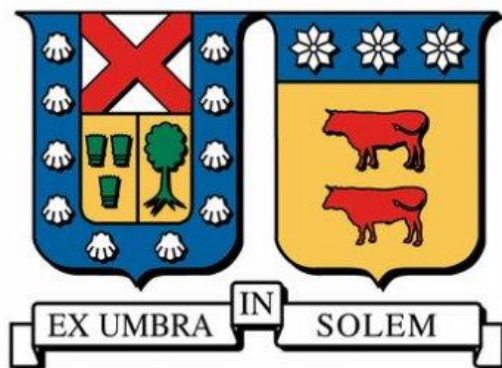


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
VALPARAÍSO – CHILE



**ESTUDIO DE FACTIBILIDAD TÉCNICA Y ECONÓMICA DE LA
IMPLEMENTACIÓN DEL CONTACT CENTER COGNITIVO DE IBM EN UNA
EMPRESA DE RETAIL EN CHILE**

NATALIA CECILIA JARA SÁNCHEZ

**MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERA CIVIL INDUSTRIAL**

PROFESOR GUÍA : RAFAEL FAVEREAU URQUIZA
PROFESOR CORREFERENTE : FREDY KRISTJANPOLLER RODRÍGUEZ

ENERO 2019

A Carolina,

*Porque mientras vivas en mis recuerdos,
jamás te habrás ido. Desde el primer día de
Universidad juntas y en la recta final sé que me
acompañas. Amiga, esto es por las dos.*

1 AGRADECIMIENTOS

A mis padres quienes siempre han sido mi principal apoyo, gracias a ellos y a su constante esfuerzo soy todo lo que soy. Gracias por creer en mi mucho más que yo misma, y por siempre apoyarme a luchar por mis sueños.

A mi familia y a todos aquellos cercanos que, aun no siendo familia, se interesaron por mi bienestar, se alegraron por mis logros y me apoyaron en las derrotas.

A mis amigas, por siempre estar presentes cuando las necesité, entregando alegría, apoyo y paciencia aun en los momentos más difíciles y de máximo estrés. Gracias por tantas risas, conversaciones y gritos, mi vida universitaria no hubiese sido lo mismo sin ustedes.

A IBM y su equipo de coaches, quienes me transmitieron su experiencia para el desarrollo de este trabajo, confiaron en mis capacidades, y fueron parte importante de mi iniciación en la vida laboral.

Al equipo de trabajo DoLab, quienes ayudaron desinteresadamente a guiar mi investigación, y que se convirtieron en grandes amigos. Gracias por hacer de la jornada laboral algo entretenido.

A mis profesores, quienes me formaron como ingeniero durante estos largos 7 años, entregándome las herramientas necesarias para desempeñarme en el mundo laboral.

2 ÍNDICE

2.1 ÍNDICE GENERAL

1	Agradecimientos	3
2	Índice	4
2.1	Índice General.....	4
2.2	Índice de figuras	9
2.3	Índice de tablas	11
3	Resumen Ejecutivo	14
4	Introducción	16
5	Problema de investigación	18
6	Objetivos.....	21
6.1	Objetivo general	21
6.2	Objetivos específicos.....	21
7	Marco teórico	22
7.1	Servicio al cliente	22
7.2	Contact Center	24
7.2.1	Evolución del Contact Center	24
7.2.2	Tipos de Contact Center.....	25
7.2.3	Tecnologías involucradas.....	26
7.2.4	Métricas de desempeño	28
7.2.5	Contact Center en Chile	30
7.3	Cloud Computing	33

7.3.1	IBM Cloud.....	35
7.4	Computación Cognitiva.....	35
7.4.1	Aplicaciones del servicio al cliente.....	37
7.4.2	IBM Watson	39
7.5	Contact Center Cognitivo	40
7.5.1	Contact Center Cognitivo de IBM	40
8	Metodología	48
8.1	Análisis de Factibilidad	48
8.1.1	Factibilidad Técnica	48
8.1.2	Factibilidad Económica.....	49
9	Descripción de la situación sin proyecto	51
10	Implementación de la solución en modalidad MVP	59
10.1	Análisis de los resultados	61
10.2	Definición de supuestos.....	64
11	Factibilidad Técnica	66
11.1	Tamaño del proyecto	66
11.2	Arquitectura de la solución.....	69
11.3	Funcionamiento y descripción del proceso de interacción.....	72
11.4	Hardware requerido	75
11.5	Descripción y dimensionamiento de los componentes.....	76
11.5.1	Watson Assistant.....	76

11.5.2	Watson Speech to Text (STT).....	80
11.5.3	Text to Speech (TTS)	81
11.5.4	IBM Voice Gateway.....	82
11.5.5	TMP.....	84
11.5.6	SigAS	86
11.5.7	IBM Cloud Functions.....	87
11.5.8	IBM Cloudant.....	88
11.5.9	IBM Cloud Object Storage.....	89
11.5.10	Reportería de datos	92
11.6	Requerimientos de personal.....	94
11.6.1	Consultoría IBM.....	95
11.6.2	Equipo de trabajo Empresa Cliente	97
11.7	Restricciones técnicas.....	99
11.7.1	Latencia en la comunicación	99
11.7.2	Conectividad.....	100
12	Factibilidad Económica.....	102
12.1	Costos de la solución	102
12.1.1	Hardware	103
12.1.2	Servicios y Aplicaciones	105
12.1.3	Administración de sistemas de TI	106

12.1.4	Soporte	107
12.1.5	Desarrollo	109
12.1.6	Comunicaciones	110
12.1.7	Costo total de propiedad.....	111
12.2	Beneficios de la solución.....	112
12.2.1	Beneficio tangible	112
12.2.2	Beneficios intangibles	114
12.3	Evaluación de la rentabilidad	117
12.3.1	Valor Actual Neto (VAN).....	117
12.3.2	Tasa interna de retorno (TIR).....	118
12.3.3	Periodo de recuperación de la inversión (PRI).....	118
12.3.4	Retorno sobre la inversión (ROI)	119
12.3.5	Análisis de los indicadores	119
12.4	Análisis de sensibilidad	120
12.5	Modelo unidimensional de sensibilización	121
12.5.1	Tasa de contención objetivo.....	121
12.5.2	Cantidad de llamadas	123
12.5.3	Tasa de crecimiento de la demanda anual por servicio	124
12.5.4	TMO de ejecutivos y asistente virtual.....	124
12.5.5	Beneficios por reducción de ejecutivos.....	126

12.6	Modelo multidimensional de sensibilización	128
13	Conclusiones	132
14	Recomendaciones y futuras líneas de investigación	134
15	Referencias	136

2.2 ÍNDICE DE FIGURAS

Figura 1: Arquitectura básica de un Contact Center.....	26
Figura 2: Distribución de ingresos por servicios tercerizados de Contact Center según sector.	32
Figura 3: Distribución de ingresos por servicios tercerizados de Contact Center según tipo de servicio.	32
Figura 4: Distribución de ingresos por servicios tercerizados de Contact Center según canal de comunicación.	33
Figura 5: Modelos de Cloud Computing.	34
Figura 6: Diagrama general de funcionamiento del Contact Center Cognitivo de IBM.....	43
Figura 7. Nivel de ventas en USD por trimestre, desde el año 2014 al 2017.....	51
Figura 8. Diagrama de arquitectura del Contact Center de la empresa	53
Figura 9: Cantidad de llamadas atendidas diariamente en el periodo de evaluación.	61
Figura 10: Distribución de llamadas ingresadas por opción de despacho..	62
Figura 11: Diagrama de contención por etapa del proceso de atención en el seguimiento de despachos.....	63
Figura 12. Línea de tiempo de la implementación de un Asistente virtual. Fuente: Elaboración Propia a partir del documento Customer Care Voice Agent (IBM, 2018)	68
Figura 13: Diagrama de arquitectura general del Contact Center Cognitivo de IBM.....	71
Figura 14. Diagrama de proceso de atención telefónica con asistente virtual.....	74
Figura 15. Diagrama de conexión del TMP con el Contact Center de la empresa de retail.	85
Figura 16. Vista principal dashboard Contact Center Cognitivo.	93
Figura 17. Vista de detalle de llamada dashboard Contact Center Cognitivo.....	94

Figura 18. Desglose del costo total de propiedad del proyecto en USD llevado a valor presente.....	111
Figura 19. Costo total de propiedad en USD mensualizado en el horizonte de evaluación.	111
Figura 20.Cantidad de minutos mensuales requeridos de ejecutivos en el horizonte de evaluación, con y sin la implementación del Contact Center Cognitivo.....	113
Figura 21. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación.....	113
Figura 22.Valor actual neto (USD) en función del porcentaje promedio ponderado de contención objetivo.	123
Figura 23. Valor actual neto (USD) en función del factor de crecimiento de la demanda por servicio.	124
Figura 24.Valor actual neto (USD) en función del TMO (min) del ejecutivo	125
Figura 25.Valor actual neto (USD) en función del TMO (min) del ejecutivo..	125
Figura 26. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante.....	127
Figura 27.Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante cada año.	127
Figura 28.Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante cada semestre	128
Figura 29. Distribución del VAN por simulación Montecarlo con 1.000 iteraciones.....	130
Figura 30. Contribución a la varianza del VAN de cada una de las variables sensibilizadas.	131

2.3 ÍNDICE DE TABLAS

Tabla 1. Tipos de consultas por opción de IVR.	54
Tabla 2. Clasificación de consultas según dificultad.....	55
Tabla 3. Porcentaje de llamadas recibidas y tasa de abandono por opción de IVR en el año móvil entre mayo 2017 y mayo 2018.	56
Tabla 4. Tipo de consultas respondidas por el chatbot implementado.....	57
Tabla 5. Proyección de llamadas a recibir en el Contact Center de la empresa en el horizonte de evaluación.	66
Tabla 6. Estimación de la contención (%) a alcanzar por opción de IVR.....	67
Tabla 7. Volumen de llamadas estimado a contener por la solución en el horizonte de evaluación.....	68
Tabla 8. Cantidad de interacciones en el canal web en el horizonte de evaluación	68
Tabla 9. Estimación de la cantidad de minutos atendidos por la solución durante el horizonte de evaluación.	69
Tabla 10. Características de modalidades de despliegue de Watson Assistant.	79
Tabla 11. Tarifa por minuto utilizado del servicio de Watson STT.....	81
Tabla 12. Especificaciones técnicas del servidor para la ejecución del servicio de TTS Nuance Vocalizer.	82
Tabla 13. Especificaciones técnicas del servidor para la ejecución del servicio IBM Voice Gateway.....	83
Tabla 14. Tarifa por minuto utilizado del servicio IVG	84
Tabla 15. Especificaciones técnicas del servidor para la instalación del software TMP. ...	86
Tabla 16. Especificaciones técnicas del servidor para la instalación del software SigAS. .	87
Tabla 17. Especificaciones del uso de IBM Cloud Functions.	87

Tabla 18. Especificaciones del uso de la base de datos Cloudant.	89
Tabla 19. Estimación de uso de Cloudant en el horizonte de evaluación	89
Tabla 20. Estimación del uso de IBM Cloud Object Storage con sus respectivos costos....	92
Tabla 21. Dedicación y costo de cada rol requerido en la célula de consultoría de IBM ...	96
Tabla 22. Presencia, dedicación y costo de cada rol requerido en la célula de consultoría de la empresa cliente.	98
Tabla 23. Especificaciones del servidor seleccionado.....	104
Tabla 24. Valor referencial en USD del servidor y disco duro seleccionados.	104
Tabla 25. Valor referencial en USD de licencias a utilizar	104
Tabla 26. Costo anual en USD por el uso de los servicios de IBM Cloud.....	105
Tabla 27. Costo anual en USD por el uso de las aplicaciones de Sixbell.	106
Tabla 28. Niveles de soporte de TI.....	107
Tabla 29. Severidad, tiempo de respuesta y cobertura del servicio de soporte de IBM Cloud.	108
Tabla 30. Severidad, tiempos de atención y solución de incidentes del servicio de soporte de Sixbell.....	109
Tabla 31. Costo anual en USD del personal requerido.....	110
Tabla 32. Beneficios anuales en USD por ejecutivos liberados en el horizonte de evaluación.	114
Tabla 33. Indicadores promedio de e-commerce entre enero y agosto del 2018.	115
Tabla 34. Resumen de los indicadores de rentabilidad obtenidos.....	120
Tabla 35. Límite superior de contención objetivo a alcanzar para el análisis de sensibilidad.	122
Tabla 36. Valor de la contención objetivo que genera un VAN igual a 0.....	122

Tabla 37. Proyección de llamadas a recibir en el Contact Center de la empresa, que genera un VAN igual a 0.....	123
Tabla 38. Detalle de los escenarios para la sensibilización.....	129

3 RESUMEN EJECUTIVO

Los centros de contacto afrontan actualmente el desafío constante de cumplir con las metas de costo impuestas por el negocio, y a la vez, con las expectativas del cliente en términos de calidad de atención. En este contexto, el *Contact Center Cognitivo* se plantea como una alternativa de autoservicio, que atiende los requerimientos de los clientes en múltiples canales mediante la figura de un asistente virtual con disponibilidad 24x7, capaz de entender, en lenguaje natural las consultas realizadas, conectarse con los sistemas de la empresa, y entregar una respuesta.

El objetivo de esta investigación es determinar la viabilidad técnica y económica de la implementación del *Contact Center Cognitivo* de IBM, en el servicio de atención al cliente de una empresa de retail en Chile. Esto se lleva a cabo mediante un análisis de factibilidad, que utiliza la metodología de evaluación de proyecto factible.

Del estudio técnico se tiene que la implementación del *Contact Center Cognitivo* no presenta limitaciones para su instalación en el *Contact Center* ya existente, encontrándose todos los servicios *cloud* a utilizar disponibles para Chile, a través de la plataforma IBM Cloud, y sin inconvenientes de capacidad dado el tamaño del proyecto. Se destaca sin embargo que, para su operación es requerido el uso de un enlace dedicado con un ancho de banda mínimo de 6 Mbps, y para su monitoreo y control, se indica la necesidad de contar con personal de la empresa que realice los ajustes necesarios a la solución para garantizar su correcto funcionamiento a lo largo del tiempo.

En el estudio económico, realizado en un horizonte de evaluación de tres años, se detallan los costos y beneficios, obteniendo por resultado un costo total de propiedad de

744.266 USD, y beneficios por ahorro en costos de 1.139.543 USD. Con esto es posible calcular indicadores de rentabilidad tales como el VAN (395.277 USD), TIR (7,89%) y ROI (53%), los que permiten concluir que el proyecto es factible económicamente.

Para finalizar se realiza un análisis de sensibilidad unidimensional y multidimensional, respecto a distintas variables de interés. Del análisis unidimensional se tiene que dos variables críticas corresponden al TMO y a la cantidad de ejecutivos contratados por políticas de la empresa, resultando en un proyecto no rentable si la empresa decide no reducir su personal del inicio al fin del horizonte de evaluación. En el análisis multidimensional, se analizan las variaciones de todas las variables definidas de interés, de manera simultánea, en tres escenarios, pesimista, más probable y optimista, obteniendo como resultado un VAN ajustado a una distribución de probabilidad beta con media 177.814 USD, esto indica que a pesar de los escenarios escogidos, el proyecto tiene una alta probabilidad de ser atractivo económicamente.

Finalmente, los resultados en su conjunto permiten catalogar la implementación del *Contact Center Cognitivo* como un proyecto factible técnica y económicamente.

4 INTRODUCCIÓN

Los centros de contacto le permiten a una compañía construir, mantener y administrar las relaciones con los clientes al entregar información, resolver problemas, quejas, y preguntas; por lo que corresponden a una herramienta fundamental para la gestión de relaciones con los clientes.

Actualmente se está viviendo un proceso de transformación digital en el que los modelos empresariales necesitan de soluciones que les ayuden gestionar estas relaciones, logrando entregar a bajo costo un servicio de alta calidad que permita comprender a sus clientes en cada interacción para así optimizar su relación con ellos. En este contexto, los centros de contacto cognitivos pueden comprender, razonar y aprender de grandes cantidades de datos, mejorando la experiencia del cliente, a la vez que entregan una alta disponibilidad de atención a un costo menor comparado con los agentes humanos.

La presente investigación tiene como objetivo evaluar la factibilidad técnica y económica de implementar, en una empresa de retail en Chile, el *Contact Center Cognitivo* de IBM, el cual corresponde a una solución capaz de atender el canal web y telefónico de la organización, mediante la figura de un asistente virtual capaz de entender, en lenguaje natural, los requerimientos de los clientes e integrarse a los sistemas de la empresa para dar respuesta a las consultas realizadas.

El estudio del proyecto de implementación consta principalmente de cuatro partes: la descripción de la situación sin proyecto, la implementación en modalidad MVP¹, el estudio

¹ Producto mínimo viable. Correspondiente a un prototipo básico funcional de un producto o servicio.

técnico y el estudio económico. En la primera parte se describe la empresa, y las características de su actual *Contact Center*, identificando métricas de interés como la cantidad de llamadas recibidas, tasa de abandono, tiempo medio de operación, y consultas más frecuentes.

En la segunda parte se detallan los resultados obtenidos en el desarrollo de un prototipo de la solución, realizado entre el 26 de mayo y el 15 de junio del presente año, cuya implementación comprende la atención, mediante el canal telefónico, de dos opciones del actual IVR de la empresa. A partir de estos resultados, se establecen distintos supuestos y parámetros para la construcción de los estudios de factibilidad técnica y económica.

En el estudio técnico, se indican los aspectos técnicos que concierten a la implementación del *Contact Center Cognitivo*, tales como su arquitectura, definición y dimensionamiento de sus componentes, requerimiento de capital humano, y algunas restricciones que se deben tener en cuenta para su correcto funcionamiento.

En el estudio económico se determinan, en primer lugar, los costos que implica llevar a cabo el proyecto, mediante el método de costo total de propiedad y, en segundo lugar, se detallan los beneficios tangibles e intangibles de su implementación, finalizando con el cálculo y análisis de indicadores de rentabilidad como VAN, TIR, y ROI que permiten concluir respecto a la factibilidad económica de realizar el proyecto.

Por último, se presenta un análisis de sensibilidad, con el propósito de evaluar el comportamiento del VAN frente a variaciones de parámetros de interés, reconociendo aquellos que impactan con mayor fuerza a la rentabilidad del proyecto.

5 PROBLEMA DE INVESTIGACIÓN

Prácticamente todas las empresas están interesadas en proporcionar información y asistencia a los existentes y futuros clientes, contando con diferentes medios para ser contactados (Aksin, Armony, & Mehrotra, 2009). En este contexto, los centros de llamados o *Call Centers* son una vía de comunicación indispensable para muchos negocios comprometidos a dar respuesta en tiempo real a la demanda creciente por servicio de sus clientes, pudiendo manejar tres necesidades claves: fácil acceso, cercanía, y obtención de información personalizada y detallada (Feinberg, Kim, Hokama, de Ruyter, & Keen, 2000).

En un escenario de continua adopción tecnológica y conectividad, la manera en la que los clientes eligen comunicarse con las organizaciones ha ido cambiando. Esto ha provocado la evolución natural del concepto de *Call Center* a *Contact Center* donde la comunicación puede ocurrir a través de múltiples canales, teniendo como foco mejorar la satisfacción y experiencia del cliente, dado que su funcionamiento y calidad de atención, tiene un impacto significativo en las actitudes del cliente hacia una empresa y sus productos (Mehrotra, Grossman, & Samuelson, 2013).

Ahora, si bien el cliente cuenta con una amplia variedad de canales para comunicarse con la compañía, es el canal de voz uno de los preferidos al momento de realizar consultas más complejas y el más crítico en términos de costos y calidad de servicio (Ovum Consulting, 2016). Los altos costos afectan principalmente a las soluciones intensivas en personal basadas en operadores telefónicos, mientras que la baja calidad en la interacción se hace presente en las soluciones de autoservicio basadas en tecnología de tono táctil y respuestas pregrabadas (Cena & Torre, 2004). Es por esto que, en la última década, se ha producido un escenario en

el que ambos sistemas coexisten con el objetivo de generar un balance entre costo y calidad de servicio, generalmente bajo el uso de sistemas de Respuesta Interactiva de Voz o *Interactive Voice Response* (IVR), mediante el cual, parte de las interacciones del agente son reemplazadas por un guion interactivo pregrabado, al cual el cliente puede responder presionando la opción del menú correspondiente a su consulta en el discador de un teléfono (sistemas *touch-tone*), o mediante su voz utilizando *Automatic Speech Recognition* (ASR), que permite al cliente seleccionar una opción ya sea diciendo el número, o bien una palabra clave previamente mencionada en el menú de opciones (Suhm & Peterson, 2001).

Si bien este último corresponde a un enfoque de interacción más humano, los problemas frecuentes en el entendimiento y las respuestas estandarizadas impactan negativamente en la experiencia de servicio del cliente, luego se genera la percepción de una paradoja en la que un sistema que intenta reproducir el estilo conversacional de interacción a menudo no satisface las necesidades y expectativas del usuario como evitar la aburrida repetición de instrucciones y navegación por distintos menús jerarquizados con exceso de niveles (Cena & Torre, 2006).

En términos administrativos, los centros de contacto son manejados generalmente como un centro de costos por lo que, minimizar los costos del personal asociado es una importante preocupación comercial (Aguir, Karaesmen, Aksin, & Chauvet, 2004). Al mismo tiempo, el mercado en el que opera es altamente competitivo y brindar un buen servicio a los clientes constituye una ventaja por sobre la competencia, por lo que cada vez más se espera que los administradores de *Contact Center*, logren entregar una alta calidad de servicio a bajos costos (Aksin, Armony, & Mehrotra, 2009).

Según Leggett (2017), el uso de Inteligencia Artificial (IA) es el núcleo de la transformación digital que mejora los costos y las capacidades de todos los aspectos de las empresas, incluido el servicio al cliente. Esto porque la IA puede ayudar a los agentes a completar (o asumir por completo) tareas repetitivas y predecibles, e interactuar con los clientes de manera autónoma mediante el uso de asistentes virtuales integrados y consistentes en todos los canales de comunicación. Esto agrega valor al proceso de atención mediante una mayor eficiencia, mayor compromiso de autoservicio, y mejor control de la experiencia del cliente.

En este contexto, las herramientas cognitivas presentan al mercado un servicio que da cobertura a los *Contact Center* en áreas de gestión del conocimiento para el agente, análisis de información, interacciones omnicanal e integración asistida al cliente por medio de asistentes virtuales, los cuales pueden transcribir llamadas en tiempo real, junto con identificar y comprender frases, gramática y relaciones entre palabras que son clasificadas y almacenadas para posteriores análisis de grandes volúmenes de datos (Leggett, 2017).

La aplicación de esta nueva tecnología a los servicios de *Contact Center* podría, en general, aportar a la gestión homogénea del conocimiento para dar servicio en la totalidad del entorno de canales, y en específico, ayudar de manera significativa a descongestionar la atención de los operadores en preguntas rutinarias, derivando a ellos sólo cuando se presenten requerimientos más complejos, beneficiando tanto a clientes como operadores y, en consecuencia, generando un mejor balance entre la relación costo – calidad en el canal telefónico. Esto sugiere como interrogante ¿es técnica y económicamente factible la implementación del *Contact Center Cognitivo* de IBM en el centro de atención al cliente de una empresa?

6 OBJETIVOS

6.1 OBJETIVO GENERAL

Determinar la factibilidad técnica y económica de la implementación del *Contact Center Cognitivo* de IBM mediante un análisis de factibilidad para su posible utilización en el servicio de atención al cliente de una empresa de retail en Chile.

6.2 OBJETIVOS ESPECÍFICOS

- Determinar la viabilidad técnica de implementar el *Contact Center Cognitivo* en el *Contact Center* de una empresa de retail en Chile, mediante un análisis de factibilidad técnica.
- Definir los costos operacionales y de capital asociados a la implementación del *Contact Center Cognitivo*.
- Cuantificar los beneficios por ahorro en costos producto de la implementación del *Contact Center Cognitivo*.
- Obtener la rentabilidad de la implementación del *Contact Center Cognitivo* en una empresa de retail de Chile.

7 MARCO TEÓRICO

7.1 SERVICIO AL CLIENTE

Previo a la existencia del teléfono, el servicio al cliente estaba limitado por las interacciones presenciales, luego con la aparición de los *Call Centers* se dio paso a la creación de un departamento de servicio al cliente como tal, capaz de entregar respuestas a las problemáticas de clientes en un horario determinado sin necesidad de acudir a ellos de manera presencial. La llegada del internet posibilitó la vía de comunicación digital permitiendo a las empresas utilizar el correo electrónico y el chat para conectarse personalmente con la mayoría de los clientes, y finalmente, el surgimiento de las redes sociales como Facebook y Twitter se consolidaron en la vida de los consumidores, y las empresas altamente orientadas al cliente vieron en estas plataformas un canal más para comercializar y conectarse con ellos.

De esta forma, la relación entre los clientes y los representantes de servicio al cliente ha cambiado, teniendo profundas consecuencias en las empresas provocando que los centros de servicios al cliente modernos deban adaptarse a las necesidades de los clientes y las nuevas tecnologías de la información. En términos de servicios se está viviendo un proceso de transformación digital que tiene como centro la experiencia del cliente o *Customer Experience*, concepto que ha sido ampliamente estudiado en la literatura. Lemon y Verohef (2016) concluyen que “la experiencia del cliente es una construcción multidimensional que se centra en las respuestas cognitivas, emocionales, conductuales, sensoriales y sociales del cliente a las ofertas de una empresa durante todo el recorrido de compra del cliente”, y reconocen la importancia de gestionar estas experiencias con el objetivo de crear relaciones duraderas entre la empresa y sus clientes.

Según Tarokh & Ghahremanloo (2007), el desafío actual de las organizaciones es gestionar las interacciones de los clientes a través de múltiples canales de comunicación, incluidos la web, los centros de llamadas, las ventas de campo y los distribuidores o las redes de socios, en sus múltiples líneas de negocios (si las tienen), con el objetivo de facilitar a los clientes hacer negocios con la organización de la forma que quieran en cualquier momento, a través de cualquier canal, en cualquier idioma o moneda y hacer que los clientes perciban que están tratando con una organización unificada que los reconoce en cada canal de contacto.

La interacción de los clientes a través de múltiples canales se conoce como multicanalidad, experiencia que ha evolucionado hacia la omnicanalidad, concepto que se entiende como la estrategia y gestión de canales que tiene como objetivo la integración y alineación de todos los canales disponibles, con el fin de brindar a los clientes una experiencia de usuario homogénea a través de los mismos (Deloitte, 2016).

Generar canales alternativos a través de medios digitales, permite extender el tiempo de atención al cliente, alcanzando horarios de servicio 24x7, minimizando la gestión de las “colas”, tanto telefónica como presencial, y atrae también una ventaja en costos, teniendo en cuenta que los canales presenciales y de voz son los más costosos al ser intensivos en capital humano (BearingPoint, 2012).

Deloitte (2016) en su reporte sobre la omnicanalidad, indica que “dar un paso más y brindar una experiencia omnicanal conlleva un conjunto importante de oportunidades para mejorar la eficiencia en su operativa y generar nuevas oportunidades de negocios, además de generar importantes mejoras en la calidad del servicio durante la venta o posterior a la misma”. Por último, reconoce que la adopción de una estrategia omnicanal implica gestionar

de manera apropiada la gran cantidad información generada través de los distintos canales, para obtener patrones que permitan conocer al cliente de manera de generar “la mejor próxima acción”, aumentando las probabilidades de éxito al momento de ofrecer un determinado servicio en el momento y lugar adecuado y por el canal de preferencia.

7.2 CONTACT CENTER

Según la Asociación Chilena de Empresas de *Call Center* (ACEC), un *Contact Center* es una plataforma de interacción multicanal, es decir, que atiende, mediante la figura de un agente, cualquier medio de contacto (teléfono, web, mensajería instantánea, etc.) con el fin de acercar, rentabilizar y optimizar las transacciones entre una empresa con sus proveedores y su actual y/o potencial mercado, en reemplazo o complemento del sistema presencial de atención.

7.2.1 EVOLUCIÓN DEL CONTACT CENTER

La figura de un centro de atención a clientes de manera no presencial tiene sus inicios en década de los 70, basando su funcionamiento en la atención telefónica por lo que eran conocidos como *Call Center*. Estos iniciaron sus operaciones con una infraestructura de hardware sencilla, donde el núcleo era una central PBX (*Private Branch Exchange*), capaz de gestionar las llamadas entrantes y salientes, acompañada de un ACD (*Automatic Call Distribution*), para realizar el filtro y asignación de las llamadas al pool de agentes disponibles, permitiendo una carga de trabajo homogénea entre estos. Posteriormente en la búsqueda de la convergencia de voz y datos, fue integrado un servicio CTI (*Computer Telephony Integration*), encargado de enviar a los agentes del *Call Center* información de los clientes, si puede obtenerse, en el momento que se produce la llamada mediante el despliegue de una ventana en la pantalla del agente.

Con el tiempo fue incrementando la demanda por atención, lo que generó la necesidad de sistemas de autoservicio para aumentar la eficiencia como el *touch-tone* (IVR's), que permite la navegación en un menú de opciones realizando pulsaciones en las teclas del teléfono para poder identificar las consultas y dar respuestas pregrabadas a los clientes.

Más adelante, el uso masivo de internet y del comercio electrónico, dieron paso a la necesidad de atender nuevos canales de comunicación, otorgando flexibilidad a la infraestructura utilizada de manera de poder atender diversos canales de comunicación dando paso al concepto de *Contact Center*.

7.2.2 TIPOS DE CONTACT CENTER

En la industria del *Contact Center* se encuentran una amplia variedad de empresas que desarrollan estos servicios, diferenciándose en términos generales, entre las que se dedican a atender únicamente a sus propios clientes (*inhouse*) y las que proveen atención de clientes a otras compañías, esta modalidad corresponde a *insourcing* cuando son prestados dentro de las instalaciones de la empresa cliente, y a *outsourcing* cuando el servicio es desarrollado en el establecimiento de la empresa proveedora.

A su vez los *Contact Centers* que proveen servicios externos, pueden hacerlo en el mercado interno (*inshore*) o bien hacia el extranjero (*offshore*).

En cuanto a la variedad de los servicios ofrecidos: atención de clientes, atención de reclamos, televentas, telemarketing, investigación de mercado, cobranza, mesas de ayuda, fidelización, entre otros, los *Contact Centers* pueden clasificarse en *inbound* (recepción de llamadas), *outbound* (salida de llamadas), y *blending* o bidireccionales (hacen y reciben llamadas) (Uribe-Echevarría & Morales, 2010).

