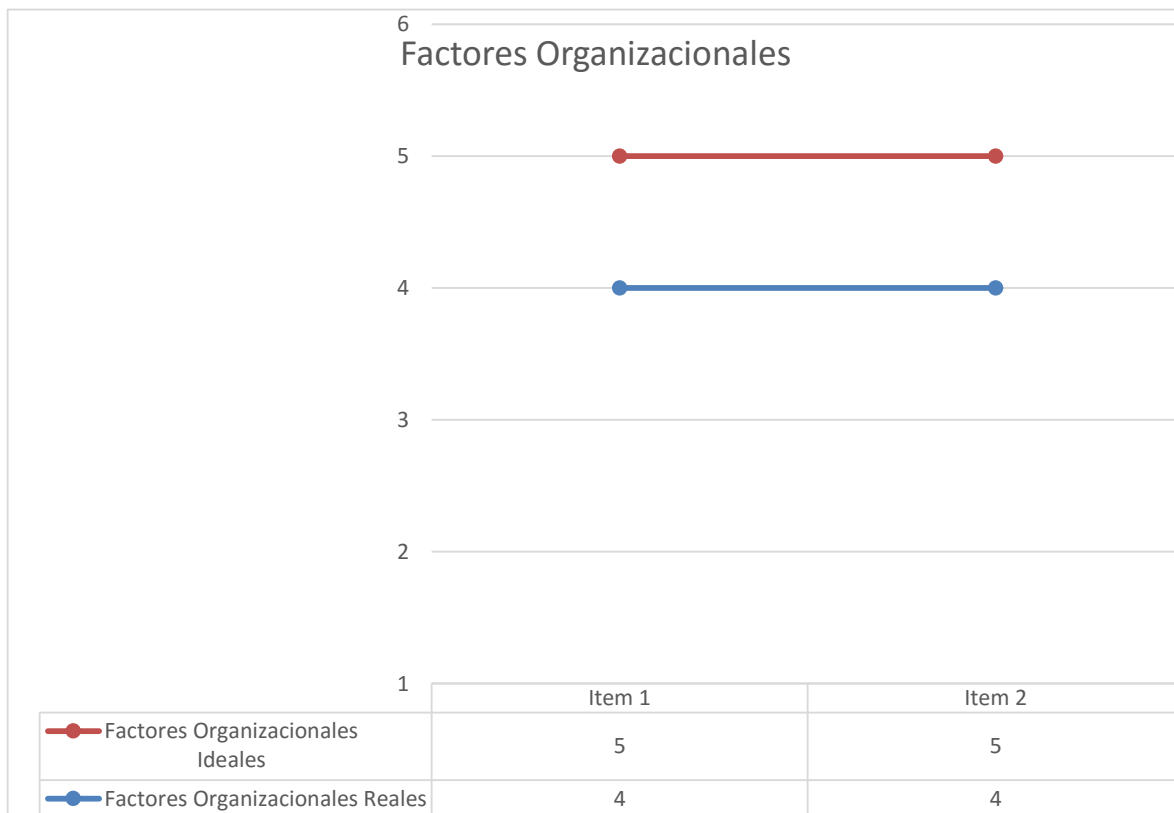
ANEXO IV

Resultados de la aplicación de la guía de observación

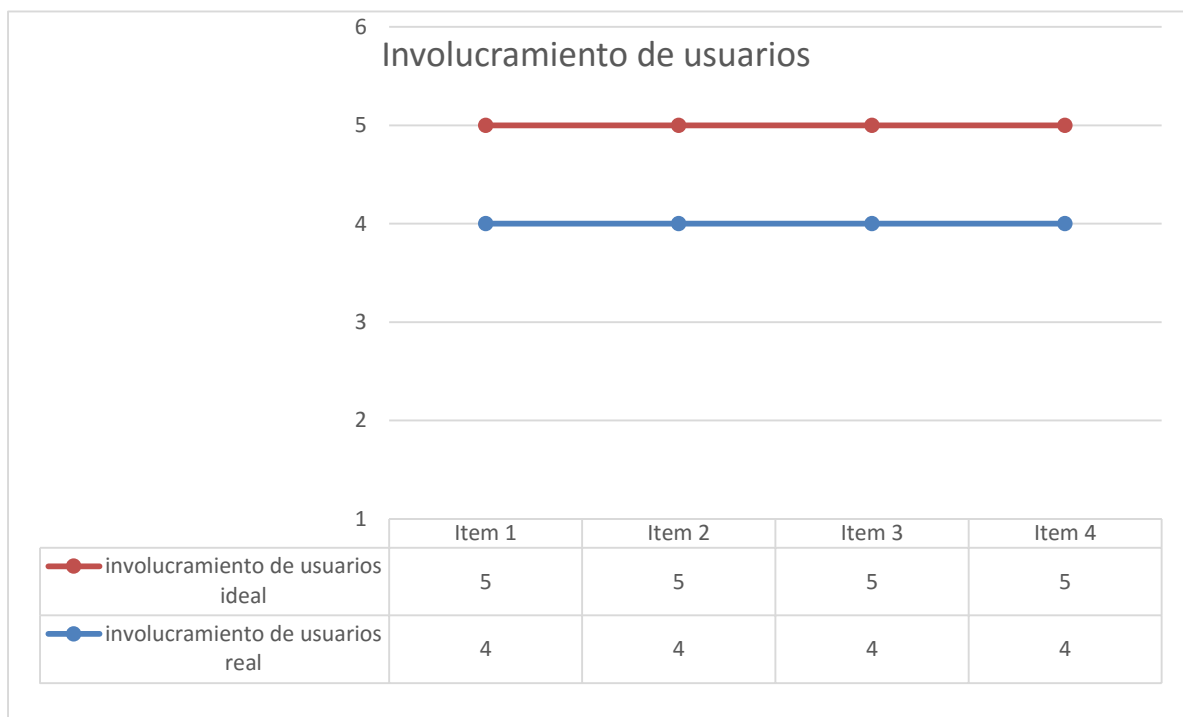
#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
1	Disponibilidad de los datos básicos requeridos para dar soporte en la toma de decisiones.		x				
2	Calidad de los datos básicos requeridos para dar soporte en la toma de decisiones.		x				
3	Acceso, integración, análisis y presentación de datos requeridos para la de toma de decisiones gerenciales			x			
