



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

“ANÁLISIS EXPLORATORIO DE METODOLOGÍAS ÁGILES Y
BIG DATA APLICADOS EN EMPRESAS ECUATORIANAS”

TRABAJO DE TITULACIÓN

Previo a la obtención del título de

MAGISTER EN CIENCIAS DE LA COMPUTACIÓN

ROSA ELIZABETH QUELAL MORA

GUAYAQUIL - ECUADOR

AÑO: 2018

AGRADECIMIENTO

Agradezco sobre todo a Dios, quien me ha guiado y llenado de bendiciones en este proceso. A mi esposo que ha sido mi sostén y apoyo, y a mi hija quien ha sido mi mayor inspiración.

Agradezco a la ESPO, por permitirme participar en este programa de maestría, y de manera especial agradezco a mi tutora, la PhD. Mónica Villavicencio, por su orientación académica, profesional y personal.

Finalmente, agradezco a las personas que participaron de este estudio por su valiosa contribución.

DEDICATORIA

Dedico este trabajo a mi esposo, quien me alentó a iniciar esta aventura, y con amor y paciencia me ayudó culminarla.

TRIBUNAL DE EVALUACIÓN

PhD. Cesar Martin

SUBDECANO DE LA FIEC

PhD. Mónica Villavicencio

DIRECTOR DEL TRABAJO DE TITULACIÓN

PhD. Luis Eduardo Mendoza

MIEMBRO PRINCIPAL DEL TRIBUNAL

Msia. EDUARDO CRUZ

MIEMBRO PRINCIPAL DEL TRIBUNAL

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de este trabajo de titulación, me corresponde exclusivamente; y doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”.

Rosa E. Quelal Mora

RESUMEN

El uso de Big Data se está incrementando en diferentes industrias; sin embargo, no resulta claro el proceso y metodologías utilizados para desarrollar proyectos de este tipo.

El presente trabajo expone dos tipos de investigación: 1) Revisión sistemática de literatura respecto a Metodologías Ágiles aplicadas a proyectos de Big Data; 2) Análisis descriptivo y cuantitativo del uso de Metodologías Ágiles y Big Data en empresas ecuatorianas.

Entre los hallazgos, encontramos evidencia del uso de Metodologías Ágiles en proyectos de Big Data; entre estas destacan Scrum, Programación Extrema y Kanban. Sin embargo, dentro de las empresas ecuatorianas el conocimiento sobre los beneficios que las Metodologías Ágiles y Big Data pueden aportar es limitado, lo cual podría ser la causa del alto porcentaje de abandono de las metodologías en mención (56%), y la falta de interés en contratar científicos de datos (83%).

No obstante, sectores económicos como el Comercial y Bancario mostraron interés y continuidad en el uso de las Metodologías Ágiles; y a modo general, las organizaciones mostraron interés en adquirir servicios y productos de Big Data desarrollados por terceros, que les permitan analizar sus operaciones, el mercado, y generar nuevas ideas en tiempo real.

ABSTRACT

The use of Big Data is increasing in different type of companies; however, the process and methodology applied to develop this kind of projects is not clear.

This work present two types of research: 1) A systematic literature review of Agile Methodologies applied to Big Data projects; 2) a descriptive and quantitative analysis of the use of Agile Methodologies and Big Data in Ecuadorian companies.

Among the findings, we find evidence of the use of Agile Methodologies in Big Data projects, where the most applied Agile Methodologies have been Scrum, Extreme Programming and Kanban. However, within Ecuadorian companies the knowledge about the benefits that Agile Methodologies and Big Data bring to them is limited, which could be the cause of the high percentage of abandonment of these kind of methodologies (56%), and the lack of interest in hiring data scientists (83%).

Nevertheless, economic sectors like Commercial and Banking showed interest and continuity in the use of Agile Methodologies; and in general, the organizations showed interest in acquiring Big Data products and services developed by third parties, in order to analyze their operations, the market, and to generate new ideas in real time.

ÍNDICE GENERAL

AGRADECIMIENTO.....	ii
DEDICATORIA.....	iii
TRIBUNAL DE EVALUACIÓN.....	iv
DECLARACIÓN EXPRESA.....	v
RESUMEN.....	vi
ABSTRACT	vii
ÍNDICE GENERAL	viii
ÍNDICE DE FIGURAS	x
ÍNDICE DE TABLAS.....	xi
INTRODUCCIÓN	xii
CAPÍTULO 1	1
GENERALIDADES	1
1.1 Antecedentes.....	1
1.2 Descripción del problema.....	2
1.3 Objetivo General	2
1.4 Objetivos Específicos.....	2
CAPÍTULO 2	3
MARCO TEÓRICO	3
2.1 Big Data.....	3
2.2 Metodologías Ágiles de desarrollo de software	5
2.3 Metodologías Ágiles aplicadas a proyectos de Inteligencia de Negocios	9
2.4 Metodologías Ágiles aplicadas a proyectos de Big Data	10
CAPÍTULO 3	11
MARCO METODOLÓGICO.....	11

3.1	Revisión Sistemática de Literatura.....	11
3.2	Estudio descriptivo y cuantitativo.....	13
CAPÍTULO 4		15
RESULTADOS Y ANÁLISIS DE DATOS		15
4.1	Resultados de la Revisión sistemática de literatura	15
4.2	Resultados del Estudio descriptivo y cuantitativo	17
CONCLUSIONES		31
RECOMENDACIONES.....		32
BIBLIOGRAFIA.....		33
Apéndice A.....		41
Apéndice B.....		42

ÍNDICE DE FIGURAS

Figura 4.1. Percepción de capacidad de análisis y toma de decisiones que ofrecen las herramientas tecnológicas de la Empresa.	19
Figura 4.2. Percepción de relevancia, precisión y oportunidad de tomar decisiones con la información disponible.	19
Figura 4.3. Metodologías Ágiles por tipo de empresa sin iniciativas Big Data.	28
Figura 4.4. Metodologías Ágiles por tipo de empresa con iniciativas Big Data en curso o en planificación.	28
Figura 4.5. Metodologías Ágiles por sector empresarial sin iniciativas Big Data.	29
Figura 4.6. Metodologías Ágiles por sector empresarial con iniciativas Big Data en curso o en planificación.	29
Figura 4.7. Metodologías Ágiles usadas en empresas con iniciativas Big Data en curso o en planificación.	30

ÍNDICE DE TABLAS

Tabla 1 Clasificación de Artículos por tipo de Investigación	15
Tabla 2 Distribución de Artículos de Metodologías Ágiles	15
Tabla 3 Distribución de Artículos de Big Data	16
Tabla 4 Publicaciones de Metodologías Ágiles aplicadas en proyectos de Big Data.....	17
Tabla 5 información general del Encuestado.....	17
Tabla 6 información general de las Empresas	18
Tabla 7 Tipo de desarrollo y cantidad de Información almacenada	19
Tabla 8 Herramientas de análisis de Datos	20
Tabla 9 Duración, Personal y Entrenamiento en Proyectos Ágiles	20
Tabla 10 Metodologías y frecuencia de liberación	21
Tabla 11 Herramientas y prácticas de Metodologías Ágiles	21
Tabla 12 Razones por las cuales se ha abandonado el uso de Metodologías Ágiles....	22
Tabla 13 Nivel de comprensión de Big Data.....	22
Tabla 14 Iniciativas de Big Data.....	23
Tabla 15 Metodología de Administración de Iniciativas Big Data	23
Tabla 16 ¿Quiénes y cómo desarrollan las Iniciativas Big Data?.....	24
Tabla 17 Estructuras de Datos y su Administración	24
Tabla 18 Motivación de uso de Big Data.....	25
Tabla 19 Aplicaciones de Big Data	26
Tabla 20 Uso de Big Data y Metodologías Ágiles según características de las organizaciones.....	27

INTRODUCCIÓN

Hoy en día, los términos Big Data y desarrollo ágil son usados en los ámbitos académico, gubernamental y empresarial; pero ¿cómo convergen estos términos?, ¿existen proyectos de Big Data desarrollados con Metodologías Ágiles?

En entornos empresariales, Big Data es considerado como un sucesor de la Inteligencia de Negocios; ambos se usan para responder preguntas analizando datos, pero se diferencian en la variedad, cantidad y velocidad de generación y análisis de datos.

En Investigaciones respecto al uso de Metodologías Ágiles en proyectos de Inteligencia de Negocios, se evidenció el aumento de la satisfacción de los interesados en un 60% y 70% [1]. Sin embargo, algunas empresas que iniciaron la implementación de proyectos de Big Data, no lo concluyeron[2]. No obstante, existe evidencia de un proceso evolutivo del uso de Metodologías Ágiles en proyectos de Big Data.

Este documento presenta el resultado de dos tipos de investigación, una revisión sistemática de literatura científica sobre la coexistencia de Metodologías Ágiles en proyectos Big Data; y un análisis descriptivo y cuantitativo del uso de Metodologías Ágiles y Big Data en empresas ecuatorianas.

El capítulo 1, contiene los antecedentes, descripción del problema, y determinación de objetivos generales y específicos.

El capítulo 2, presenta el marco teórico que permite la comprensión de los conceptos principales abordados en esta investigación.

El capítulo 3, describe los pasos metodológicos que se siguieron en la revisión sistemática de literatura y el estudio descriptivo y cuantitativo.

El capítulo 4, presentan los resultados de la revisión sistemática de literatura y el estudio descriptivo y cuantitativo.

Finalmente, se encuentran las conclusiones y recomendaciones del presente trabajo; además, los apéndices que contienen: un resumen de los artículos usados para la revisión de literatura sistemática, y el cuestionario usado para el estudio descriptivo y cuantitativo.

CAPÍTULO 1

GENERALIDADES

1.1 Antecedentes

Actualmente, vivimos en un ambiente de evolución tecnológica acelerada, la cual está cambiando constantemente nuestra forma de trabajar, incluyendo las herramientas que usamos. Entre los conceptos que han contribuido en los cambios tecnológicos empresariales de los últimos años están: Las Metodologías Ágiles y Big Data.

Por un lado, las Metodologías Ágiles aplicadas en el desarrollo de software han cambiado la manera de administrar el desarrollo y mantenimiento de los sistemas empresariales, usando como base el "Manifiesto Ágil". Las premisas del "Manifiesto Ágil" dictan: Reducir el ciclo de vida de los proyectos, generar productos con alta calidad, definir los requerimientos con claridad, incrementar la flexibilidad al cambio, y la satisfacción de los interesados del proyecto [1, 2].

Por otro lado, Big Data representa una oportunidad para el sector empresarial, ya que puede estimular el mercado, generar información para la dirección estratégica, mejorar el servicio al cliente, generar nuevos productos, y descubrir nuevos clientes y mercados. [4] Para lograrlo, Big Data usa datos de gran volumen, estructurados o no estructurados, de origen propio o externo; además, ejecuta procesamientos optimizados para solucionar los problemas de volumen y variedad [5].

Entre la literatura científica encontramos que las Metodologías Ágiles han sido aplicadas en mayor escala en empresas pequeñas y medianas [6], con un mayor aprovechamiento por la industria de desarrollo de software. No obstante, las empresas que incursionan en el uso de Big Data tienden a ser grandes. Cada vez más empresas incursionan en el uso de Metodologías Ágiles y Big Data; sin embargo, la forma idónea de fusionar estos dos conceptos aún no está definida.

1.2 Descripción del problema

Según recientes estudios realizados por Gardner, el uso de herramientas Big Data dentro de entornos empresariales se ha incrementado; aunque, el 59% de estos proyectos no llegan a su finalización [3]. Así también, este tipo de proyecto requiere una inversión económica importante, por lo que demanda un mayor control y resultados. Consecuentemente, vemos necesario el uso de una metodología que garantice la finalización del proyecto Big Data.

1.3 Objetivo General

Determinar la situación actual del uso de Metodologías Ágiles y Big Data dentro del sector empresarial ecuatoriano por medio de estrategias de ingeniería de software experimental.

1.4 Objetivos Específicos

- Clasificar publicaciones científicas de estudios de Metodologías Ágiles aplicadas a proyectos de Big Data.
- Identificar las Metodologías Ágiles más usadas en proyectos de Big Data.
- Explorar la opinión y comprensión de empresas ecuatorianas respecto al uso de Metodologías Ágiles y big dada.
- Identificar los sectores económicos ecuatorianos más interesados en el uso de Metodologías Ágiles y Big Data.

CAPÍTULO 2

MARCO TEÓRICO

El desarrollo de software ha usado principios ágiles desde el año 2001. Estos principios se han aplicado en diversidad de proyectos y ámbitos, entre ellos la inteligencia de negocios [2].

Por otro lado, conceptos como Big Data aplicada a los negocios se considera una evolución de los sistemas de inteligencia de negocios [7], se lo conoce también como BI (Business Intelligence) analítico, ciencia de datos (Data Science) y Analítica Rápida (Fast Analytics) [1].

A continuación, ampliaremos un poco más los conceptos de Big Data y Metodologías Ágiles.

2.1 Big Data

Los conceptos de Big Data se han aplicado en diferentes ámbitos, desde el científico, académico, hasta el empresarial; permitiendo así, el análisis de grandes volúmenes de datos a una velocidad antes no imaginados. [8]

En el mundo de los negocios, Big Data permite: realizar marketing más efectivo, generar nuevas oportunidades de ingreso económicos, brindar mejor servicio al cliente, mejorar la eficiencia de operaciones, avizorar estrategias competitivas sobre organizaciones rivales, entre otros beneficios [9].

Frankováa, define a Big Data como un *“gran volumen de información, a alta velocidad, con amplia diversidad, que requiere nuevas formas de procesar, para permitir mejorar la toma de decisiones, descubrir pistas relevantes que permitan optimizar los procesos”*. [2, 4]

2.1.1 Escenarios de Usos de Big Data en entornos empresariales

Podemos identificar tres tipos de aplicaciones de Big Data:

- Ciencia de datos: Campo interdisciplinario que tiene como objetivo encontrar patrones. Usa herramientas de minería de datos y análisis estadístico, algoritmos estadísticos y de

aprendizaje de máquina con la finalidad de generar modelos de datos que ayuden a describir y predecir datos [1, 10].

El procesamiento se realiza, comúnmente, en lote, de forma supervisada y con enfoque exploratorio. Entre sus etapas, normalmente, encontramos la selección, extracción y preparación de datos con un formato específico, luego el procesamiento de los datos por medio de un algoritmo experimental que busca responder una pregunta particular.

- Analítica Rápida: Este tipo de aplicación de Big Data se enfoca en adquirir y visualizar datos rápidamente [1, 11, 12] con la finalidad de modelar, estructurar y asociar información para responder a las necesidades de información.
- Inteligencia de negocios Analítica: Implementaciones tradicionales de Inteligencia de negocios usan herramientas Datawarehouse, las cuales están construidas con un enfoque relacional de datos, permitiendo un almacenamiento de datos estructurados. La inteligencia de negocios comúnmente usa fuentes de datos propias de las empresas. Sus etapas son: extraer, transformar y cargar datos para posteriormente generar información resumida que apoye a la toma de decisiones.

La inteligencia de negocios analítica añade a la inteligencia de negocios tradicional, el descubrimiento de patrones por medio de herramientas de ciencia de datos [9]. Este concepto puede también ser comprendido como una Analítica Rápida en un entorno empresarial.

Estos conceptos también podrían describirse como la evolución de Big Data dentro de una organización empresarial. Por ejemplo, una tienda de abarrotes con varias sucursales se pregunta ¿Cómo influye la distribución de los productos en percha en el recorrido y compras que realizan las personas que ingresan a sus tiendas?

Para responder esta pregunta, un grupo de científicos de datos inicia el levantamiento de información correspondiente, da un tratamiento adecuado a los datos y realiza un análisis exploratorio buscando patrones y comportamientos. Los hallazgos de este análisis, será el resultado del proceso de *ciencia de datos*.

Una vez conocidos los hallazgos, y dependiendo de la utilidad predictiva o funcional que brinden a la organización, es posible que se requieran nuevos análisis similares, donde se usará algoritmos probados experimentalmente, y optimizados de tal manera que permitan responder con rapidez a los requerimientos de información, convirtiéndose entonces en un proceso de *analítica rápida*. Sin embargo, esta información sigue generándose a demanda.

Por otro lado, si la información brindada por este proceso se demanda en temporadas, u otro tipo de regularidad; o se necesita vincular esos datos con otros propios de la organización, es posible conectar ese proceso con la *inteligencia de negocios*.

Algunos otros ejemplos de la aplicación de Big Data en empresas son:

- Identificar patrones de fraude en el manejo del stock.
- Obtener ideas que permitan tomar decisiones informadas.
- Detectar posibles fraudes en el uso de tarjetas de crédito.
- En el caso de telefonía, identificar sectores claves de inversión en infraestructura.
- Identificar relaciones y dependencias entre personas, productos, etc. que permitan predecir riesgos de enfermedades, fraudes, compras, etc.
- Mejorar y optimizar procesos de producción agrícola con la ayuda de sensores que generen información climática [9], entre otros.

2.2 Metodologías Ágiles de desarrollo de software

Las Metodologías Ágiles han cambiado la forma de administrar los proyectos de desarrollo de software. Todo empezó desde la publicación del “manifiesto ágil” en el año 2001, el cual propone las siguientes premisas:

- Reducir el tiempo del ciclo de vida
- Generar alta calidad de productos de software
- Clarificar los requerimientos
- Incrementar flexibilidad en cambios, y
- Satisfacer a los stakeholders [13, 14].