7.2.3 TECNOLOGÍAS INVOLUCRADAS

El funcionamiento de los centros de contacto se basa en la convivencia entre el capital humano y la tecnología, en este último punto se reconocen algunos de los elementos que son claves para su desarrollo. En la **Figura 1** se muestra una arquitectura simple de *Contact Center*.

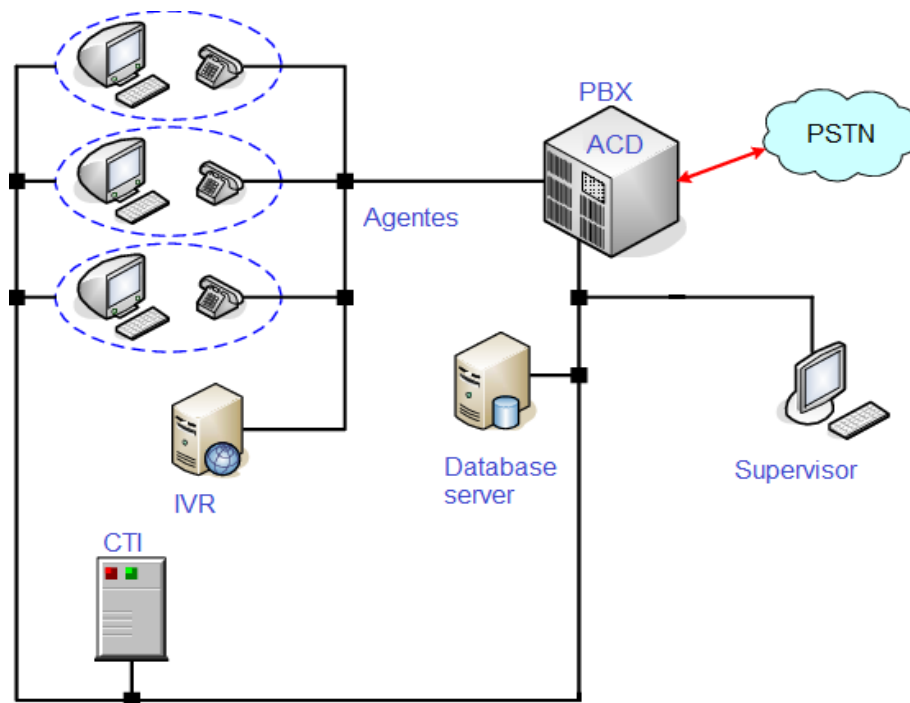


Figura 1. Arquitectura básica de un Contact Center.
 Fuente: (Chromy, Diezka, Kavacky, & Vozňák, 2011)

7.2.3.1 PRIVATE BRANCH EXCHANGE (PBX)

Corresponde a una central telefónica privada automática para gestionar las llamadas internas, entrantes y salientes sin necesidad de contar con operadores que manualmente conecten los cables para establecer la comunicación.

7.2.3.2 AUTOMATIC CALL DISTRIBUTION (ACD)

Corresponde a un software diseñado para procesar eficientemente grandes cantidades de llamadas entrantes y asegurar que sean atendidas en un tiempo determinado, filtrándolas y asignándolas a los distintos agentes disponibles normalmente organizados en diferentes grupos de trabajo (Lovera Raffo, 2008).

7.2.3.3 COMPUTER TELEPHONY INTEGRATION (CTI)

Servicio que se conecta con la PBX para lograr la interacción entre una llamada telefónica y un computador de manera coordinada permitiendo, por ejemplo, desplegar una ventana en la pantalla del agente con información del llamante, administrar colas virtuales, iniciar, manejar y rutear llamadas, entre otras funcionalidades.

7.2.3.4 INTERACTIVE VOICE RESPONSE (IVR)

Es un sistema automatizado de respuesta utilizado para brindar un menú de autonavegación en el que el agente es reemplazado por un guion pregrabado al cual el llamante responde mediante la señal que se genera al oprimir las teclas de un teléfono (DTMF o *Dual-tone multi-frequency*), o bien mediante respuestas simples especificadas (sí o no, por ejemplo).

Un IVR se puede utilizar para establecer el contacto inicial en una llamada identificando las necesidades de quien llama a través de la obtención directa de información, como por ejemplo número de RUT, número de cuenta y claves de acceso, o bien de manera indirecta mediante la navegación en el menú de opciones, lo que permite el redireccionamiento de la llamada a un agente especializado para responder a su requerimiento, o en algunos casos, responder con información pregrabada sin necesidad de

que intervenga un operador. De la misma forma puede utilizarse para operar solo en casos que no existan agentes disponibles.

Los sistemas de IVR pueden atender, a un mucho menor costo de operación, una gran cantidad de llamadas con un alto nivel de disponibilidad, pudiendo operar para respuestas automatizadas las 24 horas del día y todos los días de la semana, además es particularmente útil en el tratamiento de datos sensibles, como contraseñas y números de cuenta, acción especialmente compleja de mantener confidencial si se tratase con agentes humanos (Corkrey, 2002). Sin embargo, los sistemas de autoservicio presentan importantes problemas que disminuyen la calidad del servicio prestado, algunos de ellos son: incapacidad de abordar las diferentes necesidades de los diferentes usuarios (por ejemplo clientes insatisfechos, novatos o de alto valor), alta probabilidad de generar malos entendidos, menús jerarquizados con demasiada complejidad que retardan mucho la atención del cliente y que finalmente termina de igual forma contactado a un ejecutivo, incapacidad de enfrentar casos especiales y situaciones inesperadas, entre otros (Cena & Torre, 2004).

7.2.4 MÉTRICAS DE DESEMPEÑO

Según Contactbabel (2017), el éxito de los *Contact Center* se ha medido tradicionalmente mediante la observación de métricas relacionadas con el costo y la eficiencia, tales como: duración promedio de la llamada, porcentaje de llamadas respondidas dentro de un periodo de tiempo, velocidad promedio de respuesta, etc.

Es importante mencionar que, si bien existe un número importante de *Call Centers* que se están convirtiendo en centros de contacto multi-canal, existen escasas métricas que evalúen el rendimiento de otros canales distintos al telefónico, sin embargo se espera que en el futuro se desarrollen indicadores que puedan evaluar todos los canales utilizados para

comprender de mejor manera el desempeño del *Contact Center* en su totalidad (Belfiore, Chatterley, & Petouhoff, 2012).

A continuación, se detallan las métricas para medir el desempeño de un *Contact Center*, mencionadas en el reporte de BenchhmarkPortal (2017), que serán utilizadas en el desarrollo de este trabajo.

7.2.4.1 COSTO PROMEDIO POR LLAMADA O *COST PER CALL (CPC)*

Este indicador corresponde al costo promedio de procesar una llamada recibida o emitida en un *Contact Center*. Los costos provienen de, por ejemplo: sueldos de los operadores, costo de los equipos y usos de líneas, depreciación del inmueble en el que se trabaja (si es propio) o costo de arriendo, entre otros. Esta es una de las métricas más comúnmente revisadas y utilizadas para medir la eficiencia de un *Contact Center*, sin embargo su interpretación debe ser realizada con precaución debido a que, por ejemplo, un valor creciente de este indicador puede ser bueno si es producto de la recepción de un menor número de llamadas debido a la mejora de procesos de la compañía, lo que implica que el costo fijo de operación del *Contact Center* se prorratee en una menor cantidad de llamadas (Incoming Calls Managment Institute (ICMI), 2003)).

$$CPC = \frac{\text{Costos totales}}{\text{Total de llamadas}}$$

7.2.4.2 TIEMPO MEDIO DE OPERACIÓN (TMO) O *AVERAGE TALK TIME (ATT)*

Representa la cantidad real de tiempo invertido en comprender y resolver el problema de la persona que llama y se calcula como el tiempo promedio que le toma a un agente atender un requerimiento considerando el tiempo de conversación y en espera. Su duración dependerá

tanto de la dificultad de la consulta como de la velocidad de los sistemas de apoyo al *Contact Center* y la experiencia del agente.

$$TMO = \frac{\text{Duración de las llamadas atendidas}}{N^{\circ} \text{ de llamadas atendidas}}$$

Cuando se incluye además el tiempo de trabajo posterior a la llamada, este indicador pasa a llamarse Tiempo promedio de manejo de la llamada o *Average Call Handle Time* (AHT). Tanto el TMO como el AHT, corresponden a métricas altamente observadas dentro de un centro de contacto, al considerarse como un indicador de la habilidad y productividad del agente. Se recomienda analizar su valor en conjunto con el FCR, esto porque una atención rápida no asegura que realmente se haya dado una solución al requerimiento.

7.2.4.3 TASA DE ABANDONO PROMEDIO

El abandono corresponde a la cantidad de llamadas que se conectan al ACD, es decir son detectadas por el *Contact Center*, pero la persona que llama se desconecta voluntariamente antes de llegar a un agente o de completar un proceso dentro del IVR. La tasa de abandono mide el porcentaje de estas llamadas en comparación con las recibidas. Esta métrica está fuertemente relacionada con la velocidad promedio de atención.

$$\text{Tasa de abandono} = \frac{N^{\circ} \text{ de llamadas abandonadas}}{N^{\circ} \text{ de llamadas recibidas}}$$

7.2.5 CONTACT CENTER EN CHILE

Según Uribe-Echeverría & Morales (2010), la industria de *Contact Center* en Chile se presenta como una industria rentable y en constante crecimiento, ligado a la expansión de las comunicaciones y a los avances en la tecnología de la información. Es además intensiva en capital humano y aunque los autores reconocen la necesidad de avanzar en la

profesionalización de la industria, apuntando a ofrecer servicios más especializados y con mayor valor agregado, actualmente es un trabajo que requiere de competencias básicas y baja calificación. En este escenario, las empresas compiten por costos y los empresarios del sector reclaman mayores grados de flexibilidad laboral, tanto para reducir el precio de la mano de obra como para ajustar el proceso de trabajo al dinamismo que requiere la industria.

En términos generales, los servicios ofrecidos por estas empresas corresponden a: atención de clientes (*customer care*), atención de reclamos, toma de pedidos, telemarketing, televentas, investigación de mercado, cobranza, CRM, fidelización, mesas de ayuda, control de calidad, y gestión y administración de bases de datos. Estos pueden ser ofrecidos tanto en modalidad *inhouse*, como a terceros por medio de *outsourcing* y, según cifras de la ACEC, este último segmento de la industria representa el 35% de las posiciones existentes en Chile, mientras que el restante 65% corresponde a compañías que desarrollan estos servicios internamente. En esta misma línea el 66%, corresponde a prestaciones *inbound* donde el principal servicio entregado es la atención de clientes, mientras que los servicios *outbound*, representan un tercio de la oferta, donde los ejecutivos efectúan principalmente llamadas orientadas al telemarketing y ventas (Uribe-Echevarría & Morales, 2010).

Respecto a las cifras de la industria, existen escasos estudios públicos que caractericen los centros de contacto, existiendo datos únicamente del mercado de tercerización por lo que la caracterización de la industria se realiza usando solo este segmento.

7.2.5.1 ANÁLISIS LA INDUSTRIA

Según un estudio desarrollado por Frost & Sullivan para la ACEC (2015), el 56,3% de ingresos por servicios tercerizados de *Contact Center* se atribuyen a los servicios prestados a empresas de Telecomunicaciones, seguidos por empresas de Bancos y Finanzas (13,7%).

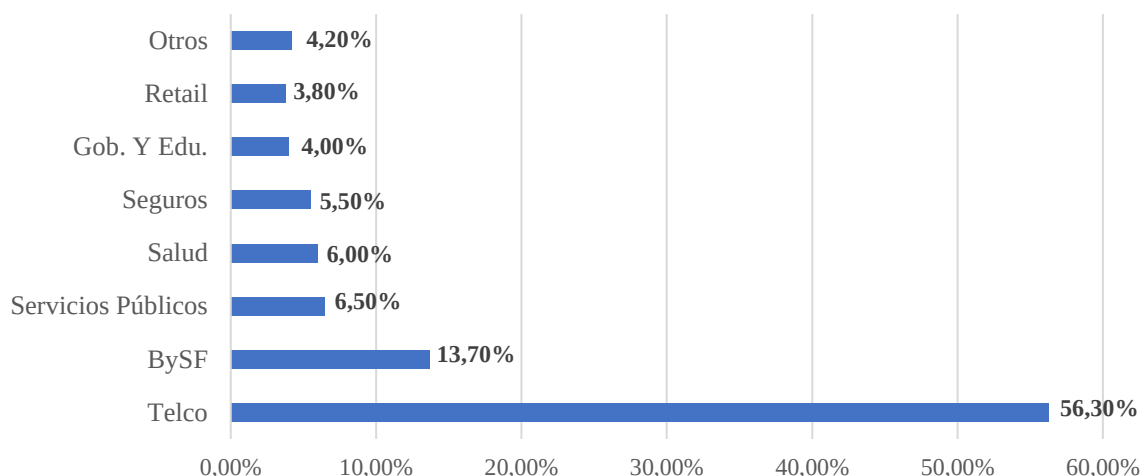


Figura 2: Distribución de ingresos por servicios tercerizados de *Contact Center* según sector.
Fuente: Elaboración propia a partir de ACEC (2015)

Analizando en detalle el tipo de servicio ofrecido, el 51,1% corresponde a atención al cliente (servicios de postventa, gestión de incidencias o averías, reclamos, comunicaciones relacionadas a logística, etc.) seguido por servicios de venta (23,8%).

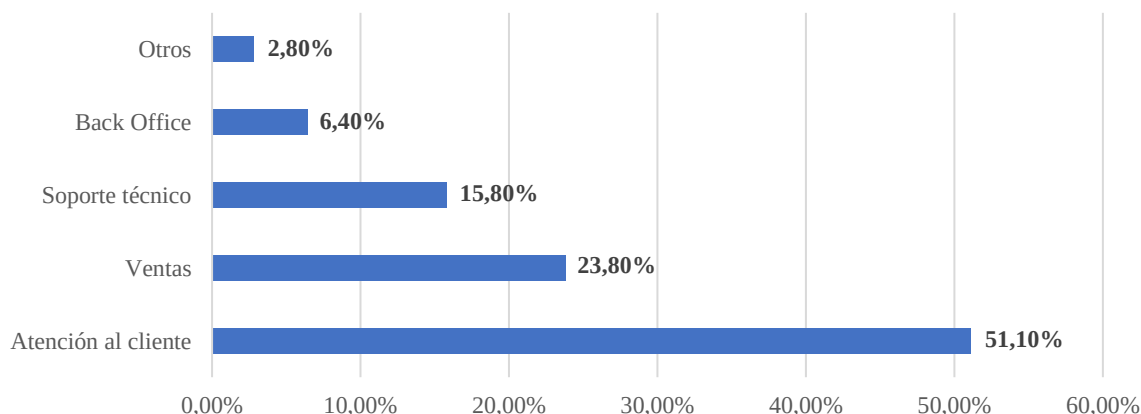


Figura 3: Distribución de ingresos por servicios tercerizados de *Contact Center* según tipo de servicio.
Fuente: Elaboración propia a partir de ACEC (2015)

Por último, en cuanto a los canales de comunicación utilizados por los clientes, el 84,7% de los ingresos por servicios prestados corresponden al canal de voz, seguido por el presencial (6,4%).

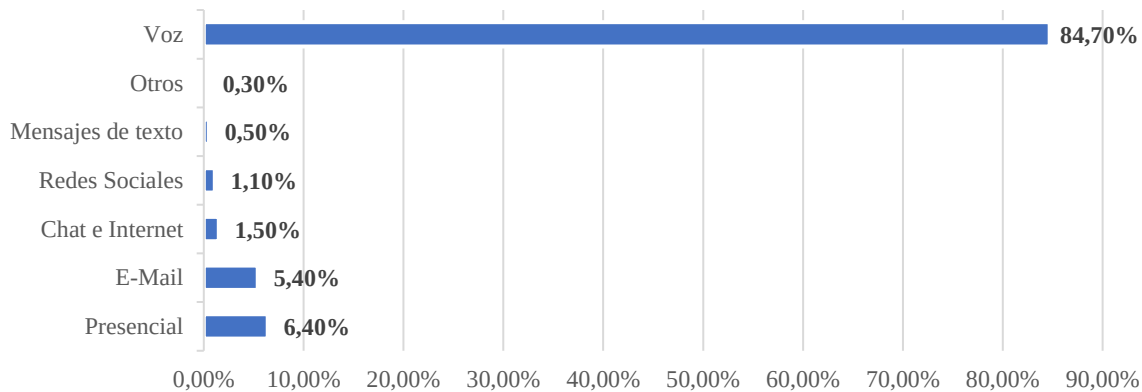


Figura 4: Distribución de ingresos por servicios tercerizados de Contact Center según canal de comunicación.
Fuente: Elaboración propia a partir de ACEC (2015)

7.3 CLOUD COMPUTING

La computación en la nube, muy a menudo denominado simplemente cloud o “la nube”, consiste en el suministro de recursos informáticos on-demand que proporciona una infraestructura dinámicamente escalable para aplicaciones, datos y almacenamiento de archivos basado en un modelo de pago por uso (IBM Cloud, s.f.).

Desde la perspectiva de una organización, es un modelo de adquisición y entrega de servicios para recursos de TI que puede ayudar a mejorar el rendimiento comercial y controlar los costos de la entrega de recursos de TI a la organización, transformando soluciones intensivas en capital a un entorno de precio variable.

El servicio basado en la nube puede ser público, privado o una combinación de ambos. Los *clouds* públicos pertenecen y son administrados por empresas que ofrecen a través de

una red pública acceso rápido a recursos informáticos asequibles, con estos servicios, los usuarios no necesitan adquirir hardware, software ni infraestructura de soporte, ya que pertenece a los proveedores y lo gestionan ellos. Mientras que un *cloud* privado se trata de una infraestructura que utiliza únicamente una única organización, ya sea gestionada internamente o por terceros y alojada interna o externamente.

Los proveedores de cloud computing ofrecen servicios que pueden ser agrupados en tres categorías: *Software as a Service* (SaaS), que corresponde a aplicaciones de software que se ocupan mediante método de pago por uso, donde los usuarios se conectan a la aplicación que se ejecuta en algún servidor distante a través la red, evitando la inversión inicial en licencias de softwares; *Platform as a Service* (PaaS), que comprende el acceso a un entorno de desarrollo en el que el cliente puede construir sus propias aplicaciones las cuales serán ejecutadas en la infraestructura del proveedor del servicio; e *Infrastructure as a Service* (IaaS), que proporciona a las empresas recursos informáticos, incluyendo servidores, redes, almacenamiento y espacio en centro de datos con pago en función del uso (IBM, 2008).



Figura 5: Modelos de Cloud Computing. Fuente: (Microsoft Azure, s.f.)

7.3.1 IBM CLOUD

IBM Cloud es una plataforma abierta de computación en la nube que combina plataforma como servicio (PaaS) con infraestructura como servicio (IaaS) e incluye un catálogo de diversos servicios en la nube, que se pueden usar para construir e implementar aplicaciones o infraestructura empresarial.

Como PaaS, proporciona a los desarrolladores acceso para crear aplicaciones mediante el uso de un conjunto de *run times* como Node.js, Python y Java. También proporciona servicios para integración, datos, seguridad y otras aplicaciones.

Como IaaS, permite a los desarrolladores controlar la infraestructura en la que se implementan sus aplicaciones. Los desarrolladores pueden seleccionar servidores bare metal, servidores virtuales, contenedores y almacenamiento en la nube, en ubicaciones de centros de datos IBM Cloud en todo el mundo. (Azraq, Chugthai, Mashhour, Nguyen, & Marcelo Dos Santos, 2018)

IBM Cloud proporciona un mercado alojado en la nube donde los proveedores de aplicaciones de todos los tamaños e industrias pueden acceder a recursos para desarrollar aplicaciones pudiendo combinar los servicios de cognitivos con otros servicios de IBM Cloud y de otros proveedores.

7.4 COMPUTACIÓN COGNITIVA

Las soluciones de computación convencionales funcionan en base al cumplimiento de reglas y lógicas para entregar respuestas lógicamente correctas a lo largo de un diagrama

rígido de decisión. Este sistema ha permitido, entre otras cosas, el manejo y procesamiento de información de tipo estructurada, pero resulta incapaz de gestionar datos de tipo no estructurados e incluso estructurados en grandes volúmenes. La computación cognitiva, a diferencia del sistema tradicional, no se centra en hacer cálculos rápidos en grandes cantidades de datos, sino que se enfoca en la exploración de estos, y mediante una serie de algoritmos de *Machine y Deep Learning*, es capaz de reconocer patrones e identificar relaciones entre elementos (como figuras o palabras) para realizar predicciones, “aprendiendo” en función de la información histórica que se le ha sido entregada al motor cognitivo en la fase de entrenamiento.

El estudio y desarrollo de los sistemas de inteligencia artificial (IA) tiene como objetivo construir sistemas informáticos capaces de realizar tareas que normalmente requieren inteligencia humana. Las máquinas basadas en IA han sido creadas para percibir su entorno y tomar acciones que optimicen su nivel de éxito. La investigación en IA utiliza técnicas de muchos campos, como la informática, la filosofía, lingüística, economía, reconocimiento del habla y psicología, que se manifiestan en aplicaciones, como sistemas de control, procesamiento del lenguaje natural, reconocimiento facial, reconocimiento de voz, análisis, coincidencia de patrones, extracción de datos y logística. La computación cognitiva se encuentra entre las subdisciplinas que dan forma a la IA, utilizando el potencial de la máquina para “simular” los procesos de pensamiento humano (Gliozzo, y otros, 2017). Así los sistemas cognitivos tienen las siguientes capacidades:

- Entender imágenes, lenguaje y otros datos no estructurados. El sistema cognitivo operacionaliza datos estructurados y no estructurados imitando el raciocinio humano.

- Razonar, captar conceptos subyacentes, formar hipótesis e inferir y extraer ideas.
- Aprender y adaptarse en función de la experiencia, interacciones y datos entregados.
- Interactuar con humanos en lenguaje natural gracias a herramientas que les permiten ver, hablar y escuchar.

La computación cognitiva genera un modelo colaborativo entre las personas y los sistemas de computación, al aprender de cada interacción y aumentar las capacidades analíticas, manejando grandes cantidades de información que para los sistemas comunes resulta inviable de procesar. Sus capacidades aplicadas a la industria se presentan como una alternativa para mejorar procesos y operaciones, tomar decisiones basadas en análisis de gran cantidad de datos y desarrollar estrategias de *Customer Engagement* (Gliozzo, y otros, 2017).

7.4.1 APLICACIONES DEL SERVICIO AL CLIENTE

Una de las principales herramientas para el *Customer Engagement* corresponde al análisis de los datos recopilados de los clientes. Los sistemas de inteligencia artificial y computación cognitiva son capaces de procesar grandes bases de datos no estructuradas generando relaciones e insights en tiempo real que resultan valiosos a la hora de desarrollar una estrategia de compromiso con el cliente. Según Gartner (2018), la inteligencia artificial (IA) impulsa el valor en función de los resultados del negocio en torno a la experiencia del cliente, la reducción de costos y la generación de ingresos, luego las aplicaciones de IA más que proyectos técnicos y tácticos, son una capacidad tecnológica que puede permitir el desarrollo de nuevas oportunidades y ayudar a alcanzar los objetivos comerciales.

Según Schwartz, Mooney & Heller (2018), las empresas usan IA para mejorar la experiencia del cliente en tres áreas principales:

- **Información:** La IA puede revelar ideas que ayuden a los empleados a tomar decisiones informadas rápidamente permitiendo desde refinar su estrategia general de *Customer experience*, hasta proporcionar respuestas en tiempo real a las consultas de los clientes.
- **Automatización:** La IA puede automatizar procesos detrás de escena, como promociones y mensajes de marketing automáticos dirigidos.
- **Interacción con el cliente:** Las empresas pueden integrar IA en dispositivos con los que los clientes pueden interactuar directamente, como chatbots y robots.

Los grandes proveedores que ofrecen servicios cognitivos empresariales incluyen una amplia gama de soluciones, la mayoría a través de internet, aprovechando los beneficios de las plataformas en la nube, como la escalabilidad, el acceso y las opciones flexibles de facturación. Estos proveedores se centran en capacidades cognitivas básicas aplicadas a soluciones industriales basadas en el procesamiento de lenguaje natural (NLP), el *Machine Learning* (ML) y las tecnologías de preguntas y respuestas (QA) (Gliozzo, y otros, 2017).

Los servicios cognitivos básicos disponibles en el mercado incluyen: servicios de conversación, *text mining*, análisis de texto, reconocimiento de imágenes, reconocimiento de voz, entre otros. Algunos de los más grandes proveedores de computación cognitiva corresponden a:

- Amazon a través de Amazon Web Services (AWS)
- Microsoft a través de MS Azure Cloud
- Google a través de Google Cloud Platform.
- IBM a través de IBM Cloud

7.4.2 IBM WATSON

IBM Watson es la oferta de computación cognitiva de IBM, soportada en IBM Cloud, que se encuentra disponible como un conjunto de interfaces de programación de aplicaciones abiertas (API²) y *Software as a Service* (SaaS) para el desarrollo de soluciones industriales.

Los servicios de Watson disponibles actualmente pueden agruparse según su funcionalidad en: *Language*, relacionadas con la comprensión de textos y herramientas conversacionales; *Speech*, asociadas a las herramientas que dan voz y oído al bot durante una interacción en las soluciones conversacionales en formato de voz; *Vision*; y *Data Insights*, que está asociado a la analítica de textos. Estos servicios deben ser entrenados por los usuarios, definiendo una base de conocimiento para su uso en particular, luego no basta con solo utilizar el servicio, sino que debe desarrollarse una estrategia de gestión del conocimiento para darle la inteligencia a cada instancia.

Juan Avilés (2016), experto en sistemas cognitivos de IBM España plantea que, “como una de las fases iniciales, la documentación del área de conocimiento en la que se va a entrenar el sistema cognitivo se “ingesta” en Watson. Dicha información contiene las respuestas y las evidencias en las que el sistema se va a basar a la hora de dar una respuesta. La selección de qué información es relevante y fiable la realiza un experto del área en concreto. Posteriormente, Watson debe ser entrenado por humanos para aprender y entender contenido específico de diferentes sectores, especialidades e idiomas. Esto implica capacitar al sistema para reconocer patrones complejos y abstractos, alimentándolo con grandes

² Conjunto de comandos y funciones que permiten a los desarrolladores crear programas específicos para ciertos sistemas operativos y/o usar funciones predefinidas para interactuar con otros programas.

cantidades de datos etiquetados y luego trabajando con expertos humanos para refinar las respuestas. A través de sucesivas rondas de información y retroalimentación, la comprensión y las respuestas de Watson mejoran”.

7.5 CONTACT CENTER COGNITIVO

Según Forrester (2018), un *Contact Center Cognitivo* (CCC), es la nueva generación de *Contact Center* que combina datos, aprendizaje automático, interfaces de conversación, omnicanalidad y otros procesos que aceleran la evolución del autoservicio y la optimización de la fuerza de trabajo. Además, al incorporar IA, personaliza la comunicación entre agente y cliente, mejora sus habilidades de interacción, y aliviana las tareas de administración.

En otras palabras, un CCC corresponde a una solución conversacional de IA que permite crear agentes que llevan a cabo conversaciones que comunican conocimiento, responden preguntas o guían a través de un proceso. Según el reporte “Conversational A.I for the Enterprise” (IBM, 2017), una vez implementada la solución, estas conversaciones se transforman en una fuente de datos para comprender qué buscan los clientes y cómo se relacionan con una organización. En última instancia, proporciona nuevos conocimientos y formas de optimizar la experiencia del cliente.

7.5.1 CONTACT CENTER COGNITIVO DE IBM

El servicio de CCC ofrecido por IBM en colaboración con Sixbell, su aliado estratégico en telefonía, hace coexistir la tecnología tradicional de *Contact Center* con Inteligencia Artificial, operando con asistentes virtuales 24x7 sostenidos por los servicios de Watson, que permiten en el contexto de una conversación, entender lo que el usuario está

diciendo y aprender sobre sus preferencias y comportamiento, adaptándose al contexto de la conversación.

Según el Solution Brief “El Contact Center Cognitivo: Transformando los sistemas de relación con los clientes” presentado por IBM, el asistente cognitivo proporciona la capacidad de interactuar de forma conversacional, contextual y documentada con el usuario de modo que:

- Entiende la intención del usuario expresada de forma natural.
- Comprende la necesidad o intención comunicativa del usuario y responde en función de ésta.
- Gestiona un diálogo complejo, manteniendo el contexto de la conversación a la vez que captura conceptos y e información específica entregada por el usuario para su reutilización en el flujo de la interacción.
- Proporciona una respuesta adaptada a las especificidades del canal de interacción (por ejemplo, utilizando tags, HTML o capacidades de texto a audio).
- Tiene una “personalidad” acorde al caso de uso y a aspectos de imagen corporativa de la empresa en la que se implementa.
- Es capaz de dirigir activamente una conversación hacia ámbitos específicos, así como también puede mostrar una actitud más reactiva.

En específico dentro de los servicios que ofrece, el *Contact Center Cognitivo*, permite levantar un asistente virtual en redes sociales, páginas web, aplicaciones de escritorio y vía telefónica. Esto acompañado de una herramienta de reportería en forma de un *dashboard* desplegado en una página web que entrega información en tiempo real de las interacciones, indicando KPI's definidos por la compañía, transcripción y audio de llamadas, número y

clasificación de llamadas (atendida, abandonada, transferida), estadísticas respecto a las intenciones comunicativas reconocidas por el asistente virtual, latencia de los servicios, entre otras cosas.

7.5.1.1 DESCRIPCIÓN DEL CASO DE USO Y FUNCIONAMIENTO

En la **Figura 6** se detalla un diagrama general del funcionamiento del *Contact Center Cognitivo*, donde el usuario interactúa con la aplicación a través del canal e interfaz implementada, por ejemplo, una ventana de chat en una página web, teléfono, servicio de mensajería instantánea, etc.

La aplicación envía el requerimiento o mensaje del usuario al asistente virtual creado en Watson Assistant, el servicio interpreta el requerimiento y avanza a través de la estructura de diálogo diseñada recopilando la información que necesita para entregar una respuesta. De manera paralela y si es requerido, se conectan otros servicios de Watson para analizar el *input* del usuario como Speech to text (utilizado para transcribir un mensaje de audio a texto) o Text to speech (utilizado para transformar mensajes de texto a audios sintetizados).