4	Facilidad de uso o habilidad de la tecnología para proporcionar el soporte necesario en la toma de decisiones del usuario del sistema			x			



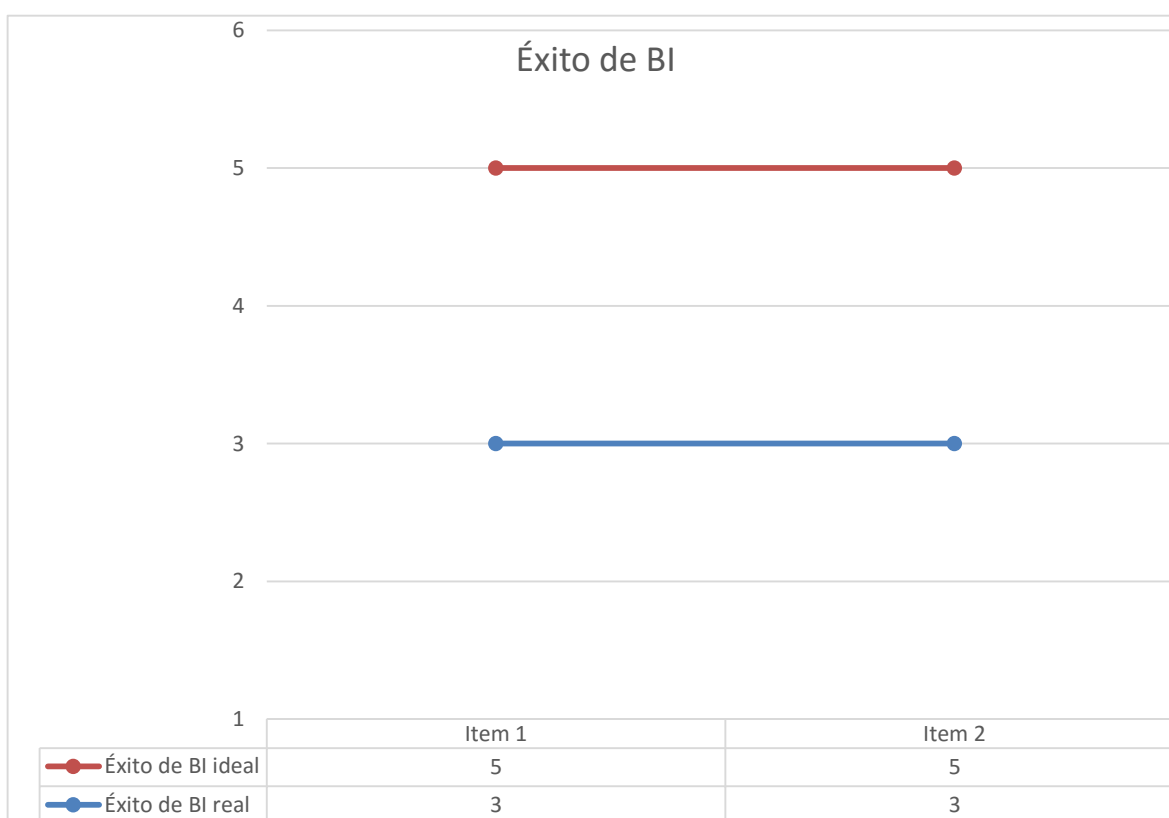
#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
1	Existencia de una cultura analítica respecto a la medición, las pruebas y la evaluación cuantitativa de evidencias para la toma de decisiones.		x				
2	Apoyo a los líderes a través del entendimiento del rol de BI en la toma de decisiones de negocios		x				