En base a la revisión literaria, las Metodologías Ágiles más usadas en proyectos Big Data son: Scrum, Programación Extrema y Kanban, las cuales se explican a continuación.

2.2.1 Scrum

Esta metodología tiene su nombre basado en un concepto deportivo del rugby, relacionado con la recuperación rápida del juego en el caso de una infracción menor [15]. Las bases de esta metodología son la flexibilidad y rapidez con las que puede integrar diferentes tipos de personas y procesos diversos continuos o traslapados [16], acepta el cambio como una constante y se prepara para responder a cambios inesperados.

Comprende tres etapas: El pre-juego, el desarrollo y el post-juego.

- El pre-juego consiste de dos fases: La planeación donde se define el sistema a desarrollar; y la arquitectura diseñada a un alto nivel.
- La etapa de desarrollo es como una caja negra donde se espera cualquier cosa. Todo puede cambiar durante el proceso de esta etapa.
- El post-juego inicia cuando se realizan entregables, en esta etapa se han aceptado las definiciones del producto y se realizan tareas de integración, pruebas y documentación.

Los roles en Scrum son:

- Scrum Master: Es el responsable de asegurar que el proyecto siga las prácticas de la metodología Scrum; interactúa con el cliente y el equipo; soluciona los problemas y mantiene al equipo productivo.
- Propietario del producto: Es el responsable oficial del proyecto; toma la decisión final respecto al producto.
- Cliente: Es el que participa en las tareas de definición del producto a desarrollar o mejorar.
- Equipo Scrum: Es el equipo conformado por programadores, arquitecto, y testers. Se caracteriza por ser auto-organizado y se enfoca en alcanzar las metas de cada sprint.

En la metodología Scrum se realiza un producto backlog, la estimación del esfuerzo y el desarrollo del producto se realizan en ciclos llamados sprints.

- En el Producto backlog se define todo lo que debe contener el producto final, el cual puede cambiar en el transcurso del proyecto.
- La estimación de esfuerzo la realiza el propietario del producto junto al equipo Scrum, y se calcula en base a la definición del producto Backlog.
- Los Sprint son ciclos cortos en los que se desarrolla software. Durante los sprint se deben realizar: una reunión de planeación, definir el sprint backlog, hacer una reunión diaria de scrum, y la reunión de revisión de Sprint [17].

2.2.2 Programación Extrema

Definida así porque es capaz de llevar sus principios y prácticas a un nivel extremo [17]. Requiere de una continua revisión y retroalimentación; permite cambiar las definiciones de los requerimientos en cualquier momento, y genera código sin características particulares de un programador, por lo que usa la programación en parejas.

Su ciclo de vida comprende cinco etapas: Exploración, planificación, iteración de entregables, producción, mantenimiento y muerte.

- La etapa de exploración se concentra en la descripción de las características que los clientes desean incluir en el producto entregable, estas son descritas en tarjetas de historias de usuario, cada historia asocia una característica. En esta etapa se prueba la posible arquitectura por medio del desarrollo de un prototipo. Esta fase puede tomar entre pocas semanas a pocos meses.
- En la Planificación se establecen las prioridades de las historias de usuario para el primer entregable; este proceso no debe demorar más de dos días. En dicha fase, los programadores calculan el esfuerzo y definen el cronograma, el cual normalmente no supera dos meses.
- La etapa de Iteración de entregables incluye varias iteraciones antes del primer entregable; para esto, se divide el alcance planificado en iteraciones de entre una a cuatro semanas, y en cada una de ellas se asignan las historias de usuario que corresponden a

dicha iteración. En la primera iteración se debe incluir toda la estructura. Las pruebas se basarán en las historias de usuario asignadas a dicha iteración.

- La fase de Producción comprende pruebas adicionales previo a la entrega del producto al cliente. Esta fase puede demorar entre una y tres semanas y puede incluir pequeños cambios de requisitos adicionales.
- La fase de mantenimiento puede incluir pequeños cambios y el ingreso de nuevos miembros al equipo, lo que permite desarrollar tareas de soporte al cliente.
- Finalmente, el proyecto muere cuando el cliente no tiene más historias para implementar.

Las personas y grupos con diferentes tipos de responsabilidades en proyectos ágiles con Programación Extrema son:

- Los Programadores desarrollan sus códigos lo más simple posible, y deben mantener una buena coordinación entre ellos.
- Los clientes escriben sus historias de usuario y definen las prioridades de implementación.
- Los testers o probadores definen las pruebas funcionales en base a las historias de usuario, y evalúan los resultados.
- Los Tracker o rastreadores evalúan y mejoran las estimaciones de tiempo y esfuerzo, y siguen el progreso de cada iteración evaluando si se ha alcanzado o no el objetivo.
- El Coach o entrenador guía a los miembros en el uso de la metodología.
- Los consultores son miembros externos con conocimientos específicos necesarios para el desarrollo del proyecto.
- El manager o gran jefe toma las decisiones principales [17].

Por otro lado, la metodología de Programación Extrema aplica algunos principios y prácticas, entre ellos están:

- Sus principios: rápida retroalimentación, simplicidad, cambio incremental, aceptación del cambio y trabajo de calidad [18].

- Sus prácticas: el juego de planificación, pequeñas entregas, diseño simple, programación en parejas, pruebas, refactorización, integración continua, código de propiedad común, sostenibilidad, clientes en sitio, metáforas y código estandarizado. [17]

2.2.3 Kanban Agile

Esta metodología está basada en los principios *lean*, tiene sus inicios en la industria japonesa de 1950, y busca la forma de mantener y mejorar la producción. En el caso de desarrollo de software con una metodología kanban, cada miembro del equipo tiene una actividad para trabajar a la vez; cuando lo termine automáticamente tendrá otra asignación. Con Kanban se tiene como objetivo comunicar las prioridades, proporcionar visibilidad del proceso y solucionar los cuellos de botella para un flujo constante de entregables [19].

Kanban tiene cinco principios básicos:

- Visualizar el flujo de trabajo
- Limitar el trabajo en curso
- Medir y gestionar el flujo de trabajo
- Hacer conocer las políticas del proceso
- Usar modelos para reconocer oportunidades [19]

El método Kanban busca eliminar el miedo al cambio, especialmente los temores iniciales de aceptar y respetar los roles, responsabilidades y liderazgo en todos los niveles de la organización [19, 20]

2.3 Metodologías Ágiles aplicadas a proyectos de Inteligencia de Negocios

Las Metodologías Ágiles no solo se han aplicado en proyectos de desarrollo de sistemas transaccionales, sino también en proyectos de Inteligencia de negocios.

Para la implementación exitosa de proyectos de Big Data con Metodologías Ágiles los proyectos deben ser pequeños y rápidos [1].

Uno de los autores que ha estudiado el uso de Metodologías Ágiles en proyectos de inteligencia de negocios es Hughes, quien definió el Data Warehousing Ágil así:

“Un conjunto de principios, comportamientos, y técnicas que permiten responder a los clientes más rápido con información de calidad, por medio de proyectos BI que inicien rápido, involucren a los altos directivos, actualicen regularmente el plan de generación de información, y aporten valor en la información producida de manera incremental” [21, 22].

Entre los aportes científicos de Hughes se encuentra un “Manifiesto ágil” para Data Warehousing, el cual incluye las siguientes premisas:

- La entrega rápida de resultados al sponsor supera la perfección técnica.
- La gestión del riesgo supera la eliminación de la incertidumbre.
- La evolución de los datos supera la reutilización del código.
- La apropiación tecnológica supera la monocultura de la infraestructura [21, 22].

Data Warehousing Ágil es una mezcla de Scrum y Programación Extrema, en el que usa Scrum para la planificación y seguimiento del proyecto, y programación extrema para la generación de código y gestión de los cambios.

2.4 Metodologías Ágiles aplicadas a proyectos de Big Data

Los estudios con encuestas respecto a Metodologías Ágiles aplicadas a grandes empresas, ha sido mayormente reportada en países de Europa y Norte América [23]. Dentro de América latina, en Argentina [24], Colombia [25, 26], México [27], y Uruguay se han presentado publicaciones de encuestas enfocadas principalmente en empresas pequeñas y medianas [28]. Por otro lado, en Chile [29], Perú [30] y Ecuador [31] se han realizado casos de estudio de su aplicación.

En Australia, Reino Unido, Estados Unidos, Japón, Alemania, India, México y Brasil se han presentado estudios del uso de Big Data en empresas grandes; [32] en países de América latino se han publicado artículos científicos referentes a Big Data mayormente en Brasil, México, Chile, Argentina, Colombia y Cuba.

No obstante, aún es limitado el estudio del uso de Metodologías Ágiles en proyectos de Big Data; en el capítulo 4 ampliamos información referente a la documentación científica existente respecto a la fusión estos dos conceptos.

CAPÍTULO 3

MARCO METODOLÓGICO

La investigación realizada en esta tesis está basada en las guías de la Ingeniería de Software Experimental. La metodología de investigación comprende 2 fases: 1) Una revisión sistemática de literatura y 2) Un estudio descriptivo y cuantitativo de la adopción de Metodologías Ágiles y Big Data, los cuales se explican a continuación:

3.1 Revisión Sistemática de Literatura.

Para el desarrollo de esta revisión sistemática de literatura usamos las guías de Petersen y Kuzniarz [33], las cuales sugieren realizar: la definición de un objetivo, la formulación de preguntas de investigación, y la definición de un procedimiento de búsqueda de información y presentación de resultados.

Objetivo de revisión sistemática:

- Clasificar publicaciones relacionadas a estudios de Metodologías Ágiles de desarrollo de software aplicadas a proyectos de Big Data.

Preguntas de investigación de revisión sistemática:

- ¿Qué estudios se han realizado en los últimos 5 años, en los que se utilicen Metodologías Ágiles de desarrollo de software en proyectos de Big Data?

Procedimiento de búsqueda y presentación de resultados de Revisión sistemática:

En la siguiente sección se detalla el procedimiento de búsqueda aplicado. Los resultados se presentan en el capítulo 4.

3.1.1 Búsqueda de artículos científicos

La búsqueda de artículos científicos se realizó en 3 bases de datos científicas: ScienceDirect, Springer, y GoogleScholar. La búsqueda se enfocó en artículos científicos publicados entre los años 2012 y 2017 en idioma inglés. El criterio de búsqueda incluyó 3 componentes:

- Metodologías Ágiles,
- Big Data,
- Tipo de investigación (aplicación a un caso de estudio).

Esta búsqueda se efectuó varias veces entre noviembre de 2016 y abril de 2017. La búsqueda se ejecutó en el contenido de los artículos científicos. El query utilizado es el siguiente:

("scrum" or "extreme programming" or "Agile Data" or "Crystal" or "kanban" or "Agile Software") and ("Big Data" or "data science" or "analytic") and ("case study")

Con Google Scholar, el query completo no generó resultados, por lo que se procedió a ejecutar varias consultas incluyendo una metodología ágil a la vez. Un ejemplo se muestra a continuación:

("case study") and ("scrum") and ("Big Data" or "data science" or "analytic")

De la misma manera, hicimos búsquedas parciales por tipo de proyecto y se repitió para todos los tipos, Un ejemplo:

("scrum" or "extreme programming" or "Agile Data" or "Crystal" or "kanban" or "Agile Software") and ("Big Data") and ("case study")

Como resultado de búsqueda obtuvimos lo siguiente:

- En la base de datos Science Direct, 108 resultados, 18 de los cuales no correspondían a artículos científicos, restando 90 artículos para el análisis.
- En la base de datos Springer, 68 resultados, 7 de los cuales no correspondían a artículos científicos, restando 61 artículos para el análisis.
- En la base de datos Google Scholar, con ayuda de las búsquedas parciales obtuvimos 14 artículos científicos para el análisis.

3.1.2 Clasificación de Artículos obtenidos

Los 165 artículos obtenidos del proceso de búsqueda fueron clasificados por: tipo de investigación, Metodologías Ágiles, y tipo de proyectos Big Data.

El proceso de clasificación empezó con la lectura de los resúmenes de los artículos, y de no encontrar información se buscó dentro del contenido.

Aunque en el filtro de búsqueda se especificó *casos de estudio*, se encontraron los siguientes tipos: entrevistas, mapeos sistemáticos, revisiones de literatura, casos de estudio, modelos conceptuales, y marcos de trabajo. El detalle de la clasificación de los 165 artículos se adjunta en el Apéndice A.

3.2 Estudio descriptivo y cuantitativo

Este estudio tiene como objetivo rescatar la opinión del encuestado respecto a las Metodologías Ágiles de desarrollo de software y big data, así como de identificar la orientación de las empresas donde laboran respecto a estos conceptos.

3.2.1 Fuentes bibliográficas

Para la elaboración del cuestionario se usaron principalmente las siguientes fuentes, [2, 34, 35, 36] las cuales corresponden a estudios respecto al uso de Metodologías Ágiles y Big Data en entornos empresariales.

3.2.2 Cuestionario

El cuestionario usado para este estudio contiene 59 preguntas distribuidas en 5 secciones:

- Sección 1: Información General del Encuestado;
- Sección 2: Información General de la Empresas;
- Sección 3: Estado Actual de los datos de las empresas y su capacidad de Análisis;
- Sección 4: Uso de Metodologías Ágiles; y
- Sección 5: Big Data, comprensión y utilización.

El cuestionario completo se adjunta en el Apéndice B.

3.2.3 Población objetivo

La población objetivo corresponde a profesionales que laboran en empresas grandes o medianas, dentro del área de tecnología de la información. Con la finalidad de encontrar al sujeto adecuado para contestar todas las interrogantes, se orientó la búsqueda a profesionales con cargos de gerencia o jefaturas tecnológicas dentro de la organización.

Para el proceso de búsqueda se utilizó una lista de las 500 empresas más rentables del Ecuador, publicado por la revista Vistazo en el año 2016 [37]. Posteriormente, se buscó e invitó a través de la red de LinkedIn a profesionales de dichas empresas que se ajusten al perfil objetivo.

3.2.4 Pre-Test

Con la finalidad de evaluar la interacción y funcionamiento en campo del cuestionario diseñado, se realizó un pre-test. El pre-test consistió en ejecutar el cuestionario con tres profesionales que cumplieran el perfil objetivo, y al finalizar el proceso compartieron sus observaciones y recomendaciones.

Entre las observaciones y recomendaciones recibidas tuvimos:

- No solicitar nombre de encuestado o nombre de empresa para mantener la confidencialidad. Dichas preguntas fueron retiradas.
- Evitar preguntas que parezcan evaluaciones académicas, ya que esto intimida a los encuestados.
- Cambiar o incluir escalas de Likert para algunas preguntas.

3.2.5 Ejecución

La recopilación de los datos para este estudio se realizó de forma presencial para un 36% de los encuestados; con el 18% se mantuvo comunicación vía Skype; y con el 48% se mantuvo comunicación vía LinkedIn para dar seguimiento al llenado del cuestionario a través de la plataforma Google Forms. Con este proceso, se obtuvo 31 cuestionarios válidos para el estudio.

CAPÍTULO 4

RESULTADOS Y ANÁLISIS DE DATOS

En este capítulo se presentan los resultados de cada una de las fases de la investigación realizada.

4.1 Resultados de la Revisión sistemática de literatura

Por medio del proceso de revisión sistemática se obtuvo 165 artículos. Los artículos encontrados incluyen varios tipos de investigación como: entrevistas, marcos de trabajo, revisiones de literatura, y modelos conceptuales [80, 86, 187, 193]. A continuación, un resumen de los hallazgos en la Tabla 1.

Tabla 1 Clasificación de Artículos por tipo de Investigación

Año	Entrevistas	Revisiones Literarias	Mapeos Sistemáticos	Caso de Estudio	Marcos de trabajo	Modelo Conceptuales
2012	3	4	1	12	6	4
2013	4	1	1	8	4	3
2014	5	4	1	15	4	5
2015	6	7	3	22	6	9
2016	14	9	6	24	11	18
2017	0	1	1	8	1	3
Total	32	26	13	89	32	42

Adicionalmente, se presenta en las tablas 2 y 3 un resumen de los resultados obtenidos por criterio de búsqueda, donde 41 artículos incluyeron información respecto a Metodologías Ágiles y 55 respecto a Big Data.

Tabla 2 Distribución de Artículos de Metodologías Ágiles

Año	Resultado por criterio de búsqueda							Resultado por análisis de contenido	
	Scrum	XP	Agile Data	Kanban	Crystal	Agile Software	Total	Metodologías Ágiles	Relación
2012	7	5	0	13	0	7	32	7	22%
2013	5	6	1	6	4	9	31	5	16%
2014	6	5	0	12	0	8	31	5	16%
2015	13	8	0	18	4	17	60	13	22%
2016	22	8	2	9	9	29	79	8	10%
2017	2	2	0	1	1	1	7	3	43%
Total	55	34	3	59	18	71	240	41	17%

Tabla 3 Distribución de Artículos de Big Data

Año	Resultado por criterio de búsqueda				Resultado por análisis de contenido	
	Big Data	Data Science	Analytic	Total	Big Data	Relación
2012	3	2	19	24	8	33%
2013	2	0	11	13	2	15%
2014	6	1	19	26	5	19%
2015	16	4	30	50	15	26%
2016	25	7	26	58	17	29%
2017	8	3	11	22	8	32%
Total	60	17	116	193	55	28%

Al clasificar los artículos según el criterio "Metodologías Ágiles", se buscó las metodologías: "Scrum", "Extreme Programming", "Crystal", y "kanban"; además, se incluyó en la búsqueda los criterios: "Agile Software" y "Agile Data", con la finalidad de obtener artículos con referencia a otras Metodologías Ágiles.