Por último, la aplicación puede interactuar con los sistemas del back-end -que corresponden a los sistemas que gestionan bases de datos- y back-office - que corresponde a la división de una empresa que atiende tareas de carácter administrativo como el servicio de atención al cliente-, en caso de que el usuario requiera información adicional, actualizar información u obtener respuesta a preguntas más complejas.

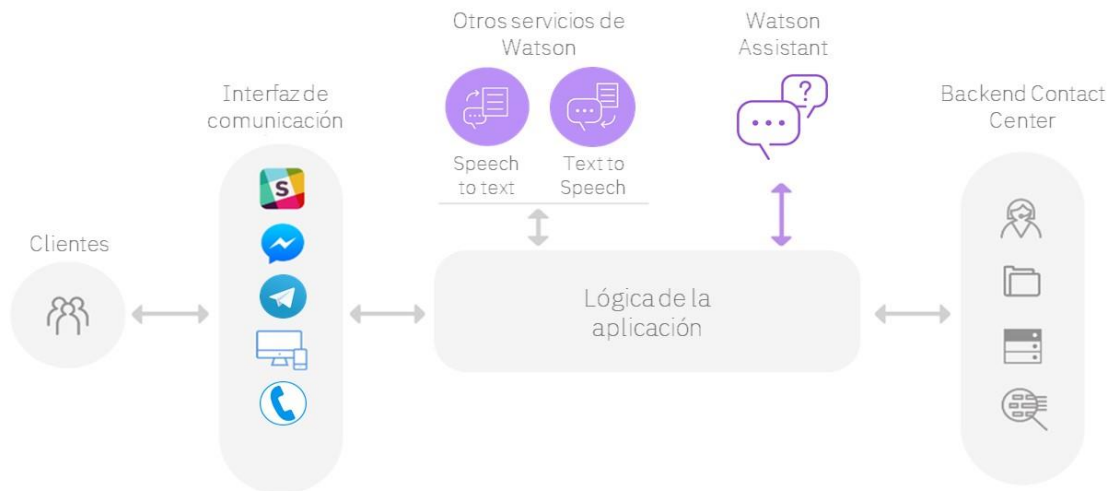


Figura 6: Diagrama general de funcionamiento del Contact Center Cognitivo de IBM.
 Fuente: Elaboración propia a partir de (IBM Cloud, s.f.)

7.5.1.2 VENTAJAS DE SU IMPLEMENTACIÓN

Un sistema de información que dote a un canal digital de la capacidad de conversar, es decir, de entender y expresarse en el lenguaje natural de las personas, comporta tal cantidad de ventajas que supone una palanca fundamental en la transformación digital de las empresas, y muy especialmente, en el ámbito de la estrategia de la relación entre una organización y sus usuarios y clientes (Díaz, Fernández, Gallego, González, & Millares, 2016). En este contexto, las ventajas que entrega el *Contact Center Cognitivo* de IBM son:

7.5.1.2.1 Mejora y eficiencia en la atención al cliente

La disponibilidad de agentes virtuales 24x7, que pueden manejar una gran cantidad de interacciones de manera simultánea y en todos los canales de comunicación, permite aumentar el nivel de servicio de atención al cliente a la vez que libera a los ejecutivos de atender consultas rutinarias, por lo que disponibiliza capital humano para enfocarse en cubrir los requerimientos más especializados y prestar una mejor atención al cliente.

En términos operativos, al conectarse de manera automatizada con los sistemas de la compañía, logra dar respuestas a requerimientos de información en un tiempo menor al que lo haría un ejecutivo que consulte manualmente los sistemas, lo que hace que los tiempos medios de atención sean inferiores para el caso de los asistentes virtuales.

También su implementación permite desplegarse en múltiples canales a la vez que evita la duplicidad en el trabajo, ya que el mismo asistente virtual soportado en Watson Assistant puede ocuparse para los diferentes canales, entregando además una experiencia consistente en la atención.

7.5.1.2.2 Adaptabilidad

Es capaz de integrarse a los canales web de comunicación más utilizados, de la misma forma gracias a la tecnología ofrecida por Sixbell, puede integrarse de manera certificada a la infraestructura telefónica de los *Contact Centers* ya operativos. Además, corresponde a tecnología que puede aplicarse tanto a servicios *inbound* como *outbound* en todos los canales de comunicación.

Puede soportarse totalmente en modalidad *cloud*, si es que la organización tiene una estrategia de negocios en esa línea, o pueden mantenerse componentes directamente en el data center para la gestión de información sensible en las instalaciones de quien contrata el servicio. Sin embargo, el mantener las componentes esenciales en la nube, hace que sea una solución escalable con poco esfuerzo tanto monetario como de gestión en comparación a los *Contact Center* tradicionales. En este contexto, la solución da un alto nivel de adaptabilidad a los periodos críticos en que la demanda por servicio es notablemente más alta que el promedio anual. En un *Contact Center* tradicional para poder cubrir esos pick de demanda se debe contratar personal extra, mientras que, con el uso de asistentes virtuales, se puede

planificar la capacidad para los periodos con mayor demanda sin incurrir en grandes costos de capacidad ociosa para los demás periodos del año.

En cuanto a los servicios ofrecidos de cara a cliente, los asistentes virtuales pueden adaptarse de manera rápida y sencilla a cambios en políticas de tratamiento al cliente guiadas por el personal de negocios.

Por otra parte, su arquitectura modular permite la personalización de los servicios ofrecidos de manera que si el cliente requiere alguna otra funcionalidad que esté disponible en IBM Cloud puede incorporarse a la solución.

7.5.1.2.3 Reportería y manejo de datos

Uno de los elementos más importantes de la relación con el cliente, a través de todos los puntos de contacto, es la posibilidad de realizar una acción en el momento adecuado de forma personalizada. La visión de la operación del *Contact Center* asistido por Watson, añade a la información tradicional de los KPIs de atención la clasificación automática de las interacciones, identificando los requerimientos tratados y respondidos. Watson puede realizar el análisis de las conversaciones que se realizan en todos los canales, aumenta el detalle del conocimiento extraído de cada una de ellas, identificando las diferentes intenciones de las interacciones, conceptos, entidades, actividades y la resolución, ofreciendo una perspectiva de operación idónea para la estrategia de gestión de los modelos de atención (Díaz, Fernández, Gallego, González, & Millares, 2016).

En este contexto, la solución da la posibilidad de generar analítica de las interacciones entre un cliente y el asistente virtual, a partir de las cuales se pueden tener *insights* de la experiencia y comportamiento del cliente en el proceso de comunicación, información que

para la gran mayoría de los *Contact Centers* tradicionales es invisible debido a que no se cuentan con las herramientas ni recursos para gestionarla. La información también puede ser analizada a posterioridad con otras herramientas debido a que cada interacción es almacenada en una base de datos, en formato texto y/o audio con su respectiva transcripción.

La solución también puede incorporar un *dashboard* que permite presentar información en tiempo real de las métricas que se configuren, entregando una herramienta potente de gestión al servicio al cliente y a la auditoría de estas interacciones.

7.5.1.3 LIMITACIONES

Si bien la implementación del *Contact Center Cognitivo* de IBM trae consigo numerosas ventajas, también existen limitaciones para su uso. Algunas relacionadas con las estructuras tecnológicas existentes y otras propias del negocio al cual se pretende implementar.

- No es recomendable mantener un *Contact Center Cognitivo* únicamente con asistentes virtuales. La solución está pensada para trabajar en conjunto con asistentes humanos que atiendan los requerimientos que los virtuales no puedan resolver, ya sea por la complejidad de la consulta o por falta de entrenamiento. De esta forma es necesario contar con ejecutivos de respaldo, para dar respuesta a las consultas más complejas, y con un equipo de desarrollo que controle el desempeño de la solución y realice ajustes o entrenamientos en caso de ser requerido.
- Actualmente, no es recomendable su instalación en servicios de atención al cliente altamente especializados como servicios de asistencia técnica, ni para servicios de venta telefónica que implique negociación, esto porque las habilidades de negociación de un ejecutivo están muy por encima de las que un asistente virtual ha logrado desarrollar. Sin

embargo, no se descarta la posibilidad que, en el tiempo, los avances en el área permitan alcanzar un nivel similar al humano de manera que sea factible implementar asistentes en estos casos de uso.

- Requiere de la consultoría constante de expertos para ir ajustando la solución a las necesidades del negocio.
- Requiere de la constante revisión de las interacciones del asistente virtual de manera de tener control del nivel de entendimiento, reentrenando los casos mal interpretados y añadiendo ejemplos para los que no han sido entendidos.
- Para el canal de voz será necesario realizar un entrenamiento especializado con audios reales de clientes de manera de aumentar la precisión de la transcripción. Esto considerando la heterogeneidad del habla en el idioma español, más en específico en Chile.

De manera adicional a estas limitaciones, debe tenerse en cuenta también que, el funcionamiento del *Contact Center Cognitivo* está condicionado a la existencia de conexión a la red. Esto debido a que la gran mayoría de sus componentes se encuentran operando en un ambiente *cloud*, por lo tanto, es necesario garantizar un enlace de alta disponibilidad con un ancho de banda adecuado para su funcionamiento. Además, para el canal telefónico, la cantidad de llamadas concurrentes que pueden ser procesadas se encuentra limitada principalmente por la tecnología del *Contact Center* ya implementado en la organización y el ancho de banda contratado para usos de servicios de voz IP.

8 METODOLOGÍA

Para llevar a cabo la investigación sobre la viabilidad de la implementación del *Contact Center Cognitivo* de IBM, se realiza un estudio de factibilidad técnica y económica bajo el esquema de Proyecto Factible, que tiene como objetivo central, el determinar la factibilidad técnica, económica y operativa de un producto o servicio a través de herramientas contables y económicas (Blanco, 2008).

8.1 ANÁLISIS DE FACTIBILIDAD

El concepto de factibilidad se refiere a la disponibilidad de los recursos necesarios para llevar a cabo los objetivos y metas señaladas. Este estudio sirve para recopilar datos relevantes sobre el desarrollo de un proyecto y en base a ello tomar la mejor decisión, para dar continuidad al desarrollo o implementación (Santos, 2008).

El análisis de factibilidad a desarrollar contempla dos aspectos:

- Factibilidad Técnica
- Factibilidad Económica

En este punto es importante mencionar que, dentro de un análisis de factibilidad, un sistema debe ser factible desde ambos puntos de vista para ser implementado, así si es factible técnica, pero no económicamente para la organización o inversionistas, este no puede ser instaurado (Santos, 2008).

8.1.1 FACTIBILIDAD TÉCNICA

El objetivo del estudio técnico consiste en analizar y verificar la viabilidad técnica de las diferentes alternativas para la implementación del proyecto. A partir del mismo, se

determinan los costos de inversión requeridos, y los costos de operación que intervienen en el flujo de caja que se realiza en el estudio económico – financiero (Santos, 2008).

Para este caso, el estudio se centra en un análisis tecnológico de todos los factores y componentes que son necesarios para la implementación de la solución, así en términos generales se abarcan los siguientes puntos:

- Tamaño del proyecto
- Arquitectura y funcionamiento de la solución
- Descripción y dimensionamiento de la tecnología utilizada
- Restricciones técnicas de operación

8.1.2 FACTIBILIDAD ECONÓMICA

La evaluación económica, pretende determinar cuál es el monto de los recursos económicos necesarios para la realización del proyecto, cuál será el costo total de la operación, y otra serie de criterios que comparan flujos de beneficios y costos; esta evaluación permite determinar si conviene llevar a efecto un proyecto, o sea, si es o no rentable (Sánchez, 2003). Por lo que, en este capítulo, se detallaran los costos y beneficios que implica la implementación del *Contact Center Cognitivo*.

En específico los proyectos orientados a tecnología contribuyen con el desarrollo de los procesos del negocio, por lo que el tipo de beneficios que se pueden obtener son ahorro en costos, aumento en la productividad, disponer de la información para una correcta toma de decisiones, mejor atención al cliente, entre otros. Estos deben ser cuantificados para identificar si los costos del proyecto son o no justificados (Sucely, 2008).

Junto con los beneficios, deben detallarse todos los costos asociados al funcionamiento de la solución indicando el periodo en el horizonte de evaluación en que se incurren. Para esto se utiliza la metodología TCO (Costo Total de Propiedad o *Total Cost of Ownership*), que corresponde a un modelo que ayuda a las empresas a analizar todos los costos relacionados con la adquisición, desarrollo y uso de componentes de Tecnologías de Información (IT) a lo largo de su ciclo de vida, estos costos se agrupan en:

- Hardware
- Software
- Administración de sistemas
- Soporte
- Desarrollo
- Comunicación

Posterior a la determinación de costos y beneficios, se realiza un flujo de fondos de manera de obtener indicadores que permitan concluir respecto a la rentabilidad del proyecto, como el Valor Actual Neto (VAN), la Tasa Interna de Retorno (TIR), Periodo de Recuperación de la Inversión (PRI) y Retorno sobre la Inversión (ROI).

9 DESCRIPCIÓN DE LA SITUACIÓN SIN PROYECTO

Como se ha mencionado, en este trabajo se evalúa la implementación del *Contact Center Cognitivo* de IBM en una empresa de retail en Chile que, según datos presentados en su memoria anual, corresponde a una de las empresas líderes en la industria de retail con una participación de mercado del 23,2%. Su principal negocio es la venta al detalle de vestuario, accesorios y productos para el hogar a través de los distintos formatos de tienda por departamento, acompañado por un fuerte negocio financiero a través del crédito directo a sus clientes por medio de la tarjeta de la tienda. Asimismo, maneja el banco asociado y participa en la propiedad de centros comerciales. A finales del año 2017 la tienda cuenta con 46 sucursales a lo largo del país, 18 de ellas se encuentran en la región Metropolitana.

En la **Figura 7** se muestra su nivel de ventas trimestral en dólares desde el año 2014, el cual presenta una marcada estacionalidad con un alza en las ventas en el segundo y cuarto trimestre, esto se puede asociar principalmente a la campaña de ventas conocida como *cyberday* tradicionalmente realizada en el segundo trimestre, y a las ventas de navidad en el último trimestre.

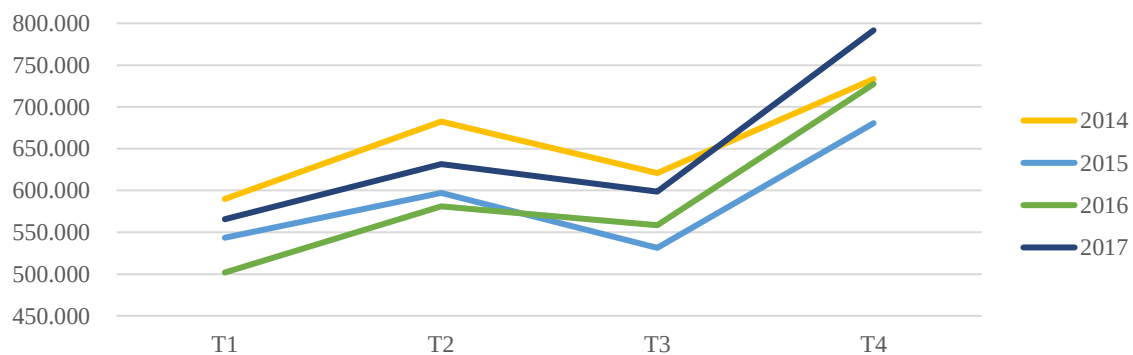


Figura 7. Nivel de ventas en USD por trimestre, desde el año 2014 al 2017.
Fuente: Elaboración propia.

La empresa maneja un servicio de *Contact Center* interno que opera con 120 ejecutivos orientados a responder consultas del área financiera, retail y fono compras por medio de los canales web (redes sociales y página web) y, con especial énfasis, en el canal telefónico. Para esta implementación no se toma en cuenta el *Contact Center* del área financiera, por decisiones propias de la empresa, luego para las consultas restantes, el servicio al cliente opera con 60 ejecutivos y procesa el 51% del volumen de llamadas recibidas por todo el *Contact Center*. El costo empresa por mes por ejecutivo corresponde a USD 1.500, y el costo por llamada atendida en el *Contact Center* corresponde a \$170.

La jornada laboral de los ejecutivos corresponde a turnos de 8 horas, de manera que el servicio de atención al cliente esté disponible de lunes a viernes desde las 9:00 a las 21:00 horas y sábados de 10:00 a 22:00 horas. La productividad promedio de los ejecutivos corresponde a un 72%, con un ausentismo del 20%, esto significa que los ejecutivos se presentan a trabajar de manera normal el 80% de los días laborales y dedican el 72% de su jornada exclusivamente a contestar llamadas de clientes. El tiempo medio de operación en la atención de una llamada es de 4,5 minutos.

El servicio de atención al cliente consta de un IVR que presenta 4 opciones: horarios y ubicación de tiendas; despachos; cambios y garantías; y otras consultas. Este IVR no se utiliza como herramienta de contención, lo que implica que no entrega mensajes previamente grabados con información y, por lo tanto, el 100% de las llamadas atendidas pasan a ejecutivo.

En la **Figura 8**, se muestra un diagrama de arquitectura del actual *Contact Center* de la empresa, donde el proceso de atención comienza con la llamada de un cliente, que es recibida por la PBX, la que luego se comunica con el IVR para hacer la solicitud del ingreso

del RUT del cliente. Una vez que el cliente ingresa su RUT, se relata el menú de opciones disponibles, cuando el cliente digita una de ellas, la PBX se conecta con el sistema de administración de *Call Center*, el que asigna la llamada al grupo de agentes adecuado para su atención. Posterior a esto, uno de los agentes suscritos al grupo asignado contesta la llamada atendiendo al cliente. Durante este proceso el agente registra en un sistema CRM, el motivo de la llamada, la descripción y el resultado del requerimiento, indicando si ha sido resuelto o si se requiere hacer un seguimiento. De manera adicional, el agente puede consultar los sistemas de la empresa para obtener información que sea necesaria para dar respuesta al cliente. Una vez atendidas todas las consultas del cliente, se finaliza la llamada.

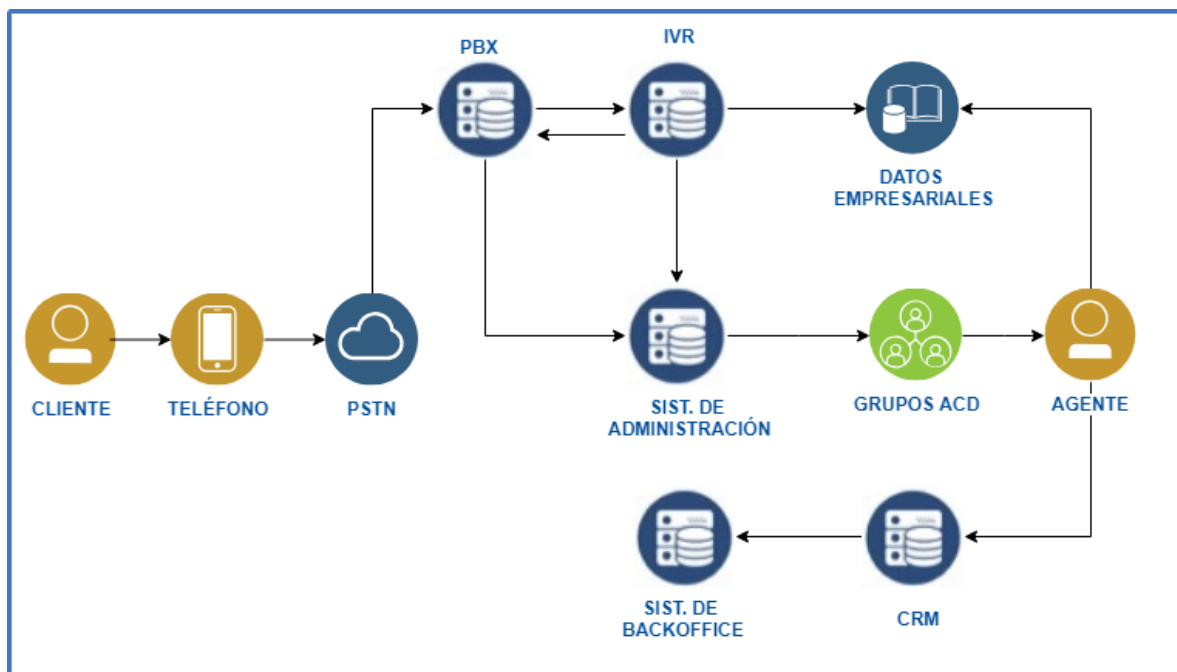


Figura 8. Diagrama de arquitectura del Contact Center de la empresa.
 Fuente: Elaboración propia.

En la **Tabla 1** se presenta de manera general los tipos de consultas que ingresan por cada opción, estas consultas son clasificadas según dificultad para la posterior definición del alcance de la solución. La dificultad baja corresponde a consultas que se categorizan como de pregunta y respuesta, esto es, que tiene una respuesta predeterminada siempre y que no requiere consultas a sistemas ni bases de datos. La dificultad media comprende consultas que, en general, necesitan de preguntas complementarias dentro del proceso y que, además requieren algún tipo de integración con bases de datos o sistemas web únicamente para lectura de información. Por último, la dificultad alta corresponde a requerimientos que para dar respuesta se debe modificar información en alguno de los sistemas de la empresa o bien es necesario un trato con un humano. Esta clasificación puede verse en la **Tabla 2**. De esta manera, se entiende por preguntas automatizables a aquellas de dificultad baja y media.

Tabla 1.
Tipos de consultas por opción de IVR.

Opción de IVR	Tipo de consulta
Horarios y ubicación	Ubicación (94%); horarios de tienda (6%)
Despachos	Seguimiento de despacho (39%); cambio de dirección de entrega (7%); cambio de fecha de entrega (11%); definición de hora de entrega (11%) información de retiro en tienda (7%); precios de despacho (3%); disponibilidad de despacho (3%); modificación del tipo de despacho (14%)
Cambios y garantías	Política de devolución (4%); política de cambio (19%); cobertura de garantía (11%); garantía extendida (3%); producto fallado (63%)
Otras consultas	Financiero (62%); confirmación de compra (3%); reclamos (18%); recuperar boleta (1%); anular compra (3%); problema compra online (6%), información compra online (7%)

Tabla 2.
Clasificación de consultas según dificultad.

<i>Dificultad</i>	<i>Tipo de consulta</i>
Alta	Cambio de dirección de entrega; cambio de fecha de entrega; modificación del tipo de despacho; producto fallado; recuperar boleta, anular compra, problema de compra online
Media	Seguimiento de despacho; información de retiro en tienda; confirmación de compra; reclamos
Baja	Definición de hora de entrega; precios de despacho; disponibilidad de despacho; política de devolución; política de cambio; garantía extendida; cobertura de garantía; información compra online

En la **Tabla 3**, se indica el total de llamadas que ingresan al servicio de atención al cliente en el año móvil comprendido entre mayo 2017 y mayo 2018, con sus respectivos porcentajes de ingreso y tasa de abandono por opción discada. De aquí se obtiene que el número de llamadas recibidas corresponde a 879.230, y las opciones consultadas con mayor frecuencia corresponden a otras consultas (60,7%) y seguimientos de despacho (25,0%). La opción otras consultas corresponde en su mayoría (62%) a consultas financieras que deben ser transferidas a dicha área, tales como problemas con la tarjeta de crédito, ciclos de facturación, solicitud de apertura de cuenta, entre otras. El 48% restante de las llamadas corresponden a requerimientos de tipo post venta, como reclamos, problemas y anulación de compra online, y otros. La empresa espera que el volumen de llamadas crezca un 10% anual, esto basado en datos históricos del *Contact Center*.

Tabla 3.
Porcentaje de llamadas recibidas y tasa de abandono por opción de IVR en el año móvil entre mayo 2017 y mayo 2018.

<i>Opción IVR</i>	<i>Cantidad de llamadas</i>	<i>Ingreso por opción (%)</i>	<i>Abandono por opción (%)</i>
Horarios y ubicación	13.699	1,56	15,6
Despachos	221.331	25,17	9,3
Cambios y Garantía	111.698	12,70	10,0
Otras consultas	532.502	60,56	10,8
Total	879.230	100	11,4

Como política de la empresa, se deben tener los registros en audio del 100% de las llamadas y se auditan el 1,2% de ellas para llevar un control del comportamiento tanto del cliente como del ejecutivo. La auditoría se lleva a cabo de manera manual, esto es, un ejecutivo designado escucha todos los audios a revisar, evalúa la interacción y en caso de ser requerido, transcribe la conversación.

Se reconoce que los momentos críticos en términos de demanda por servicio corresponden al periodo *cyber* y post *cyber*, típicamente entre mayo y junio, en el que se registran alrededor del doble de los requerimientos que se reciben en periodo normal. Para el periodo en que se realiza la actividad, la mayoría de los contactos corresponden a consultas de cómo comprar y de confirmación de la compra realizada a través de la página web, mientras que para periodo post *cyber*, los motivos principales son el seguimiento de despacho y problemas de entrega de sus productos, tales como: cambios de fecha o dirección de entrega, productos equivocados, despachos retrasados, entre otros.

En estos periodos, se ve superada la capacidad instalada del *Contact Center* lo que afecta fuertemente al nivel de servicio ofrecido al cliente que tiene que esperar por largos

periodos en línea y que en el 40% de los casos prefiere abandonar la llamada, esto es un alto porcentaje comparado con el 11,4% de abandono promedio anual.

En cuanto a las comunicaciones a través del canal web, la empresa implementó a finales del año 2017, la solución de IBM de asistente virtual cognitivo en modalidad chatbot. Este consiste en una ventana de chat mediante la que se pueden realizar consultas en lenguaje natural y obtener una respuesta de manera automática. En la **Tabla 4** se presentan los tipos de consultas a nivel general atendidas por el asistente virtual.

Tabla 4.
Tipo de consultas respondidas por el chatbot implementado.

<i>Categoría</i>	<i>Tipo de consulta</i>
Información general	Horario y ubicación de tiendas, atención <i>call center</i> , transferencia a ejecutivo, información página web, información boleta, reclamos, programas de fidelidad, entre otras.
Tarjeta de crédito	Solicitud de apertura, estado de cuenta, olvido de clave, facturación y cobros indebidos, problemas asociados.
Servicio de venta	Cómo comprar, confirmación de compra, problemas de compra online, catálogo de productos, consulta de precio, consulta de stock, medios de pago.
Servicio de post venta	Despachos, retiro en tienda, garantías, entrega de producto equivocado, asistencia de armado de productos, anulación de compra, devolución y/o cambio de producto.
Conversación cotidiana	Capacidades del asistente, hora y fecha actual, chistes, conocimiento de personajes, preguntas personales (edad, gustos, estado civil), entre otras.

Actualmente el asistente está capacitado para reconocer alrededor de 130 intenciones comunicativas, correspondientes a las distintas variaciones y especificaciones de los tipos de consulta señalados. Las cinco consultas más realizadas a la plataforma corresponden a asistencia en la compra (36%), estado de la tarjeta de crédito de la tienda (25%), seguimiento de despacho (10%), información respecto a canales de comunicación (6%), y consulta sobre medios de pago (4%). En cuanto a la cantidad de interacciones, en el año móvil comprendido entre mayo del 2017 y mayo del 2018 se registra un total de 1.858.380, con una proyección³ anual de crecimiento del 15%.

³ Porcentaje estimado por la empresa en función del crecimiento histórico del *e-commerce*.

10 IMPLEMENTACIÓN DE LA SOLUCIÓN EN MODALIDAD MVP

Un producto mínimo viable o MVP (por sus siglas en inglés) corresponde a un prototipo de un nuevo producto o servicio que posee las características mínimas y suficientes para demostrar la funcionalidad y el valor que tendrá el producto final. Su desarrollo permite a un equipo recopilar la cantidad máxima de información validada sobre los clientes y generar un bucle de retroalimentación para guiar el desarrollo futuro, a la vez que proporciona herramientas para estimar los beneficios que podrían obtener, tanto los clientes como los desarrolladores, con la adquisición del producto final (Ries, 2009).

A través del programa IBM Do Lab, se implementó en la modalidad MVP, un prototipo de la solución para la empresa de retail descrita en la sección anterior, abarcando únicamente el canal telefónico. Esto porque no existe experiencia previa de la implementación en este medio, y porque es el canal más crítico en términos de volumen de contacto y en integración con los sistemas de la empresa.

Este prototipo, tiene el objetivo de evaluar de manera preliminar la factibilidad y efectividad de la expansión del chatbot, ya implementado en el sitio web de la empresa, al canal telefónico. De esta manera, se busca obtener valores de la tasa de contención, llamadas concurrentes, llamadas totales, confianza y entendimiento en la transcripción de audio de los clientes, entre otras cosas que sirvan como precedente para el dimensionamiento de la solución de *Contact Center Cognitivo* a implementar.

El desarrollo del MVP comprende el uso gratuito por parte de la empresa de un asistente virtual que contesta vía telefónica las opciones del IVR de horario, ubicación de tiendas y despachos, durante el periodo *cyber* y *post cyber* comprendido entre el 26 de mayo

y el 15 de junio del presente año. Esto fue especialmente escogido así para medir la solución en el momento de mayor concurrencia y demanda por servicio.

La habilitación de este prototipo de solución implica para el cliente que, al momento de llamar al centro de contacto de la empresa y discar una de las opciones incorporadas en el alcance, se dirige su consulta directamente al asistente virtual, quien puede entender en lenguaje natural su requerimiento y entregar una respuesta. Además, si el llamante prefiere hablar con un ejecutivo, o el asistente virtual no pudo entregar una respuesta, existe la opción de transferencia a un agente humano para atender el requerimiento. Para todas las demás opciones del IVR no implementadas en el MVP, el procedimiento de atención no varía al descrito en la situación sin proyecto.

En detalle para la opción de despacho, solo se habilita el seguimiento del pedido, dejando fuera del alcance del MVP todos aquellos temas relacionados a cambios de dirección, entregas de producto equivocado, productos dañados, entre otras cosas. Así cuando el cliente disca la opción de despachos, el asistente virtual informa que puede realizar el seguimiento, a lo que el cliente deberá responder ingresando el número de su orden de compra, para obtener la información del estado del despacho. En caso de que exista algún inconveniente con la entrega de esta información debido a un estado de despacho atrasado o cancelado, el cliente es transferido de manera inmediata a un ejecutivo.

En concordancia con lo mencionado anteriormente, si entrando a la opción de seguimiento de despacho el cliente manifiesta un problema respecto a la recepción de su pedido (producto equivocado, producto dañado, cambio de dirección de entrega, entre otras), se le ofrece la alternativa de transferencia con un ejecutivo, esto solo por estar fuera del

alcance del prototipo implementado, por lo que, en la operación de la solución completa, estos casos debiesen tener un tratamiento adecuado.

En cuanto a la opción de horarios y ubicación de tiendas, el cliente podrá consultar por la información de cualquier tienda presente en el país.