#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
1	Conocimiento de los usuarios respecto a las capacidades, funcionalidades y limitaciones del sistema		x				
2	Compromiso de los usuarios con el éxito del sistema		x				
3	Participación de los usuarios en la evaluación del sistema		x				
4	Contribución de los usuarios al continuo mejoramiento del sistema sugiriendo/revisando las mejoras		x				



#	Comportamiento a Evaluar	5	4	3	2	1	Observaciones
1	Eficiencia y funcionamiento del proceso de BI según lo planificado			x			
2	Satisfacción con el retorno de inversiones en BI			x			



ANEXO V: Resultados de la aplicación del cuestionario

PREGUNTAS

RESPUESTAS

7

7 respuestas



RESUMEN

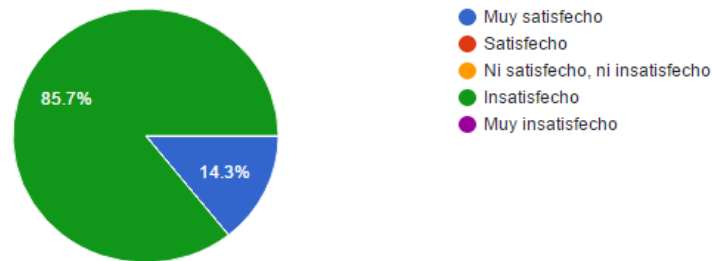
INDIVIDUAL

Se aceptan respuestas



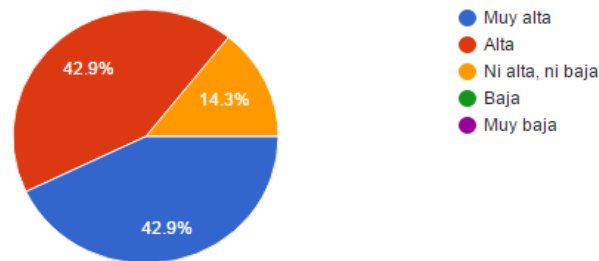
1. ¿Según su criterio, cuál es su grado de satisfacción con el tiempo invertido en la generación de reportes?

(7 respuestas)



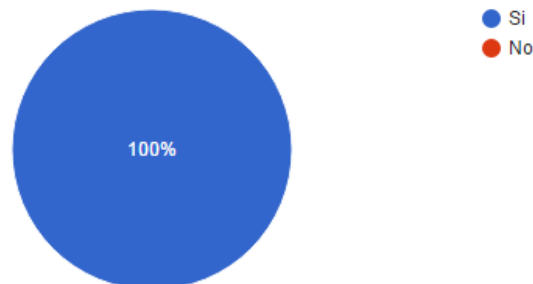
2. ¿Según su criterio, cuál es el grado de importancia que tiene para su organización el uso de soluciones tecnológicas de Inteligencia de negocios?