Entre los artículos resultantes se obtuvo estudios del uso de Metodologías Ágiles aplicados a distintos campos como: mercadotecnia, innovación, desarrollo de productos, logística, producción, entre otros. Ejemplo de estos casos son los artículos referentes a procesos industriales y logísticos [52,199], los cuales mencionan el uso de "Kanban" y "Crystal". En la Tabla III se presenta un resumen de lo indicado.

Por otro lado, al clasificar los artículos según el criterio de interés "Big Data", consideramos buscar expresiones como: "Big Data", "Data Science", y "Analytic".

Del universo de artículos obtenidos en la búsqueda, sólo 5 artículos incluyen ambos criterios ("Big Data" y "Metodologías Ágiles"). De estos artículos resultantes, cuatro se enfocan principalmente en Big Data, y sólo uno menciona el uso de Metodologías Ágiles (sin brindar muchos detalles al respecto). Uno de estos artículos [173] se distingue por la comparación y análisis de metodologías aplicadas a proyectos de Big Data; sin embargo, no define el o los tipos de proyectos Big Data comparados; además, de que fueron desarrollados en un ambiente académico con datos ficticios.

A continuación, se presenta un resumen de los artículos que incluyeron los conceptos de "Big Data" y "Metodologías Ágiles".

Tabla 4 Publicaciones de Metodologías Ágiles aplicadas en proyectos de Big Data

Artículo	Año	Ámbito	Título	Metodología Ágil Aplicada	Técnica de Big Data Aplicada
[71]	2015	Médico	<i>Practical use of medical terminology in curriculum mapping</i>	<i>Programación Extrema</i>	Análisis visual y de asociación de curriculum
[81]	2015	Desarrollo de Software	<i>Selecting Video Analytics using Cognitive Ergonomics: A Case Study for Operational Experimentation</i>	No se especifica	Inteligencia de Video
[184]	2016	Académico - Escolar	<i>Big Data Comes to School: Implications for Learning, Assessment, and Research</i>	No se especifica	Minería de datos de educación
[106]	2017	Desarrollo de Software	<i>Creating an invalid defect classification model using text mining on server development</i>	Scrum	Minería de texto y Aprendizaje de máquina supervisado
[173]	2017	Académico – Universitario	<i>Comparing Data Science Project Management Methodologies via a Controlled Experiment</i>	Scrum - Kanban	Análisis geográfico, Algoritmos de aprendizaje de máquina, y minería de reglas de asociación.

4.2 Resultados del Estudio descriptivo y cuantitativo

Se obtuvieron 31 cuestionarios válidos. A continuación, se presentan los resultados divididos por las secciones del cuestionario.

4.2.1 Sección 1: Información General del Encuestado

La tabla 5 muestra la demografía de los encuestados. La mayoría de ellos son hombres con edad entre los 35 y 44 años; poseen una maestría como máximo grado de educación; y ocupan un cargo de liderazgo dentro de la empresa que laboran.

Tabla 5 información general del Encuestado

GÉNERO	
Femenino	19,35%
Masculino	80,65%
EDAD DE ENCUESTADO	
25-34 años de edad	16,13%
35-44 años de edad	61,29%
45-54 años de edad	19,35%
55-64 años de edad	3,23%
MÁXIMO GRADO DE EDUCACIÓN DEL ENCUESTADO	
4th nivel	64,51%
3rd nivel	29,03%
Otros	6,45%
CARGO DENTRO DE LA ORGANIZACIÓN	
Director / Gerente / Jefe TI	29%
Líder de Proyecto TI	23%
Arquitecto de Software	13%
Analista de TI	29%
Ingeniero de pruebas de Software	6%

4.2.2 Sección 2: Información General de la Empresa

La tabla 6 presenta la información demográfica de las organizaciones en las cuales trabajan los encuestados.

Tabla 6 información general de las Empresas

TAMAÑO DE EMPRESAS	
Grandes	87,10%
Medianas	12,90%
EDAD DE EMPRESAS	
6-10 años	12,90%
11-15 años	6,45%
16-20 años	12,90%
21-25 años	3,23%
>25 años	64,52%
TIPO DE ORGANIZACIÓN	
Privada	54,84%
Pública	38,71%
Pública & Privada	6,45%
TIPO DE MERCADO QUE ATIENDE	
Mercado minorista	61,29%
Mercado mayorista	35,48%
Mercado minorista & mayorista	3,23%
ACTIVIDAD ECONÓMICA DE EMPRESAS	
Servicio Publico / Gobierno Local / Gobierno Central	32,26%
Agro industrial	16,13%
Banca / Finanzas / Seguros	16,13%
Comercio al por mayor y menor	16,13%
Educación	6,45%
Fabricación	3,22%
Militar	3,22%
Salud	3,22%
Telecomunicaciones	3,22%

Como se puede observar, la mayoría de organizaciones son grandes (con más de 200 empleados), y con larga trayectoria (más de 25 años). Un gran porcentaje de estas organizaciones se enfoca en el mercado minorista.

Los sistemas que se desarrollan son principalmente de procesamiento de transacciones (e.g. nómina o inventario); aplicaciones en tiempo real (e.g. sistemas de gestión de la Producción); y de gestión de la información (e.g. BI). Además, el 65% de las organizaciones tiene más de cinco años de información almacenada lista para analizar (ver tabla 7).

Tabla 7 Tipo de desarrollo y cantidad de Información almacenada

Aplicaciones de procesamiento de Transacciones	80,65%
Aplicaciones en tiempo real	67,74%
Aplicaciones de administración de la Información	64,52%
Aplicaciones Web o Carros de compra	41,94%
Aplicaciones en dispositivos	35,48%
Sistemas multimedia	32,26%
EDAD DE INFORMACIÓN EN AÑOS	
<=5 años	35,48%
6-10 años	38,71%
11-15 años	12,90%
16-20 años	9,68%
>20 años	3,23%

4.2.3 Sección 3: Estado actual de sus Datos y capacidad de Análisis

Los encuestados reportaron estar de medianamente a totalmente satisfechos con las herramientas que usan para analizar datos, y del tiempo que toma obtener la información (ver figuras 4.1 y 4.2).

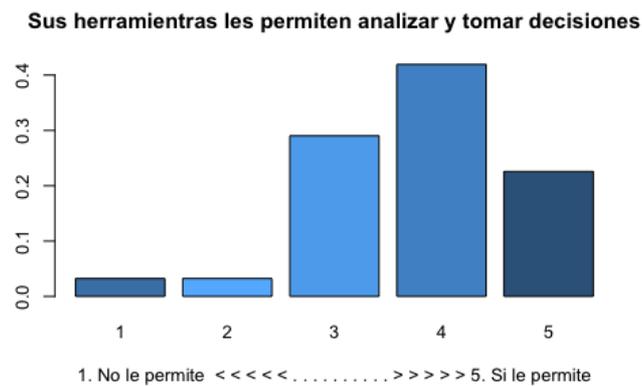


Figura 4.1. *Percepción de capacidad de análisis y toma de decisiones que ofrecen las herramientas tecnológicas de la Empresa.*

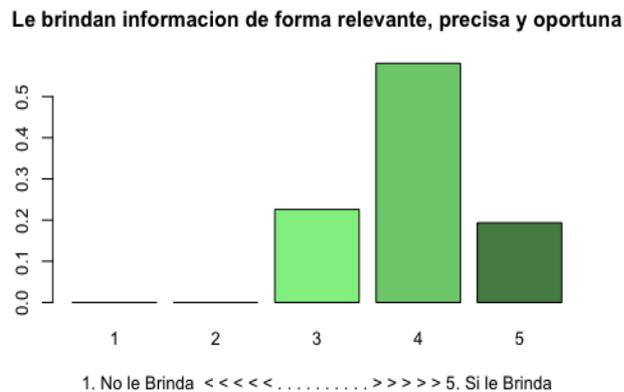


Figura 4.2. *Percepción de relevancia, precisión y oportunidad de tomar decisiones con la información disponible.*

Por otro lado, la tabla 8 muestra que el 61,28% de las organizaciones tienen 5 o menos personas realizando análisis de información; y el 74% usa hojas electrónicas para esta actividad.

Tabla 8 Herramientas de análisis de Datos

PERSONAS QUE REALIZAN ANÁLISIS DE INFORMACIÓN	
<5 personas	61,28%
6-10 personas	16,13%
11-15 personas	3,23%
16-20 personas	3,23%
>20 personas	16,13%
HERRAMIENTAS DE ANÁLISIS DE DATOS GENERAL	
Hojas Electrónicas	74,19%
BI	58,06%
Reportes Impresos	51,61%
Software estadístico	51,61%
Cuadros de Mando	41,94%
Otros	9,68%

4.2.4 Sección 4: Metodologías Ágiles

El 48% de los encuestados reportaron no haber usado Metodologías Ágiles. El 52% que reportó haber usado Metodologías Ágiles, compartieron hechos y su percepción de utilidad. A continuación, exponemos los hallazgos (ver tabla 9).

- El 88% de los proyectos ágiles duraron menos de un año; el 82% han involucrado 9 o menos personas en el equipo ágil; y el 69% no recibió capacitación referente a la metodología ágil usada.

Tabla 9 Duración, Personal y Entrenamiento en Proyectos Ágiles

DURACIÓN DE PROYECTOS ÁGILES	
<3 meses	25%
3-6 meses	38%
7-12 meses	25%
1-3 años	6%
>3 años	6%
PERSONAL EN PROYECTOS ÁGILES	
1-3 personas	38%
3-9 personas	44%
10-20 personas	6%
>20 personas	12%
TIEMPO DE ENTRENAMIENTO EN METODOLOGÍAS ÁGILES	
Menos de un día	6,25%
2-3 días	12,50%
4-6 días	12,50%
No tuvieron entrenamiento	68,75%

- La tabla 10 muestra que el 68,75% de los encuestados usaron Scrum; el 25% Kanban; y el 18,75% Programación Extrema. El 94% de los proyectos ágiles tuvieron liberaciones mensuales o en menor rango de tiempo.

Tabla 10 Metodologías y frecuencia de liberación

METODOLOGÍAS ÁGILES	
Scrum	68,75%
Kanban	25,00%
Programación Extrema	18,75%
Desarrollo de Software Adaptativo	6,25%
Crystal	6,25%
Metodología de desarrollo dinámico de Sistemas	6,25%
Desarrollo dirigido por características	6,25%
FRECUENCIA DE LIBERACIÓN	
Cada semana o menos	31%
Cada 2 semanas	6%
Cada mes	57%
Cada 3 a 6 meses	6%

- La práctica que consideran más valiosa es “Equipos auto-organizados”, y las herramientas más usadas son las de administración de proyectos ágiles (ver tabla 11).

Tabla 11 Herramientas y prácticas de Metodologías Ágiles

PRÁCTICAS ÁGILES	
Equipos auto-organizados	41,94%
Ciclos de desarrollo cortos	32,26%
Fácil acceso a usuarios expertos	29,03%
Construcciones Frecuentes	22,58%
Mejoramiento Reflexivo	19,35%
Comunicación Frecuente e informal	16,13%
Programación en pareja	12,90%
HERRAMIENTAS PARA LA ADMINISTRACIÓN ÁGIL	
Administración de proyectos ágil	68,75%
Pruebas funcionales automatizadas	56,25%
Sistemas de seguimiento de Errores	56,25%
Software de administración de la configuración	50,00%
Pruebas unitarias automatizadas	37,50%
Plataforma de integración de Herramientas	31,25%
Construcciones automatizadas	25,00%
Documentación de código automatizado	25,00%
Pruebas de regresión automatizadas	18,75%
HERRAMIENTAS USADAS PARA LA ADMINISTRACIÓN DE PROYECTOS ÁGILES	
Agilo for Scrum	31,25%
VersionOne	25,00%
Targetprocess	18,75%
Atlassian Jira + Agile	12,50%
SprintGround	12,50%
Pivotal Tracker	12,50%
Active Collact	6,25%

DIAGRAMAS DE DISEÑO USADOS

Diagramas de Flujo	32,26%
Diagramas de Clase	32,26%
Diagramas de Actividad	29,03%
Diagramas de flujo de trabajo	25,81%
Diagramas de secuencias	22,58%
Diagramas entidad relación	19,35%
Diagramas de casos de uso	19,35%
Diagramas de estado	19,35%
Especificación de casos de uso	16,13%

- El 56% de los encuestados que dijeron haber usado Metodologías Ágiles, han abandonado su uso, las razones se listan en la tabla 12.

Tabla 12 Razones por los cuales se ha abandonado el uso de Metodologías Ágiles

Ágiles o no, necesitamos administrar el alcance del proyecto, la planificación, los presupuestos, las estrategias y los informes. Agile no es la respuesta.	31%
Las rígidas políticas, procesos y metodologías de nuestra empresa no se ajustan a las metodologías Agile. Agile no va a funcionar aquí.	13%
Las Metodologías Ágiles sólo son para usuarios experimentados, motivados y con capacidad de organizarse solos, sin necesidad de una supervisión cercana. Así que Agile no es para nosotros.	6%
Las Metodologías Ágiles sólo funciona para pequeños proyectos y nuestros proyectos son grandes.	6%
Las Metodologías Ágiles requieren la ubicación conjunta de los participantes del proyecto y nuestro personal está geográficamente disperso.	6%

4.2.5 Sección 5: Big Data

- El 55% de los encuestados dijo “Comprendo lo que significa Big Data, pero no a nivel de detalle”; sin embargo, pocos consideran que Big data puede ser usado cuando se “requiere mejorar el tiempo de procesamiento” (ver tabla 13).

Tabla 13 Nivel de comprensión de Big Data

NIVEL DE AUTO COMPRESIÓN DE BIG DATA	
No he oído hablar de Big Data	9,68%
He oído hablar sobre Big Data, pero no sé lo que significa	6,45%
Comprendo lo que significa Big Data, pero no a nivel de detalle	54,84%
Comprendo en detalle lo que significa Big Data	29,03%
NIVEL DE COMPRESIÓN REAL DE CONCEPTOS DE BIG DATA	
Big Data es un término usado para describir conjuntos de datos muy grandes de origen interno y externo de las organizaciones.	64,52%
Big Data es un término usado para describir conjuntos de datos complejos como los provenientes de medios sociales, video, fotos, texto no estructurado, o los recolectados por dispositivos.	61,29%
La manipulación de conjuntos de datos muy grandes a menudo requiere de aplicaciones que se ejecuten en paralelo en decenas, cientos o incluso miles de servidores.	45,16%

Big Data se usa cuando se requiere realizar procesamiento rápido para solucionar problemas de volumen y variedad de datos.	41,94%
Big Data se usa cuando se requiere mejorar tiempo de procesamiento.	35,48%

- Tal como se observa en la tabla 14, el 61% de las organizaciones no han incurrido aún en proyectos de Big Data; pero el 39% si, ya sea en curso o en etapa de planeación.

Tabla 14 Iniciativas de Big Data

No tienen iniciativas en progreso o planeadas	61,29%
Tienen iniciativas en etapa de planificación	22,58%
Tienen iniciativas en curso	16,13%
ETAPA DE PROYECTOS BIG DATA	
Está en etapa de desarrollo	31,25%
Se ha desarrollado el plan de trabajo	18,75%
Se ha identificado el grupo de datos a analizar	18,75%
Se están evaluando los resultados obtenidos	18,75%
Se está en el proceso de recolección de datos	12,50%

Con relación a las organizaciones que indicaron tener iniciativas Big Data en curso o en etapa de planeación, compartimos los siguientes hallazgos.

- El 56% de las empresas comparten el liderazgo de los proyectos de Big Data entre áreas de negocio y TI.
- El 53% para administrar sus proyectos Big Data usa o usará las mismas Metodologías empleadas para desarrollar software (ver tabla 15).

Tabla 15 Metodología de Administración de Iniciativas Big Data

LIDERAZGO INTERNO DE PROYECTOS BIG DATA	
Colaboración entre las áreas de negocios y TI	56,25%
Principalmente dirigido por TI	18,75%
Principalmente dirigido por las áreas de negocios	18,75%
No está definido	6,25%
¿COMO PIENSA MEDIR EL ÉXITO DE SUS INICIATIVAS BIG DATA?	
Cualitativamente con datos de desempeño empresarial	82,35%
Cuantitativamente con datos de desempeño empresarial	52,94%
Cuantitativamente con datos de desempeño de TI	64,71%
Cualitativamente con datos de desempeño de TI	5,88%
No está definido	5,88%
¿ADMINISTRA SUS INICIATIVAS BIG DATA USANDO LA MISMA METODOLOGÍAS DE ADMINISTRACIÓN DE PROYECTOS QUE PARA EL DESARROLLO DE SOFTWARE?	
Si, la misma	52,94%
No, otra	35,29%
No está definido	11,76%

- La tabla 16 muestra quiénes y cómo se desarrollan los proyectos de Big Data. El 50% de los proyectos Big Data son administrados principalmente por terceros bajo supervisión de personal de la empresa; el 71% usa o usará herramientas desarrolladas por terceros; y el 50% usa o usará servicios de vendedores de productos para el análisis de datos.
- El 83% de las empresas con iniciativas Big Data en curso o en planeación, no planean contratar científicos de datos.
- El 47% de los encuestados no conocen el monto presupuestado del proyecto.