Es importante mencionar que, si bien el cliente ingresa por opciones delimitadas en cuanto al contexto de la consulta, éste puede preguntar sobre cualquier otro tema (que esté dentro del alcance) y obtener una respuesta independiente de la opción por la cual haya ingresado. Así si marca la opción de horarios y ubicación, puede igualmente consultar por el seguimiento de un despacho.

10.1 ANÁLISIS DE LOS RESULTADOS

Durante el periodo en que estuvo habilitado el prototipo, se registró un total de 36.133 llamadas, donde el 96,1% de las consultas fue por seguimiento de despacho y solo el 3,9% restante fue por horarios y ubicación de tiendas.

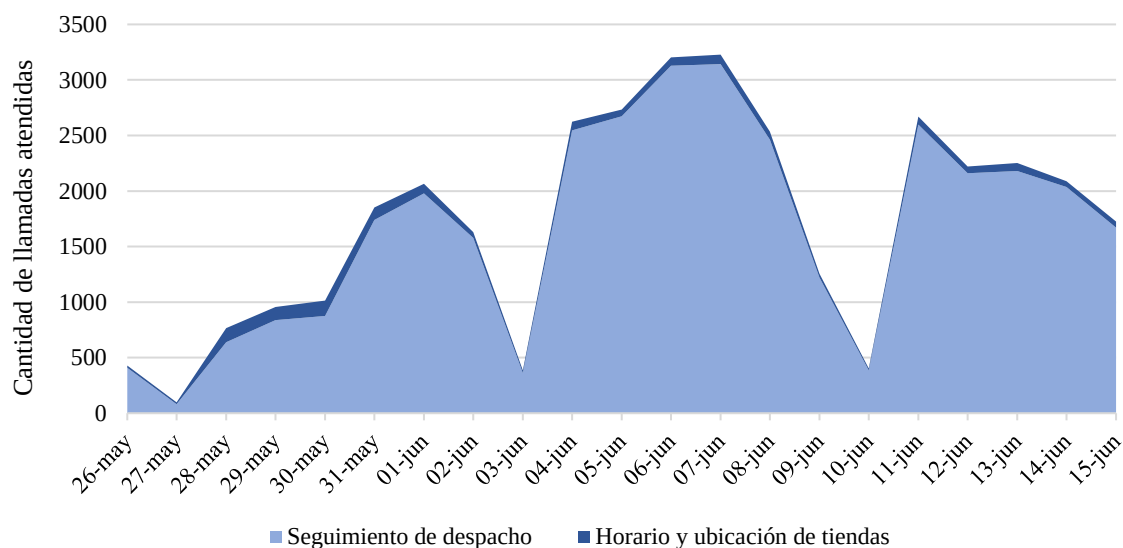


Figura 9: Cantidad de llamadas atendidas diariamente en el periodo de evaluación.
 Fuente: Elaboración propia

En la **Figura 9** se indica la distribución de las llamadas según el día, alcanzándose un pick de 3.142 llamadas el día 7 de junio, día en el que además se registró la más alta tasa de concurrencia correspondiente a 15 llamadas simultáneas. El día 3 de junio se presentó una anomalía en la cantidad de llamadas recibidas, esto por una falla en el funcionamiento de uno de los servidores. El día 10 de junio igualmente se presenta un decrecimiento de las llamadas, pero en dicha oportunidad fue producto de una baja programada del servicio a petición de la empresa cliente.

Del total de llamadas ingresadas por la opción de despacho, el 52% corresponde a requerimientos de seguimiento de despacho, y del 48% restante de las consultas ingresadas a la opción, el 39% corresponden a llamadas en las que el cliente decide terminarlas sin siquiera realizar una consulta, el 46% a consultas que el asistente virtual reconoció apropiadamente, pero que no tenía definido un tratamiento para dar respuesta, y solo el 15% corresponde consultas que el asistente no entendió, producto de ruido ambiental o falta de entrenamiento. Esto se indica en la **Figura 10**.

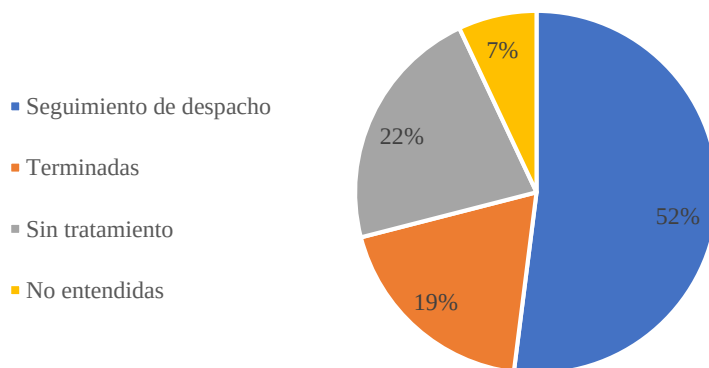
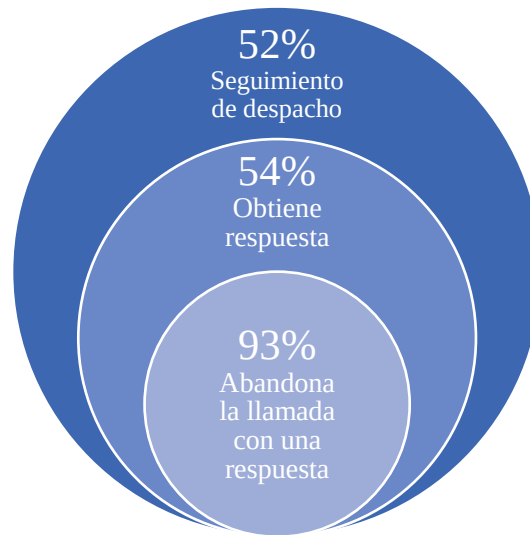


Figura 10: Distribución de llamadas ingresadas por opción de despacho.
 Fuente: Elaboración propia.

Como se muestra en la

Figura 11, del 52% de llamadas de seguimiento de despacho, solo el 54% obtuvo respuesta a su requerimiento, esto producto del desconocimiento por parte del cliente de su número de orden de compra y por despachos atrasados o cancelados, casos que según la empresa deben ser tratados por un ejecutivo en el desarrollo del MVP. Finalmente, del total de clientes que se les entregó el estado de su despacho, el 7% solicitó una transferencia a un



ejecutivo a pesar de tener una respuesta.

*Figura 11: Diagrama de contención por etapa del proceso de atención en el seguimiento de despachos.
Fuente: Elaboración propia*

De esta manera, en términos generales se obtiene una tasa de contención del 46% del total de llamadas atendidas. Esto es considerado como la contención simple debido a que, en términos de costos de operación para la empresa, una llamada contenida es aquella que no es atendida por un ejecutivo. Sin embargo, una medida de la contención a nivel más estricto indica un 26% de contención, considerando solo los casos en que se dio una respuesta a la consulta del cliente y este no manifestó la intención de ser redirigido a un ejecutivo.

Respecto de la consulta por horarios y ubicación de tiendas, el 80% de los casos fue respondido, dejando el restante 20% a llamadas en las que no se realizaban consultas, o bien si se realizaban el asistente no estaba entrenado para resolverla o fue incapaz de entender producto del ruido ambiental, por lo que terminaron con una transferencia a un ejecutivo o con el corte de la llamada por parte del cliente.

Respecto al tiempo medio de operación (TMO) se tiene un valor⁴ de 1,65 minutos para la opción de seguimiento de despacho, y un valor de 0,87 minutos para la opción de horarios y ubicación. Considerando el total de las llamadas atendidas, el TMO corresponde a 1,14 minutos, lo que representa el 25,3% del TMO de los ejecutivos en la situación sin proyecto.

10.2 DEFINICIÓN DE SUPUESTOS

En base a los resultados obtenidos del desarrollo del MVP, se realizan una serie de supuestos necesarios para el dimensionamiento de la solución en su implementación completa, por esto es importante detallar algunas observaciones del prototipo y cómo ajustaras para la realización final.

Respecto a la tasa de contención, esta depende de factores relacionados al alcance de la solución, el entrenamiento del asistente virtual y a la preferencia de los clientes que se comunican al centro de contacto de la empresa. De esta forma, las gestiones a realizar para un aumento más probable de la tasa de contención deben estar enfocadas en aumentar el alcance y el entrenamiento. Al ser relajada la limitante del prototipo de responder únicamente

⁴ Para el cálculo de este valor se han descontado los casos en los que el cliente no realiza un requerimiento, consulta otras cosas, o bien no ingresa la información requerida para consultar los sistemas. Esto para que sea comparable con los tiempos medios de operación de los ejecutivos.

casos de seguimiento de despachos, se permite accionar sobre el 47% de las llamadas que no clasificaron dentro de esa opción, específicamente dentro de ese grupo, se estima que se puede entregar respuesta a la totalidad de casos cuya intención comunicativa fue identificada pero no se tenía un protocolo definido para su tratamiento. Además, en términos de contención más estrictas, se puede entregar una respuesta al requerimiento del cliente aun en casos en que el estado del despacho se encuentre en un estado conflictivo (por ejemplo, atrasado), con un protocolo adecuado de reprogramación. Por esto, de manera conservadora puede estimarse una tasa de contención final cercana al 60%.

En cuanto a la tasa de contención de la opción de horarios y ubicación de tiendas, se reconoce que puede aumentarse el entrenamiento para la mejor comprensión de las consultas, pero dado que la tasa es bastante alta, no se considera un aumento sustancial para la estimación de la contención en la solución total implementada.

En cuanto a la concurrencia de llamadas, si bien se obtuvo un pick de demanda de 15 llamadas simultáneas, la infraestructura del *Contact Center* ya implementada en la empresa, permite un total de 30 llamadas concurrentes, luego al abarcar todas las opciones del IVR, es importante no disminuir esa capacidad y dimensionar las componentes necesarias para la concurrencia aceptada actualmente.

11 FACTIBILIDAD TÉCNICA

Por medio de este estudio se realiza la evaluación de los aspectos técnicos que se requieren para determinar la factibilidad de la implementación del *Contact Center Cognitivo* de IBM, específicamente en el servicio de atención al cliente de una empresa de retail en Chile. Cabe mencionar que, si bien se dimensionará para dicho caso de uso, las componentes técnicas a evaluar aplican para la generalidad de la implementación de la solución en otras industrias.

En este capítulo se define el tamaño del proyecto, la arquitectura y funcionamiento de la solución, la descripción y dimensionamiento de las componentes, y el requerimiento de capital humano para soporte y mantenimiento. Además, se estudian las restricciones técnicas asociadas a la implementación para el correcto funcionamiento de la solución.

11.1 TAMAÑO DEL PROYECTO

Para el posterior dimensionamiento de las capacidades de los servicios e infraestructura, es de vital importancia detallar el tamaño del proyecto a evaluar. En la **Tabla 5** se muestra la proyección, realizada por la empresa de retail, de la cantidad de llamadas a recibir en el horizonte de evaluación del proyecto.

Tabla 5.
Proyección de llamadas a recibir en el Contact Center de la empresa en el horizonte de evaluación.

<i>Opción IVR</i>	<i>Año 0</i>	<i>Año 1</i>	<i>Año 2</i>	<i>Año 3</i>
Horarios y ubicación	13.699	15.069	16.576	18.233
Despachos	221.331	243.464	267.811	294.592
Cambios y Garantía	111.698	122.868	135.155	148.670
Otras consultas	532.502	585.752	644.327	708.760
Total	879.230	967.153	1.063.868	1.170.255

Además, en base a la situación visualizada en la implementación del MVP, en conjunto con la visión del área de consultoría de IBM y el área de servicio al cliente de la empresa de retail, se realiza una estimación de la contención a alcanzar para cada una de las opciones actuales que involucra el IVR, esto se muestra en la **Tabla 6**.

Tabla 6.
Estimación de la contención (%) a alcanzar por opción de IVR

Opción IVR	Contención objetivo
Horarios y ubicación	85 %
Despachos	60 %
Cambios y Garantía	30 %
Otras consultas	25 %
Contención promedio ponderado⁵	35,38%

La estimación fue realizada considerando que se habilitará la solución para atender el 90% de todos los requerimientos que fueron clasificados con dificultad baja y media. Por la realización de proyectos similares, IBM estima que la contención objetivo se alcanza en un periodo de 6 meses, siguiendo la línea de tiempo de desarrollo indicada en la **Figura 12**. Sin embargo, dado que ya fue realizado el MVP, el punto inicial en que se sitúa este proyecto es en las 6 semanas, siendo necesario 1 mes para la migración a producción y otros 3 meses para alcanzar la solución completa.

En la **Tabla 7** se muestra la cantidad de llamadas que se estima serán atendidas por el *Contact Center Cognitivo*.

⁵ Promedio ponderado en función del porcentaje del total de llamadas que representa cada opción del IVR.

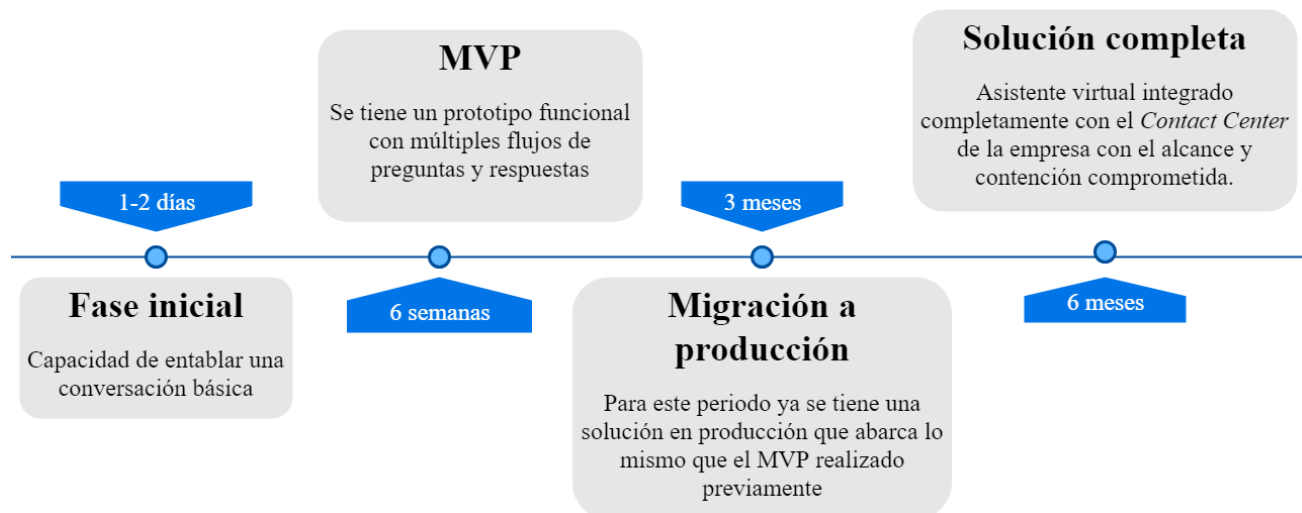


Figura 12. Línea de tiempo de la implementación de un Asistente virtual.
Fuente: Elaboración Propia a partir del documento *Customer Care Voice Agent* (IBM, 2018)

Tabla 7.
Volumen de llamadas estimado a contener por la solución en el horizonte de evaluación.

Opción IVR	Año 1	Año 2	Año 3
Horarios y ubicación	12.809	14.090	15.498
Despachos	146.078	160.687	176.755
Cambios y Garantía	36.860	40.547	44.601
Otras consultas	146.438	161.082	177.190
Total	342.185	376.404	414.044

Respecto al canal web, se pretende mantener las funcionalidades descritas en la situación sin proyecto y por lo tanto solo se considera en términos de volumen el crecimiento anual de las interacciones, el cual corresponde a un 15%, según lo proyectado por la empresa, en función del crecimiento del *e-Commerce*. La proyección de interacciones se encuentra en la **Tabla 8**.

Tabla 8.
Cantidad de interacciones en el canal web en el horizonte de evaluación

	Año 1	Año 2	Año 3
Cantidad de interacciones	2.137.137	2.457.708	2.826.364

Otra métrica importante de detallar a nivel de tamaño corresponde a la concurrencia. Como fue descrito en la situación sin proyecto, el *Contact Center* de la empresa permite 30 llamadas de manera simultánea y, si bien en la implementación del MVP se obtuvo un pick de 15 llamadas concurrentes, se espera que en la implementación en producción este número aumente debido a que estará habilitado para contestar todas las opciones, por lo tanto, se considera apropiado dimensionar para el nivel de demanda actual, esto es 30 llamadas concurrentes.

Para el canal telefónico será necesario también dimensionar en términos de los minutos atendidos por la solución. Esta cantidad depende de manera directa del volumen y duración de las llamadas atendidas junto con el porcentaje de contención estimado para cada opción. En la **Tabla 9** se detalla el consumo estimado de minutos, tomando en consideración el mayor TMO medido en el MVP, que corresponde a la opción de despachos con una duración de 1,65 minutos.

Tabla 9.
Estimación de la cantidad de minutos atendidos por la solución durante el horizonte de evaluación.

Opción IVR	Año 1	Año 2	Año 3
Horarios y ubicación	21.134	23.248	25.572
Despachos	241.029	265.133	291.646
Cambios y Garantía	60.820	66.902	73.592
Otras consultas	241.623	265.785	292.364
Total	564.606	621.067	683.173

11.2 ARQUITECTURA DE LA SOLUCIÓN

En la **Figura 13**, se presenta un diagrama de la arquitectura general del servicio de *Contact Center Cognitivo* de IBM, éste se enfoca en mostrar las componentes necesarias para su funcionamiento y los flujos de comunicación entre ellas. Por simplicidad no son mencionadas partes de la lógica de la integración, dado que no contribuye directamente al

cumplimiento de los objetivos de este trabajo de título. Cabe también mencionar que la solución está diseñada para conectarse, sin mayores modificaciones, a la infraestructura ya existente del *Contact Center* de una empresa, por lo que no se considera en el estudio las componentes de *back-office* y de telefonía y ya existentes en el centro de contacto de la empresa. En la misma línea, para el canal telefónico IBM trabaja en conjunto con la empresa Sixbell, quienes toman el rol de la integración de la parte cognitiva con la parte telefónica, por lo que algunas de sus aplicaciones son utilizadas para dichos fines.

Es importante mencionar que el *Contact Center Cognitivo* de IBM tiene una arquitectura de tipo modular, lo que implica que existe una base de servicios necesarios para su funcionamiento y sobre estos se pueden incorporar otros que la empresa que lo implementa desee tener, así, pueden añadirse servicios cognitivos de reconocimiento de imágenes (*visual recognition*) o de procesamiento de textos (*Discovery*) para la lectura de manuales de servicios o cualquier tipo de base de datos no estructurada con el objetivo de entregar una mayor base de conocimiento al asistente virtual.

Esta arquitectura está diseñada para una implementación en modalidad *cloud* de la solución, ejecutándose solo los componentes estrictamente necesarios directamente en el data center de la empresa.

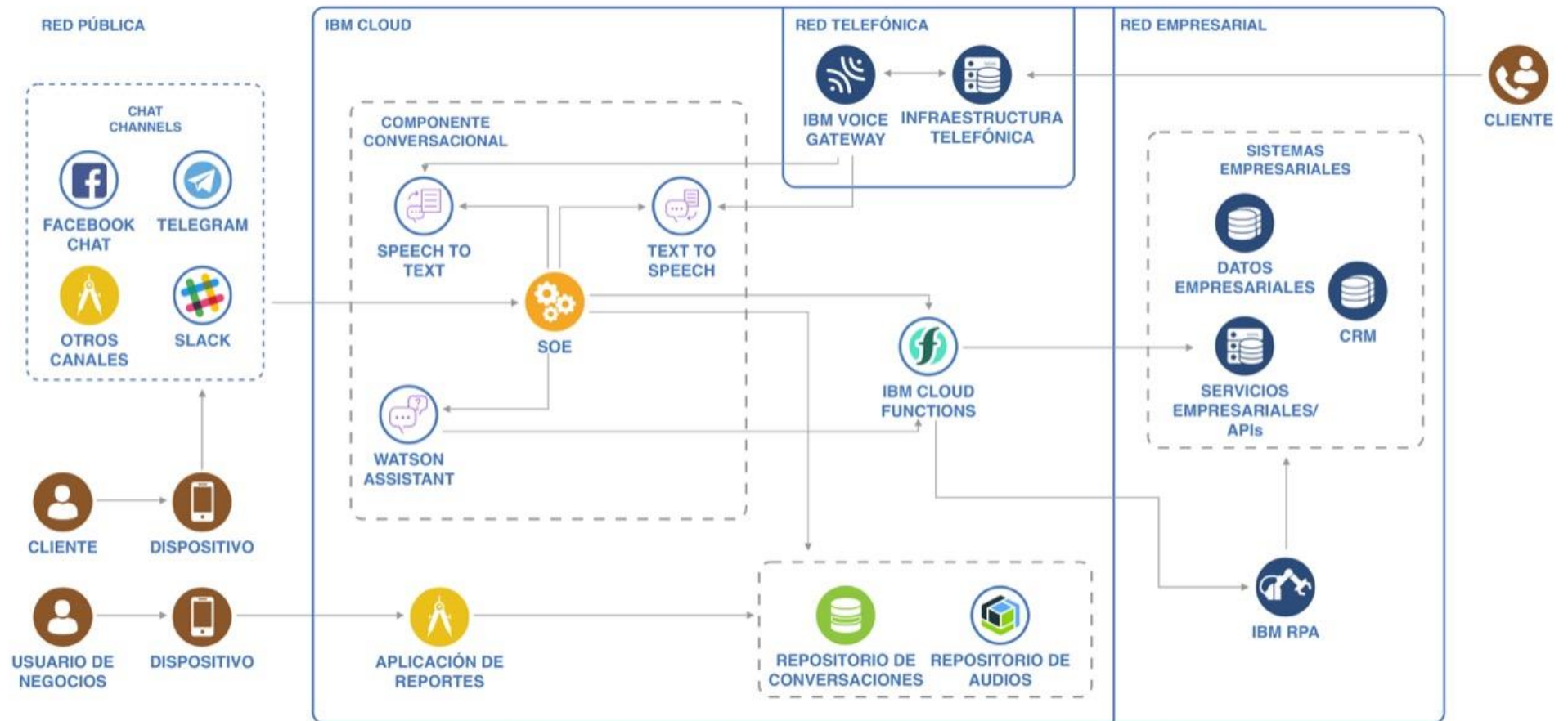


Figura 13: Diagrama de arquitectura general del Contact Center Cognitivo de IBM.
Fuente: Elaboración Propia

11.3 FUNCIONAMIENTO Y DESCRIPCIÓN DEL PROCESO DE INTERACCIÓN

Siguiendo el diagrama presentado en la **Figura 13**, el proceso de interacción con el asistente virtual comienza con el contacto por parte del cliente a través de alguno de los canales. Si la conexión se establece mediante el canal telefónico, la llamada ingresa a la infraestructura telefónica ya existente en el *Contact Center* de la empresa y, a través de las aplicaciones implementadas por Sixbell, se transforma la llamada proveniente de la red pública o PSTN a un protocolo de audio digital para telefonía (SIP), de manera de establecer la comunicación con el servicio de IBM Voice Gateway, quien habilita la conexión con los servicios de Watson. Luego, el audio del cliente es transformado a texto para entregarlo a Watson Assistant, que corresponde al servicio donde está creado el asistente virtual, éste responde al requerimiento en formato texto, que a continuación es traspasado a audio para ser escuchado por el cliente.

Durante todo el proceso, interfiere un orquestador (*Service Orchestrator Engine* o SOE) que conecta los servicios previamente mencionados y que, si es necesaria la integración con algún sistema de la empresa o realizar alguna consulta a un servicio externo, maneja la ejecución de las acciones requeridas llamando al servicio IBM Cloud Functions, quien consultará las APIs conectadas a los sistemas de la empresa.

Además, el servicio orquestador será el responsable de almacenar en una base de datos las conversaciones en formato de audio y de texto para su posterior análisis y reportería.

Una vez entregada la respuesta al requerimiento, si el cliente decide ser transferido a un ejecutivo, su intención es reconocida por Watson Assistant, quien envía un comando de transferencia para ser interpretado por los sistemas telefónicos que finalmente enrutan la

llamada a la central de agentes humanos. De lo contrario, si el cliente no tiene más consultas, se finaliza el proceso.

Ahora si el cliente se contacta vía chat, el proceso es ligeramente más sencillo donde el orquestador reconoce el canal de comunicación, toma la información, le da el formato apropiado y envía el mensaje a Watson Assistant para obtener una respuesta. De la misma manera que en el canal telefónico, el orquestador puede ejecutar acciones para la integración con los sistemas de la empresa, a la vez que almacena las conversaciones en una base de datos para su posterior reportería.

En términos operacionales, el proceso de atención de una llamada por medio del asistente virtual se indica en la **Figura 14**, donde el asistente procesa los requerimientos del cliente, consulta sistemas en caso de ser necesario, entrega la respuesta a la consulta del cliente, y finaliza la conversación con el respaldo de la transcripción y grabación de la llamada. Si durante este proceso el cliente manifiesta la intención de comunicarse con un ejecutivo, el asistente transfiere la llamada. De la misma forma si durante el proceso no es capaz de entender la consulta del cliente después de tres intentos, transfiere la llamada a un ejecutivo para que pueda dar respuesta. Las atenciones por el canal web siguen el mismo procedimiento, pero no incorpora la opción automática de transferencia a un ejecutivo, entregando sólo el número telefónico del *Contact Center* para que el cliente se comunique directamente.

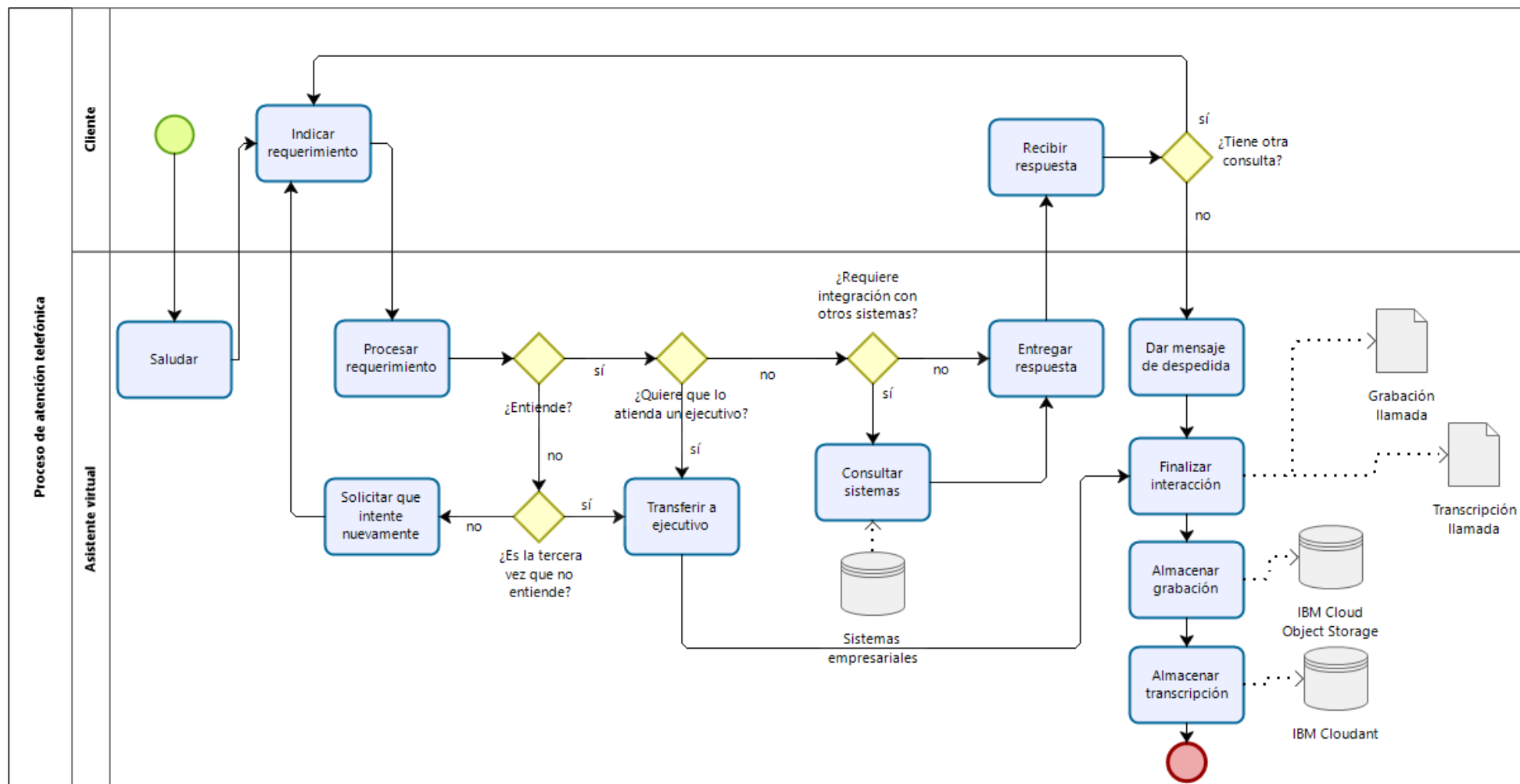


Figura 14. Diagrama de proceso de atención telefónica con asistente virtual.
Fuente: Elaboración propia.

11.4 HARDWARE REQUERIDO

Un servidor corresponde a un computador que atiende las peticiones que le hacen otros computadores denominados clientes. En esta relación, el servidor es quien centraliza los diferentes recursos y aplicaciones con que se cuenta, por lo que será la infraestructura en la que se ejecutarán todos los procesos, softwares y servicios necesarios para el funcionamiento de la solución. Dicho esto, un servidor puede ser físico, virtual o *cloud*. Los físicos o *bare metal* corresponde a la máquina en sí como hardware puro, que están disponibles exclusivamente para un cliente (dedicados); los virtuales están basados en la partición de un servidor en distintas máquinas virtuales, así cada virtualización tiene asignado una determinada capacidad de almacenamiento y procesamiento. Por último, los servidores *cloud* corresponden a servidores virtuales respaldados en una infraestructura redundante, lo que significa que están alojados en una combinación de servidores que se distribuyen la ejecución de las distintas instancias, esto lo hace más flexible y tolerante a las fallas de hardware.

Para la implementación del *Contact Center Cognitivo* de IBM, es necesario la utilización de servidores en los cuales se ejecuten los servicios. Para el caso de los servicios *cloud* se utilizan servidores en la nube, que en general vienen incluidos dentro de la suscripción, mientras que para el caso de las componentes telefónicas se utilizan servidores virtualizados en el *data center* de la organización que implementa la solución, a esta modalidad se le conoce como *on premise*.