(7 respuestas)



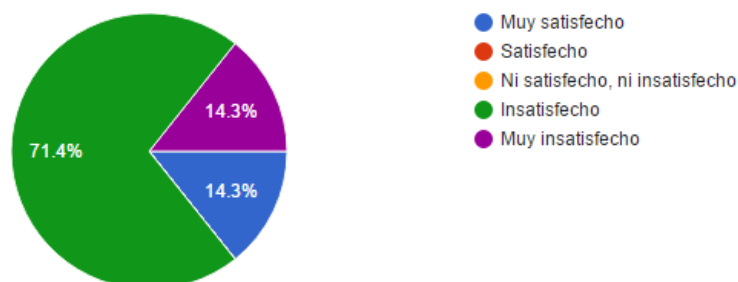
3. ¿Utiliza su organización actualmente alguna solución tecnológica de Inteligencia de negocios? Si su respuesta es "NO" ir a la pregunta 6.

(7 respuestas)



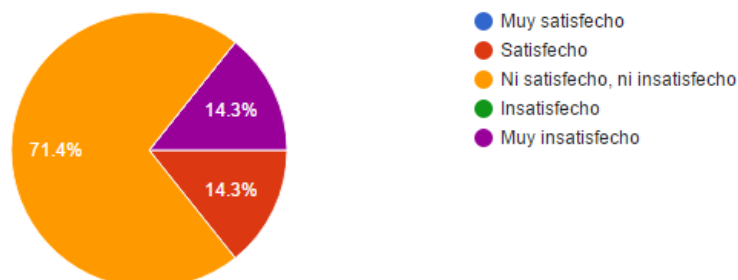
4. ¿Según su criterio, cuál es su grado de satisfacción con respecto a costo de licenciamiento, mantenimiento y soporte de las herramientas de inteligencia de negocios en su organización?

(7 respuestas)



5. ¿Según su criterio, cuál es el grado de satisfacción con respecto a la inteligencia de negocios realizada por su empresa?

(7 respuestas)



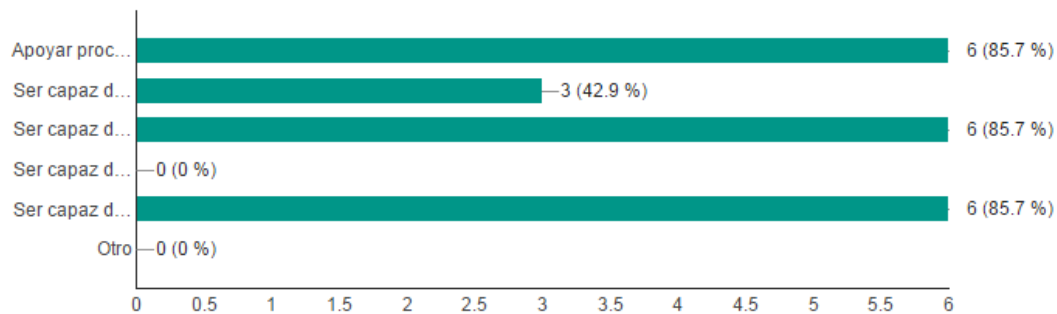
6. ¿Cuáles serían los tres aspectos que en su opinión representan mayores oportunidades estratégicas, basadas en inteligencia de negocios y soluciones Big Data?

(7 respuestas)



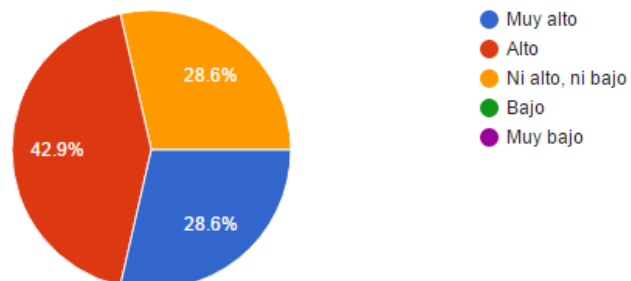
7. ¿Cuáles serían los tres aspectos que considera Usted más prometedores y diferenciales en su negocio, basados en inteligencia de negocios y soluciones Big Data?