Tabla 16 ¿Quiénes y cómo desarrollan las Iniciativas Big Data?

¿COMO HA ADMINISTRADO SUS INICIATIVAS BIG DATA?	
Principalmente con terceras partes bajo nuestra dirección y supervisión	50,00%
En mayoría internamente, con ayuda de terceras partes	25,00%
Todo internamente	25,00%
¿COMO OBTIENE LAS HABILIDADES DE CIENCIA DE DATOS DENTRO DE SU ORGANIZACIÓN?	
Desarrollado por terceras partes	70,59%
Desarrollado internamente	29,41%
No lo sé aun	17,65%
¿COMO PLANEA REALIZAR ANÁLISIS DE DATOS?	
Con la ayuda de vendedores de Productos	50,00%
Formando analistas de datos	33,33%
Contratando analistas de datos	27,78%
Con la ayuda de empresas con experiencia	16,67%
Con pasantes o recién graduados de estudios Doctorales	11,11%
Con pasantes o recién graduados de estudios de maestrías	0,00%
HERRAMIENTAS DE ANÁLISIS DE DATOS CON BIG DATA	
No lo sé aun	41,18%
Software especializado en análisis	29,41%
Productos de visualización / Cuadros de Mando	29,41%
Paquete estadístico o matemático	23,53%
Software de análisis espacial o geográfico	5,88%

- El 76,47% de las organizaciones desean procesar “Datos Transaccionales”; y el 94,12% pretenden usar “Bases de datos Relacionales / Data Warehouse” en sus iniciativas Big Data (ver tabla 17).

Tabla 17 Estructuras de Datos y su Administración

TIPO DE ESTRUCTURA DE DATOS A TRATAR CON TECNOLOGÍAS BIG DATA		
	Actualmente	En 3 años
Datos Transaccionales	76,47%	0,00%
Datos Geoespaciales	52,94%	23,53%
Datos de Imágenes	41,18%	17,65%
Data de sensores en máquinas o dispositivos	35,29%	35,29%
Datos de redes Sociales	23,53%	47,06%
Correos / documentos (no estructurados)	23,53%	41,18%
Clickstream	17,65%	35,29%
Datos Científicos	12,50%	29,41%

Voz / Audio	0,00%	41,18%
¿QUÉ ESTRUCTURA DE DATOS ES DE SU INTERÉS EN PROYECTOS BIG DATA?		
Bases de datos Relacionales / Data Warehouse		94,12%
Imágenes		47,06%
XML		41,18%
Archivos Plano		35,29%
Textos no estructurados		17,65%
Formato Json		17,65%
Web Semantica		11,76%
Audio / Video		11,76%
Series-Logs		6,25%

- El 70,59% de los encuestados dijeron que con sus iniciativas Big Data buscan “analizar datos desde diferentes fuentes internas y externas”; La mayoría de las organizaciones quieren analizar sus Operaciones y el Mercado; y su principal motivación es “Obtener nuevas ideas” (82,35%) (ver tabla 18).

Tabla 18 Motivación de uso de Big Data

¿QUE LO HA HECHO CONSIDERAR BIG DATA?		
	Actualmente	En 3 años
Analizar datos de diferentes fuentes	70,59%	29,41%
Analizar datos de tamaño no conocido	52,94%	0,00%
Analizar nuevos tipos de datos	41,18%	17,65%
Analizar datos de transmisión y comunicación	41,18%	11,76%
Analizar datos 1TB - 100TB	31,25%	29,41%
Analizar datos <1TB	23,53%	29,41%
Analizar datos > 1PB	5,88%	23,52%
Analizar datos 100TB - 1PB	0,00%	35,29%
ANÁLISIS DE FUNCIONES DEL NEGOCIO		
	Análisis Actual	Enfoque de Big Data
Operaciones	67,74%	62,50%
Análisis del Mercado	61,29%	56,25%
Servicio al cliente	51,62%	43,75%
Tecnología de la Información	51,61%	37,50%
Administración de Productos	41,94%	37,50%
Recursos Humanos	38,71%	18,75%
Administración de Riesgo	25,81%	18,75%
Administración del Fraude	25,81%	12,50%
E-Commerce	16,13%	12,50%
Otros	12,90%	6,25%
Marketing Digital	3,23%	43,75%
¿QUE BENEFICIOS ANALÍTICOS ESPERA OBTENER DE USAR BIG DATA?		
Obtener nuevas ideas		82,35%
Información del mercado e ideas		64,71%
Analizar datos en menor tiempo		52,94%
No lo sé		5,88%

- La función analítica de mayor importancia para los encuestados es la “Visualización de Datos”. Seguida por el análisis de redes sociales y análisis de algoritmos. Además, la mayoría de los encuestados dijo que sus aplicaciones Big Data se integrarán con “Aplicaciones de negocio” (ver tabla 19).
- Además, el 88,24% de las empresas encuestadas están abordando desafíos de procesamiento de datos “En tiempo real”; el 58,82% busca “Integrar datos propios con otros provenientes de distintas fuentes como redes sociales”; y, el 47% de las empresas encuestadas dijeron usar o que usarán “Nube Privada” para sus iniciativas Big Data; y el 29% “almacenamiento local”.

Tabla 19 Aplicaciones de Big Data

¿CUALES FUNCIONES ANALÍTICAS SON DE MAYOR IMPORTANCIA PARA USTED?	
Visualización de Datos	82,35%
Análisis de redes Sociales	41,18%
Análisis de algoritmos Avanzados	41,18%
Ejecución rápida de algoritmos	41,18%
Análisis de sentimientos	29,41%
Aprendizaje Automático	23,53%
Análisis de Texto	23,53%
Ejecutar algoritmos con más datos	17,65%
Modulación of Tópicos	11,76%
SUS APLICACIONES DE BIG DATA SERÁN INDEPENDIENTES O SE INTEGRARÁN O EMBEBERÁN EN OTROS SISTEMAS IMPORTANTES PROPIOS COMO:	
Aplicaciones del Negocio	58,82%
Aplicaciones de Procesos	11,76%
Aplicaciones de reglas de Negocio	11,76%
No se integrará con otros sistemas	41,18%
¿QUE DESAFÍOS DE TRATAMIENTO DE DATOS SE ESTÁN ABORDANDO CON BIG DATA?	
Uso de datos en Tiempo Real	88,24%
Integrar variedad de datos (propios, redes sociales, etc.)	58,82%
Acceder a un mayor rango de datos históricos	47,06%
Acceso a datos con mayor granularidad	47,06%
Comprender datos no estructurados	23,53%

4.2.6 Análisis de Resultados

De la muestra seleccionada para este trabajo de investigación, podemos notar que la tercera parte (36%) no han usado Metodologías Ágiles ni tienen proyectos de Big Data en curso o planeación; y que alrededor del 25% de los encuestados han usado Metodologías Ágiles y tienen proyectos de Big data. Sin embargo, más de la mitad de ellos abandonó el uso de metodologías

ágiles. La principal razón del abandono de las Metodología Ágiles es que no consideran Ágil sea para administrar sus proyectos.

La tabla 20 resume el uso de Metodologías Ágiles y Big Data según las características de las empresas.

Tabla 20 Uso de Big Data y Metodologías Ágiles según características de las organizaciones.

	No hay iniciativas de Big Data				Si hay iniciativas Big Data en curso o en planificación			
	No han usado M. Ág.	Si han usado M. Ág.	Siguen usando M. Ág.	Dejaron de usar M. Ág.	No han usado M. Ág.	Si han usado M. Ág.	Siguen usando M. Ág.	Dejaron de usar M. Ág.
EDAD DE EMPRESAS								
6-10 años	3,23%	-	-	-	-	9,68%	3,23%	6,45%
11-15 años	-	6,45%	-	6,45%	-	-	-	-
16-20 años	-	3,23%	3,23%	-	6,45%	3,23%	-	3,23%
21-25 años	-	3,23%	3,23%	-	-	-	-	-
>25 años	32,26%	12,90%	6,45%	6,45%	6,45%	12,90%	6,45%	6,45%
Total	35,49%	25,81%	12,91%	12,90%	12,90%	25,81%	9,68%	16,13%
TIPO DE ORGANIZACIÓN								
Privada	9,68%	12,90%	6,45%	6,45%	12,90%	19,35%	9,68%	9,68%
Pública	25,81%	6,45%	-	6,45%	-	6,45%	-	6,45%
Pública & Privada	-	6,45%	6,45%	-	-	-	-	-
Total	35,49%	25,80%	12,90%	12,90%	12,90%	25,80%	9,68%	16,13%
ACTIVIDAD ECONÓMICA DE EMPRESAS								
Servicio Publico / Gobierno Local o Central	22,58%	6,45%	3,23%	3,23%	-	3,23%	-	3,23%
Agro industrial	3,23%	3,23%	3,23%	-	6,45%	3,23%	-	3,23%
Banca / Finanzas / Seguros	3,23%	6,23%	6,45%	-	-	6,45%	6,45%	-
Comercio al por mayor y menor	3,23%	3,23%	-	3,23%	6,45%	3,23%	3,23%	-
Educación	-	3,23%	-	3,23%	-	3,23%	-	3,23%
Fabricación	-	3,23%	-	3,23%	-	-	-	-
Militar	3,23%	-	-	-	-	-	-	-
Salud	-	-	-	-	-	3,23%	-	3,23%
Telecomunicaciones	-	-	-	-	-	3,23%	-	3,23%
Total	35,50%	25,80%	12,91%	12,92%	12,90%	26%	9,68%	16,13%

Las organizaciones privadas han mostrado mayor interés en el uso de Metodologías Ágiles y Big Data, así también mayor continuidad en su uso (ver figura 4.3 y 4.4).

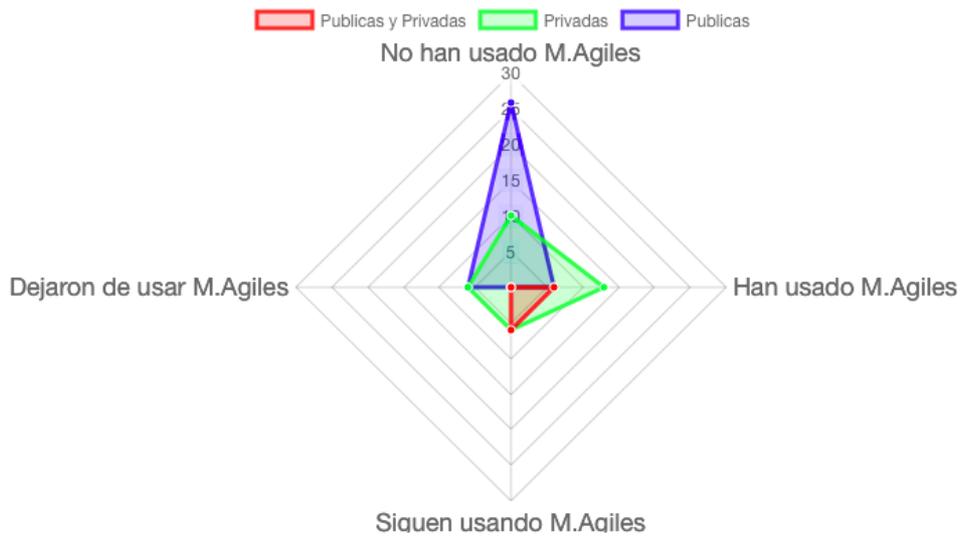


Figura 4.3. Metodologías Ágiles por tipo de empresa sin iniciativas Big Data.

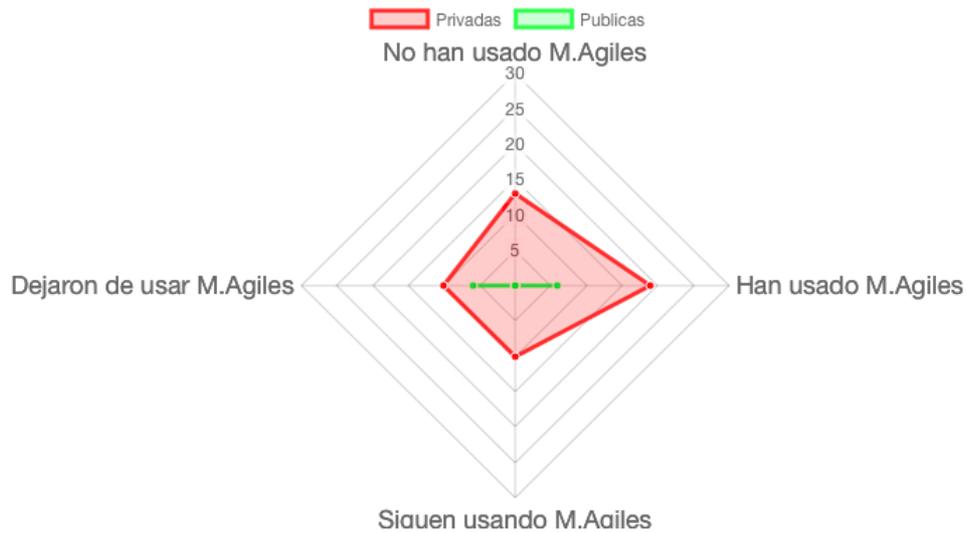


Figura 4.4. Metodologías Ágiles por tipo de empresa con iniciativas Big Data en curso o en planificación.

Adicionalmente, podemos identificar que las empresas con interés y continuidad en el uso de Metodologías Ágiles y Big Data se encuentran en los sectores de Comercio y Banca (ver figura 4.5 y 4.6).

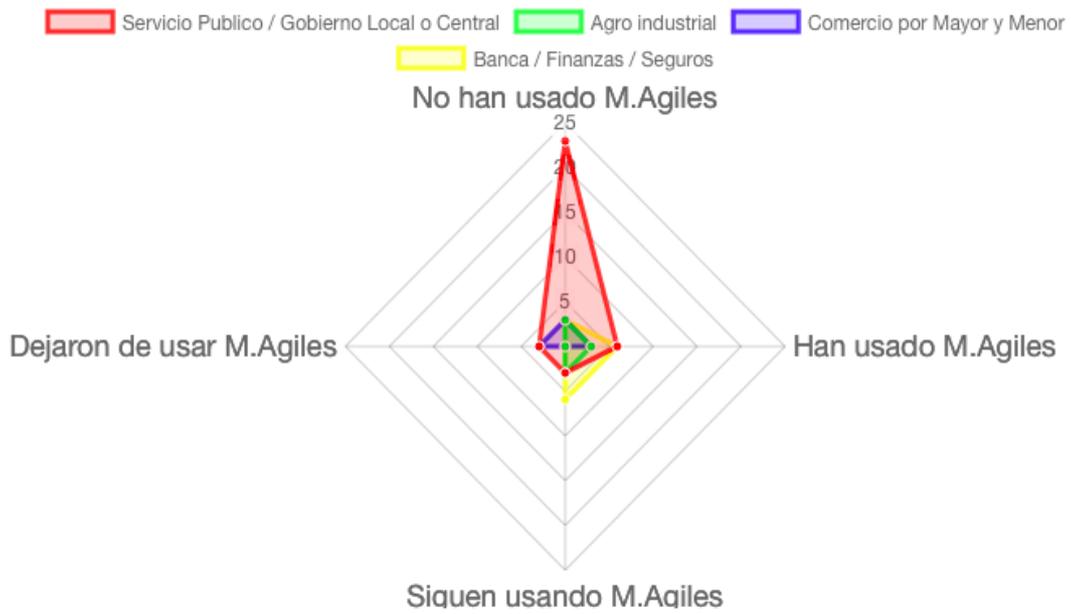


Figura 4.5. Metodologías Ágiles por sector empresarial sin iniciativas Big Data.

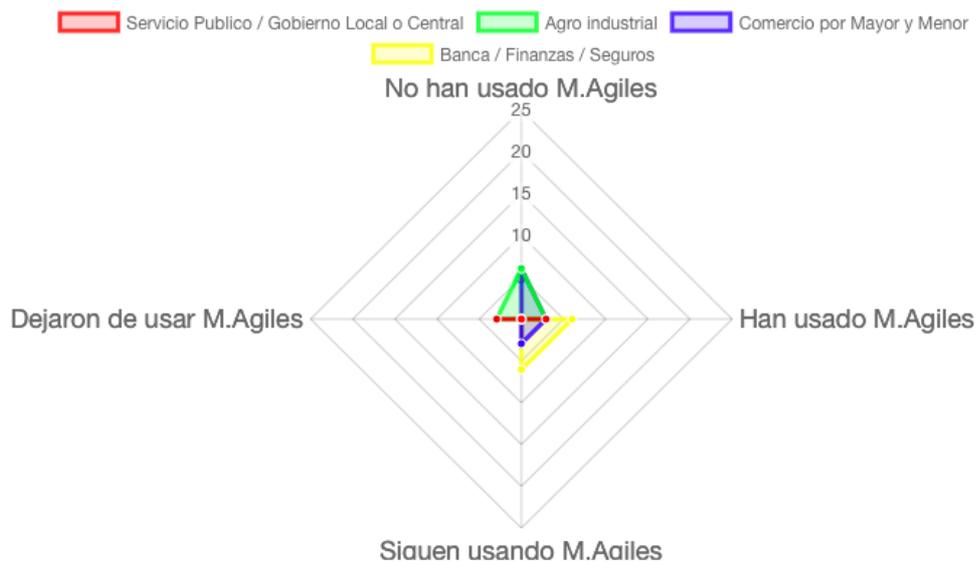


Figura 4.6. Metodologías Ágiles por sector empresarial con iniciativas Big Data en curso o en planificación.