Las especificaciones técnicas de cada servidor a utilizar son detalladas junto con las características y dimensionamiento de cada componente en la sección que sigue.

11.5 DESCRIPCIÓN Y DIMENSIONAMIENTO DE LOS COMPONENTES

11.5.1 WATSON ASSISTANT

Watson Assistant es una plataforma de computación conversacional de inteligencia artificial para la creación de asistentes virtuales soportada en IBM Cloud. A través del reconocimiento de lenguaje natural, permite implementar soluciones conversacionales escalables, privadas, seguras e integradas en una gran variedad de canales (IBM, 2017).

En el contexto del desarrollo del *Contact Center Cognitivo*, Watson Assistant corresponde al servicio que “crea” al asistente virtual que se comunicará a través de todos los canales, aquí se configura su base de conocimiento, personalidad, respuestas a entregar, y todo lo necesario para la correcta interacción con el cliente.

Algunos de los conceptos necesarios a conocer para la comprensión del funcionamiento de este servicio son:

- **Intención:** Se define como la detección de la necesidad de información de un usuario en un sistema conversacional.
- **Entidad:** Objeto de negocio o información concreta comunicada que, combinado con la intención, puede dar lugar a respuestas específicas.
- **Diálogo:** Corresponde al diseño, regla y lógica mediante las cuales se habilita a un agente virtual a comprender y procesar lenguaje y otros elementos de interacción, para generar una respuesta apropiada. En términos de estructura, un diálogo se compone por un conjunto de nodos condicionados en los cuales se ejecuta una determinada secuencia de pasos, que permiten recoger datos, solicitar la acción de otros servicios, entregar una respuesta, entre otras cosas.

- **Espacio de trabajo:** Se define como el marco en el que se desarrolla un determinado diálogo con sus intenciones y entidades propias, por lo que se podría visualizar cada uno de los espacios de trabajo como un asistente diferente.

Dicho esto, su funcionamiento consiste en identificar intenciones comunicativas e información específica entregada en una frase comunicada por un cliente, para posteriormente moverse dentro de la estructura de diálogo creada por los desarrolladores y entregar una respuesta al cliente.

Watson Assistant tiene una interfaz gráfica de usuario que admite tanto a desarrolladores como a personal no técnico a colaborar en la construcción de soluciones, esto es particularmente valioso para dar la visión de negocio y personalidad de la compañía al asistente virtual (IBM, 2017).

En términos técnicos, Watson Assistant corresponde a un servicio que es consumido a través de la interfaz de usuario o la API del Servicio de Cloud, que se ejecuta en algún servidor de IBM Cloud, el cual puede ser público o privado dependiendo de las necesidades del negocio. La suscripción al servicio incluye el uso de estos servidores, por lo que no deben ser dimensionados aparte.

Watson Assistant tiene tres modalidades de despliegue tipo *cloud* asociadas al uso comercial que se le dará, las cuales se detallan a continuación:

- **Lite:** En esta modalidad, la implementación se realiza en una infraestructura compartida en alguna de las localizaciones de servidores públicos de IBM Cloud. Luego, todos los clientes y sus usuarios consumen el servicio desde la misma plataforma tecnológica, a esto se le conoce como *multi-tenant*. Este despliegue está orientado a

personas que quieran aprender a usar la herramienta, por lo tanto, su uso es gratuito y limitado en la cantidad de interacciones a realizar.

- **Estándar:** Al igual que la modalidad Lite, la implementación es *multi-tenant* en una infraestructura compartida en algún servidor público de IBM Cloud, la diferencia se encuentra en que en este despliegue no existen restricciones en la cantidad de consultas a la API. Esta modalidad está orientada al desarrollo de proyectos de tamaño mediano a grande que no impliquen el manejo de información altamente sensible (como contraseñas bancarias).
- **Premium:** En esta modalidad, la implementación es igualmente en una infraestructura compartida, pero que tiene una instancia del servicio para cada cliente, luego todos los usuarios consumen instancias propias del servicio en una misma plataforma tecnológica, a esto se le conoce como *single tenant*. Además, en conjunto con el aislamiento a nivel de cálculo, en este despliegue se realiza un cifrado de datos (en tránsito y en reposo) para la seguridad de estos mismos. Por lo tanto, este plan está orientado al desarrollo de proyectos de negocio que requieran una alta seguridad de la información.

En la **Tabla 10** se detallan a modo de comparación las características de cada modalidad de despliegue, estas contribuyen a la elección del plan a seleccionar para la implementación de la solución.

Tabla 10.
Características de modalidades de despliegue de Watson Assistant.

Tipo de plan	Llamadas de API al mes	Espacios de trabajo	Intenciones	Entidades
Lite	Hasta 10.000	5	100	25
Estándar	Ilimitadas	20	2000	1000
Premium	Ilimitadas	50	2000	1000

Las llamadas a la API corresponden, en términos sencillos, a la cantidad de veces que el asistente virtual enviará un mensaje al cliente en el transcurso de la conversación. Por lo visualizado en el actual chat desplegado en la página web, la cantidad de llamadas a la API en dicho canal corresponde a 2 en promedio, mientras que, para el canal telefónico este número asciende a 5 según lo observado en el MVP. Para cualquiera de los dos casos, es indispensable contar con un plan ilimitado en términos de llamadas a la API, de lo contrario la solución no sería capaz de resolver los requerimientos una vez alcanzada la cuota de 10.000 mensajes al mes.

El número de intenciones y entidades a reconocer está directamente relacionado con las características del negocio del retail y, en consecuencia, con la variedad de las preguntas que deban ser respondidas. Para estimar la cantidad de intenciones y entidades a utilizar se toma en cuenta el chat ya implementado en la página web de la empresa, el cual cuenta con 136 intenciones y 50 entidades.

Por las características mencionadas, se decide que el plan más adecuado para la implementación corresponde al estándar. Este tiene un costo de 0,00288 USD / llamada de API.

11.5.2 WATSON SPEECH TO TEXT (STT)

Speech to Text corresponde a un servicio de Watson disponible en IBM Cloud, que permite transcribir mensajes de voz a texto de alta o baja calidad en una amplia variedad de formatos (MP3, WAV, FLAC, y otros), combinando información sobre la gramática y la estructura del lenguaje con el conocimiento de la composición de la señal de audio. El servicio utiliza modelos estadísticos e IA para determinar el resultado más preciso indicando el nivel de confianza de su transcripción, el cual puede ir aumentando a medida que se le entrena con un mayor volumen de audios. También permite crear, entrenar y alojar modelos de lenguaje personalizados para casos de uso más específicos (IBM, 2017)

El dimensionamiento del uso de este servicio se realiza en base a la cantidad de minutos de transcripción requeridos, así se tienen las siguientes modalidades de despliegue:

- **Lite:** Esta modalidad corresponde a una implementación *multi-tenant* en un servidor público de IBM Cloud. Ofrece un límite de transcripción de 100 minutos.
- **Estándar:** Al igual que la modalidad Lite, la implementación es *multi-tenant* en un servidor público de IBM Cloud. Ofrece una base de consumo de 1.000 minutos gratuitos y luego se paga por cada minuto adicional de transcripción según precios diferenciados por tramo. Esta modalidad permite, de manera opcional, generar un modelo de transcripción customizado al caso de uso del cliente, esto es, un modelo que esté entrenado exclusivamente por audios reales proporcionados por la organización de manera de que el servicio sea capaz de transcribir correctamente terminologías específicas del negocio.
- **Premium:** Corresponde a una implementación con iguales características que la Estándar, pero en modalidad *single tenant*, por lo que su uso está orientado a empresas

que requieran aislamiento de datos en el tratamiento de información sensible de sus clientes.

Debido a la cantidad de minutos a procesar y a la característica adicional de modelos customizados para el caso de uso, se decide implementar un plan estándar, cuyo costo se muestra en la **Tabla 11**. La modalidad premium es descartada debido a que la empresa no considera necesario el aislamiento de datos a ese nivel.

Tabla 11.
Tarifa por minuto utilizado del servicio de Watson STT.

<i>Tramo de minutos</i>	<i>Precio minuto estándar (USD/minuto)</i>	<i>Adicional por modelo customizado (USD/minuto)</i>
1 – 250.000	0,023	0,0345
250.001 – 500.000	0,0173	0,0345
500.001 – 1.000.000	0,0144	0,0345
1.000.000+	0,0115	0,0345

11.5.3 TEXT TO SPEECH (TTS)

Corresponde al servicio necesario para comunicar los mensajes de Watson Assistant a través del canal telefónico pasándolos de texto a voz. IBM cuenta con un servicio de Watson TTS, sin embargo, la calidad de la voz en español es baja comparada con la alternativa en inglés, esto hace que la solución se vea afectada en términos de experiencia del cliente y por tanto se descarta el uso de este servicio al menos hasta que se tengan nuevas versiones más humanizadas de voz.

Como alternativa se presenta el uso de Nuance Vocalizer, que corresponde al servicio de TTS con mayor participación en el mercado de voces sintetizadas para IVR, y con gran experiencia en el desarrollo de voces en más de 40 idiomas, esto les ha permitido desarrollar

una voz más humanizada y natural, contando con alrededor de 100 voces diferentes, tanto en género femenino como masculino. Este servicio además entrega al usuario un alto nivel de flexibilidad, permitiendo tener el control sobre el volumen, la velocidad, el tono y el énfasis, con el cual se reproduce el mensaje (Nuance, 2014).

En particular para el *Contact Center Cognitivo* se utiliza, por requerimientos de la empresa de retail, la voz femenina en español neutro, esta corresponde a Paulina dentro del catálogo de Nuance Vocalizer. Este servicio se ejecuta *on premise*, y los requerimientos del servidor necesarios para su funcionamiento se encuentran en la **Tabla 12**.

Tabla 12.
Especificaciones técnicas del servidor para la ejecución del servicio de TTS Nuance Vocalizer.

Sistema Operativo	Windows Server 2012
CPU	1vCPU
Memoria RAM	4GB
Disco Duro 1	100GB
Disco Duro 2	100GB

La comercialización de la aplicación se realiza a modo de licencia perpetua por concurrencia de llamadas y por voz a utilizar, así según una cotización realizada por Sixbell, el costo de la licencia para el uso de la voz en español neutro (Paulina) corresponde a 500 USD/llamada concurrente, de esta forma, el costo a perpetuidad sería de 15.000 USD.

11.5.4 IBM VOICE GATEWAY

IBM Voice Gateway (IVG) es un servicio de conectividad que conecta Watson Assistant, Watson Speech to Text y Watson Text to Speech con puntos finales SIP, comúnmente utilizados en sistemas telefónicos. De esta manera, se utiliza para hacer que los asistentes virtuales activos de Watson Assistant estén disponibles por teléfono, de modo que

los usuarios finales puedan sostener una conversación con el asistente creado. La herramienta permite el uso de servicios de voz a texto o de texto a voz de Watson (que es la opción predeterminada), o también la conexión a servicios de voz de otros proveedores.

Este servicio se implementa en modalidad *on premise* junto con los otros componentes de integración telefónica.

La tarifa de uso de este servicio está asociado a la cantidad de minutos utilizados, existiendo dos modalidades:

- **Lite:** Corresponde a un plan de prueba con 100 minutos gratuitos para uso del servicio.
- **Estándar:** Corresponde a un plan sin restricción de consumo de minutos, en el que se cuenta con 1.000 minutos gratuitos y luego se paga por cada minuto adicional según tramos de precios diferenciados. De manera adicional se debe cancelar un valor fijo por el número de conexiones concurrentes deseadas de 23,00 USD/por conexión simultánea.

Por las condiciones de los planes detallados y la cantidad de minutos a procesar, se opta por la contratación de un plan en modalidad estándar. En la **Tabla 13** se detallan las características necesarias del servidor para la ejecución de este servicio, y en la

Tabla 14 se detallan los costos del plan.

Tabla 13.
Especificaciones técnicas del servidor para la ejecución del servicio IBM Voice Gateway.

Sistema Operativo	Linux Ubuntu 16.04
CPU	2vCPU
Memoria RAM	8GB
Disco Duro	200GB

Tabla 14.
Tarifa por minuto utilizado del servicio IVG

<i>Tramo de minutos</i>	<i>Precio (USD/minuto)</i>
1 – 249.000	0,069
249.001 – 499.000	0,0575
499.001 – 999.000	0,046
999.000+	0,0345

11.5.5 TMP

El TMP o Telecom Media Portal, es un software desarrollado por Sixbell para reemplazar a IVR convencionales tecnológicamente inferiores en cuanto a características y funcionalidades. Provee medios para generar una interacción con el usuario final, permite reproducir y grabar audio, reconocer opciones ingresadas por el usuario mediante DTMF o voz, enviar y recibir fax, permitir la creación y manejo de conferencias, entre otras cosas (Bravo Abarca, 2012).

Además, en la comunicación de voz a través de internet (VoIP), el TMP actúa también como un enlace que traduce la información del protocolo⁶ utilizado en una red inicial, al protocolo usado en la red de destino, para los múltiples protocolos de telefonía existentes. A esta funcionalidad se le conoce como Gateway y hace que pueda conectarse a casi cualquier infraestructura de *Contact Center* que posea una organización sin realizar mayores modificaciones.

⁶ Un protocolo de comunicaciones se define como un conjunto de reglas y/o de normas que permiten que dos o más equipos se comuniquen entre ellos. Estas reglas involucran cómo tiene que hacerse el intercambio de información (señalización) y el traspaso de información de los datos de voz a entregar.

En el contexto de la solución ofrecida por IBM, el TMP se utiliza para realizar la comunicación de la red de *Contact Center*, ya implementado en la organización, con el *Contact Center Cognitivo*, permitiendo pasar del protocolo de comunicación H323 (utilizado en la organización) al protocolo SIP, que es requerido para su funcionamiento. Esto a la vez que da la posibilidad de recibir y transferir llamados a un agente de la central en caso de ser necesario, y de grabar estas llamadas para posterior reportería o respaldos de seguridad. En la **Figura 15**, se muestra un diagrama⁷ de conexión del TMP con la actual infraestructura del *Contact Center* de la empresa y con el *Contact Center Cognitivo*.

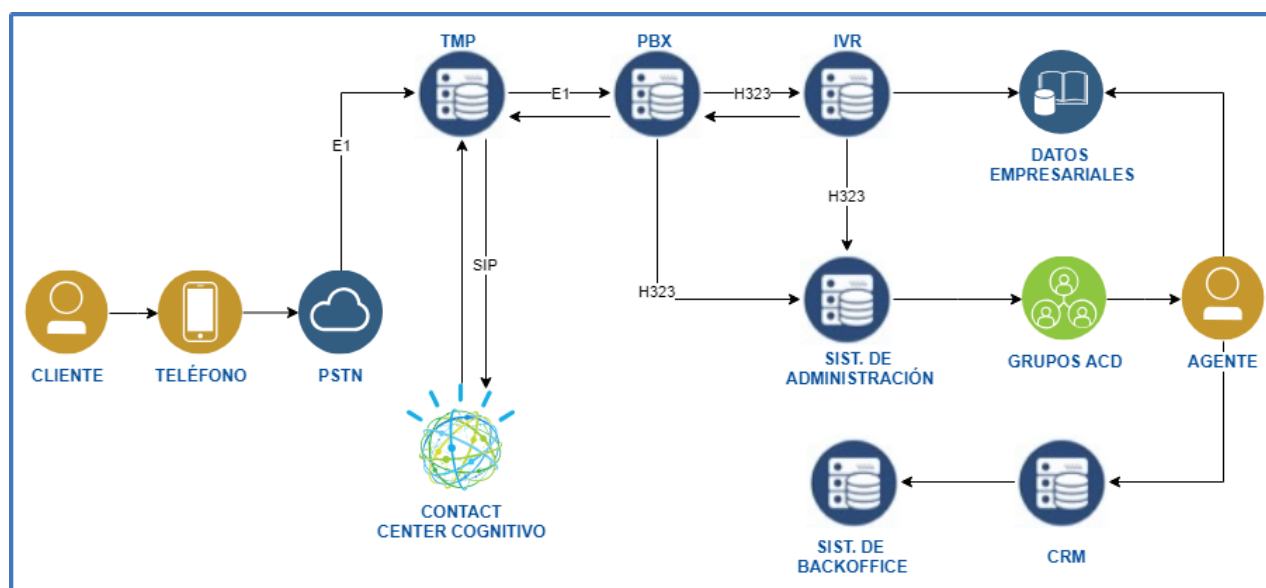


Figura 15. Diagrama de conexión del TMP con el Contact Center de la empresa de retail.
Fuente: Elaboración propia.

El TMP debe ser instalado en un servidor cuyas características guardan relación con la concurrencia de las llamadas, esto porque entre más llamadas simultáneas se procesen, se consume una mayor cantidad de memoria RAM y CPU, por lo tanto, el dimensionamiento

⁷ Los comentarios añadidos en las flechas corresponden a los protocolos de conexión. La E1 corresponde a un formato de transmisión digital de datos.

del servidor a utilizar se realiza en torno a la concurrencia de llamadas que se desea atender.

En la **Tabla 15**, se detallan las características requeridas para un servidor de manera que se pueda atender las 30 llamadas concurrentes que se requieren dado el tamaño del proyecto.

Tabla 15.
Especificaciones técnicas del servidor para la instalación del software TMP.

Sistema Operativo	Linux Red hat 6.9
CPU	8vCPU
Memoria RAM	12GB
Disco Duro 1	500GB
Disco Duro 2	500GB

11.5.6 SIGAS

El SigAS o Signaling Application Server, es un software desarrollado por Sixbell para manejar el intercambio de información entre los puntos involucrados en la red telefónica de manera de configurar, controlar y finalizar cada llamada telefónica (Bravo Abarca, 2012). De esta forma el SigAS es capaz de, entre otras cosas, indicar el número de destino y procedencia de la llamada, informar respecto de la configuración de llamada (por ejemplo: descolgado, tono de marcado, ocupado y números de redirecciones), y almacenar información relevante del cliente recuperada durante el proceso (por ejemplo, el número de RUT a través de DTMF), de manera de hacerla llegar a la central de ejecutivos, pasando por el mundo telefónico.

Al igual que el TMP, este software se dimensiona en función a la cantidad de llamadas concurrentes procesadas, así para 30 llamadas concurrentes será requerido un servidor con las características indicadas en la **Tabla 16**.

Tabla 16.
Especificaciones técnicas del servidor para la instalación del software SigAS.

Sistema Operativo	Linux Red hat 6.9
CPU	6vCPU
Memoria RAM	12GB
Disco Duro	500GB

11.5.7 IBM CLOUD FUNCTIONS

IBM Cloud Functions es una plataforma de función como servicio (FaaS) que ejecuta funciones como respuesta a sucesos entrantes, soportando cualquier lenguaje de programación ya sea de manera directa o a través de la carga de otros lenguajes (IBM, 2017).

Cloud Functions se factura por segundo y por gigabyte de memoria utilizada en cada una de las ejecuciones, de esta forma el usuario solo paga por lo que usa, en función del tiempo durante el cual el código da servicio a solicitudes, con un redondeo de hasta 100 milisegundos.

En el contexto de la implementación del *Contact Center Cognitivo*, se utiliza para crear acciones, que corresponde una parte de código que realiza tareas específicas, para consultar APIs que se conectan a los sistemas de la organización de manera de rescatar información necesaria para dar respuesta al requerimiento del cliente atendido por el asistente virtual. En la **Tabla 17** se detallan las especificaciones del uso de esta plataforma.

Tabla 17.
Especificaciones del uso de IBM Cloud Functions.

Tiempo de ejecución de acciones	500ms
Memoria requerida	256MB
Ejecuciones al mes	5.000.000
Costo mensual	10,63 USD

11.5.8 IBM CLOUDANT

IBM Cloudant es una base de datos como servicio (DBaaS) NoSQL, que permite el almacenamiento e indexación de documentos JSON para la gestión de datos en múltiples cargas de trabajo. Su propuesta de DBaaS trae consigo una reducción de costos en materia de administración, almacenamiento, desarrollo y servidores, en comparación con las alternativas de almacenamiento tradicionales (IBM, 2017).

El precio se basa en el rendimiento y almacenamiento requerido, de esta forma existen los siguientes planes asociados:

- **Lite:** Es gratuito e incluye una cantidad fija de capacidad de procesamiento y almacenamiento para fines de desarrollo y evaluación. La capacidad de rendimiento aprovisionada se fija en 20 búsquedas por segundo, 10 escrituras por segundo y 5 consultas por segundo, y el almacenamiento de datos está limitado a 1 GB.
- **Estándar:** Ofrece capacidad de rendimiento configurable y almacenamiento sin restricción, con precios que se escalan a medida que cambian los requisitos del proyecto para el cual se utiliza. El precio del plan estándar se basa en dos factores: la capacidad de rendimiento que se asigna y la cantidad de datos que se almacenan en la instancia.

En el contexto del *Contact Center Cognitivo*, Cloudant se utiliza para almacenar los registros de las interacciones realizadas. Estos documentos se agrupan principalmente en dos categorías, los correspondientes a los detalles de la conversación, donde se tiene información respecto a la transcripción de los mensajes, las intenciones y entidades reconocidas, los nodos visitados, fecha y hora de la transacción, entre otras cosas; y los asociados a la interacción telefónica, estos registros contienen información del número telefónico en el cual se realizó la llamada, duración de la llamada, y otros datos que fueron necesarios para realizar el ruteo

de ésta. Así, estos documentos corresponden a la fuente de datos para herramientas de reportería implementadas que presentan información consolidada de hasta un año de historia. Para periodos posteriores, la información deberá ser migrada a una base de datos histórica, que será consultada solo en caso de necesitarse un registro en específico.

Considerando los requerimientos de la aplicación de reportería y el almacenamiento de un año de historia, los requerimientos de la base de datos se presentan en la **Tabla 18**. En la **Tabla 19** se proyecta el consumo de la base de datos en el horizonte de evaluación y su costo implicado, considerando que cada contacto telefónico implica un registro de tamaño 15 kB y cada contacto escrito uno de 4 kB.

Tabla 18.
Especificaciones del uso de la base de datos Cloudant.

<i>Característica</i>	<i>Especificación</i>	<i>Costo</i>
Almacenamiento	Variable por año	1,15 USD/GB
Búsquedas por segundo	100	0,2875 USD/búsqueda por seg
Escrituras por segundo	50	0,575 USD/escritura por seg
Consultas por segundo	5	5,75 USD/consulta por seg

Tabla 19.
Estimación de uso de Cloudant en el horizonte de evaluación

	<i>Año 1</i>	<i>Año 2</i>	<i>Año 3</i>
Almacenamiento (GB)	5,6	7,4	9,2
Costo (USD)	93,15	94,30	96,6

11.5.9 IBM CLOUD OBJECT STORAGE

“Cloud Object Storage es un formato para almacenar datos no estructurados en la nube que tiene las características de ser elástico, flexible y escalable para soportar un crecimiento ilimitado de datos. La arquitectura almacena y administra los datos como objetos en

comparación con otros tipos de almacenamiento que los manejan como bloques o archivos jerárquicos” (Rouse, 2017).

La principal ventaja del almacenamiento de objetos es que hace que los datos sean más resistentes a desastres o fallas de hardware porque están altamente distribuidos, por lo que aún están disponibles incluso si fallan varios nodos. También es mucho más barato en comparación con el almacenamiento tradicional porque el almacenamiento de objetos se almacena máquinas o servidores virtuales con alta capacidad para ser escaladas.

“Por sus características de lectura y escritura, este tipo de almacenamiento es más adecuado para datos estáticos y almacenamiento de copias de seguridad. (...) El caso de uso menos deseable para el almacenamiento de objetos son las aplicaciones y los entornos con altas tasas de transacción, esto porque los sistemas de almacenamiento de objetos no son lo suficientemente consistentes para los sistemas en tiempo real, como las bases de datos transaccionales. El almacenamiento de objetos no proporciona una garantía de que una solicitud de lectura devolverá la versión más reciente de los datos. Además, la tecnología no siempre es adecuada para aplicaciones con altas demandas de rendimiento” (Rouse, 2017).

IBM Cloud Object Storage es la propuesta de almacenamiento de objetos en la nube de IBM, que se puede utilizar para almacenar datos no estructurados como videos, audios, imágenes, documentos, etc. Dependiendo de las necesidades del negocio, puede utilizarse como un repositorio para respaldo y recuperación, repositorio de contenido de multimedia, *data lake* para análisis, y como un servicio de almacenamiento para aplicaciones nativas de la nube.

El uso de este servicio tiene asociado métricas de cargo según la cantidad de almacenamiento y la cantidad de consultas a realizar, teniendo disponible una gran cantidad de planes según las características de la base de datos que se pretenda implementar. Para este caso se estudian los siguientes planes:

- **Lite:** Corresponde a un plan gratuito para almacenar hasta 25 GB con una capacidad de 2.000 escrituras y 20.000 lecturas por mes.
- **Estándar:** Corresponde a un plan con almacenamiento y métodos de petición ilimitados, el costo corresponde a 0,022 USD GB/mes almacenado; 0,006 USD por cada 1.000 escrituras; 0,005 USD por cada 10.000 lecturas.
- **Archive:** Esta modalidad se utiliza para el almacenamiento de datos a los que raramente se accede, por lo que para hacer consultas a la base de datos se debe realizar previamente una restauración de los datos a una modalidad estándar. El costo de almacenamiento 0,002 USD GB/mes, y el de restauración de 0,02 USD GB/mes.

En el caso del *Contact Center Cognitivo*, IBM Cloud Object Storage cumple la funcionalidad de almacenamiento de los audios de las llamadas atendidas por el asistente virtual, y de base de datos histórica como respaldo de los archivos almacenados en Cloudant una vez pasado un año de antigüedad.

Para este propósito se utiliza una modalidad Estándar para el almacenamiento de los audios, que estarán disponibles para consulta en la aplicación de reportería durante 1 año, y luego de este periodo se migrarán a una modalidad Archive. De la misma forma, se tiene una instancia Archive para el traspaso de los datos almacenados en Cloudant una vez transcurrido un año, quedando a criterio de la empresa el periodo de tiempo que serán mantenidos.

El dimensionamiento del repositorio de audios se realiza en base al tamaño de las grabaciones, el cual por lo obtenido en el MVP corresponde en promedio a 3 MB. Para el dimensionamiento de las bases de datos de respaldos de información, se consideran los datos acumulados de almacenamiento de 1 año de historia en Cloudant detallados en la **Tabla 19** y el total de audios almacenados en el año anterior en modalidad Archive. En la **Tabla 20** se muestra la estimación del uso del servicio con los costos asociados a cada modalidad.

Tabla 20.
Estimación del uso de IBM Cloud Object Storage con sus respectivos costos.

	<i>Acumulado Año 1</i>	<i>Acumulado Año 2</i>	<i>Acumulado Año 3</i>
Estándar (audio)(GB)	408,3	659,6	899,9
Costo Estándar (USD)	65,6	103,1	139,1
Archive (audio) (GB)	0	408,3	1067,9
Archive (texto) (GB)	0	7,4	9,3
Total Archive (GB)	0	415,8	1.077,1
Costo Archive (USD)	0	0,8	2,2

11.5.10 REPORTERÍA DE DATOS

Como ha sido mencionado, los datos almacenados producto de las interacciones de los clientes con el asistente virtual deben ser presentados en alguna herramienta de reportería que permita visualizar de manera rápida y en tiempo real el comportamiento de la solución. Para ello fue desarrollado un *dashboard* que permite a los usuarios de la empresa que implementa el servicio tener información respecto a la cantidad de llamadas procesadas, la tasa de contención alcanzada, el tiempo medio de operación del asistente virtual, el tiempo

de respuesta promedio en la atención, la calidad de las llamadas⁸, las intenciones comunicativas más reconocidas y los tiempos de latencia de los servicios. Además, a través del *dashboard* se puede acceder al detalle de cada una de las llamadas atendidas por el asistente virtual, con su grabación y transcripción respectiva, para un año de historia, por medio de la conexión con las bases de datos Cloudant y IBM Cloud Object Storage definidas anteriormente. En la **Figura 16** se muestra la vista principal del *dashboard*, y en la **Figura 17** la vista de detalle de la llamada.

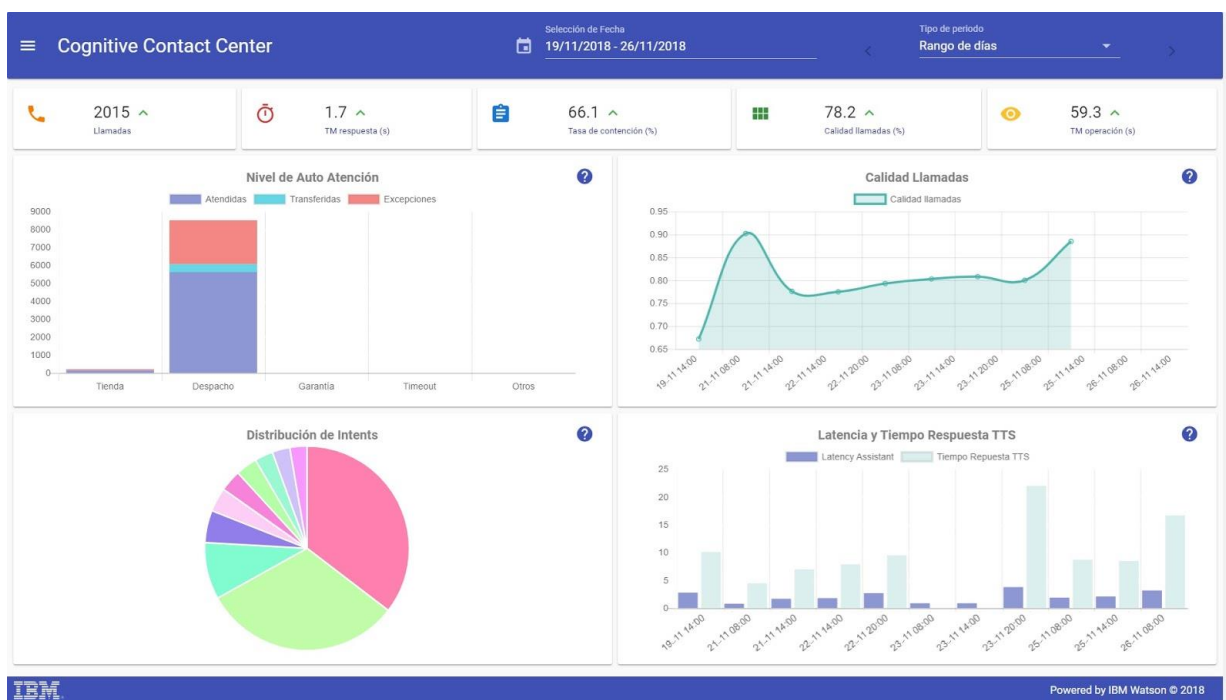


Figura 16. Vista principal dashboard Contact Center Cognitivo.