(7 respuestas)



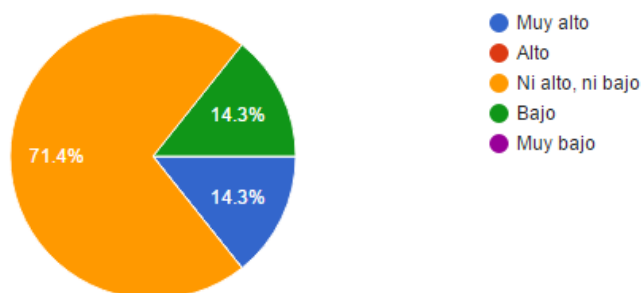
8. ¿Según su criterio, cuál considera que es el nivel de apoyo organizacional que existe para implementar proyectos de inteligencia de negocios y Big Data?

(7 respuestas)



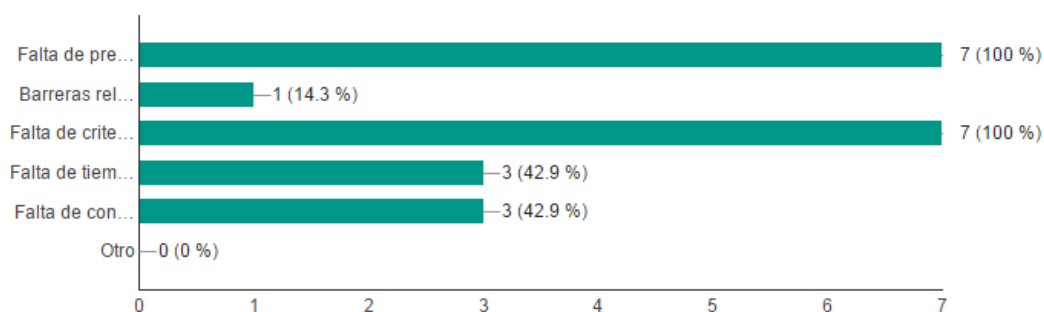
9. ¿Según su criterio, cuál considera que es el porcentaje de resistencia al cambio por parte de los funcionarios para implementar proyectos de inteligencia de negocios y Big Data?

(7 respuestas)



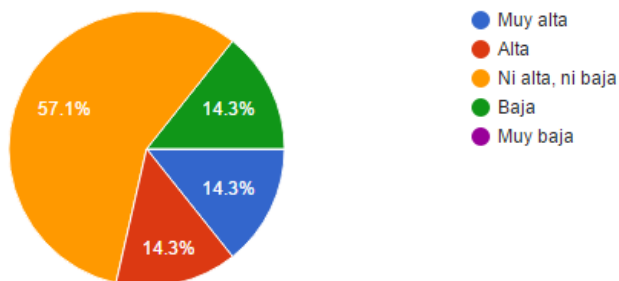
10. ¿Cuáles considera que serían las tres principales barreras para implementar proyectos de inteligencia de negocios y Big Data?

(7 respuestas)



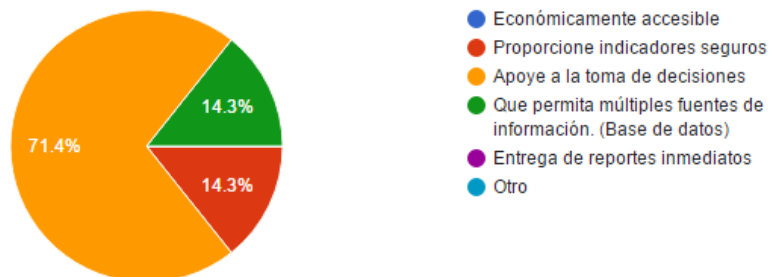
11. ¿Según su criterio, cuál considera usted que es la capacidad de los recursos tecnológicos que estarán disponibles para implementar proyectos de inteligencia de negocios y Big Data?