La figura 4.7 muestra que las Metodologías Ágiles más usadas en las organizaciones con iniciativas Big Data en curso o en planeación son: Scrum, Kanban y Programación Extrema.

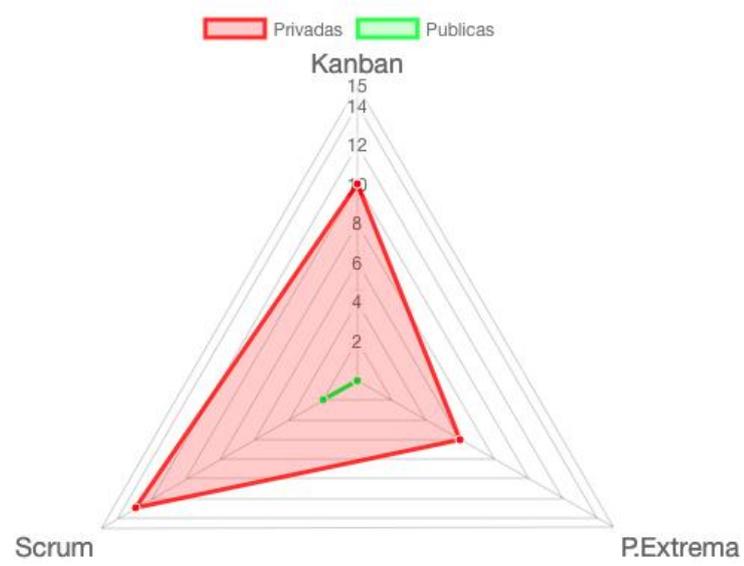


Figura 4.7. Metodologías Ágiles usadas en empresas con iniciativas Big Data en curso o en planificación.

CONCLUSIONES

La revisión sistemática realizada nos evidencia el uso ocasional de Metodologías Ágiles en proyectos de Big Data, específicamente en el análisis de imágenes y videos, análisis de archivos, uso de algoritmos de aprendizaje de máquina, y análisis de información geográfica.

Por medio del estudio descriptivo y cuantitativo realizado en una muestra de organizaciones ecuatorianas, pudimos conocer que:

- En el 52% de las organizaciones se han usado Metodologías Ágiles, abandonando su uso la mayoría de los casos (56%). Además, el 69% no recibió capacitación.
- Solo el 39% tienen proyectos de Big Data en curso o en etapa de planeación; para ello usan herramientas desarrolladas por terceros (71%); y no han considerado contratar científicos de datos (83%).
- Entre los mayores intereses respecto a Big Data están: analizar datos desde diferentes fuentes internas y externas; analizar las operaciones y el mercado; y obtener nuevas ideas de los datos procesados. Además, buscan visualizar esta información en tiempo real, y de forma integrada con sus aplicaciones del negocio.
- Las organizaciones privadas han mostrado mayor interés y continuidad en el uso de Metodologías Ágiles y Big Data, especialmente los sectores de Comercio y Banca.

Según la revisión sistemática de literatura y el estudio descriptivo y cuantitativo, las Metodologías Ágiles más usadas son Scrum, Kanban y Programación Extrema.

RECOMENDACIONES

Se recomienda realizar una revisión y categorización de los principios, valores, prácticas y herramientas de las Metodologías Ágiles más usadas, para evaluar su factibilidad de uso en proyectos Big Data, en sus etapas de gestión de requerimientos, desarrollo y pruebas.

Debido a que las iniciativas de Big Data de mayor interés de las empresas en los cuales no se requieren datos internos, se enfocan en el estudio del mercado a través del análisis de datos de diferentes fuentes en tiempo real, se sugiere realizar experimentaciones para este tipo de proyectos con Metodologías Ágiles.

BIBLIOGRAFIA

- [1] V. C. Deanne Larsona, "A review and future direction of agile, business intelligence, analytics and data science," *International Journal of Information Management*, 2016.
- [2] M. D. o. b. Patrícia Franková, Peter Balco, "Agile project management approach and its use in big data management," *Procedia Computer Science*, vol. 83, pp. 576 – 583, 2016.
- [3] Dinesh Batra and R. Dahiya, "Adapting Agile Practices for Analytics Projects," 2016.
- [4] B. M. L. D., "The importance of 'big data': a definition," Stamford, CT: Gartner, 2012.
- [5] E. Isitor, C. Stanier, "Defining Big Data", 2016.
- [6] L.R. Vijayarath, D. Turk, "Agile Software development: A survey of early adopters," *Journal of Information Technology Management*, vol. 19, no. 2, 2008.
- [7] S. Debotoli, "Comparing Business Intelligence and Big Data Skills," *Business & Information Systems Engineering*, vol. 8, no. 51, 2014.
- [8] E. Grill, Müller, M. & Mansmann, U., "Health—exploring complexity: an interdisciplinary systems approach HEC2016,"
- [9] N. Khan, I. Yaqoob, "Big Data: Survey, Technologies, Opportunities, and Challenges", *The Scientific World Journal*, vol. 2014, ID 712826, pp. 18, 2014
- [10] R. Schutt, & O'Neil, C., "Doing data science: straight talk from the frontline," O'Reilly Media, Inc., 2013.
- [11] S. Jarr, "Fast data and the new enterprise data architecture," O'Reilly Publishing, 2015. 7
- [12] F. Halpern, "Next-generation analytics and platforms for business success," tDWI research report, 2015.
- [13] C. Y. Hsieh, & Chen, C. T., "Patterns for continuous integration builds in cross-platform agile software development," *Journal of Information Science & Engineering*, vol. 31, no. 3, pp. 897–924, 2015.
- [14] N. Kaleshovska, Josimovski, S., Pulevska-Ivanovska, L., Postolov, K., & Janevski, Z., "The contribution of scrum in managing successful software development projects," *Economic Development/Ekonomiski Razvoj*, vol. 17, no. 1/2, pp. 175–194, 2015.
- [15] International Rugby Board (2012, mayo 15). Ley 20, Scrum. En *Leyes del juego de Rugby* [Internet], Disponible desde <http://www.irblaws.com> [Acceso Junio 1, 2013].
- [16] Takeuchi, H. & Nonaka, I. The new product development game. *Harvard Business Review*, Ene-Feb, 137-146, 1986.
- [17] A. Navarro, J.D. Fernández, J. Morales, "A review of agile methodologies for software development," *Prospect*. Vol. 11, No. 2, July - december de 2013, págs. 30-39.
- [18] K. Beck, *Extreme Programming Explained: Embrace Change* [1a ed.]. Addison Wesley, Stoughton, 1999.
- [19] M. Khaled Yacoub, M. Abdel Athim Mostafa, A. Bahaa Farid, "A New Approach for Distributed Software Engineering Teams Based on Kanban Method for Reducing Dependency", *Journal of Software*, 2016.
- [20] D. J. Anderson, "Kanban and the end of methodology," *Lean-Kanban University*, North America Pittsburgh October 2013 Release 1.0, Disponible: <http://cmmiconferences.com/wp-content/uploads/2013/10/Kanban%20and%20the%20End%20of%20Methodology.pdf>
- [21] R. Hughes, "Agile Data Warehousing Project Management: Business Intelligence Systems Using Scrum," Morgan Kaufmann, 2013.
- [22] R. M. Troy J. Mueller, "Data Warehouse and Business Intelligence Implementation: Agile and Iterative Themes," *Issues in Information Systems*, vol. 14, no. 1, pp. 129-132, 2013.
- [23] K. Dikert, M. Paasivaara, C. Lassenius, "Challenges and success factors for large-scale agile transformations: A systematic literature review," *The Journal of Systems and Software*, vol. 119, pp. 87-108, 2016.
- [24] J. Gutiérrez, R. Borillo, "Análisis de implementación de prácticas ágiles en Argentina," *2da Conferencia Agile-Spain*, Castellón de la Plana, 2011.
- [25] L.E. Peláez, A. Toro, J.F. López, A. Ramírez, "Software development process characterization in Colombia: A view from SMEs producers," *Revista Académica e Institucional (UCP)*, pp.89-98, 2012.
- [26] P. Pedroza, "Choosing a Development Methodology from the Advantages of an Agile Methodology and a Robust Model like CMMI-DEV 1.3," *Barranquilla*, 2013.
- [27] C. Enriquez Ramírez, "Análisis empírico sobre la adopción de las metodologías ágiles en los equipos de desarrollo de software en empresas mexicanas," *Congreso Nacional e Internacional de la Información y Computación (ANIEI)*, pp.48-55, 2012.
- [28] F. Núñez, G. Villa, P. Oríz, "Uso de técnicas de estimación en el cálculo de la duración de proyectos de TIC en Uruguay," *Prisma Tecnológico*, Vol. 6, n.O 1, pp.40-43, 2015.
- [29] Colegio de Ingenieros de Chile, «Metodologías Ágiles en las Empresas Aumenta Productividad y Mejoran Clima Laboral.» 12 08 2013. [En línea]. Available: <http://www.ingenieros.cl/metodologias-agiles-en-las-empresas-aumenta-productividad-y-mejoran-clima-laboral/>. [Último acceso: 11 05 2016].
- [30] F. Ticona, "Scrum methodology for software development and project management in small and medium enterprises city Juliaca," *Revista Científica Investigación Andina*, Vol. 13, n.1, 2014.

- [31] M.L. Sánchez-Gordón, R.V. O'Connor, "Understanding the gap between software process practices and actual practice in very small companies," *Software Quality Journal*, vol. 24, no. 3, pp. 549-570, 2016.
- [32] S. Siddiqui, D. Gupta, "Big Data Process Analytics: A Survey", *International Journal of Emerging Research in Management & Technology*, vol. 3, issue 7, 2014.
- [33] S. V. Kai Petersen, Ludwik Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Information and Software Technology*, vol. 64, pp. 1-18, 2015.
- [34] D. Strode, "The agile methods: an analytical comparison five agile methods and an investigation of their target environment," Unpublished Master of Information Science thesis. Massey University, Palmerston North, New Zealand, 2005.
- [35] «State of Agile Development Survey Results | VersionOne». [online]. Disponible: http://www.versionone.com/state_of_agile_development_survey/2015/. [Last Access: 08-2017].
- [36] "Report of the 2015 Big Data Survey", United Nations Statistics Division, 2016.
- [37] *Revista Vistazo* (20 Oct 2016). TDT, 500 mayores empresas del Ecuador [consultado 09-2017]. Disponible: <http://www.vistazo.com/seccion/proyectos-especiales/500empresas>.
- [38] M. V. aura L. Willems, "Classification of articles and journals on project control and earned value management," *International Journal of Project Management*, vol. 33, no. 7, pp. 1610-1634, 2015.
- [39] F. D. Mahdi Fahmideh Gholami, Graham Low, Ghassan Beydoun, "Cloud migration process—A survey, evaluation framework, and open challenges," *Journal of Systems and Software*, vol. 120, pp. 31- 69, 2016.
- [40] M. S. Rodrigo Reyes Levalle, Shimon Y. Nof, "Collaborative production line control for collaborative supply networks," *IFAC Proceedings Volumes*, vol. 45, no. 6, pp. 487-492, 2012.
- [41] A. C. FENO Mahenina Remiel, Alain FERRARINI, "Conceptual design and simulation of an automotive body shop assembly line," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 760-765, 2014.
- [42] D. G. Cristina Domènech, Marek Pękala, "Decreasing Kd uncertainties through the application of thermodynamic sorption models," *Science of The Total Environment*, vol. 527–528, pp. 297- 305, 2015.
- [43] T. T. arcello Colledani, Anath Fischer, Benoit lung, Gisela Lanza, Robert Schmitt, József Váncza, "Design and management of manufacturing systems for production quality," *CIRP Annals - Manufacturing Technology*, vol. 63, no. 2, pp. 773-796, 2014.
- [44] E. B. Stelian Brad, "Enhancing SWOT Analysis with TRIZ-based Tools to Integrate Systematic Innovation in Early Task Design," *Procedia Engineering*, vol. 131, pp. 616-625, 2015.
- [45] A. H. Tugrul U. Daim, Shawn Reutiman, Brennan Hughes, Ujjal Pathak, Wayne Bynum, Ashok Bhatla, "Exploring the communication breakdown in global virtual teams," *International Journal of Project Management*, vol. 30, no. 2, pp. 199-212, 2012.
- [46] J. J. P. Ali Reza Montazemi, Hamed Qahri Saremi, Yongbin Wei, "Factors of stickiness in transfers of know-how between MNC units," *The Journal of Strategic Information Systems*, vol. 21, no. 1, pp. 31-57, 2012.
- [47] T. A. b. H. Hosseini Nasab, H. Khademi Zare, "Finding a probabilistic approach to analyze lean manufacturing," *Journal of Cleaner Production*, vol. 29-30, pp. 73-81, 2012.
- [48] F. Q. B. d. S. Shirley Cruz, Luiz Fernando Capretz, "Forty years of research on personality in software engineering: A mapping study," *Computers in Human Behavior*, vol. 46, pp. 94-113, 2015.
- [49] R. L. David Würfel, Stephan Diehl, "Grounded requirements engineering: An approach to use case driven requirements engineering," *Journal of Systems and Software*, vol. 117, pp. 645- 657, 2016.
- [50] T. S. M. Nicolli S.R. Alves, Manoel G. de Mendonça, Rodrigo O. Spínola, Forrest Shull, Carolyn Seaman, "Identification and management of technical debt: A systematic mapping study," *Information and Software Technology*, vol. 70, pp. 100-121, 2016.
- [51] J. K. Alfredo Rivera, "Identifying the State of the Project Management Profession," *Procedia Engineering*, vol. 145, pp. 1386- 1393, 2016.
- [52] S. G. A. Kannan Govindan, Helena Carvalho, V. Cruz-Machado, "Impact of supply chain management practices on sustainability," *Journal of Cleaner Production*, vol. 85, 15, pp. 212-225, 2014.
- [53] Y. C. Nima Jafari Navimipour, "Knowledge sharing mechanisms and techniques in project teams: Literature review, classification, and current trends," *Computers in Human Behavior*, vol. 62, pp. 730- 742, 2016.
- [54] S. R. F. Fruggiero, Y. Ouazene, R. Macchiaroli, V. Guglielmi, "Incorporating the Human Factor within Manufacturing Dynamics," *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 1691-1696, 2016.
- [55] M. S. Nazmi Saeb Jarrar, "Innovation in entrepreneurial organisations: A platform for contemporary management change and a value creator," *The British Accounting Review*, vol. 46, no. 1, pp. 60-76, 2014.
- [56] M. J.-. Kaczmarek, "Integrating Lean and Green Paradigms in Maintenance Management," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 4471-4476, 2014.
- [57] A. S. Paolo Tonella, Francis Palma, "Interactive requirements prioritization using a genetic algorithm," *Information and Software Technology*, vol. 55, no. 1, pp. 173-187, 2013.
- [58] D. S. C. Claudia de O. Melo, Fabio Kon, Reidar Conradi, , "Interpretative case studies on agile team productivity and management," *Information and Software Technology*, vol. 55, no. 2, pp. 412-427, 2013.
- [59] J. D. Sherif Mostafa, Hassan Soltan, "Lean Maintenance Roadmap," *Procedia Manufacturing*, vol. 2, pp. 434-444, 2015.