⁸ Definida en función del nivel de confianza con el que el servicio de STT hace la transcripción de los mensajes y la cantidad de llamadas que tuvieron que ser transferidas por problemas de entendimiento.

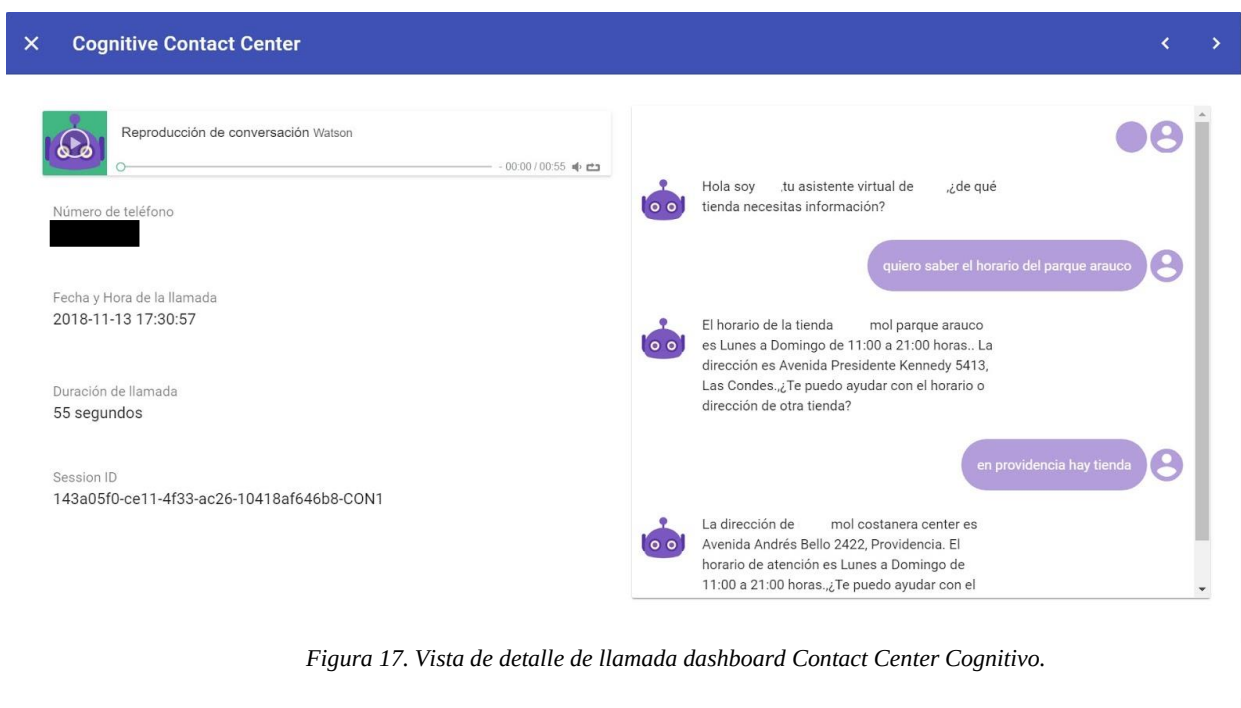


Figura 17. Vista de detalle de llamada dashboard Contact Center Cognitivo.

Para el funcionamiento del *dashboard* se requiere del uso de un Kit de Desarrollo de Software (SDK) para Node.js correspondiente a la plataforma como servicio en la cual se ejecutará el código de desarrollo para que el *dashboard* sea operativo. Este servicio se encuentra disponible en IBM Cloud, y se paga según la cantidad de instancias y memoria requerida. Para esta herramienta es necesario 6 instancias de 512 MB de memoria, lo que implica un costo de 143,69 USD mensuales.

11.6 REQUERIMIENTOS DE PERSONAL

Como ha sido mencionado, el *Contact Center Cognitivo* está diseñado para realizar un trabajo en conjunto con los ejecutivos, apuntando en este caso de uso, a atender solo aquellos requerimientos de complejidad baja y media. Para lograr esto, y que sea sostenible en el tiempo, se requiere constantemente de personal que mida el progreso de la solución realizando ajustes en caso de ser necesario.

El esquema de trabajo propuesto por IBM consiste en entregar las herramientas suficientes para que el personal de la empresa cliente pueda tomar un rol activo en la construcción y ajustes de la solución, por lo que se requieren equipos de trabajo de ambas partes para que trabajen en conjunto. A continuación, se define el perfil y responsabilidades de cada uno de los integrantes por equipo de trabajo.

11.6.1 CONSULTORÍA IBM

La labor de IBM en el proyecto consiste en el diseño, construcción y puesta en marcha de la solución, a la vez que realiza un traspaso de conocimiento al equipo de la empresa cliente con el objetivo de capacitarlos para el mantenimiento y ajustes de la solución en el tiempo. De esta forma se tienen los siguientes roles:

- **Project Manager (PM):** Es responsable de liderar el avance del proyecto, velando por el alcance de los hitos de entrega con la calidad ofrecida en la propuesta.
- **Consultor Cognitivo (CSC):** Es responsable de definir los requerimientos cognitivos del proyecto, a través de levantamientos de información con personal clave de la empresa cliente, con el fin de acotar el dominio de conocimiento.
- **Arquitecto Cognitivo (CSA):** Es el responsable de definir la arquitectura de la solución y velar por su correcta implementación. Define las implementaciones de hardware, software y la integración con sistemas de la empresa cliente si corresponde.
- **Especialista Cognitivo (CTE):** Este cargo se divide generalmente en dos posiciones, abarcando la parte cognitiva y de desarrollo. En su rol cognitivo realiza las capacitaciones a los clientes sobre el uso de las herramientas involucradas en la solución; modela e implementa el dominio de conocimiento a través de Watson Assistant; y lidera el

proceso de entrenamiento cognitivo, durante el cual el asistente virtual aprende a responder preguntas basadas en el dominio de conocimiento definido. En su rol de desarrollador es el responsable de configurar las APIs, servicios de orquestación y cualquier otro elemento que sea relevante para soportar los servicios de IBM Cloud involucrados, de manera de desplegar correctamente la solución en los canales deseados.

En la **Tabla 21**, se muestra la dedicación en horas hombre por mes de cada uno de los roles que componen la célula de consultoría junto al costo que implica, donde el costo fijado por IBM por hora de cada perfil corresponde a 2UF. Respecto a la duración, normalmente para proyectos de implementación de asistentes virtuales, IBM contempla una consultoría de 6 meses. Para este proyecto en particular se considera 1 mes de trabajo para el traspaso de lo desarrollado durante el MVP a producción; 3 meses de trabajo dedicados a cubrir el alcance total del proyecto; 1 mes de prueba con clientes para realizar ajustes; y 1 mes para finalizar, entregar documentación y hacer entrega final al cliente.

Tabla 21.
Dedicación y costo de cada rol requerido en la célula de consultoría de IBM

Rol	Dedicación (HH/mes)	Costo (UF/mes)
Project Manager	16	32
Consultor Cognitivo	40	80
Arquitecto Cognitivo	40	80
Especialista Cognitivo ⁹	320	640

⁹ Considerando 2 recursos, uno que se despeña en el rol cognitivo y otro en el rol de desarrollador.

11.6.2 EQUIPO DE TRABAJO EMPRESA CLIENTE

En el desarrollo del proyecto, la empresa cliente tiene la responsabilidad de definir el alcance de la solución, transmitiendo la lógica del negocio al equipo de IBM. Esto es fundamental para el correcto entendimiento de la situación sin proyecto y para la creación de un asistente virtual en consecuencia con esta lógica, con los objetivos estratégicos de la empresa y con su imagen corporativa.

Además, es responsabilidad de la empresa facilitar los contenidos requeridos para crear la base de conocimiento del asistente virtual (preguntas frecuentes con sus respectivas respuestas, protocolos de atención, audios de atención de clientes, entre otras cosas), y la certificación del funcionamiento a lo largo de la etapa de desarrollo.

Dicho esto, es necesario contar con los siguientes roles:

- **Project Manager (PM):** Responsable de la relación comercial entre IBM y la empresa cliente. Proporciona la administración general de los recursos contratados en IBM y la logística de en torno a la entrega de los avances.
- **Jefe de servicio (JS):** Debe facilitar el acceso a las decisiones, definiciones y contenidos necesarios junto con asegurar la participación de personal de la empresa requerido para completar las tareas que implica el desarrollo de la solución.
- **Administrador Técnico (AT):** Responsable de gestionar los aspectos técnicos y de seguridad de la implementación de la solución, habilitando el hardware, servicios web y sistemas empresariales para integrarse con el *Contact Center Cognitivo*.
- **Consultor de Negocio (CN):** Corresponde a personal que maneja la lógica del negocio y que es responsable de la definición de los requerimientos funcionales de la solución y de la validación y certificación de estos en las etapas de pruebas. También

debe facilitar los contenidos requeridos para el desarrollo de la base de conocimientos del asistente.

- **Usuario Cognitivo (UC):** Corresponde a quien, durante el transcurso del proyecto, se forma en el uso y construcción de la solución, por lo que cumple el rol de experto (por parte de la empresa) en el dominio de la solución general, y en específico de la herramienta Watson Assistant. Es responsable del reentrenamiento y actualización de la solución durante todo el horizonte de evaluación del proyecto.

En la **Tabla 22**, se muestra la dedicación en horas hombre por mes de cada uno de los roles, junto al costo que implica. Además, se muestra el periodo de tiempo en que será necesario cada rol, donde el periodo de consultoría corresponde al tiempo en que se contratarán los servicios de IBM.

Tabla 22.
Presencia, dedicación y costo de cada rol requerido en la célula de consultoría de la empresa cliente.

Rol	Presencia	Dedicación (HH/mes)	Costo¹⁰(\$/mes)
Project Manager	Consultoría	16	\$ 216.218
Jefe de servicio	Consultoría	32	\$ 432.436
Administrador Técnico	Consultoría	40	\$ 356.678
Consultor de Negocio	Consultoría	80	\$ 561.557
Usuario Cognitivo	Permanente	160	\$ 1.088.145
Total			\$ 2.665.033

¹⁰ Considerando un sueldo de Ingeniero Civil Industrial para el cargo de PM y JS con 3 o más años de experiencia y con 1 año de experiencia para el CN. Para el AT, se considera un sueldo de Ingeniero en ejecución Informática con 3 o más años de experiencia, y para el UC, un sueldo de Ingeniero Civil Informático con 1 año de experiencia. Estos son basados en el Estudio Nacional de sueldos de Ingenieros 2017.

11.7 RESTRICCIONES TÉCNICAS

11.7.1 LATENCIA EN LA COMUNICACIÓN

Cuando un usuario se conecta mediante el canal telefónico, se establece una comunicación de voz sobre internet (VoIP), donde el cliente habla por un micrófono, el audio se convierte en un flujo digitalizado y se comprime en "paquetes" de datos de duración entre 10 y 40 milisegundos cada uno. Esos paquetes individuales se transmiten secuencialmente a través de la estructura del *Contact Center*, y en el proceso pueden generarse retrasos o pérdida de paquetes de audio, lo que compromete la calidad de la interacción (Barbier, 2018).

La latencia corresponde al tiempo de retardo en la entrega de los paquetes de voz y generalmente se mide en milisegundos desde el momento en que el hablante pronuncia una palabra hasta que el oyente realmente escucha la palabra, esto provoca que los participantes perciban intervalos o pausas irregulares dentro de la conversación, lo que incita a verificar constantemente que la conexión siga operativa, mediante por ejemplo “¿aló, me escucha?”. Este tipo de frases pueden generar confusiones a lo largo del flujo de comunicación, pudiendo sacar al asistente virtual de contexto y en consecuencia generar un problema de entendimiento que baje considerablemente la percepción de calidad de la atención terminando en una transferencia a un ejecutivo.

Según Percy (2010), para una comunicación de buena calidad, se debe tener una latencia inferior a los 200 milisegundos, sin embargo, por lo observado en el MVP la percepción del usuario de la calidad de la comunicación se deteriora a medida que la latencia supera los 2 segundos, donde en promedio el cliente espera hasta 2,5 segundos antes de verificar que esté siendo realmente escuchado, por lo que una restricción de operación para

el canal telefónico consiste en mantener la latencia en el mínimo posible, llegando a un límite máximo de 2,5 segundos.

La latencia viene generada principalmente por dos fuentes, la red telefónica y la conexión a internet de la organización, es en esta última en la que se pueden concentrar los esfuerzos para mantener al mínimo los niveles de latencia, asegurando el ancho de banda adecuado para la transmisión de datos y paquetes de voz.

11.7.2 CONECTIVIDAD

Dado que la solución es soportada en un ambiente *cloud*, la restricción técnica más importante a analizar es la conexión a internet para poder acceder a los servicios. Al ser necesaria una conexión permanente y de alta calidad, este tipo de conexiones se realizan mediante la contratación de un enlace dedicado.

“Un enlace dedicado consiste en un servicio de enlace de internet para cada cliente, que no se comparte con otros usuarios y que opera en forma continua. Técnicamente, es una conexión por fibra óptica directa desde cada empresa, hasta la central del proveedor o hasta un *backbone* de internet. Esto garantiza un enlace más confiable y en muchos casos simétrico, es decir, misma velocidad de subida y de bajada (...) al ser éste un servicio principalmente orientado a empresas o instituciones, los diferentes proveedores se preocupan de otorgar un servicio mucho más robusto en comparación a un servicio convencional de internet de banda ancha, por lo que en general se puede contar con un soporte personalizado con personal capacitado, una red más segura, monitoreo de tráfico en tiempo real y rápida restauración del servicio frente a eventuales caídas.” (Pronto Comunicaciones Ltda., 2017)

El parámetro clave para la contratación de un enlace dedicado es el ancho de banda requerido, el cual corresponde a la medida de datos y recursos de comunicación disponible o consumida en una red de comunicación, indicada generalmente en megabits por segundo (Mbps) (Quiñonez, 2016). En otras palabras, el ancho de banda es la cantidad de información o de datos que se puede enviar a través de una conexión de red en un período de tiempo dado.

En la operación del canal telefónico del *Contact Center Cognitivo*, el ancho de banda se dimensiona en función del tamaño de los datos a enviar, considerando que se deberá enviar además de los paquetes de voz de la llamada, archivos de texto con información relevante de esta (transcripción del mensaje, intención comunicativa detectada, número de quien realiza la llamada, etc.) y la grabación completa de la llamada para ser almacenada en la base de datos de reportería. De esta forma según las estimaciones realizadas por Sixbell, cada llamada simultánea requiere de 200 kbps, luego considerando las 30 llamadas concurrentes, se llega a un ancho de banda de 6 Mbps, el cual garantiza una latencia máxima de 2,5 segundos en la comunicación.

12 FACTIBILIDAD ECONÓMICA

A través de este estudio, se permite identificar si los costos de las componentes detalladas en el estudio técnico se justifican con los beneficios que se pueden obtener con la solución propuesta. Así, en la primera parte de este capítulo se especifican los costos y beneficios que implica la implementación del *Contact Center Cognitivo*, finalizando con el cálculo de indicadores como VAN, TIR y ROI, para evaluar la rentabilidad asociada.

12.1 COSTOS DE LA SOLUCIÓN

Para realizar el análisis de los costos se utiliza la metodología TCO (Costo Total de Propiedad o *Total Cost of Ownership*), esta corresponde a un modelo que ayuda a las empresas a analizar todos los costos relacionados con la adquisición, desarrollo y uso de componentes de Tecnologías de Información (IT) a lo largo de su ciclo de vida

“El TCO proporciona un informe de costos que se incurren por adquisición, mantenimiento, configuración, contratación, administración, entre otros. Siendo una de las herramientas usadas como medidor financiero para la toma de decisiones en proyectos informáticos. Clasificado como un método para calcular los distintos tipos de costos, apoyando tanto a usuarios como a gerentes de organizaciones.” (Mendoza, 2009)

El Modelo de TCO utilizado en el desarrollo de este estudio corresponde al propuesto por The Gartner Group, de su clasificación se consideran relevantes para este trabajo los siguientes costos:

- **Hardware:** Inversiones en activos fijos y cuotas de alquiler con opción de compra.
Puede incluir las estaciones de trabajo, redes, teléfonos y otras infraestructuras.
- **Software:** Para este trabajo, corresponde al pago por uso de los servicios y aplicaciones requeridas para el funcionamiento de la solución.
- **Administración de sistemas:** Gastos por personal propio y pago por servicios externos para la administración de la infraestructura TI.
- **Soporte:** Gastos por personal propio y pago por servicios externos para dar soporte a los usuarios.
- **Desarrollo:** Gastos de mano de obra para el diseño, consultoría, pruebas, documentación y mantenimiento.
- **Comunicaciones:** Gasto por arriendo de líneas, servicios de acceso remoto, Web y enlaces de conexión.

12.1.1 HARDWARE

En este ítem se contempla el costo asociado a la adquisición de un servidor para la virtualización en 4 máquinas que cumplen con los requerimientos detallados en el estudio técnico para la instalación del TMP, Sigas, IVG y TTS de Nuance.

El servidor escogido corresponde a un HP ProLiant BL460c Gen9 Server Blade con las características indicadas en la **Tabla 23**. Dado que dicho servidor no cuenta con disco duro incluido se debe adquirir uno con una capacidad de 1,8 TB. Los costos referenciales asociados se detallan en la **Tabla 24**.

Tabla 23.
Especificaciones del servidor seleccionado.

Procesador	Intel® Xeon® E5-2670 v3 (12 core, 2.3 GHz, 30MB, 120W)
Número de procesadores	2
Cantidad de núcleos disponibles	20
RAM	128GB (4x32GB)
Disco duro	No incluido
Garantía	3 años por falla de partes y piezas

Tabla 24.
Valor referencial en USD del servidor y disco duro seleccionados.

Producto	Costo (USD)
HP ProLiant BL460c Gen9 Server Blade	\$ 7.209
HP 1.8TB SAS 12G Enterprise 10K SFF (2.5in) SC 3yr WTY 512e Digitally Signed Firmware HDD	\$1.145

Fuente: Hewlett Packard Enterprise¹¹

Adicionalmente para el funcionamiento del servidor se debe contar con las licencias del sistema operativo, Linux Red hat 6.9 y Windows Server 2012 para la operación de las máquinas virtuales, y con la licencia de la plataforma VMware vSphere para la virtualización. Los costos de lo anterior se detallan en la **Tabla 25**, estos deberán ser cancelados una sola vez.

Tabla 25.
Valor referencial en USD de licencias a utilizar

Producto	Costo (USD)
VMware vSphere Essentials Kit	\$ 560
Windows Server 2012 Standar	\$ 824
Linux Red hat 6.9	Gratuito

¹¹ Cotización disponible en <https://buy.hpe.com/ocspunchout?model=813192-B21&country=us&locale=en> (Configuration id: 1114077879-01)

De manera adicional, Sixbell cobra una tarifa asociada a la instalación y puesta en marcha de sus aplicaciones, este costo corresponde a 3.000 USD, que deberá ser cancelado una sola vez al comienzo del proyecto.

12.1.2 SERVICIOS Y APLICACIONES

En este ítem se detallan los costos de los componentes asociados a la operación del *Contact Center Cognitivo* correspondiente a los servicios de IBM Cloud y de las aplicaciones desarrolladas por Sixbell.

12.1.2.1 SERVICIOS DE IBM CLOUD

Los servicios de IBM Cloud tienen una estructura de costo en función del uso de cada uno de ellos. En la **Tabla 26** se muestran los valores correspondientes al uso de cada uno de los servicios de IBM Cloud a utilizar para cada año en el horizonte de evaluación, considerando el tamaño del proyecto definido en la sección 11.1, y los costos estipulados en el estudio técnico.

Tabla 26.
Costo anual en USD por el uso de los servicios de IBM Cloud.

Servicio	Año 1	Año 2	Año 3
Watson Assistant	\$ 16.665	\$ 19.577	\$ 22.242
Speech to Text	\$ 28.005	\$ 35.021	\$ 38.592
IBM Voice Gateway	\$ 41.887	\$50.306	\$ 54.591
IBM Cloud Functions	\$ 128	\$ 128	\$ 128
Cloudant	\$ 1.124	\$ 1.145	\$ 1.159
IBM Cloud Object Storage	\$ 112	\$ 178	\$ 221
SDK for Node.js	\$ 1.724	\$ 1.724	\$ 1.724
Total IBM Cloud	\$ 89.654	\$ 108.078	\$ 118.657

12.1.2.2 APLICACIONES DE SIXBELL

Las aplicaciones utilizadas en la solución que son otorgadas por Sixbell corresponden al TMP, Sigas y Text to Speech, el uso de estas tiene un costo asociado a la cantidad de minutos procesados, de esta manera la propuesta de Sixbell se realiza en un precio variable por minuto que contempla el uso de las tres aplicaciones mencionadas en conjunto, así el costo por minuto corresponde a 0,06 USD. En función a este valor, en la **Tabla 27** se presenta el costo anual asociado al uso de la tecnología de Sixbell.

Tabla 27.
Costo anual en USD por el uso de las aplicaciones de Sixbell.

Servicio	Año 1	Año 2	Año 3
Text to Speech			
Telecom Media Portal	\$ 29.943	\$ 37.264	\$ 40.990
Signaling Application Server			

Estos valores contemplan la capacidad de atender 30 llamadas concurrentes, y no considera el costo del hardware en el que se implementa cada aplicación. Además, tiene como restricción la contratación de las aplicaciones por un periodo mínimo de 24 meses.

12.1.3 ADMINISTRACIÓN DE SISTEMAS DE TI

En la administración de los sistemas de TI se requiere de un rol con la responsabilidad de implementar, configurar, mantener, monitorizar, documentar y asegurar el correcto funcionamiento de un sistema informático, o algún aspecto de éste (Miranda, 2005). Así las actividades en la administración de sistemas corresponden al monitoreo proactivo de los servidores físicos y virtuales, resolución de problemáticas del sistema operativo, asistencia técnica de las plataformas, seguimiento de alertas, aplicación de actualizaciones, verificación

y solución de problemas de seguridad y escalamiento de incidentes a otros grupos de soporte cuando corresponda.

La administración de sistemas de TI también considera el monitoreo de enlaces y comunicaciones, verificando los tiempos de respuesta y dando aviso al proveedor de red en caso de que exista algún inconveniente.

Para el cumplimiento de estas tareas se considera necesario un Operador de Sistemas de TI, cuyo sueldo bruto promedio corresponde a 1691 USD mensuales¹².

12.1.4 SOPORTE

Una vez que la solución se encuentra operando, se hace necesario contar con un servicio de soporte para las distintas aplicaciones, de manera de hacer frente a modificaciones que deban ser realizadas en algunos de los servicios, correcciones de errores o bajas del sistema. En la **Tabla 28** se describen los distintos niveles de soporte de TI.

Tabla 28.
Niveles de soporte de TI

<i>Nivel de soporte</i>	<i>Descripción</i>
T1	El usuario inicia la búsqueda de la solución al problema. Comunica síntomas al técnico para que éste pueda identificar el error y ofrecer acciones a ser aplicadas para solucionarlo.
T2	Se llevan a cabo tareas más complejas como recuperaciones de datos, reinicios de fábrica o cambio de equipos y configuración de redes inalámbricas.
T3	Técnicos de soporte TI que están altamente capacitados para gestionar los problemas que no pudieron ser solucionados en los niveles 1 y 2. Además, tienen capacidad para tomar la decisión de reemplazar equipos de manera definitiva.

¹² Consultado en tusalarario.org (Chile) para un operador con 3 años de experiencia.

Para la implementación del *Contact Center Cognitivo*, se estipula que la empresa de retail se hace cargo del nivel T1 de soporte, siendo esta la responsable de hacer el sondeo inicial cuando exista algún inconveniente o baja en el sistema. Para los servicios *Cloud*, IBM es responsable de entregar soporte tanto a nivel T2 como T3, mediante la apertura de ticket de soporte¹³, donde la empresa de retail deberá indicar el grado de severidad del requerimiento siguiendo lo indicado en la **Tabla 29**. Finalmente, para el soporte de las aplicaciones de Sixbell, son ellos quienes cubren los niveles T2 y T3, este último con posible derivación si se concluye que la causa raíz es hardware. Los tiempos de respuesta del soporte de Sixbell se presentan en la **Tabla 30**.

Tabla 29.
Severidad, tiempo de respuesta y cobertura del servicio de soporte de IBM Cloud.

<i>Severidad</i>	<i>Definición de severidad</i>	<i>Tiempo de respuesta</i>	<i>Cobertura</i>
1	Impacto empresarial crítico/servicio inactivo: Indica incapacidad para acceder a los servicios lo que resulta en un impacto crítico en las operaciones. Se registra el caso dentro de las 24 horas de haberse detectado la falla.	Dentro de 1 hora.	24 x 7
2	Impacto empresarial significativo: un servicio o una función del servicio está severamente restringido en su uso, en un periodo crítico del negocio.	Dentro de 2 horas hábiles.	Lunes a viernes (horario hábil).
3	Impacto empresarial menor: se presenta una falla en un servicio o funcionalidad, pero se puede utilizar y el problema no representa un impacto crítico en las operaciones.	Dentro de 4 horas hábiles.	Lunes a viernes (horario hábil).

¹³ Correspondiente al levantamiento de una solicitud de soporte en la página web. Esta solicitud es revisada por el área de IBM responsable a la cuenta del cliente para entregar una respuesta en los plazos comprometidos.

4	Impacto empresarial mínimo: corresponde a una consulta o solicitud no técnica.	Dentro de 1 día hábil.	Lunes a viernes (horario hábil).
---	---	------------------------	----------------------------------

Fuente: Elaboración Propia a partir de IBM Cloud and SaaS support (2018)

Tabla 30.

Severidad, tiempos de atención y solución de incidentes del servicio de soporte de Sixbell.

<i>Severidad</i>	<i>Atención de incidente</i>	<i>Solución de incidente</i>
Menor	Hasta 2 horas en horario 24x7	Hasta 18 horas en horario hábil
Mayor	Hasta 2 horas en horario 24x7	Hasta 9 horas en horario hábil

Fuente: Elaboración Propia a partir de información entregada por Sixbell.

Además, Sixbell realiza un monitoreo a sus aplicaciones, de manera de detectar de forma inmediata la caída de los servicios o condiciones anómalas que podrían convertirse en fallas del sistema. Este servicio tiene un costo de 8 UF o 326 USD mensuales, y se considera como el único costo asociado dado que los servicios de soporte previamente descritos se encuentran incluidos dentro del precio de las aplicaciones.

En lo que refiere a la infraestructura, se estima un costo anual del 10% sobre la inversión inicial por concepto de imprevistos.

En base a lo anterior, el costo anual por conceptos de soporte corresponde a 4.745 USD cada año en el horizonte de evaluación.

12.1.5 DESARROLLO

En este ítem se estudian los costos asociados a la contratación de la consultoría de IBM necesaria para diseñar y poner en marcha la solución. De la misma forma se detallan los valores asociados al personal interno de la empresa de retail que deberá estar disponible

durante el proceso de implementación de la solución y durante todo el tiempo de funcionamiento de esta misma, de manera de tener un control sobre el desempeño, y la capacidad de generar modificaciones cuando sea necesario.

En la **Tabla 31** se detallan los costos anuales en personal durante todo el horizonte de evaluación del proyecto, considerando 6 meses de consultoría de IBM y de dedicación de un equipo de trabajo de la empresa de retail, de este último equipo se considera que será necesario contar con un rol de Usuario Cognitivo con dedicación de 80 horas mensuales de manera permanente, para monitoreo y control; y 160 horas dos veces al año, con el objetivo de realizar un mantenimiento previo a los puntos de mayor demanda anual (Junio y Diciembre).

Tabla 31.
Costo anual en USD del personal requerido.

Personal requerido	0-6 meses	6-12 meses	Año 2	Año 3
Consultoría IBM (6 meses)	\$ 198.482	-	-	-
Equipo empresa retail ¹⁴	\$ 23.991	\$ 6.531	\$ 11.428	\$ 11.428

12.1.6 COMUNICACIONES

Este ítem corresponde al desembolso realizado por concepto de conexión del enlace dedicado de 10 Mbps, correspondiente al mínimo ofrecido. La compañía Level 3, ofrece esta velocidad a un precio de 1.510 USD mensuales, por un periodo de tres años. Este corresponde a un costo recurrente fijo en el tiempo, con un valor anual de 18.000 USD.

¹⁴ Equipo compuesto por los roles detallados en la sección de Requerimientos de personal del estudio técnico, durante los primeros 6 meses y luego continua sólo con el rol de Usuario Cognitivo durante todo el horizonte de evaluación.

12.1.7 COSTO TOTAL DE PROPIEDAD

Una vez detallados todos los puntos atingentes al costeo del proyecto, se tiene que el costo total de propiedad, llevado a valor presente, corresponde a 744.267 USD. Donde el costo por concepto de servicios y aplicaciones representa el 51,4% del total, seguido por el costo de desarrollo que corresponde a un 31,1%, esto se muestra en la **Figura 18**. El detalle mensual para cada año en el horizonte de evaluación se muestra en la **Figura 19**.