(7 respuestas)



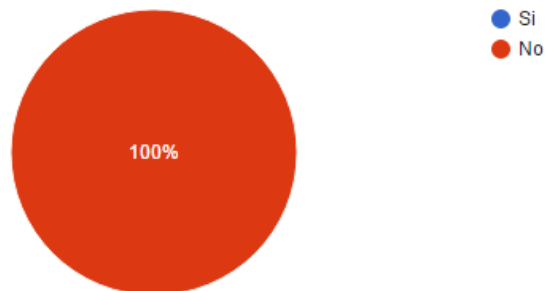
12. ¿Según su criterio, al momento de obtener una solución de inteligencia de negocios, cuál sería el aspecto más importante a considerar?

(7 respuestas)



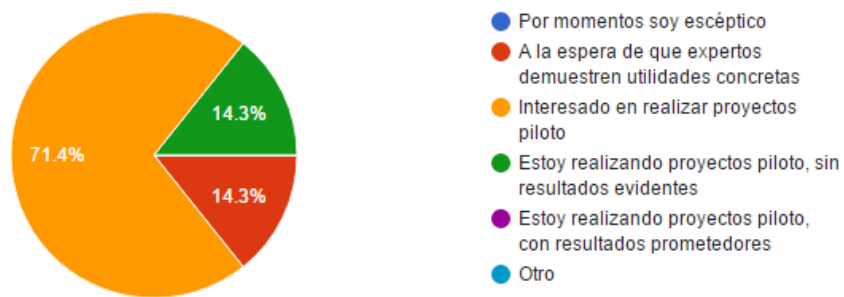
13. ¿Ha tenido su organización experiencia directa en el desarrollo de proyectos de Big Data?

(7 respuestas)



14. De las siguientes posibles posiciones ante los proyectos de inteligencia de negocios y Big Data, ¿Cuál reflejaría mejor su actual situación?

(7 respuestas)



ANEXO VI: Matriz del Modelo diseñado

Universidad Católica Andrés Bello.
Vicerrectorado Académico.
Estudios de Postgrado
Área de Ingeniería.
Postgrado en Sistemas de Información.

MATRIZ DEL MODELO DE IMPLEMENTACIÓN DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS CON BIG DATA PARA LA BANCA EN LÍNEA

A continuación se presenta la matriz que permitirá aplicar el Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea. La presentación de la matriz se divide en tres secciones, la primera corresponde al encabezado que contiene los datos de la entidad financiera (Nombre de la organización), los datos del evaluador (Nombre y cargo) y los datos de la evaluación (Fecha de la evaluación, áreas involucradas y responsable de dicha área).

La segunda sección corresponde a la matriz propiamente, la cual contiene siete (7) atributos a evaluar, a través de cinco (5) subatributos en cada uno, con un encabezado que corresponde a la medición, indicador, puntuación obtenida, puntuación máxima, % sobre la puntuación máxima obtenida, % sobre la puntuación máxima de cada subatributo y subtotales.

Finalmente en la tercera sección se presenta un resumen de los resultados obtenidos, calculados en base al porcentaje de cumplimiento por atributo y la suma del total de los mismos.

DATOS DE LA ENTIDAD FINANCIERA	
Nombre de la organización	
DATOS DEL EVALUADOR	
Nombre del evaluador	
Cargo en la organización:	
DATOS DE LA EVALUACION	
Fecha de evaluación:	
Áreas involucradas	
Responsable del área:	



Métricas

[Ver indicadores»](#)

Valor				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Gestión de multicanalidad	¿Es la gestión de multicanalidad una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?		0.00%	2.86%
Segmentación avanzada de clientes	¿Es la segmentación avanzada de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?		0.00%	2.86%
Fidelización de clientes	¿Es la fidelización de clientes una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?		0.00%	2.86%
Análisis de riesgos	¿Es el análisis de riesgos una oportunidad percibida al implementar proyectos de inteligencia de negocios con Big Data?		0.00%	2.86%
Anonimizar los datos y compartir estadísticos	¿Anonimizar los datos que se tienen de los clientes y compartir estadísticas con otros sectores, permitiría generar nuevas oportunidades de negocios?		0.00%	2.86%
			0.00%	14.29%