- [60] M. Z. Ali Azadeh, Mohammad Abdollahi, Saeid Noury, Shabnam Farahmand, "Leanness assessment and optimization by fuzzy cognitive map and multivariate analysis," *Expert Systems with Applications*, vol. 42, no. 15, 16, pp. 6050-6064, 2015.
- [61] D. Sui, "Looking through Hågerstrand's dual vistas: towards a unifying framework for time geography," *Journal of Transport Geography*, vol. 23, pp. 5-16, 2012.
- [62] H. A. Seher Arslankaya, "Maintenance Management and Lean Manufacturing Practices in a Firm Which Produces Dairy Products," *Procedia - Social and Behavioral Sciences*, vol. 207, pp. 214-224, 2015.
- [63] I. B. Debprotim Dutta, "Managing a Big Data project: The case of Ramco Cements Limited," *International Journal of Production Economics*, vol. 165, pp. 293-306, 2015.
- [64] X. D. Qian Shi, Jian Zuo, George Zillante, "Mobile Internet based construction supply chain management: A critical review," *Automation in Construction*, vol. 72, Part 2, pp. 143-154, 2016.
- [65] J. B. Albertus van 't Ooster, Eldert J. van Henten, Silke Hemming, "Model-based analysis of skill oriented labour management in a multi-operations and multi-worker static cut rose cultivation system," *Biosystems Engineering*, vol. 135, pp. 87-102, 2015.
- [66] S. G. Eva Klenk, Willibald A. Günthner, "Operating Strategies for In-Plant Milk-Run Systems," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 1882-1887, 2015.
- [67] A. P. F. De Felice, "Optimization of Manufacturing System through World Class Manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 741-746, 2015.
- [68] A. D. Alexander Bucksch, Hannah Schneider, Nirav Merchant, Joshua S. Weitz, "Overcoming the Law of the Hidden in Cyberinfrastructures," *Trends in Plant Science*, vol. 22, no. 2, pp. 117-123, 2017. ^[SEP]
- [69] M. G. Dan Tofan, Paris Avgeriou, Wes Schuitema, "Past and future of software architectural decisions – A systematic mapping study," *Information and Software Technology*, vol. 56, no. 8, pp. 850-872, 2014. ^[SEP]
- [70] J. K. R. A. Mohd Yazid, M.S. Awaluddin, Emelia Sari, "Pattern Recognition on Remanufacturing Automotive Component as Support Decision Making Using Mahalanobis-taguchi System," *Procedia CIRP*, vol. 26, pp. 258-263, 2015. ^[SEP]
- [71] D. S. Martin Komenda, Jan Švancara, Christos Vaitsis, Nabil Zary, Ladislav Dušek, "Practical use of medical terminology in curriculum mapping," *Computers in Biology and Medicine*, vol. 63, pp. 74-82, 2015. ^[SEP]
- [72] I. M. F. Rafaela Mantovani Fontana, Paula Andrea da Rosa Garbuio, Sheila Reinehr, Andreia Malucelli, "Processes versus people: How should agile software development maturity be defined?," *Journal of Systems and Software*, vol. 97, pp. 140-155, 2014.
- [73] V. M. J. Rafaela Mantovani Fontana, Sheila Reinehr, Andreia Malucelli, "Progressive Outcomes: A framework for maturing in agile software development," *Journal of Systems and Software*, vol. 102, pp. 88-108, 2015. ^[SEP]
- [74] M. J. E. F.J. Domínguez-Mayo, M. Mejías, M. Ross, G. Staples, "Quality evaluation for Model-Driven Web Engineering methodologies," *Information and Software Technology*, vol. 54, no. 11, pp. 1265-1282, 2012.
- [75] N. O. N. Jazdi, M. Weyrich, "Quantification of the quality characteristics for the calculation of software reliability," *IFAC- PapersOnLine*, vol. 49, no. 30, pp. 1-5, 2016. ^[SEP]
- [76] J. M. Eveliina Lindgren, "Raising the odds of success: the current state of experimentation in product development," *Information and Software Technology*, vol. 77, pp. 80-91, 2016. ^[SEP]
- [77] R. L. Nozomi Ikeya, Dave Randall, "Recovering the emergent logic in a software design exercise," *Design Studies*, vol. 33, no. 6, pp. 611-629, 2012. ^[SEP]
- [78] T. V. Roy Suddaby, Yves Gendron, "Reflexivity: The role of embedded social position and entrepreneurial social skill in processes of field level change," *Research in Organizational Behavior*, vol. 36, pp. 225-245, 2016. ^[SEP]
- [79] E. T. Myoungkyu Song, "Reusing metadata across components, applications, and languages," *Science of Computer Programming*, vol. 98, no. 4, pp. 617-644, 2015. ^[SEP]
- [80] C. T. M. Iiro Harjunoski, Peter Bongers, Pedro M. Castro, Sebastian Engell, Ignacio E. Grossmann, John Hooker, Carlos Méndez, Guido Sand, John Wassick, "Scope for industrial applications of production scheduling models and solution methods," *Computers & Chemical Engineering*, vol. 62, no. 5, pp. 161-193, 2014. ^[SEP]
- [81] W. P. Nina M. Berry, Circe Lunkins, James Vega, Robin Landry, Leonel Garciga, "Selecting Video Analytics using Cognitive Ergonomics: A Case Study for Operational Experimentation," *Procedia Manufacturing*, vol. 3, pp. 5245-5252, 2015. ^[SEP]
- [82] E. N. Shigeru Hosono, Yoshiki Shimomura, "Servitization Methodology in ICT Service System Design," *Procedia CIRP*, vol. 47, pp. 18-23, 2016. ^[SEP]
- [83] S. V. Lieslot Danneels, "Simple rules strategy to transform government: An ADR approach," *Government Information Quarterly*, vol. 32, no. 4, pp. 516-525, 2015. ^[SEP]
- [84] J. S. S. Ashkan Negahban, "Simulation for manufacturing system design and operation: Literature review and analysis," *Journal of Manufacturing Systems*, vol. 33, no. 2, pp. 241-261, 2014. ^[SEP]
- [85] S. U. K. Sikandar Ali, "Software outsourcing partnership model: An evaluation framework for vendor organizations," *Journal of Systems and Software*, vol. 117, pp. 402-425, 2016. ^[SEP]

- [86] C. U. Muhammad Sulayman, Emilia Mendes, Stefan Seidel, "Software process improvement success factors for small and medium Web companies: A qualitative study," *Information and Software Technology*, vol. 54, no. 5, pp. 479-500, 2012. ^[1]_[SEP]
- [87] J. D. F. Mahut, M. Bricogne, B. Eynard, "Survey on Product-Service System applications in the automotive industry," *IFAC- PapersOnLine*, vol. 48, no. 3, pp. 840-847, 2015. ^[1]_[SEP]
- [88] A. S. Angappa Gunasekaran, "Sustainability of manufacturing and services: Investigations for research and applications," *International Journal of Production Economics*, vol. 140, no. 1, pp. 35-47, 2012. ^[1]_[SEP]
- [89] D. I. K. S. Yngve Lindsjörn, Torgeir Dingsøy, Gunnar R. Bergersen, Tore Dybå, "Teamwork quality and project success in software development: A survey of agile development teams," *Journal of Systems and Software*, vol. 122, pp. 274-286, 2016. ^[1]_[SEP]
- [90] P. C. P. Arash Azadegan, Abouzar Zangouinezhad, Kevin Linderman, "The effect of environmental complexity and environmental dynamism on lean practices," *Journal of Operations Management*, vol. 31, no. 4, pp. 193-212, 2013.
- [91] S. B. Behzad Esmailian, Ben Wang, "The evolution and future of manufacturing: A review," *Journal of Manufacturing Systems*, vol. 39, pp. 79-100, 2016.
- [92] A. A. Areti Ampatzoglou, Alexander Chatzigeorgiou, Paris Avgeriou, "The financial aspect of managing technical debt: A systematic literature review," *Information and Software Technology*, vol. 64, pp. 52-73, 2015.
- [93] I. M. Rami Bahsoon, Nour Ali, T.S. Mohan, Nenad Medvidović, "The future of software engineering IN and FOR the cloud," *Journal of Systems and Software*, vol. 86, no. 9, pp. 2221-2224, 2013.
- [94] R. Wendler, "The maturity of maturity model research: A systematic mapping study," *Information and Software Technology*, vol. 54, no. 12, pp. 1317-1339, 2012.
- [95] M. K. Malte Brettel, Niklas Friederichsen, "The Relevance of Manufacturing Flexibility in the Context of Industrie 4.0," *Procedia CIRP*, vol. 41, pp. 105-110, 2016.
- [96] T. S. Fausto Guaragni, Kristin Paetzold, "Traditional and Agile Product Development in a Hyperconnected World: Turning Weaknesses into Strengths," *Procedia CIRP*, vol. 52, pp. 62-67, 2016.
- [97] A. V. Evelyne Vanpoucke, Kenneth K. Boyer, "Triggers and patterns of integration initiatives in successful buyer-supplier relationships," *Journal of Operations Management*, vol. 32, no. 1-2, pp. 15-33, 2014.
- [98] C. U. Ohad Barzilay, "Understanding reuse of software examples: A case study of prejudice in a community of practice," *Information and Software Technology*, vol. 56, no. 12, pp. 1613-1628, 2014.
- [99] C. A. I. Emilio Serrano, "Validating viral marketing strategies in Twitter via agent-based social simulation," *Expert Systems with Applications*, vol. 50, pp. 140-150, 2016.
- [100] A. J. R. Eksa Kilfoyle, Laura D. MacDonald, "Vernacular accountings: Bridging the cognitive and the social in the analysis of employee-generated accounting systems," *Accounting, Organizations and Society*, vol. 38, no. 5, pp. 382-396, 2013.
- [101] A. S. G. Fabian Fagerholm, Hanna Mäenpää, Jürgen Münch, "The RIGHT model for Continuous Experimentation," *Journal of Systems and Software*, vol. 123, pp. 292-305, 2017.
- [102] D. M. F. Henning Femmer, Stefan Wagner, Sebastian Eder, "Rapid quality assurance with Requirements Smells," *Journal of Systems and Software*, vol. 123, pp. 190-213, 2017.
- [103] A. T. d. A. Isaac Pergher, "A multi-attribute decision model for setting production planning parameters," *Journal of Manufacturing Systems*, vol. 42, pp. 224-232, 2017.
- [104] Y. P. Ray Y. Zhong, Fan Xue, Ji Fang, Weiwu Zou, Hao Luo, S. Thomas Ng, Weisheng Lu, Geoffrey Q.P. Shen, George Q. Huang, "Prefabricated construction enabled by the Internet-of-Things," *Automation in Construction*, vol. 76, pp. 59-70, 2017.
- [105] P. L. Ben Clegg, Steve Govette, Joel Logue, "Transformation of a small-to-medium-sized enterprise to a multi-organisation product- service solution provider," *International Journal of Production Economics*, 2017.
- [106] P. L. Yihsung Su, Yue-Shi Lee, Show-Jane Yen, "Creating an invalid defect classification model using text mining on server development," *Journal of Systems and Software*, vol. 125, pp. 197- 206, 2017.
- [107] S. K. homas G. Schmitt, Kathryn E. Stecke, Fred W. Glover, Mark A. Ehlen, "Mitigating disruptions in a multi-echelon supply chain using adaptive ordering," *Omega*, vol. 68, pp. 165-198, 2017.
- [108] D.-P. S. Chung-Y ee Lee, "Ocean container transport in global supply chains: Overview and research opportunities," *Transportation Research Part B: Methodological*, vol. 98, pp. 442-474, 2017.
- [109] S. S. Richard Vidgen, David B. Grant, "Management challenges in creating value from business analytics," *European Journal of Operational Research*, 2017.
- [110] E. Vezzetti, Violante, M.G. & Marcolin, "A benchmarking framework for product lifecycle management (PLM) maturity models," *The International Journal of Advanced Manufacturing Technology*, vol. 71, no. 5, pp. 899-918, 2014.
- [111] Strode, "A dependency taxonomy for agile software development projects," *Information Systems Frontiers*, vol. 18, no. 1, pp. 23-46, 2016.
- [112] E. A. Roghanian, "A fuzzy model for achieving lean attributes for competitive advantages development using AHP-QFD- PROMETHEE," *Journal of Industrial Engineering International*, 2014.

- [113] M. A. Almomani, Abdelhadi, A., Mumani, A., "A proposed integrated model of lean assessment and analytical hierarchy process for a dynamic road map of lean implementation," *The International Journal of Advanced Manufacturing Technology*, vol. 72, no. 1, pp. 161-172, 2014.
- [114] A. K. GUPTA, T.K. , "A review of designing machine tool for leanness," *Sadhana*, vol. 37, no. 2, pp. 241-259, 2012.
- [115] D. Battleson, West, B., Kim, J., "Achieving dynamic capabilities with cloud computing: an empirical investigation," *European Journal of Information Systems*, vol. 25, no. 3, pp. 209-230, 2016.
- [116] L. Cao, Mohan, K., Ramesh, B., "Adapting funding processes for agile IT projects: an empirical investigation," *European Journal of Information Systems*, vol. 22, no. 2, pp. 191-205, 2013.
- [117] I. T. Cil, Y.S., "An ANP-based assessment model for lean enterprise transformation," *The International Journal of Advanced Manufacturing Technology*, vol. 64, no. 5, pp. 1113-1130, 2013.
- [118] F. Chen, Nunamaker, J.F., Briggs, R.O., "An Application of Focus Theory to Project Management Processes," *Group Decision and Negotiation*, vol. 23, no. 5, pp. 961-978, 2014.
- [119] S. B. Mehrjoo, M., "An application of principal component analysis and logistic regression to facilitate production scheduling decision support system: an automotive industry case," *Journal of Industrial Engineering International*, 2013.
- [120] A. Anvari, Zulkifli, N., Sorooshian, S. , "An integrated design methodology based on the use of group AHP-DEA approach for measuring lean tools efficiency with undesirable output," *The International Journal of Advanced Manufacturing Technology*, vol. 70, no. 9, pp. 2269-2186, 2014.
- [121] R. S. Anholon, A.T., "Analysis of critical processes in the implementation of lean manufacturing projects using project management guidelines," *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 9, pp. 2247-2256, 2016.
- [122] S. Ranjan, Jha, V.K. & Pal, P. , "Application of emerging technologies in ERP implementation in Indian manufacturing enterprises: an exploratory analysis of strategic benefits," *The International Journal of Advanced Manufacturing Technology*, vol. 88, no. 1, pp. 369-380, 2017.
- [123] T. Menzies, Nichols, W., Shull, F., "Are delayed issues harder to resolve? Revisiting cost-to-fix of defects throughout the lifecycle," *Empirical Software Engineering*, pp. 1-33, 2016.
- [124] P. Arcaini, Holom, RM. & Riccobene, E, "ASM-based formal design of an adaptivity component for a Cloud system," *Formal Aspects of Computing*, vol. 28, no. 4, pp. 567-595, 2016.
- [125] A. Rago, Marcos, C. & Diaz-Pace, J.A., "Assisting requirements analysts to find latent concerns with REAssistant," *Automated Software Engineering*, vol. 23, no. 2, pp. 219-252, 2016.
- [126] M. Heikkilä, Bouwman, H., Heikkilä, J. , "Business model metrics: an open repository," *Information Systems and e-Business Management*, vol. 14, no. 2, pp. 337-366, 2016.
- [127] A. M. Bykov, Malkov, M.A., Raymond, J.C., "Collisionless Shocks in Partly Ionized Plasma with Cosmic Rays: Microphysics of Non- thermal Components," *Space Science Reviews*, vol. 178, no. 2, pp. 599-632, 2013.
- [128] S. Debortoli, Müller, O. & vom Brocke, J., "Comparing Business Intelligence and Big Data Skills," *Business & Information Systems Engineering*, vol. 6, no. 5, pp. 289-300, 2014.
- [129] S. Kodama, T., "Demand articulation in the open-innovation paradigm," *Journal of Open Innovation: Technology, Market, and Complexity*, 2015.
- [130] W. Zhang, Nie, L., Jiang, H., "Developer social networks in software engineering: construction, analysis, and applications," *Science China Information Sciences*, vol. 57, no. 12, pp. 1-23, 2014.
- [131] Y. Khojasteh-Ghamari, "Developing a framework for performance analysis of a production process controlled by Kanban and CONWIP," *Journal of Intelligent Manufacturing*, vol. 23, no. 1, pp. 61-71, 2012.
- [132] Y. Rezgui, Marks, A., AL-Hajri, H., "Developing an ISDM Adoption Decision Model Using Delphi and AHP," *Arabian Journal for Science and Engineering*, vol. 39, no. 4, pp. 2799-2815, 2014.
- [133] R. Seethamraju, "Enterprise systems and demand chain management: a cross-sectional field study," *Information Technology and Management*, vol. 15, no. 3, pp. 151-161, 2014.
- [134] B. B. N. T. de França, G.H. , "Experimentation with dynamic simulation models in software engineering: planning and reporting guidelines," *Empirical Software Engineering*, vol. 21, no. 3, pp. 1302-1345, 2016.
- [135] A. Buckley, "Getting IT to work for marketing: Exploring collaboration between marketing and IT functions for the delivery of marketing innovation," *Journal of Direct, Data and Digital Marketing Practice*, vol. 16, no. 4, pp. 285-307, 2015.
- [136] M. Janssen, Konopnicki, D., Snowdon, J.L., "Driving public sector innovation using big and open linked data (BOLD)," *Information Systems Frontiers*, 2017.
- [137] A. W. Baur, "Harnessing the social web to enhance insights into people's opinions in business, government and public administration," *Information Systems Frontiers*, 2016.
- [138] P. M. Felipe Besson, Fabio Kon, Dejan Milojevic, "Bringing Test- Driven Development to web service choreographies," *Journal of Systems and Software*, vol. 99, pp. 135-154, 2015.
- [139] S. B. Viaene, S., "How IT enables business model innovation at the VDAB," *Journal of Information Technology Teaching Cases*, vol. 3, no. 2, pp. 78-87, 2013.