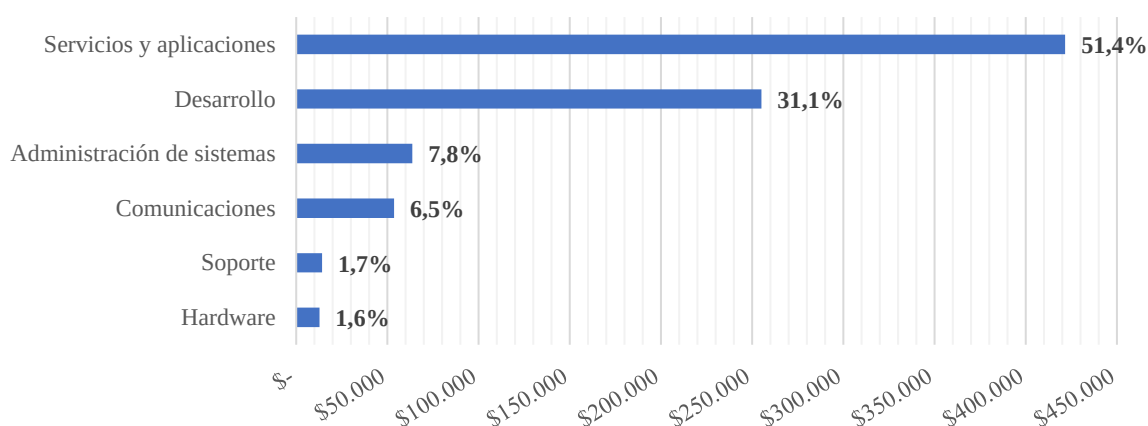


Figura 18. Desglose del costo total de propiedad del proyecto en USD llevado a valor presente.
Fuente: Elaboración propia

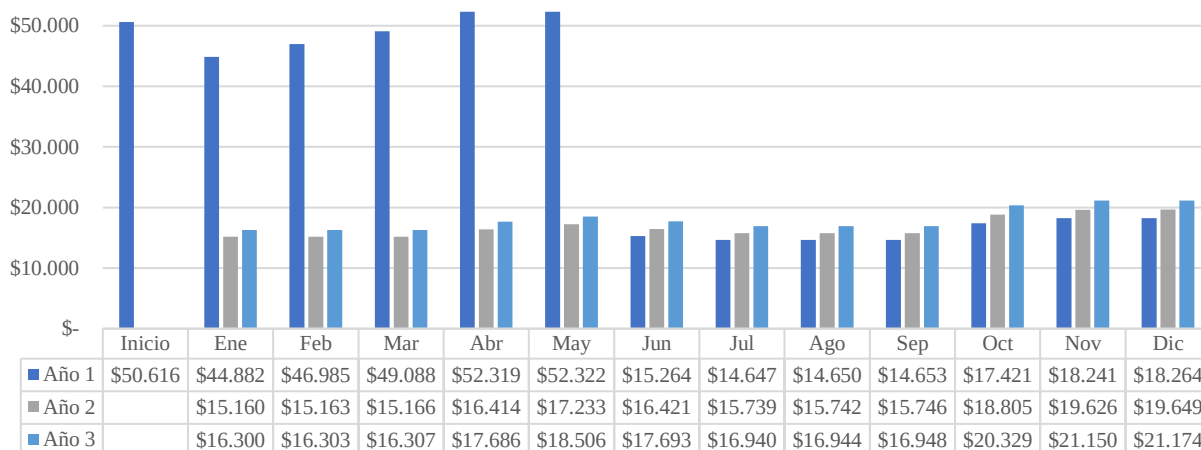


Figura 19. Costo total de propiedad en USD mensualizado en el horizonte de evaluación.
Fuente: Elaboración propia.

12.2 BENEFICIOS DE LA SOLUCIÓN

Los beneficios obtenidos en la implementación de un proyecto pueden ser de dos tipos, tangibles e intangibles. Los beneficios tangibles son aquellos que están asociados con ahorros monetarios o materiales, y los beneficios intangibles se refieren a todo aquello que agregue valor significativamente, pero que no se puede poner en términos monetarios.

En la implementación del *Contact Center Cognitivo*, el beneficio tangible corresponde a la reducción del tiempo empleado en la atención de clientes, mientras que los beneficios intangibles corresponden al incremento de la disponibilidad del servicio al cliente, mejora en el proceso de auditoría, alineamiento con la estrategia de transformación digital de la empresa, y acceso a información para analítica de la interacción con clientes.

12.2.1 BENEFICIO TANGIBLE

El uso del *Contact Center Cognitivo* tiene asociado beneficios por ahorro en costos de los agentes que son liberados debido a la contención alcanzada por la solución. Este ahorro se calcula en función de las llamadas que son atendidas por el asistente virtual, más en específico, en los minutos que hubiesen sido requeridos si estas fueran atendidas por un ejecutivo. De esta forma, y considerando la cantidad de minutos que está disponible un ejecutivo para atender llamadas durante un mes, se puede obtener la cantidad de ejecutivos que son liberados mensualmente.

En la **Figura 20** se muestra la cantidad de minutos mensuales que serían requeridos en el horizonte de evaluación con y sin el *Contact Center Cognitivo* (CCC), donde la diferencia corresponde a los minutos que son evitados a los ejecutivos producto de la implementación de la solución.

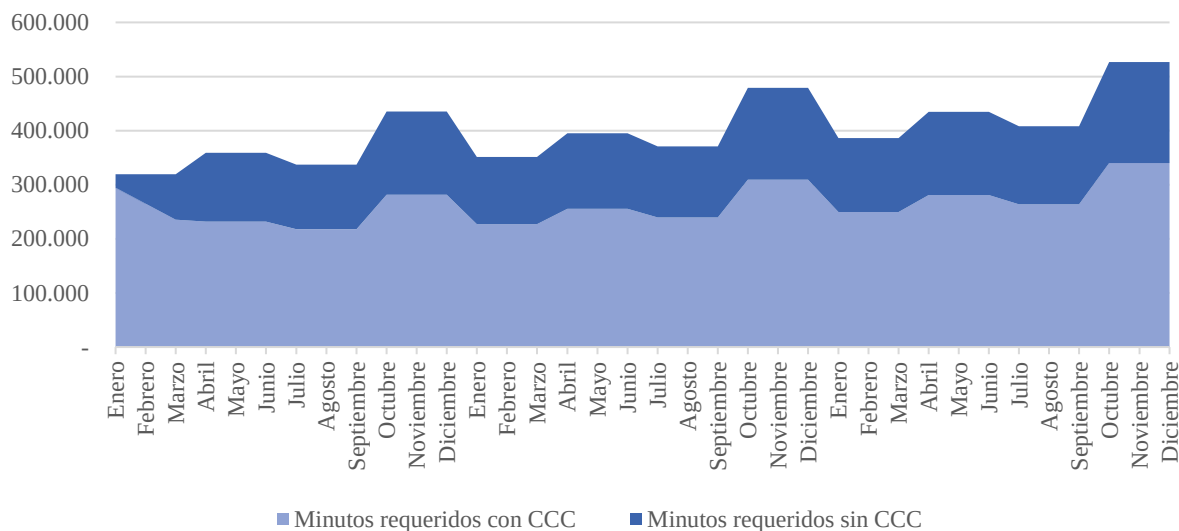


Figura 20. Cantidad de minutos mensuales requeridos de ejecutivos en el horizonte de evaluación, con y sin la implementación del Contact Center Cognitivo. Fuente: Elaboración propia.

La cantidad mensual de ejecutivos necesarios y liberados se muestra en la **Figura 21**, esto se traduce en un beneficio por ahorro en costos, el cual es presentado de manera anual en la **Tabla 32**, considerando que el costo empresa de un ejecutivo corresponde a 1.500 USD mensuales.

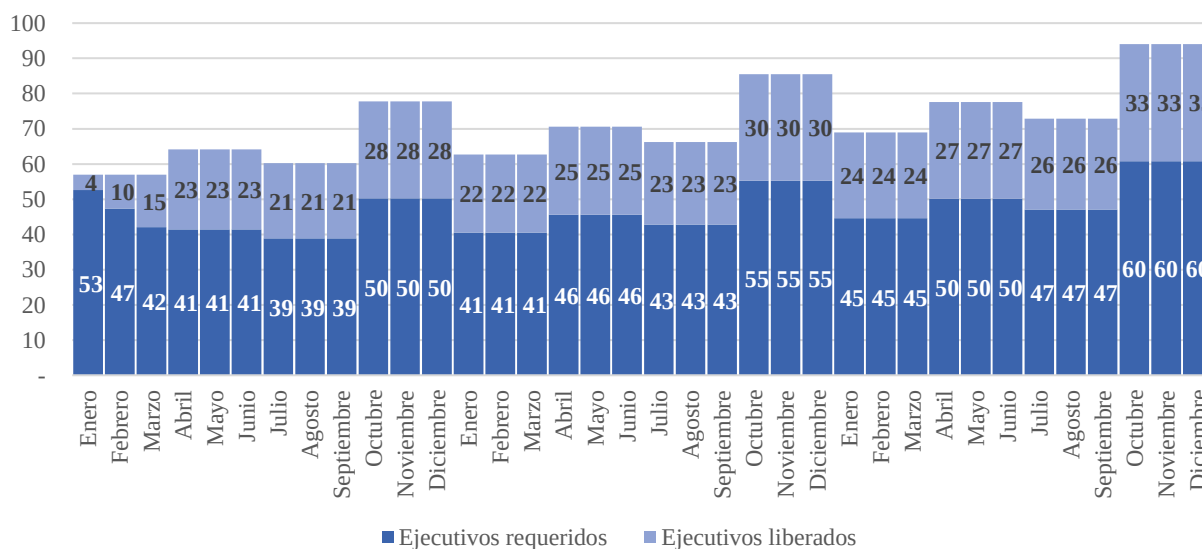


Figura 21. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación. Fuente: Elaboración propia.

Tabla 32.
Beneficios anuales en USD por ejecutivos liberados en el horizonte de evaluación.

	<i>Año 1</i>	<i>Año 2</i>	<i>Año 3</i>
Beneficio por ahorro en costos	\$ 364.650	\$ 453.806	\$499.185

12.2.2 BENEFICIOS INTANGIBLES

La implementación del *Contact Center Cognitivo* tiene una serie de beneficios que no pueden incorporarse directamente al estudio económico debido a la dificultad de llevarlos a una base monetaria. Sin embargo, es importante tenerlos en consideración al momento de tomar la decisión final de llevar o no a cabo el proyecto. A continuación, se detallan los tres beneficios intangibles principales a obtener si se implementa la solución.

Mejora en el proceso de atención a clientes:

Uno de los beneficios de contar con asistentes virtuales corresponde a la disponibilidad 24x7 de atención a clientes, con capacidad para manejar múltiples interacciones de manera simultánea, en este caso tanto vía chat como por teléfono, permitiendo aumentar el nivel de servicio de atención al cliente y a un tiempo de operación menor comparado con el de los ejecutivos. Además, la implementación del *Contact Center Cognitivo* entrega una alternativa de contacto distinta a la telefónica, permitiendo descongestionar el proceso de atención al cliente en el *Contact Center* durante los periodos de alta demanda.

Para el canal web, el asistente virtual está disponible para consultas y para guía en el proceso de compra, de esta forma permite una mayor agilidad en el acceso a la información en su plataforma *e-commerce*. En esta misma línea, la existencia del chat en la página web se asocia con los indicadores mostrados en la **Tabla 33**.

Tabla 33.
Indicadores promedio de e-commerce entre enero y agosto del 2018.

Indicador	Valor
Porcentaje de rebote	3,17%
Duración media de la sesión	7,66 minutos
Tasa de conversión	1,78%

Ocorre rebote cuando en un sitio web se produce una sesión de una sola página, esto es, que el usuario abandona el sitio a los pocos segundos sin haber interactuado. El porcentaje de rebote histórico previo a la implementación del chat en la página web era de 6,55%, por lo que se obtuvo una disminución del 51,58% del rebote implicando que los usuarios están navegando más a través del sitio.

La duración de la sesión en la página web corresponde al tiempo de interacción del usuario con la página web, el promedio histórico de duración de una sesión en el sitio previo a la implementación era de 4,96 minutos, obteniéndose un aumento del 54,47%.

Finalmente, la tasa de conversión en *e-commerce* se conoce como el porcentaje de usuarios que realiza una acción específica, en este caso comprar. La tasa de conversión promedio del sector en Chile corresponde a 2,2% (Cámara de Comercio de Santiago, 2017), y en específico la de la empresa, previo a la implementación, era de 0,92% lo que significa que las personas que efectivamente compran en el sitio web aumentaron en un 93,03%.

Estrategia de transformación digital:

El sector retail chileno se ha volcado con gran interés hacia el comercio electrónico, debido a las grandes ventajas que éste representa. Este interés queda de manifiesto con el gran evento de comercio electrónico *cyberday*, que se realiza generalmente en mayo de cada año (Diario TI, 2017). En particular la empresa de retail estudiada manifiesta ser una empresa a la vanguardia en materia digital y de comercio electrónico, siendo uno de los principales portales de *e-Commerce* de Chile y Perú, premiado por el Instituto Latinoamericano de Comercio y por la Cámara de Comercio de Lima el año 2017; y por la Cámara de Comercio de Santiago en los años 2015 y 2016.

Estos hitos reflejan el compromiso que la compañía ha tomado con la transformación digital donde, según lo indicado en su memoria anual, los canales digitales seguirán creciendo, desarrollándose y adaptándose a las necesidades de los clientes, incorporando en su *e-Commerce* una gama de productos más amplia y una mejor funcionalidad de las herramientas tecnológicas que lo soportan, debido al importante rol que juega este canal en la estrategia de omnicanalidad y en el desarrollo de una gestión más eficiente, ambos puntos correspondientes a los principales objetivos de su plan estratégico.

Es por esto que la incorporación del *Contact Center Cognitivo* en el servicio de atención al cliente corresponde a un beneficio en términos de alineamiento con su estrategia empresarial hacia la transformación digital y la búsqueda de la omnicanalidad, teniendo la opción de ser pioneros en el sector en términos de contactos telefónicos atendidos por un asistente virtual.

Analítica de las interacciones:

La implementación del *Contact Center Cognitivo* trae consigo la creación de una gran base de datos de las transcripciones de las llamadas atendidas por el asistente virtual, esto da la posibilidad de generar analítica de las interacciones para el estudio de las preferencias de los clientes, sus consultas más frecuentes, la forma de comunicarse, entre otras cosas. Según (Ransbotham & Kiron, 2018) las herramientas de analítica están impulsando el *Customer Engagement*, más aun, el análisis de datos corresponde a la principal fuente de ventaja competitiva para las empresas.

Si bien con el *Contact Center Cognitivo* se forma la base de datos con transcripciones e intenciones comunicativas detectadas en cada contacto, será labor de la empresa disponer de recursos y herramientas para hacer gestión sobre estos datos y así aprovechar el beneficio latente.

12.3 EVALUACIÓN DE LA RENTABILIDAD

Una vez obtenidos los costos y beneficios de la implementación, es necesario hacer una evaluación económica del proyecto en dónde se podrá ver y comprobar si el proyecto es factible monetariamente a través del cálculo de algunos índices como: VAN, TIR, ROI y PRI.

12.3.1 VALOR ACTUAL NETO (VAN)

El VAN consiste en actualizar a un valor presente los flujos futuros que va a generar el proyecto, descontados a la tasa definida por la empresa (9,2% anual), para permitir la comparación con montos actuales de dinero. Este indicador se calcula como:

$$VAN = \sum_{t=1}^n \frac{R_t}{(1+i)^n} - I_0$$

Donde R_t corresponde a los flujos en el periodo t , i la tasa de descuento esperada, e I_0 la inversión inicial.

Así, los criterios de decisión del resultado del proyecto dependen de si el VAN es mayor o menor a cero. Un valor positivo significa que los flujos futuros llevados a valor presente corresponden a ganancias, y por tanto es rentable su ejecución, mientras que uno negativo indica pérdidas y en consecuencia no es rentable llevar a cabo el proyecto.

12.3.2 TASA INTERNA DE RETORNO (TIR)

La TIR corresponde a la tasa de descuento que iguala el VAN a 0, y se utiliza como una medida de valor de rendimiento interno de la empresa expresado en porcentaje, y comparable con una tasa de interés. Para el análisis de viabilidad del proyecto, la TIR debe ser comparada con la tasa de descuento de la empresa, que representa el costo de oportunidad de la inversión, así si la TIR es mayor a la tasa de descuento significa que el proyecto es aceptable.

12.3.3 PERIODO DE RECUPERACIÓN DE LA INVERSIÓN (PRI)

Corresponde al período que tarda en recuperarse la inversión inicial a través de los flujos generados por el proyecto, luego se considera que se ha recuperado la inversión en el momento en que los flujos acumulados obtenidos son positivos. Este indicador se calcula como:

$$PRI = t + \frac{|F_t|}{F_{t+1}}$$

Donde t corresponde periodo en que se obtiene el último flujo acumulado negativo, y F_t el flujo acumulado en el periodo t .

12.3.4 RETORNO SOBRE LA INVERSIÓN (ROI)

El ROI es uno de los indicadores más utilizados al momento de evaluar proyectos informáticos o de tecnología, y corresponde a la razón entre el beneficio o utilidad obtenida y la inversión realizada, para generar una estimación de cuán favorable será la inversión.

Normalmente el cálculo utiliza el valor de la inversión inicial, sin embargo, para proyectos de tecnología que operan principalmente en modalidad *cloud*, la inversión no está representada únicamente en gastos iniciales, sino que también tiene componentes de costo a largo plazo de operación en la nube. Por lo que, para el cálculo del ROI, se considera la siguiente fórmula (ISACA, 2012):

$$ROI = \frac{\text{Beneficios} - TCO}{TCO}$$

El valor obtenido corresponde al número de veces en que se recupera lo invertido en el proyecto, así si el ROI es negativo, implica que el proyecto resulta en pérdidas; si es igual a cero, significa que no se ha perdido ni generado ganancias; y por último, si el valor es positivo se tienen ganancias equivalentes a un determinado porcentaje de lo invertido.

12.3.5 ANÁLISIS DE LOS INDICADORES

En la **Tabla 34** se presentan los valores obtenidos de los indicadores previamente descritos, considerando la tasa de descuento de la empresa de retail correspondiente a 9,2% anual.

Tabla 34.
Resumen de los indicadores de rentabilidad obtenidos.

<i>Indicador</i>	<i>Valor</i>
VAN	395.277 USD
TIR	7,89 %
PRI	14,7 meses
ROI	53%

Por los valores obtenidos se puede decir que el proyecto producirá ganancias por sobre la rentabilidad exigida, dado que tiene un VAN mayor a cero, con una tasa interna de retorno mayor a la mínima esperada. La inversión será recuperada en un periodo de 14,7 meses y se obtendrá un retorno del 53% del valor del costo total de propiedad, por lo que se concluye que es un proyecto factible económicamente.

12.4 ANÁLISIS DE SENSIBILIDAD

Durante la evaluación de un proyecto existe un grado de incertidumbre asociado a la definición parámetros necesarios para su construcción, dado que estos obedecen a una lógica de proyección sujeta a supuestos. Por esto resulta de utilidad poder flexibilizar estos parámetros para ver las potenciales variaciones que existentes en distintos escenarios.

El análisis de sensibilidad corresponde a un método que permite medir cuan sensible es la evaluación a variaciones de uno o más parámetros de interés, calculando su impacto principalmente en el valor actual neto.

Para este trabajo, los parámetros a sensibilizar corresponden a la tasa de contención objetivo del proyecto, la cantidad y porcentaje de crecimiento anual de las llamadas, y el TMO de los ejecutivos. Estos son escogidos debido a su influencia indirecta en los costos de los servicios y aplicaciones que representan el 51,4% del costo total de propiedad. No se

estudian variaciones en los costos por uso de los servicios debido a la estabilidad de los precios de IBM Cloud a lo largo del tiempo y a que, al momento de acordar la ejecución del proyecto, se compromete a la mantención de un precio fijo por la duración de éste.

12.5 MODELO UNIDIMENSIONAL DE SENSIBILIZACIÓN

Este modelo determina hasta donde puede modificarse el valor de una variable para que el proyecto siga siendo rentable, manteniendo las demás variables constantes. De esta forma se realiza la variación de manera independiente de cada uno de los parámetros previamente descritos, encontrando el valor al cual el proyecto presenta un VAN igual o menor a cero.

12.5.1 TASA DE CONTENCIÓN OBJETIVO

La tasa de contención objetivo corresponde a la meta, en porcentaje, de atención de llamadas mediante el asistente virtual a alcanzar, según el plan de implementación, en el cuarto mes de ejecución del proyecto. El caso base del proyecto corresponde al indicado en la **Tabla 6**, y la sensibilización se realiza teniendo como punto de partida el escenario en que la contención objetivo es la alcanzada durante el MVP, luego se realiza un crecimiento del 5% para cada una de las opciones hasta alcanzarse la contención objetivo del caso base para cada opción. Pasado el caso base, se aumenta en un 5% cada opción hasta llegar al límite indicado en la **Tabla 35**. De manera adicional se encuentra la contención objetivo que hace que el VAN sea igual a cero, esto mediante el uso del complemento Solver Excel, teniendo como restricción una contención mayor a la alcanzada durante el MVP, estos datos se presentan en la **Tabla 36**.

Tabla 35.
Límite superior de contención objetivo a alcanzar para el análisis de sensibilidad.

Opción IVR	Contención objetivo
Horarios y ubicación	99 %
Despachos	75 %
Cambios y Garantía	45 %
Otras consultas	40 %
Contención promedio ponderado¹⁵	50,37%

Tabla 36.
Valor de la contención objetivo que genera un VAN igual a 0.

Opción IVR	Contención objetivo
Horarios y ubicación	80 %
Despachos	26 %
Cambios y Garantía	0 %
Otras consultas	17,5 %
Contención promedio ponderado	18,43%

En la **Figura 22** se muestra la curva del VAN alcanzado para los distintos promedios ponderados de contención analizados. A partir de esto se puede decir que, comprometida sobre una contención promedio de 18,43% es posible considerar la implementación del *Contact Center Cognitivo*.

¹⁵ Promedio ponderado en función del porcentaje del total de llamadas que representa cada opción del IVR.

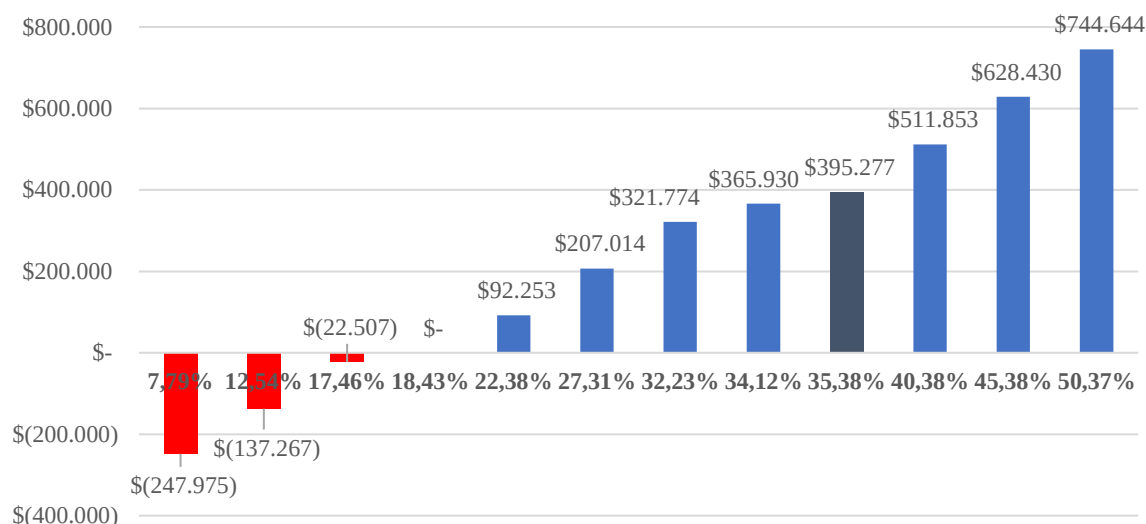


Figura 22. Valor actual neto (USD) en función del porcentaje promedio ponderado de contención objetivo.
Fuente: Elaboración propia.

12.5.2 CANTIDAD DE LLAMADAS

Uno de los parámetros que define el tamaño de un *Contact Center* corresponde a la cantidad de llamadas recibidas anualmente. Al sensibilizar este valor, lo que se espera es encontrar el tamaño mínimo del *Contact Center*, en términos de llamadas, tal que sea rentable la implementación de un asistente virtual.

Así mediante el uso del complemento Solver Excel, se obtiene la cantidad de llamadas, en el año previo a la implementación de la solución y su proyección en el horizonte de evaluación, que genera un VAN igual a cero, esto se muestra en la **Tabla 37**.

Tabla 37.
Proyección de llamadas a recibir en el *Contact Center* de la empresa, que genera un VAN igual a 0.

Opción IVR	Año 0	Año 1	Año 2	Año 3
Horarios y ubicación	12.924	14.217	15.638	17.202
Despachos	82.982	91.280	100.408	110.449
Cambios y Garantía	90.756	99.831	109.814	120.796
Otras consultas	120.215	132.237	145.461	160.007
Total	306.877	337.565	371.322	408.454

Esto significa que para que la implementación no sea rentable, las llamadas que recibe actualmente el *Contact Center* de la empresa deberían ser menor a un 34,9% de las establecidas al comienzo de este trabajo.

12.5.3 TASA DE CRECIMIENTO DE LA DEMANDA ANUAL POR SERVICIO

Para estimar la demanda anual por servicio, fue necesario hacer una proyección en base a los porcentajes de crecimiento anual estimados por la empresa. En este punto se hace variar ese porcentaje con el objetivo de determinar su impacto en el valor actual neto. Los resultados se muestran en la **Figura 23**, de donde se tiene que el proyecto deja de ser rentable si las llamadas decrecen anualmente en un 30%.

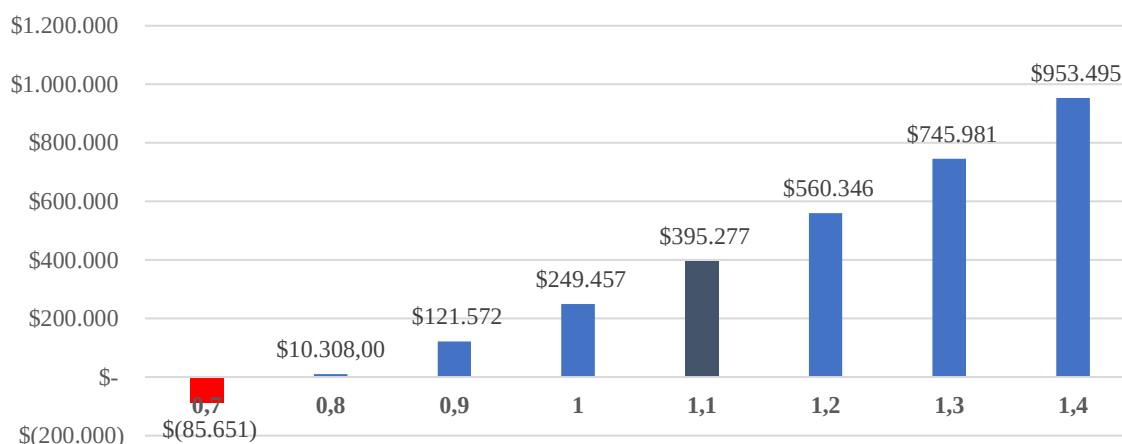


Figura 23. Valor actual neto (USD) en función del factor de crecimiento de la demanda por servicio.
Fuente: Elaboración propia.

12.5.4 TMO DE EJECUTIVOS Y ASISTENTE VIRTUAL

El tiempo medio de operación (TMO) corresponde a un indicador de desempeño que impacta directamente en los beneficios esperados y en los costos de operación de la solución, por lo que se realiza una sensibilización tanto para el TMO de los ejecutivos como el del asistente virtual. En la **Figura 24**, **Error! Reference source not found.** se muestran los

resultados obtenidos para variaciones del TMO de los ejecutivos entre 2,5 y 6 minutos, obteniéndose un VAN igual a cero cuando el TMO es de 2,93 minutos.

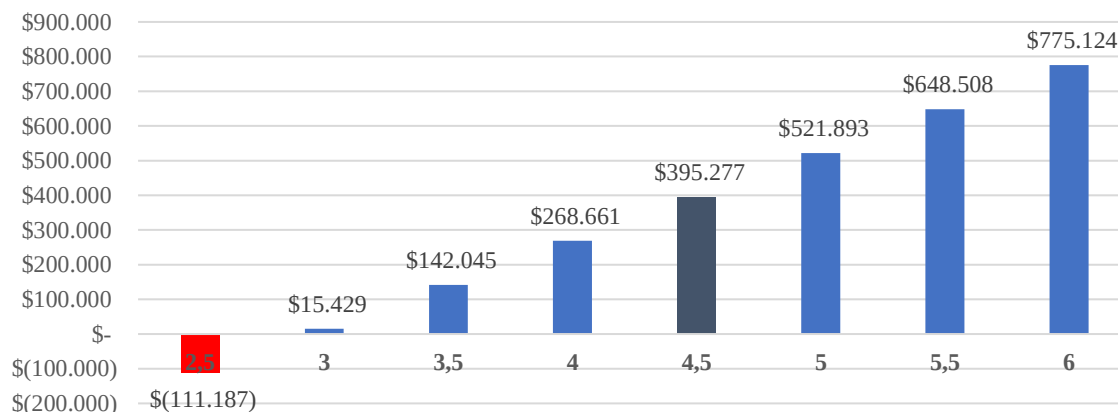


Figura 24. Valor actual neto (USD) en función del TMO (min) del ejecutivo.
Fuente: Elaboración propia.

Para el TMO del asistente virtual se realiza el mismo procedimiento de variación para valores entre 1 y 5 minutos, en este caso se tiene un VAN igual a cero cuando el TMO del asistente virtual es igual a 3,89 minutos, esto se muestra en la **Figura 25**.

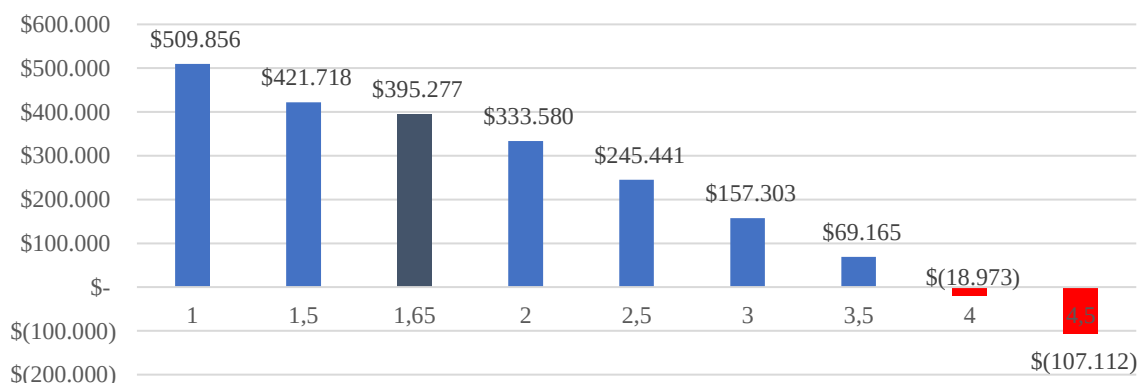


Figura 25. Valor actual neto (USD) en función del TMO (min) del ejecutivo.
Fuente: Elaboración propia.

Además, en un escenario de variación del TMO del asistente virtual y los ejecutivos en manera conjunta, se tiene que el VAN es igual a cero cuando el TMO es de 3,03 y 1,79 minutos respectivamente.