Volumen				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Volumen de la información almacenada	¿El volumen o tamaño de las bases de datos actuales se dimensiona de petabytes a exabytes, y en un futuro próximo, podría crecer hasta zettabytes?		0.00%	2.86%
Almacén de datos	¿Existe un entorno de almacén de datos que contiene un repositorio de todos los datos generados o adquiridos?		0.00%	2.86%
Datos frío o de bajo contacto	¿Hay una cantidad significativa de datos de frío o de bajo contacto que no está siendo analizada para derivar el conocimiento del negocio?		0.00%	2.86%
Desperdicio de datos	¿Se eliminan cantidades de datos debido a la incapacidad de procesamiento o procesamiento?		0.00%	2.86%
Retos técnicos por volumen de datos	¿Se plantean retos técnicos y económicos para almacenar, buscar, compartir, analizar y visualizar datos utilizando los métodos tradicionales debido al volumen de datos?		0.00%	2.86%
			0.00%	14.29%

Variedad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Tipos de datos	¿El tipo de datos que requiere ser procesado es transaccional o histórico?		0.00%	2.86%
Formato del contenido	¿El formato de datos puede variar, incluyendo datos estructurados, semiestructurados y no estructurados?		0.00%	2.86%
Origen de datos	¿Los datos son generados por sistemas transaccionales?		0.00%	2.86%
Predecir el tipo de datos	¿El contenido y la estructura de datos a ser procesado no se pueden anticipar o predecir?		0.00%	2.86%
Datos no incluidos en análisis de BI	¿Existe una gran variedad de datos que nunca han sido incluidos para los análisis de inteligencia de negocios por su complejidad y que podrían generar valor?		0.00%	2.86%
			0.00%	14.29%

Velocidad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Velocidad de generación de datos	¿La velocidad de generación de datos es grande?	4	2.86%	2.86%
Incremento en la velocidad de genera	¿La velocidad de generación de datos tiende a incrementar en el tiempo?	4	2.86%	2.86%
Velocidad de los sistemas de BI	¿Están las tecnologías y los métodos tradicionales de BI abrumados, ya no son adecuados para manejar datos que lleguen en tiempo requerido?	3	2.14%	2.86%
Cambios de requerimiento de BI	¿BI Está cambiando rápidamente y debe ser respondido de inmediato?	4	2.86%	2.86%
Satisfacción en los tiempos de BI	¿No está satisfecha la organización con los tiempos de generación de reportes para hacer inteligencia de negocios?	3	2.14%	2.86%
			12.86%	14.29%

Veracidad				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
La autenticidad o exactitud de los datos	¿La autenticidad o exactitud de los datos es desconocida?		0.00%	2.86%
Ambigüedad en los datos	¿Los datos incluyen información ambigua?		0.00%	2.86%
Complejidad de los datos	¿No está claro si los datos están completos?		0.00%	2.86%
Características ACID	¿No es necesario que el sistema de BI cumpla con las características ACID?		0.00%	2.86%
Características del teorema de CAP	¿Cumpliría un sistema distribuido con los requerimientos de BI que satisfaga solo dos de los siguientes atributos al mismo tiempo: Consistencia Fuerte, Alta Disponibilidad y Tolerancia a Particionamiento?		0.00%	2.86%
			0.00%	14.29%

Gobierno				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Seguridad y privacidad	¿Se tiene claro que datos se pueden almacenar y tener acceso? ¿Qué datos deben ser cifrados durante el movimiento o en reposo? ¿Quién está autorizado para ver los datos en bruto y los puntos de vista?	4	2.86%	2.86%
Normalización de los datos	¿Existen normas que rigen los datos?	4	2.86%	2.86%
Tiempo en el cual los datos están disponibles	¿Están los datos disponibles en un plazo que permita medidas que deben tomarse en el momento oportuno?	3	2.14%	2.86%
Propiedad de los datos	¿La solución tiene acceso y permiso apropiado utilizar los datos?	4	2.86%	2.86%
Usos permitidos	¿Está permitido el uso de los datos disponibles para la solución de inteligencia de negocios con Big Data planteada?	4	2.86%	2.86%
			13.57%	14.29%