- [140] N. Evangelopoulos, Zhang, X. & Prybutok, V., "Latent Semantic Analysis: five methodological recommendations," *European Journal of Information Systems*, vol. 21, no. 1, pp. 70-86, 2012.
- [141] J. Teixeira, Robles, G. & González-Barahona, J.M., "Lessons learned from applying social network analysis on an industrial Free/Libre/Open Source Software ecosystem," *Journal of Internet Services and Applications*, 2015.
- [142] T. Menzies, Yang, Y., Mathew, G., "Negative results for software effort estimation," *Empirical Software Engineering*, 2016.
- [143] F. A. O. Zeleti, A., "Open data value capability architecture," *Information Systems Frontiers*, 2016.
- [144] A. A. Kalenkova, van der Aalst, W.M.P., Lomazova, I.A., "Process mining using BPMN: relating event logs and process models," *Software & Systems Modeling*, 2015.
- [145] F. P. Ounnar, P., "Pull control for job shop: holonic manufacturing system approach using multicriteria decision-making," *Journal of Intelligent Manufacturing*, vol. 23, no. 1, pp. 141-153, 2012.
- [146] S. U. R. Khan, Lee, S.P., Dabbagh, M., "RePizer: a framework for prioritization of software requirements," *Frontiers of Information Technology & Electronic Engineering*, vol. 17, no. 8, pp. 750-765, 2016.
- [147] F. Boumahdi, Chalal, R., Guendouz, A., "SOA^(+d): a new way to design the decision in SOA—based on the new standard Decision Model and Notation (DMN)," *Service Oriented Computing and Applications*, vol. 10, no. 1, pp. 35-53, 2016.
- [148] R. K. Bandi, Rao, R.A. & Gunupudi, L., "State of the academic field of IS in India," *Information Technology and Management*, vol. 15, no. 3, pp. 163-175, 2014.
- [149] S. Singh, Olugu, E.U., Musa, S.N., "Strategy selection for sustainable manufacturing with integrated AHP-VIKOR method under interval-valued fuzzy environment," *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 1, pp. 547-563, 2016.
- [150] E. Vezzetti, Alemanni, M. & Macheda, J., "Supporting product development in the textile industry through the use of a product lifecycle management approach: a preliminary set of guidelines," *The International Journal of Advanced Manufacturing Technology*, vol. 79, no. 9, pp. 1493-1504, 2015.
- [151] D. Peiris, Sun, L., Patel, A., "Systematic medical assessment, referral and treatment for diabetes care in China using lay family health promoters: protocol for the SMARTDiabetes cluster randomised controlled trial," *Implementation Science*, 2015.
- [152] R. A. D. Rao, R., "Technology assimilation through conjunctures – a look at IS use in retail," *Information Systems Frontiers*, vol. 17, no. 1, pp. 31-50, 2015.
- [153] V. Blondeau, Etien, A., Anquetil, N., "Test case selection in industry: an analysis of issues related to static approaches," *Software Quality Journal*, 2016.
- [154] S. Jing, Niu, Z. & Chang, P.C., "The application of VIKOR for the tool selection in lean management," *Journal of Intelligent Manufacturing*, pp. 1-12, 2015.
- [155] N. Asadi, Fundin, A. & Jackson, M., "The Essential Constituents of Flexible Assembly Systems: A Case Study in the Heavy Vehicle Manufacturing Industry," *Global Journal of Flexible Systems Management*, vol. 16, no. 3, pp. 235-250, 2015.
- [156] A. L. Elbanna, H.C., "The formation of technology mental models: the case of voluntary use of technology in organizational setting," *Information Systems Frontiers*, vol. 17, no. 1, pp. 95-108, 2015.
- [157] R. Dubey, Gunasekaran, A., Childe, S.J., "The impact of big data on world-class sustainable manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 1, pp. 631-645, 2016.
- [158] J. M. Tien, "The next industrial revolution: Integrated services and goods," *Journal of Systems Science and Systems Engineering*, vol. 21, no. 3, pp. 257-296, 2012.
- [159] M. Hummel, Rosenkranz, C. & Holten, R., "The Role of Communication in Agile Systems Development," *Business & Information Systems Engineering*, vol. 5, no. 5, pp. 343-355, 2013.
- [160] R. R. Jain, B., "The roles of contextual elements in post-merger common platform development: an empirical investigation," *European Journal of Information Systems*, vol. 24, no. 2, pp. 159-177, 2015.
- [161] L. Ben Othmane, Chehraz, G., Bodden, E., "Time for Addressing Software Security Issues: Prediction Models and Impacting Factors," *Data Science and Engineering*, pp. 1-18, 2016.
- [162] M. T. Liukkonen, TN., "Toward decentralized intelligence in manufacturing: recent trends in automatic identification of things," *The International Journal of Advanced Manufacturing Technology*, vol. 87, no. 9, pp. 2509-2531, 2016.
- [163] R. Thakurta, "Understanding requirement prioritization artifacts: a systematic mapping study," *Requirements Engineering*, pp. 1-36, 2016.
- [164] J. A. I. Eirín Lodgaard, Inger Gamme, Silje Aschehoug, "Barriers to Lean Implementation: Perceptions of Top Managers, Middle Managers and Workers," *Procedia CIRP*, vol. 57, pp. 595-600, 2016.
- [165] U. P. Abelein, B., "Understanding the Influence of User Participation and Involvement on System Success – a Systematic Mapping Study," *Empirical Software Engineering*, vol. 20, no. 1, pp. 28-81, 2015.
- [166] Y. Zhan, Tan, K.H., Li, Y., "Unlocking the power of big data in new product development," *Annals of Operations Research*, pp. 1-19, 2016.

- [167] R. Y. Zhong, Lan, S., Xu, C., "Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 1, pp. 5-16, 2016.
- [168] M. V. Mäntylä, Itkonen, J. & Iivonen, J., "Who tested my software? Testing as an organizationally cross-cutting activity," *Software Quality Journal*, vol. 20, no. 1, pp. 145-172, 2012. ^[1]_{SEP}
- [169] S. I. Saltz J., Connors C., "A Framework for Describing Big Data Projects," *Business Information Systems Workshops*, vol. 263, pp. 183-195, 2017.
- [170] B. Ramesh, Cao, L., Kim, J., "Conflicts and complements between eastern cultures and agile methods: an empirical investigation," *European Journal of Information Systems*, vol. 26, no. 2, pp. 206- 235, 2017.
- [171] R. Jongeling, Sarkar, P., Datta, S., "On negative results when using sentiment analysis tools for software engineering research," *Empirical Software Engineering*, pp. 1-42, 2017. ^[1]_{SEP}
- [172] R. C. M. Das, D. R. Campbell, G. Agrawal and R. Ramnath, "Towards methods for systematic research on big data," presented at the 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, 2015.
- [173] S. J. S. I. S. K. C., "Comparing Data Science Project Management Methodologies via a Controlled Experiment," presented at the 50th Hawaii International Conference on System Sciences, Hawaii, 2017.
- [174] W. A. G. S. S., "Vignettes in the two-step arrival of the internet of things and its reshaping of marketing management's service- dominant logic," *Journal of Marketing Management*, vol. 33, no. 1- 2, 2017.
- [175] R. R., "CAP and Cloud Data Management," *Computers*, vol. 45, pp. 43-49, 2012. ^[1]_{SEP}
- [176] C. K. M. Kristin L. Cullen-Lester, Dorothy R. Carter, "Incorporating social networks into leadership development: A conceptual model and evaluation of research and practice," *The Leadership Quarterly*, vol. 28, no. 1, pp. 130-152, 2017.
- [177] C. Kuo M-H., D. Moa B., Hu W., "Design and Construction of a Big Data Analytics Framework for Health Applications," presented at the 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), 2015.
- [178] G. Y. H. H., "Discovery Informatics: AI Opportunities in Scientific Discovery," in "2012 AAAI Fall Symposium Series," FS-12-03, 2012. ^[1]_{SEP}
- [179] Y. Z. Kim Hua Tan, Guojun Ji, Fei Ye, Chingter Chang, "Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph," *International Journal of Production Economics*, vol. 165, pp. 223-233, 2015.
- [180] M. T. K. Salih Tutun, Jun Zhuang, "New framework that uses patterns and relations to understand terrorist behaviors," *Expert Systems with Applications*, vol. 78, pp. 358-375, 2017. ^[1]_{SEP}
- [181] I. B. Mondher Feki, Samuel Fosso Wamba, "Big Data Analytics- enabled Supply Chain Transformation: A Literature Review " presented at the 2016 49th Hawaii International Conference on System Sciences (HICSS) (2016), Koloa, HI, USA, 2016.
- [182] F. Milicchio, Rose, R., Bian, J., Min, J., & Prosperi, M. , "Visual programming for next-generation sequencing data analytics. BioData Mining," *BioData Mining*, vol. 9, 2016. ^[1]_{SEP}
- [183] R. S. M. S. G. M.E., "Developer toolchains for large-scale analytics: Two case studies," presented at the 2015 IEEE International Conference on Big Data (Big Data), 2015.
- [184] C. B. K. M., "Big Data Comes to School: Implications for Learning, Assessment, and Research " *AERA Open*, vol. 2, no. 2, pp. 1-19, 2016. ^[1]_{SEP}
- [185] J. M. A. James W. Jones, Bruno Basso, Kenneth J. Boote, Richard T. Conant, Ian Foster, H. Charles J. Godfray, Mario Herrero, Richard E. Howitt, Sander Janssen, Brian A. Keating, Rafael Munoz-Carpena, Cheryl H. Porter, Cynthia Rosenzweig, Tim R. Wheeler, "Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science," *Agricultural Systems*, 2016.
- [186] D. H. Susan G. Mason, Aaron Wells, Amit Jain, Thomas Wuerzer, Alark Joshi, "An experiment-based methodology to understand the dynamics of group decision making," *Socio-Economic Planning Sciences*, vol. 56, pp. 14-26, 2016. ^[1]_{SEP}
- [187] D. K. Alfredo Rivera, "Creating a New Project Management Model through Research," *Creating a New Project Management Model through Research*, vol. 145, pp. 1370-1377, 2016.
- [188] K. C. Xiaofeng Wang, Oisín Cawley, "'Leagile' software development: An experience report analysis of the application of lean approaches in agile software development," *Journal of Systems and Software*, vol. 85, no. 6, pp. 1287-1299, 2012.
- [189] Y. Baghdadi, "A framework for social commerce design," *Information Systems*, vol. 60, pp. 95-113, 2016.
- [190] J. E. L. Paulo Mendes Jr., Antônio Márcio Tavares Thomé, "A maturity model for demand-driven supply chains in the consumer product goods industry," *International Journal of Production Economics*, vol. 179, pp. 153-165, 2016.
- [191] M. R. Victor Chang, Yulin Yao, Yen-Hung Kuo, Chung-Sheng Li, ^[1]_{SEP}"A resiliency framework for an enterprise cloud," *International Journal of Information Management*, vol. 36, no. 1, pp. 155-166, 2016. ^[1]_{SEP}
- [192] V. C. Deanne Larson, "A review and future direction of agile, business intelligence, analytics and data science," *International Journal of Information Management*, vol. 36, no. 5, pp. 700-710, 2016. ^[1]_{SEP}

- [193] R. R. Nachiappan Subramanian, "A review of applications of Analytic Hierarchy Process in operations management," *International Journal of Production Economics*, vol. 138, no. 2, pp. 215-241, 2012. ^[1]_{SEP}
- [194] V. N. Dragan Stankovic, Miodrag Djordjevic, Dac-Buu Cao, "A survey study of critical success factors in agile software projects in former Yugoslavia IT companies," *Journal of Systems and Software*, vol. 86, no. 6, pp. 1663-1678, 2013. ^[1]_{SEP}
- [195] S. G. Roman Trubka, Oliver Lade, Chris Pettit, "A web-based 3D visualisation and assessment system for urban precinct scenario modelling," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 117, pp. 175-186, 2016. ^[1]_{SEP}
- [196] F. S. P. Amadeu Silveira Campanelli, "Agile methods tailoring – A systematic literature review," *Journal of Systems and Software*, vol. 110, pp. 85-100, 2015. ^[1]_{SEP}
- [197] J. P. Jessica Díaz, Juan Garbajosa, "Agile product-line architecting in practice: A case study in smart grids," *Information and Software Technology*, vol. 56, no. 7, pp. 727-748, 2014. ^[1]_{SEP}
- [198] W. W. Christina Feilmayr, "An analysis of ontologies and their success factors for application to business," *Data & Knowledge Engineering*, vol. 101, pp. 1-23, 2016. ^[1]_{SEP}
- [199] K. G. Susana G. Azevedo, Helena Carvalho, V. Cruz-Machado, "An integrated model to assess the leanness and agility of the automotive industry," *Resources, Conservation and Recycling*, vol. 66, pp. 85- 94, 2012. ^[1]_{SEP}
- [200] T. N. David Sammon, John McAvoy, "Analysing ISD performance using narrative networks, routines and mindfulness," *Information and Software Technology*, vol. 56, no. 5, pp. 465-476, 2014. ^[1]_{SEP}
- [201] C. S. K. B. Vijaya Ramnath, G. Riyaz Mohamed, K. Venkataraman, C. Elanchezhian, S. Sathish, "Analysis of Occupational Safety and Health of Workers by Implementing Ergonomic Based Kitting Assembly System," *Procedia Engineering*, vol. 97, pp. 1788-1797, 2014. ^[1]_{SEP}
- [202] A. H.-H. Arno Scharl, Alistair Jones, Daniel Fischl, Ruslan Kamolov, Albert Weichselbraun, Walter Rafelsberger, "Analyzing the public discourse on works of fiction – Detection and visualization of emotion in online coverage about HBO's Game of Thrones," *Information Processing & Management*, vol. 52, no. 1, pp. 129-138, 2016. ^[1]_{SEP}

Apéndice B

Cuestionario de Big Data y Métodos de desarrollo ágil

Esta encuesta ha sido diseñada para proporcionar un punto de referencia para las empresas que buscan entender el estado de las iniciativas de Big Data Data y Métodos de desarrollo ágil.

La encuesta no debería tardar más de 30 minutos en completarse. Los encuestados permanecerán anónimos y las respuestas, no serán compartidas o identificadas.

Sección 1: Información General del Encuestado

Por favor, marque con un visto ✓ en el círculo

1. Nombre Completo (*): _____
2. Género (*):

<input type="radio"/> Femenino	<input type="radio"/> Masculino
--------------------------------	---------------------------------
3. Edad del encuestado (*):

<input type="radio"/> 24 años o menos	<input type="radio"/> 45 a 54 años
<input type="radio"/> 25 a 34 años	<input type="radio"/> 55 a 64 años
<input type="radio"/> 35 a 44 años	<input type="radio"/> 65 años o más
4. · Máximo grado de educación alcanzado (completado) (*):
 - Bachiller
 - Tercer nivel
 - Cuarto nivel
5. Ciudad en la que labora (*):

<input type="radio"/> Quito	<input type="radio"/> Durán
<input type="radio"/> Guayaquil	<input type="radio"/> Machala
<input type="radio"/> Cuenca	<input type="radio"/> Manta
<input type="radio"/> Ambato	<input type="radio"/> Otra (especifique): _____
6. Cargo:
 - Gerente de la empresa
 - Programador
 - Analista
 - Aseguramiento de calidad
 - Jefe de proyectos de TI
 - Arquitecto de software
 - Gerente de Desarrollo
 - Otro (especifique): _____

Sección 3: Estado Actual de sus Datos y capacidad de Análisis

Por favor, marque con un visto ✓ en el círculo

14. ¿En qué medida las herramientas que posee su organización permiten hacer análisis de datos para la toma de decisiones?

No Permite			Si Permite
1	2	3	4	5
<input type="radio"/>				

15. ¿En qué medida el acceso a información relevante, precisa y oportuna permite la toma de decisiones?

No Permite			Si Permite
1	2	3	4	5
<input type="radio"/>				

16. ¿Para qué funciones empresariales se realizan análisis de datos en su empresa? (Marque todas las opciones que correspondan)

- Comercio electrónico, eBusiness, Operaciones digitales
- Marketing Directo y Digital
- Gestión de fraude
- Análisis de Clientes y Mercado
- Servicio al cliente
- Desarrollo y Gestión de Productos
- Tecnologías de la Información
- Operaciones
- Gestión de riesgos
- Recursos humanos
- Otro (especifique): _____

17. ¿Aproximadamente cuántos empleados de su empresa están dedicados a la analítica, el modelado y la minería de datos (sin incluir los informes de rutina)?

- <= 5
- 6-10
- 11-15
- 16-20
- > 20

Sección 4: Metodologías Ágiles

Nota: Complete esta sección si recientemente ha aplicado Metodologías Ágiles, en caso contrario continúe con la pregunta 59.