Personas				
Subatributo	Medición	Ítem	% Puntuación Obtenida	% Puntuación Máxima
Apoyo de la alta gerencia	¿Existe aceptación por parte de los interesados y la alta gerencia, y están dispuestos a invertir en el proyecto?		0.00%	2.86%
Científicos o especialistas de datos	¿Tiene científicos o especialistas de datos disponibles para generar conocimientos y valor significativo a partir de los datos?		0.00%	2.86%
Personal con conocimientos técnicos en Big Data	¿Existe personal con conocimiento técnico sobre herramientas y tecnologías de Big Data, o que puede ser capacitado?		0.00%	2.86%
Consumidores de datos	¿Se tiene identificado a todos los posibles consumidores de los datos procesados?		0.00%	2.86%
Tipo de procesamiento requerido	¿El procesamiento debe ser realizado en tiempo real, casi en tiempo real? (Lo Contrario sería procesamiento por lotes)		0.00%	2.86%
			0.00%	14.29%

ATRIBUTO	Puntuación obtenida	Puntuación Máxima	% Sobre la puntuación máxima obtenida	% Sobre la puntuación máxima
VALOR		4		14.285
VOLUMEN		4		14.285
VARIEDAD		4		14.285
VELOCIDAD		4		14.285
VERACIDAD		4		14.285
GOBIERNO		4		14.285
PERSONAS		4		14.285
Total		28		99.995

Recomendaciones:

ANEXO VII

Constancias de validación de los instrumentos

Caracas, 24 de Febrero de 2016

CONSTANCIA DE VALIDACION

Yo, Pedro Nolasco Bonillo Ramos, titular de la cedula de identidad N° 10.868.538, de Profesión Doctor en Ciencias de la Computación, ejerciendo actualmente como evaluador, en la Universidad Católica Andrés Bello, por medio de la presente hago constar que he revisado con fines de validación los instrumentos (cuestionario y Guía de Observación), del Trabajo de Grado de Maestría del estudiante William Eduardo Bastidas Fuentes, cédula de identidad nro. 19.671.995, quien opta por el título de Magíster en Sistemas de Información, titulado “Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela”, a los efectos de su validación.

Sin otro particular al cual hacer referencia, se despide cordialmente,

Dr. Pedro Nolasco Bonillo Ramos

C.I. 10.868.538

Caracas, 09 de julio de 2016

CONSTANCIA DE VALIDACION

Yo, Gloria Aponte, titular de la cedula de identidad N° 4.964.695, de Profesión Ingeniero Químico y Magister en Sistemas de Información, ejerciendo actualmente como evaluador, en la Universidad Católica Andrés Bello, por medio de la presente hago constar que he revisado con fines de validación los instrumentos (cuestionario y Guía de Observación), del Trabajo de Grado de Maestría del estudiante William Eduardo Bastidas Fuentes, cédula de identidad nro. 19.671.995, quien opta por el título de Magíster en Sistemas de Información, titulado “Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela”, a los efectos de su validación.

Sin otro particular al cual hacer referencia, se despide cordialmente,

Ing. Gloria Aponte, MSc.

C.I. 4.9646.695

Caracas, 09 de julio de 2016

CONSTANCIA DE VALIDACION

Yo, Maria Esther Remedios, titular de la cedula de identidad N° 5.530.488, de Profesión Licenciada en Computación y Magister en Sistemas de Información, ejerciendo actualmente como evaluador, en la Universidad Católica Andrés Bello, por medio de la presente hago constar que he revisado con fines de validación los instrumentos (cuestionario y Guía de Observación), del Trabajo de Grado de Maestría del estudiante William Eduardo Bastidas Fuentes, cédula de identidad nro. 19.671.995, quien opta por el título de Magíster en Sistemas de Información, titulado “Modelo de Implementación de un Sistema de Inteligencia de Negocios con Big Data para la Banca en Línea en Venezuela”, a los efectos de su validación.

Sin otro particular al cual hacer referencia, se despide cordialmente,

Lic. Maria Esther Remedios, MSc.

C.I. 5.530.488