18. ¿Cuál fue la duración del proyecto (desde el inicio hasta la puesta en producción) en el que se aplicaron Metodologías Ágiles? Si el proyecto aún está en desarrollo por favor estime la duración.
- < 3 meses
 - 3-6 meses
 - 7-12 meses
 - 1-3 años
 - > 3 años
19. ¿Cuántas personas participaron en el proyecto?
- < 3
 - 3 - 9
 - 10 – 20
 - > 20
20. ¿Qué métodos ágiles de desarrollo se utilizaron en el proyecto? Puedes seleccionar más de una caja.
- Kanban
 - Desarrollo de Software Adaptativo (ASD)
 - Crystal
 - Programación Extrema (XP)
 - Metodología de Desarrollo de Sistemas Dinámicos (DSDM)
 - Desarrollo dirigido por características (FDD)
 - Scrum
 - Otro (especifique): _____

21. Indique cuál/cuales diagramas de diseño utilizaron y su utilidad:

	Nada	Útil	Muy	Útil
	1	2	3	4	5
Diagramas entidad-relación (ERD)	<input type="radio"/>				
Diagramas de clases	<input type="radio"/>				
Diagramas de casos de uso	<input type="radio"/>				
Especificaciones de casos de uso	<input type="radio"/>				
Diagramas de actividad	<input type="radio"/>				
Diagramas de estados	<input type="radio"/>				
Diagramas de flujo	<input type="radio"/>				
Diagramas de flujo de trabajo	<input type="radio"/>				
Diagramas de secuencia	<input type="radio"/>				
Otros (especifique):					
_____	<input type="radio"/>				
_____	<input type="radio"/>				

22. Qué prácticas ágiles utilizaron en el proyecto y cuál fue su utilidad:

	Nada	Útil	Muy	Útil
	1	2	3	4	5
Fácil y constante acceso al cliente/usuarios expertos	<input type="radio"/>				
Construcciones frecuentes	<input type="radio"/>				
Ciclos de desarrollo cortos	<input type="radio"/>				
Programación en pareja	<input type="radio"/>				
Análisis retrospectivo	<input type="radio"/>				
Comunicación frecuente e informal	<input type="radio"/>				
Equipo de auto-organización	<input type="radio"/>				
Otros (especifique):					
_____	<input type="radio"/>				
_____	<input type="radio"/>				

23. Para cada herramienta utilizada en el proyecto, indique si fue obligatorio y su nivel de utilidad.

	Obligatorio	Nada	Útil	Muy	Útil
		1	2	3	4	5
Pruebas funcionales automatizadas	<input type="radio"/>					
Pruebas unitarias automatizadas	<input type="radio"/>					
Pruebas de regresión automatizadas	<input type="radio"/>					
Construcción automatizada	<input type="radio"/>					
Documentación de código automatizada	<input type="radio"/>					
Sistema de seguimiento de errores	<input type="radio"/>					
Refactorización	<input type="radio"/>					
Gestión ágil de proyectos	<input type="radio"/>					
Gestión de configuración de Software	<input type="radio"/>					
Plataformas de integración	<input type="radio"/>					
Otros (especifique):						
_____	<input type="radio"/>					
_____	<input type="radio"/>					

24. ¿Con qué frecuencia se lanzó una nueva versión del software?

- Diariamente
- Semanalmente
- Mensualmente
- Otro (especifique): _____

25. ¿Qué herramientas de gestión de proyectos ágiles utiliza/ha utilizado?

	Ha Utilizado	Utiliza Actualmente
Active Collab	<input type="radio"/>	<input type="radio"/>
Agilo for Scrum	<input type="radio"/>	<input type="radio"/>
Atlassian Jira + Agile	<input type="radio"/>	<input type="radio"/>
Pivotal Tracker	<input type="radio"/>	<input type="radio"/>
SprintGround	<input type="radio"/>	<input type="radio"/>
Targetprocess	<input type="radio"/>	<input type="radio"/>
VersionOne	<input type="radio"/>	<input type="radio"/>
Otros (especifique):		
_____	<input type="radio"/>	<input type="radio"/>
_____	<input type="radio"/>	<input type="radio"/>

26. ¿Aproximadamente cuántos días de capacitación formal de métodos de desarrollo ágil por año se proporcionan para cada desarrollador en su organización?
- No se ha brindado capacitación en este tema
 - < 1 día
 - 1-3 días
 - 4-10 días
 - > 10 días
27. Indique las razones por las cuales no ha utilizado o ha dejado de utilizar Metodologías Ágiles
- Las Metodologías Ágiles sólo son para usuarios experimentados, motivados y con capacidad de organizarse solos, sin necesidad de una supervisión cercana. Así que Agile no es para nosotros.
 - Las rígidas políticas, procesos y metodologías de nuestra empresa no se ajustan a las metodologías Agile. Agile no va a funcionar aquí.
 - Las Metodologías Ágiles sólo funciona para pequeños proyectos y nuestros proyectos son grandes.
 - Las Metodologías Ágiles requieren la ubicación conjunta de los participantes del proyecto y nuestro personal está geográficamente disperso.
 - Ágiles o no, necesitamos administrar el alcance del proyecto, la planificación, los presupuestos, las estrategias y los informes. Agile no es la respuesta.
 - Otros (especifique):

Sección 5: Big Data

Por favor, marque con un visto ✓ en el círculo **O**.

28. De las siguientes afirmaciones, escoja la que mejor se ajusta a su realidad.

- No he oído hablar de Big Data.
- He oído hablar sobre Big Data, pero no sé qué significa
- Comprendo lo que significa Big Data, pero no a nivel de detalle
- Comprendo en detalle lo que significa Big Data

Nota: Si escogió la primera opción ("No he oído..."), por favor continúe con la sección 5.

29. Indique su grado de acuerdo o desacuerdo con las afirmaciones siguientes:

"Totalmente en desacuerdo (1)", Desacuerdo (2), Neutral (3), De acuerdo (4) "Totalmente de acuerdo (5)".

	1	2	3	4	5
Big Data es un término usado para describir conjuntos de datos muy grandes de origen interno y externo de las organizaciones.	<input type="radio"/>				
Big Data es un término usado para describir conjuntos de datos complejos como los provenientes de medios sociales, video, fotos, texto no estructurado, o los recolectados por dispositivos.	<input type="radio"/>				
Big Data se usa cuando se requiere realizar procesamiento rápido para solucionar problemas de volumen y variedad de datos.	<input type="radio"/>				
Big Data se usa cuando se requiere mejorar tiempo de procesamiento.	<input type="radio"/>				
La manipulación de conjuntos de datos muy grandes a menudo requiere de aplicaciones que se ejecuten en paralelo en decenas, cientos o incluso miles de servidores.	<input type="radio"/>				

Iniciativas

30. ¿Tiene alguna iniciativa Big Data en curso o en planificación?

- No hay iniciativas en curso o previstas
- Sí, iniciativas en etapa de planificación
- Sí, iniciativas en curso

Nota: Si escogió la primera opción ("No hay iniciativas en curso o previstas..."), por favor continúe con la Sección 5.

31. ¿En qué etapa se encuentran las iniciativas de Big Data en su organización? (Marque solo uno)

- Ha elaborado un plan de trabajo para su proyecto de Big Data
- Ha identificado el grupo de datos que necesita analizar
- Cuenta con un proceso de recopilación de los datos que necesita analizar
- Se encuentra en la etapa de desarrollo de su iniciativa
- Se encuentra evaluando los resultados obtenidos del proyecto Big Data

32. ¿Sus iniciativas de Big Data van enfocadas a? (Marque todo lo que corresponda)

- Comercio electrónico, eBusiness, Operaciones digitales
- Marketing Directo y Digital

- Gestión de fraude
- Análisis de Clientes y Mercado
- Servicio al cliente
- Desarrollo y Gestión de Productos
- Tecnologías de la Información
- Operaciones
- Gestión de riesgos
- Recursos humanos
- Otro (especifique): _____

33. ¿La iniciativa de Big Data es liderado por? (Marque solo uno)

- Área de negocios, con un mínimo de soporte de TI
- Colaboración entre área de negocios y TI
- Principalmente dirigido por TI, con mínima involucramiento áreas de negocios
- Otro (especifique): _____

34. ¿Qué beneficios tangibles espera lograr con sus iniciativas Big Data? (Clasifique del 1 al 7 de lo más importante al menos importante)

- _____ Mejorar la experiencia del cliente
- _____ Aumentar ventas
- _____ Generar Productos y servicios de mayor calidad
- _____ Innovar los productos
- _____ Operar con más eficiencia
- _____ Realizar una mejor toma de decisiones basada en hechos
- _____ Reducir el Riesgo

35. ¿Cómo piensa medir el éxito de sus iniciativas Big Data? (Marque todo lo que corresponda)

- Con métricas cuantitativas atadas al desempeño empresarial
- Con métricas cualitativas vinculadas al desempeño empresarial
- Con métricas cuantitativas atadas al desempeño de TI
- Con métricas cualitativas vinculadas al rendimiento de TI
- No está definido aún
- Otro (especifique): _____

Datos

36. ¿Qué lo lleva a considerar Big Data? (Marque todo lo que corresponda)

	Actualmente	En 3 años
Analizar conjuntos de datos < 1TB	<input type="radio"/>	<input type="radio"/>
Analizar conjuntos de datos 1TB - 100TB	<input type="radio"/>	<input type="radio"/>
Analizar conjuntos de datos 100TB - 1PB	<input type="radio"/>	<input type="radio"/>
Analizar conjuntos de datos > 1PB	<input type="radio"/>	<input type="radio"/>
Analizar nuevos tipos de datos (texto, relaciones, series de tiempo)	<input type="radio"/>	<input type="radio"/>
Analizar los datos de transmisión / comunicación	<input type="radio"/>	<input type="radio"/>
Analizar datos de diversas fuentes	<input type="radio"/>	<input type="radio"/>

37. ¿Qué tipo de datos está considerando tratar con tecnologías Big Data? (Marque todo lo que corresponda)

	Actualmente	En 3 años
Datos Transaccionales	<input type="radio"/>	<input type="radio"/>
Clickstream (Recorrido del usuario en su paso por su sitio web)	<input type="radio"/>	<input type="radio"/>
Contenido no estructurado de correo electrónico, documentos de oficina, etc.	<input type="radio"/>	<input type="radio"/>
Voz / Datos de audio	<input type="radio"/>	<input type="radio"/>
Datos de Redes Sociales (Facebook, Twitter, etc.)	<input type="radio"/>	<input type="radio"/>
Sensores / Datos de Maquinarias / Datos de Dispositivos	<input type="radio"/>	<input type="radio"/>
Localizaciones / Datos geoespaciales	<input type="radio"/>	<input type="radio"/>
Datos científicos / genómicos	<input type="radio"/>	<input type="radio"/>
Datos de Imágenes (videos / fotos)	<input type="radio"/>	<input type="radio"/>

38. ¿Hacia qué dominios de datos se orientan sus iniciativas de Big Data? (Marque todo lo que corresponda)

	Actualmente	En 3 años
Datos del Cliente / Prospecto de cliente	<input type="radio"/>	<input type="radio"/>
Transacciones del cliente	<input type="radio"/>	<input type="radio"/>
Datos del canal	<input type="radio"/>	<input type="radio"/>
Datos del mercado y de la competencia	<input type="radio"/>	<input type="radio"/>
Datos de productos	<input type="radio"/>	<input type="radio"/>
Datos de servicio	<input type="radio"/>	<input type="radio"/>
Datos de la cadena de suministro	<input type="radio"/>	<input type="radio"/>
Detección de fraude	<input type="radio"/>	<input type="radio"/>
Datos específicos de la industria (especifique):	<input type="radio"/>	<input type="radio"/>

39. ¿Qué estructuras de datos son de particular interés en sus iniciativas de Big Data? (Marque todo lo que corresponda)
- | | |
|--|--|
| <input type="radio"/> Archivos Planos | <input type="radio"/> Web Semántica |
| <input type="radio"/> Base de datos Relacional / Datawarehouse | <input type="radio"/> XML |
| <input type="radio"/> Texto no estructurado | <input type="radio"/> Multimedia |
| <input type="radio"/> Serie cronológica / Log de ejecución | <input type="radio"/> Otros (especifique): _____ |
| <input type="radio"/> Gráficos | |

40. ¿Qué desafíos de tratamiento de los datos se están abordando con Big Data? (Marque todo lo que corresponda)
- Integrar una variedad de datos (Propios, de redes sociales, estructurados, no estructurados)
 - Acceder a un mayor rango de datos históricos
 - Uso de datos en tiempo real
 - Comprender datos no estructurados
 - Acceso a datos con mayor granularidad
 - Otros (especifique): _____

Plataformas

Las plataformas tecnológicas de Big Data tienen funcionalidades para la gestión de datos y para el procesamiento analítico. Las siguientes preguntas se refieren a sus prioridades para estas capacidades.

41. ¿Qué funciones de administración de datos son las más importantes para usted? (Marque todo lo que corresponda)
- Almacenamiento de alta capacidad
 - Procesamiento de alto rendimiento
 - Integración de datos
 - Aceleración / Escalado de bases de datos relacionales
 - Gestión y búsqueda de texto no estructurado
 - Otros (especifique): _____
42. ¿Donde planea ejecutar sus aplicaciones Big Data? (marque solo uno)
- Localmente
 - Nube pública
 - Nube privada
 - No lo sé aún
43. ¿Qué enfoques de gestión de datos está considerando? (Marque todo lo que corresponda)
- | | |
|---|--|
| <ul style="list-style-type: none"> <input type="radio"/> Archivo distribuido (ej. Hadoop, Grid) (especifique): _____ <input type="radio"/> Relacional Especializado (ej. Data Warehouse, BD Espaciales / Columnarias / En memoria) (especifique): _____ | <ul style="list-style-type: none"> <input type="radio"/> Relacional Tradicional (ej. SQL Server / Oracle) (especifique): _____ <input type="radio"/> No lo sé <input type="radio"/> Otro (especifique): _____ |
|---|--|

Analítica de datos

44. ¿Qué beneficios analíticos espera obtener del uso de Big Data? (Marque todo lo que corresponda)
- Analizar datos en un tiempo significativamente menor al empleado actualmente
 - Obtener modelos y patrones que brinden nuevos conocimientos respecto a la organización
 - Descubrir información respecto al mercado y obtener nuevas ideas
 - No lo sé aún
 - Otros (especifique): _____
45. ¿Qué tipo de análisis desea realizar? (Marque todo lo que corresponda)
- Análisis y alertas en tiempo real
 - Análisis de texto
 - Análisis de relaciones
 - Análisis de conjuntos de datos dispares
 - Análisis de conjuntos de datos externos
 - Evaluación nuevos algoritmos analíticos
 - Otros (especifique): _____
46. ¿Qué tipos de productos de Análisis de Datos está usando o considerando usar? (Marque todo lo que corresponda)
- Paquetes estadísticos o matemáticos para analizar datos y generar modelos. (Ej. SAS, R, Matlab) (especifique): _____
 - Productos de visualización de datos / Cuadros de mando. (Ej. Tableau, Spotfire) (especifique): _____
 - Herramientas de Análisis personalizadas
 - No lo sé aún
 - Otros (especifique): _____
47. ¿Qué funciones analíticas son más importantes para usted? (Marque todo lo que corresponda)
- Algoritmos avanzados de análisis
 - Visualización de datos
 - Aprendizaje automático

- Ejecutar algoritmos más rápidamente
- Análisis de texto
- Análisis de redes sociales
- Ejecutar algoritmos existentes con conjuntos de datos mucho más grande
- Otros (especifique): _____

48. ¿Planea utilizar aplicaciones de terceros o desarrollar su propia aplicación para el análisis de datos? (Marque todo lo que corresponda)

- Desarrollado por terceros
- Desarrollado internamente
- Otros (especifique): _____

49. ¿Qué lenguajes / herramientas de programación va a utilizar para el desarrollo? (Marque todo lo que corresponda)

- Lenguajes tradicionales (por ejemplo, Java, C, C ++)
- SQL
- Lenguajes de secuencias de comandos (por ejemplo, Python, Perl)
- Bibliotecas de código abierto Lenguajes / bibliotecas específicas del producto
- No lo sé aún
- Otros (especifique): _____

Gente

50. ¿Está contratando o planea contratar a Científicos de Datos?

- Si
- no

51. ¿Cómo obtiene las habilidades de Ciencia de Datos en su organización? (Marque todo lo que corresponda)

- Formación de análisis de datos para los empleados de la empresa.
- Contratación de profesionales con experiencia en análisis de datos.
- Contratación de pasantes o recién graduados de doctorados en matemáticas, ciencias e ingeniería.
- Contratación de pasantes o recién graduados de maestría en matemáticas, ciencias e ingeniería.
- Contratación de empresa de servicios profesionales con la experiencia analítica de Big Data
- Uso de los servicios profesionales de los vendedores de productos
- Otro (especifique): _____

52. ¿Qué tan difícil es para el área de TI encontrar Científicos de Datos?

Nada difícil	Algo difícil	Muy difícil	Extremadamente difícil
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

53. ¿Qué tan difícil es encontrar gerentes y ejecutivos de negocios que puedan identificar y aprovechar las oportunidades de negocio con el uso de herramientas de Big Data?

Nada difícil	Algo difícil	Muy difícil	Extremadamente difícil
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Ejecución

54. ¿Cómo ha gestionado sus iniciativas de Big Data? (Marque solo uno)
- Todo internamente
 - Mayormente interno, con alguna ayuda de terceros
 - Principalmente terceros bajo nuestra dirección y supervisión
 - Todos son terceros con supervisión mínima
 - No lo sé
 - Otro (especifique):

55. ¿Administra usted sus iniciativas en Big Data utilizando los mismos estándares de Gestión de Proyectos y Ciclo de Vida de proyectos de desarrollo de aplicaciones?
- Sí, el mismo
 - No, diferente
 - No sé
56. ¿Sus aplicaciones de Big Data son independientes o están estrechamente integradas o incrustadas a otros sistemas importantes? (Marque solo uno)
- Integrados a aplicaciones empresariales (ERP, CRM)
 - Integrados a Procesos empresariales (BPM)
 - Integrados a Reglas de negocio (BRE)
 - No se integra a otros sistemas
 - Integrados a otros sistemas (especifique):

57. ¿Cuál es el presupuesto total para las iniciativas de Big Data? (Marque solo uno)
- < \$ 100.000
 - \$ 100,000 - \$ 1 Millón
 - \$ 1 millón - \$ 10 millones
 - > \$ 10 millones
 - Lo Desconozco
58. ¿Cuál crees que es la mayor oportunidad para usar Big Data en tu empresa?
- _____
- _____
- _____
59. Por favor, agregue cualquier otro comentario que usted considere una contribución a este estudio.
- _____
- _____
- _____