Este análisis permite identificar que el tiempo medio de operación es el parámetro más crítico, ya que, si bien desde una perspectiva de variación independiente, se permite un aumento del TMO del asistente virtual de hasta un 136% y una disminución del TMO de los ejecutivos de hasta un 35%, de manera conjunta basta un aumento del 8% del TMO del asistente virtual con una disminución del 33% del TMO del ejecutivo, para que el proyecto deje de ser rentable, por lo que se recomienda que cualquier mejora a realizar que acelere los procesos de atención, sea aplicada de manera que se refleje también en los tiempos de operación del asistente virtual.

12.5.5 BENEFICIOS POR REDUCCIÓN DE EJECUTIVOS

Para el caso base se tiene como supuesto que la empresa cuenta con la flexibilidad suficiente para adaptarse a las variaciones de requerimientos de personal mes a mes mediante, por ejemplo, trabajadores *part-time*. Dado que los beneficios monetarios dependen únicamente del ahorro en costos por concepto de ejecutivos liberados, resulta interesante sensibilizar respecto a sus políticas en temas de recursos humanos. Así se presentan tres escenarios:

1. **Personal constante a lo largo de todo el horizonte de evaluación:** Esto significa que la empresa no pretende hacer una reducción de los ejecutivos contratados, por lo que la implementación del *Contact Center Cognitivo* tendrá beneficios únicamente en términos de no aumentar la dotación de personal en el horizonte de evaluación. En este caso se tienen beneficios monetarios sólo en los meses en que la cantidad de ejecutivos requeridos es superior a 60. Esto se muestra en la **Figura 26**.

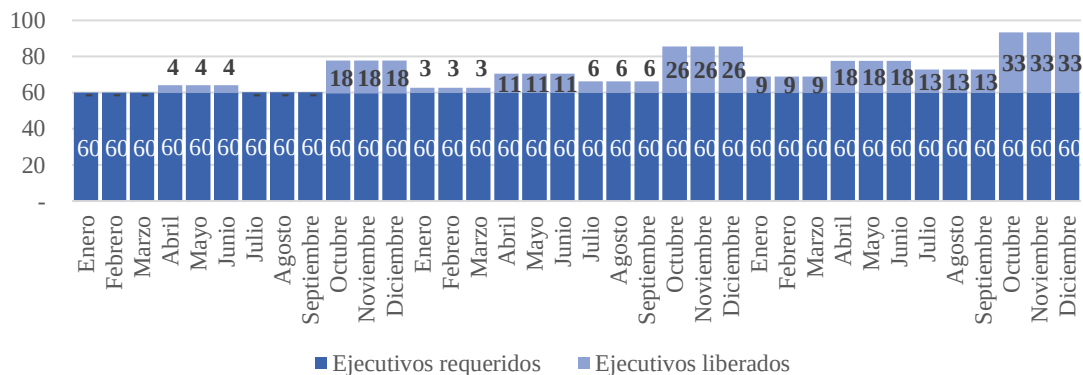


Figura 26. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante. Fuente: Elaboración propia.

Para este escenario se obtiene un VAN de -218.862 USD, por lo que no resulta rentable implementar la solución bajo esta condición.

2. **Dimensionamiento del personal constante para cada año:** Corresponde al escenario en que la empresa dimensiona cuántos ejecutivos requiere para cada año y los mantiene durante todo ese periodo sin variaciones. En este caso se debe seleccionar el máximo valor de ejecutivos que se estima se requerirán cada mes en el periodo de un año y mantenerlo constante, esto se muestra en la **Figura 27**.

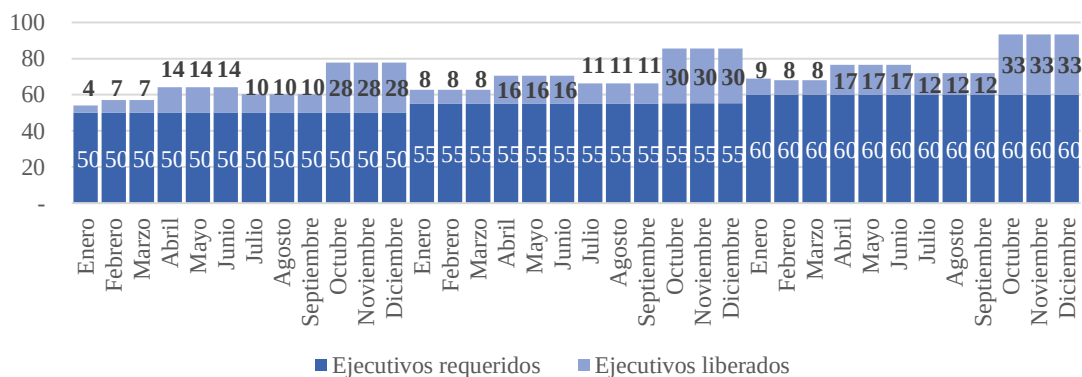


Figura 27. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante cada año. Fuente: Elaboración propia.

En este caso se tiene un VAN igual a 959,43 USD, con un ROI del 0%, lo que corresponde a un punto de indiferencia respecto a la realización del proyecto.

3. **Dimensionamiento del personal constante para cada semestre:** Este escenario mantiene la misma lógica que el anterior, pero por un periodo de tiempo de 6 meses. Esto se muestra en la **Figura 28**.

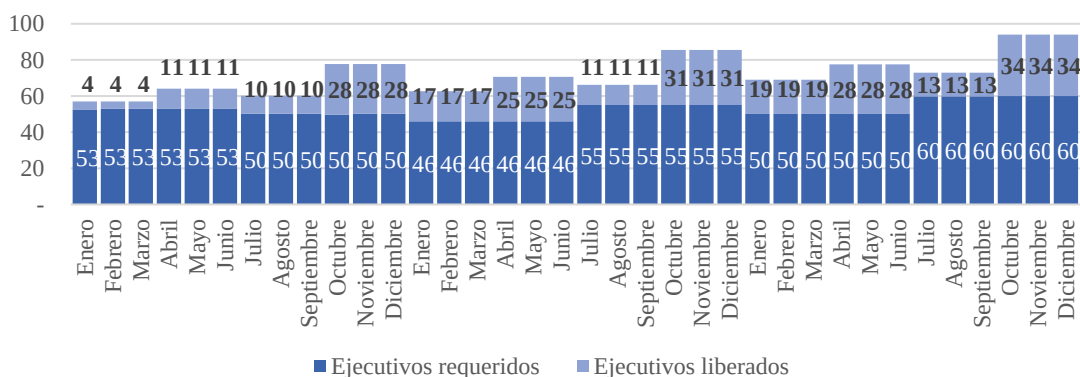


Figura 28. Cantidad de ejecutivos requeridos y liberados por mes en el horizonte de evaluación bajo el escenario de personal constante cada semestre. Fuente: Elaboración propia.

Para este escenario se tiene un VAN igual a 139.954 USD, y un ROI de 19%. Por lo que, si bien esta alternativa es menos rentable que la de la situación base, corresponde a una opción económicamente factible.

12.6 MODELO MULTIDIMENSIONAL DE SENSIBILIZACIÓN

En el modelo multidimensional, se proponen distintos escenarios que consisten en modificar el valor de una o más variables de manera simultánea, obteniendo como resultado la rentabilidad esperada del proyecto.

En general, son escogidos tres tipos de escenarios, el pesimista, más probable y optimista. En el pesimista se modifican las variables a analizar en su situación negativa, en

el más probable se mantiene la situación base, y en el optimista se modifican las variables en su situación positiva. En ambos casos se obtiene como resultado las consecuencias de estas modificaciones reflejados en los montos y/o indicadores financieros que se consideran decisorios al momento de evaluar un proyecto.

La construcción de los escenarios para este análisis se muestra en la **Tabla 38**, donde la tasa de contención objetivo por opción, la cantidad de llamadas en el año previo a la implementación de la solución, la tasa de crecimiento anual de llamadas, y la cantidad de interacciones web, se varían en +/-10%. La tasa de crecimiento anual de interacciones web se varía en +/-5%; el TMO de los ejecutivos en +/- 1 minuto; y para el TMO del asistente virtual, se varía en +/- 0,8 minutos, de manera de llegar en el escenario optimista al mínimo¹⁶ TMO obtenido en el MVP. Además, se contemplan los escenarios de realización de beneficios con las políticas de dimensionamiento de personal anual, semestral, y flexible mensual (caso base).

Tabla 38.
Detalle de los escenarios para la sensibilización.

<i>Variables</i>	<i>Pesimista</i>	<i>Más probable</i>	<i>Optimista</i>
Cantidad de llamadas (año 0)	701.307	879.230	967.153
Cantidad de interacciones web (año 0)	2.044.218	1.858.380	1.672.542
Tasa de crecimiento anual (llamadas)	1,00	1,10	1,20
Tasa de crecimiento anual (web)	1,2	1,15	1,10
TMO ejecutivos (minutos)	3,50	4,50	5,50
TMO asistente virtual (minutos)	2,43	1,65	0,87
Beneficios por reducción de ejecutivos	959,43 USD	139.954 USD	744.266 USD
	80% (H/U); 50%	85% (H/U); 60%	90% (H/U);
Tasa de contención objetivo ¹⁷	(Desp.); 20% (Gar);	(Desp.); 30% (Gar);	70% (Desp.); 40%
	15% (Otros)	25% (Otros)	(Gar); 35% (Otros)

¹⁶ Correspondiente al promedio de los tiempos de atención de la opción de horario y ubicación de tiendas.

¹⁷ Por opción de IVR.

En el desarrollo de este análisis se utiliza el modelo de simulación Montecarlo, que consiste en un método de ensayos que simula diferentes situaciones inciertas ingresando modelos probabilísticos de las variables a sensibilizar. Para esto se utiliza el Software *Crystal Ball*.

En la **Figura 29** se presentan los resultados de la simulación para 1.000 iteraciones, donde se tiene un VAN que se ajusta a una distribución de probabilidad beta con media 177.814 USD, en un rango de -80.134 USD y 468.216 USD. La gran disminución del VAN con respecto al caso base se explica por la importancia de los beneficios por ahorro en costos en este indicador, dado que corresponden al único beneficio monetario cuantificable de la implementación del *Contact Center Cognitivo*.

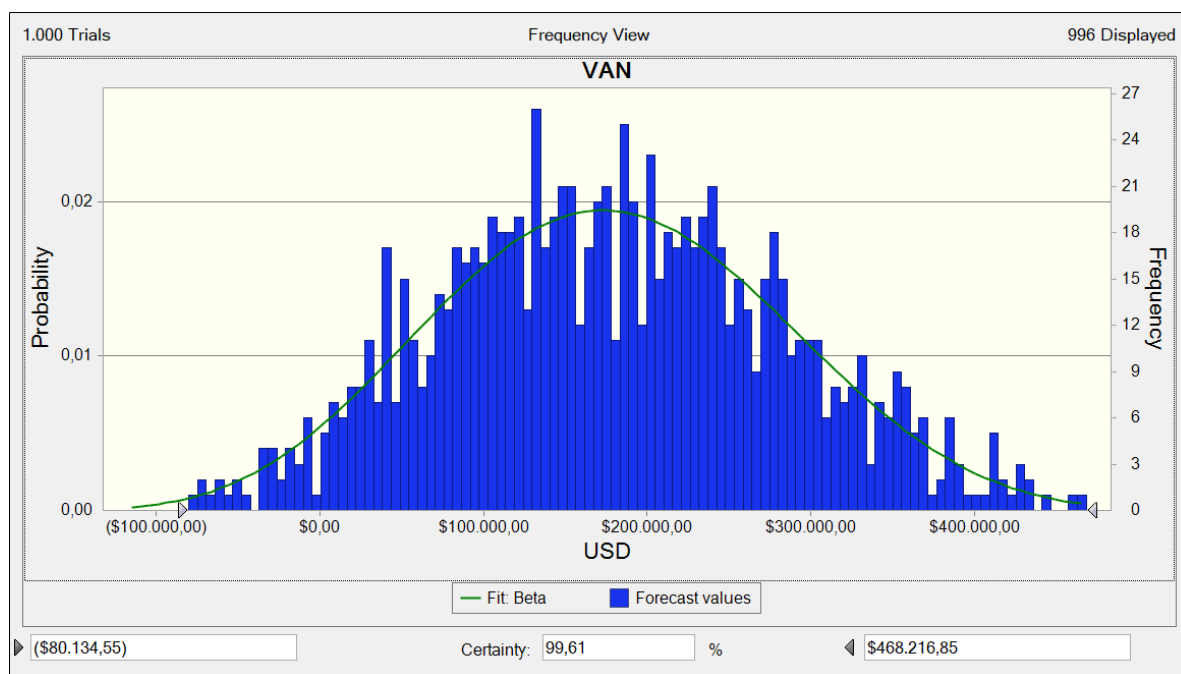
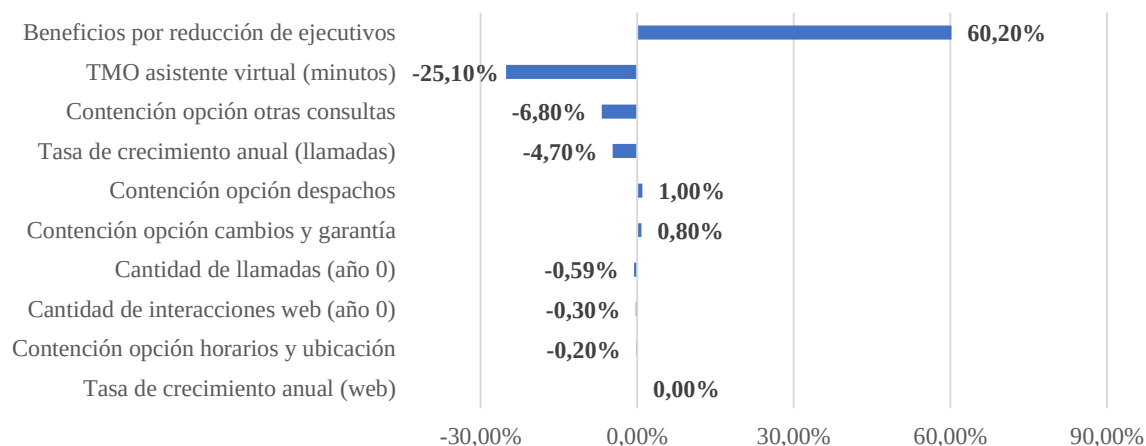


Figura 29. Distribución del VAN por simulación Montecarlo con 1.000 iteraciones.

Además, en la

Figura 30 se presenta la contribución de cada variable a la varianza de la distribución obtenida, donde las que más aportan corresponden a variables críticas por su gran impacto en el valor actual neto, estas son el beneficio por reducción de ejecutivos (60,2%) y el TMO



del asistente virtual (-25,10%).

Figura 30. Contribución a la varianza del VAN de cada una de las variables sensibilizadas. Fuente: Elaboración propia.

Finalmente se puede concluir que, al someter la evaluación a un análisis de sensibilidad multivariado, el proyecto continúa siendo rentable, presentando una probabilidad inferior al 1% de tener un VAN menor a cero.

13 CONCLUSIONES

El estudio de factibilidad que se llevó a cabo fue realizado a través de un análisis técnico y económico, con el propósito de determinar la viabilidad de la implementación del *Contact Center Cognitivo* de IBM en una empresa de retail en Chile.

En cuanto al estudio técnico, se obtuvo que la implementación del *Contact Center Cognitivo* no presenta limitaciones para su instalación en el *Contact Center* ya existente, superando las incompatibilidades entre los protocolos de comunicación SIP y H323 mediante las aplicaciones de Sixbell. En relación a los servicios *cloud* a utilizar, estos se encuentran disponibles para Chile a través de la plataforma IBM Cloud, sin encontrarse ningún inconveniente, en términos del dimensionamiento de los servicios necesarios, que pudiese afectar el desarrollo de la solución. Respecto a las restricciones de operación, se establece necesario el uso de un enlace dedicado con un ancho de banda mínimo de 6 Mbps para la transmisión de datos y paquetes de voz, con una latencia garantizada en la comunicación inferior a los 2,5 segundos. De manera adicional, se indica la necesidad de contar con personal de la empresa encargado de realizar labores de seguimiento, control y ajustes a la solución para garantizar su correcto funcionamiento a lo largo del tiempo. En base a esto se concluye que el proyecto es viable técnicamente, por tanto, es posible la implementación del *Contact Center Cognitivo* en la empresa en estudio.

En el estudio económico, fueron calculados los costos totales de la implementación de la solución propuesta a través del método de costo total de propiedad, obteniendo que, en un horizonte de evaluación de 3 años, el TCO del proyecto corresponde a 744.266 USD, donde el 51,4% se atribuye al costo por uso de servicios y aplicaciones. Posterior a esto se calculan

los beneficios por ahorro en costos de agentes liberados, que llevados a valor presente corresponden a 1.139.543 USD. Este se reconoce como el único beneficio monetario asociado a la implementación.

Una vez obtenidos los costos y beneficios, se calculan indicadores para medir la rentabilidad del proyecto, obteniéndose que los flujos generados por el proyecto son positivos después de 14,7 meses de su puesta en marcha, presentando un VAN de 395.277 USD, una TIR de 7,89% y una rentabilidad del 53% sobre el costo total de propiedad.

Con el objetivo de medir cuan sensible es la evaluación a variaciones de uno o más parámetros de interés, se realizó un análisis de sensibilidad unidimensional y multidimensional. En el primero, se estudia de manera independiente la variación de la tasa de contención objetivo, cantidad de llamadas, tasa de crecimiento de la demanda anual por servicio, TMO de ejecutivos y asistente virtual, y beneficios por reducción de ejecutivos; presentando para cada caso, el punto de indiferencia que corresponde a cuando se alcanza un VAN igual a cero. De aquí se obtiene que la variable con menor holgura de variación corresponde al TMO. Para el caso de los beneficios por reducción de ejecutivos se concluye que, si la empresa decide no reducir ni reubicar su personal contratado, el proyecto no resulta rentable presentando un VAN negativo. Sin embargo, si cuenta con la posibilidad de dimensionar su personal anual o semestralmente, sí es económicamente factible llevar a cabo el proyecto.

En el análisis de sensibilidad multidimensional, se realiza una simulación mediante el método Montecarlo, asignando probabilidades a tres escenarios posibles (pesimista, más probable, y optimista) donde se varían de manera simultánea todas las variables sensibilizadas en el análisis unidimensional. Como resultado final se tiene un VAN ajustado

a una distribución de probabilidad beta con media 177.814 USD, esto indica que a pesar de los escenarios escogidos, el proyecto tiene una alta probabilidad de ser atractivo económicamente.

Finalmente, se concluye que la implementación del *Contact Center Cognitivo* de IBM en la empresa de retail en estudio, es un proyecto viable tanto técnica como económicamente.

14 RECOMENDACIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

El *Contact Center Cognitivo* es un servicio que puede presentar numerosas ventajas asociadas a su implementación en los servicios de atención al cliente, siendo comprobada su rentabilidad, a través de este trabajo, en una importante empresa de retail en Chile. Es por esto que se recomienda explorar sobre otras industrias en las cuales podría implementarse, tomando en consideración el volumen de llamadas que maneja el servicio de atención al cliente, dado que como se observó en el análisis de sensibilidad, existe una cantidad mínima de llamadas procesadas para que los costos de implementación se justifiquen con los beneficios del proyecto. Además, deberá tenerse en cuenta la dificultad de los requerimientos de los clientes, clasificándolos y decidiendo en concordancia la contención objetivo a alcanzar.

A priori, se recomienda su estudio en industrias de telecomunicaciones, debido a la alta demanda por servicio existente en el sector, la prolongada duración de la llamada, y los altos costos laborales de sus agentes, ya que se requiere de ejecutivos con un grado de conocimientos técnicos para brindar soporte (Genesys, 2007).



De manera adicional también se propone estudiar la implementación en las industrias de servicios públicos, salud, y financiera, este último revisando si las medidas de seguridad son suficientes para garantizar la protección de la información.

Aplicable a todas las industrias mencionadas, se recomienda complementar los estudios técnico y económico, con encuestas de satisfacción del servicio a los clientes atendidos por el asistente virtual, de manera de tener conocimiento de la recepción de la solución en cuanto a la calidad de la atención, y actuar en consecuencia para obtener mejoras.

15 REFERENCIAS

BenchmarkPortal. (2017). *Industry Benchmark Report for Contact Centers*.

BenchmarkPortal LLC.

Cena , F., & Torre, I. (2004). Increasing Performances and Personalization in the Interaction with a Call Center System. *Proceedings of the 9th international conference on Intelligent user interfaces*. Funchal, Madeira, Portugal.

Aguir, S., Karaesmen, F., Aksin, O. Z., & Chauvet, F. (2004). The impact of retrials on call center performance. *OR Spectrum*, 26(3), 353-376. doi:10.1007/s00291-004-0165-7

Aksin, Z., Armony, M., & Mehrotra, V. (2009). The Modern Call Center: A Multi-Disciplinary Perspective on Operations Management Research. *Production and Operations Management*, 16(6). doi:10.1111/j.1937-5956.2007.tb00288.x

Asociación Chilena de Empresas de Call Center. (2015). *Extracto cifras de la industria 2015: ACEC*. Recuperado el Junio de 2018, de ACEC: <http://www.acec.cl/noticias/extracto-cifras-industria-2015/>

Azraq, A., Chugthai, S., Mashhour, A., Nguyen, D., & Marcelo Dos Santos, R. (2018). *Enhancing the IBM Power Systems Platform with IBM Watson Services*. IBM Redbooks.

Barbier, L. (2018). *Experiencing VoIP Issues? Here's How to Improve Call Quality Fast*. Recuperado el Noviembre de 2018, de <https://aircall.io/blog/voip-issues/>

BearingPoint. (2012). *The Digital Evolution Journey of the Contact Centre*. BearingPoint Ltd.

Belfiore, B., Chatterley, J., & Petouhoff, N. (2012). *The Impact of Technology on Contact Center Performance*. BenchmarkPortal.

Blanco, A. (2008). *Formulación y Evaluación de Proyectos* (7ma ed.). Caracas: Fondo Editorial Tropykos.

Bravo Abarca, F. (2012). *Sistema experto de configuración para sistemas telefónicos a través de diagramas*. Universidad de Chile, Santiago. Recuperado el Octubre de 2018, de http://repositorio.uchile.cl/bitstream/handle/2250/111326/cf-bravo_fa.pdf;sequence=1

Brodie, R. J., Hollebeek, L. D., Jurić, B., & Ilić, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *ournal of service research*, 14(3), 252-271.

Cámara de Comercio de Santiago. (2017). *Informe sobre navegación y tasas de conversión en ecommerce por sectores*. Cámara de Comercio de Santiago. Recuperado el 14 de Diciembre de 2018, de https://www.netrica.com/wp-content/uploads/2018/04/Report_Netrica_Industria_Chile_S22017.pdf

Cena, F., & Torre, I. (Mayo de 2006). Adapting the interaction in a call centre system. *Interacting with Computers* , 18, no 3, 478-506. doi:10.1016/j.intcom.2005.11.007

Chromy, E., Diezka, J., Kavacky, M., & Vozňák, M. (2011). Markov models and their use for calculations of important traffic parameters of contact center. *WSEAS Transactions on Communications*(10), 341-350. Recuperado el 15 de Diciembre de 2018, de https://www.researchgate.net/figure/Contact-center-architecture_fig1_253241633

- Contactbabel. (2017). *The 2017-18 UK Contact Centre Decision-Maker's Guide* (15th ed.).
- Corkrey, R. &. (2002). Interactive voice response: Review of studies 1989–2000. *Behavior Research Methods, Instruments, & Computers*, 34(3), 342-353.
- Deloitte. (2016). *En la búsqueda de la Omnicanalidad: El cliente en el centro nuevamente*. VisiónDeloitte.
- Diario TI. (26 de Julio de 2017). *Ripley lanza “Eva”, asistente virtual para retail habilitada por IBM Watson*. Obtenido de <https://diarioti.com/ripley-lanza-eva-asistente-virtual-para-retail-habilitada-por-ibm-watson/105141>
- Díaz, J., Fernández, R., Gallego, J., González, J., & Millares, J. (2016). *El contact center cognitivo. Transformando los sistemas de relación con los clientes*. IBM Watson.
- Feinberg, R. A., Kim, I., Hokama, L., de Ruyter, K., & Keen, C. (2000). Operational determinants of caller satisfaction in the call center. *International Journal of Service Industry Management*, 11(2), 131-141. doi:10.1108/09564230010323633
- Gartner. (2018). *Use artificial intelligence to foster business growth*. Recuperado el Septiembre de 2018, de Gartner.com: <https://www.gartner.com/en/information-technology/insights/artificial-intelligence>
- Genesys. (2007). *Estrategias de servicio al cliente para el sector de telecomunicaciones*. Genesys. Recuperado el 17 de Diciembre de 2018, de <https://genbin.genesys.com/old/resources/brochures/customer-service-strategies-for-the-communications-industry-strategy-guide-latin-america.pdf>

Gliozzo, A., Ackerson, C., Bhattacharya , R., Goering , A., Jumba , A., Kim , S., . . . Ribas, M. (2017). *Building Cognitive Applications with IBM Watson Services: Volume 1 Getting Started*. IBM Redbooks.

IBM. (2008). IBM Perspective on Cloud Computing. The "next big thing" or "another fad"? Estados Unidos. Obtenido de Ibm.com.

IBM. (2017). *Conversational A.I. for the Enterprise*. IBM Watson.

IBM. (2017). *IBM Cloud Functions*. Obtenido de <https://console.bluemix.net/openwhisk/>

IBM. (2017). *IBM Cloudant*. Obtenido de <https://www.ibm.com/cl-es/marketplace/database-management>

IBM. (2017). *Watson AI Assistant*. Obtenido de IBM: <https://www.ibm.com/watson/ai-assistant/>

IBM. (2017). *Watson Speech to Text*. Obtenido de IBM: <https://www.ibm.com/watson/services/speech-to-text/>

IBM. (2018). *Cloud and SaaS support*. Recuperado el 7 de Diciembre de 2018, de IBM Support: <https://www-01.ibm.com/support/docview.wss?uid=ibm10731179>

IBM. (2018). *Customer Care Voice Agent*. Obtenido de https://www.ibm.com/watson/assets/duo/pdf/Customer_Care_Voice_Agent.pdf

IBM Cloud. (s.f.). *IBM: ¿Qué es cloud computing? - España*. Recuperado el Septiembre de 2018, de Ibm.com: <https://www.ibm.com/cloud-computing/es-es/learn-more/what-is-cloud-computing/>

IBM Cloud. (s.f.). *Watson Assistant documentation: About*. Recuperado el Septiembre de 2018, de [Console.bluemix.net:
https://console.bluemix.net/docs/services/conversation/index.html?cm_mc_uid=69613642934115396326948&cm_mc_sid_50200000=88056511539632694815&cm_m
c_sid_52640000=26572821539632694825#about](https://console.bluemix.net/docs/services/conversation/index.html?cm_mc_uid=69613642934115396326948&cm_mc_sid_50200000=88056511539632694815&cm_mc_sid_52640000=26572821539632694825#about)

Incoming Calls Managment Institute (ICMI). (2003). *Call Center Metrics: Key Performance Indicators (KPIs)*.

ISACA. (2012). *Calcular el ROI de la nube: Desde la perspectiva del cliente*. ISACA. Recuperado el 8 de Diciembre de 2018, de [http://www.academia.edu/11598036/Calcular_el_ROI_de_la_nube_Desde_la_persp
ectiva_del_cliente](http://www.academia.edu/11598036/Calcular_el_ROI_de_la_nube_Desde_la_perspectiva_del_cliente)

K-12 Blueprint. (s.f.). *Digital Transformation*. Obtenido de K-12 Blueprint Web site: <https://www.k12blueprint.com/sites/default/files/POC-Overview.pdf>

Leggett, K. (16 de Junio de 2017). *How AI Will Transform Customer Service: The Present And Future Of Artificial Intelligence In Contact Centers*. Forrester, Research database. Obtenido de Forrester Research database

Lemon, K. N., & Verhoef, P. C. (Noviembre de 2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing: AMA/MSI Special Issue*, 80, 69-96. doi:10.1509/jm.15.0420

Lovera Raffo, A. J. (2008). *Diseño e implementación de un sistema experto para una plataforma de contact center*. Lima: Pontificia Universidad Católica del Perú.

Mehrotra, V., Grossman, T. A., & Samuelson, D. A. (2013). Call and Contact Centers. *Gass S.I., Fu M.C. (eds) Encyclopedia of Operations Research and Management Science*, 144-150.

Mendoza, J. (2009). *Metodología para el cálculo de costo total de propiedad para sistemas informáticos*. Guatemala: Universidad de San Carlos de Guatemala.

Microsoft Azure. (s.f.). *¿Qué es SaaS? Software como servicio*. Recuperado el Septiembre de 2018, de Azure.microsoft.com: <https://azure.microsoft.com/es-es/overview/what-is-saas/>

Miranda, C. (2005). *Sistemas informáticos y redes locales*. Ediciones Paraninfo, S.A. Recuperado el 13 de Diciembre de 2018, de [https://books.google.es/books?id=jWvPAgAAQBAJ&pg=PA93&dq=Administrador+de+sistemas+quien&hl=es&sa=X&ved=0ahUKEwiNoP37tNXyAhWGthQKHV7fAP0Q6AEIMjAC#v=onepage&q=Administrador%20de%20sistemas%20quien&false](https://books.google.es/books?id=jWvPAgAAQBAJ&pg=PA93&dq=Administrador+de+sistemas+quien&hl=es&sa=X&ved=0ahUKEwiNoP37tNXyAhWGthQKHV7fAP0Q6AEIMjAC#v=onepage&q=Administrador%20de%20sistemas%20quien&=false)

Nuance. (2014). *Nuance Vocalizer Data Sheet*.

Ovum Consulting. (2016). *Get It Right : Deliver the Omni-Channel Support Customers Want What*. Recuperado el 1 de Junio de 2018, de <https://az766929.vo.msecnd.net/document-library/boldchat/pdf/en/boldchat-whitepaper-ovum-logmein.pdf>

Percy, A. (2010). *Understanding Latency in IP Telephony*. Recuperado el 18 de Noviembre de 2018, de <http://www.ktl.elf.stuba.sk/~chromy/SpS2/referaty/49%20-%20Understanding%20Latency%20in%20IP%20Telephony.pdf>

Project Management Institute, Inc. (2013). *Guía de los fundamentos para la dirección de proyectos (Guía del PMBOK)* (5ta ed.). Newtown Square, PA: Project Management Institute, Inc.

Pronto Comunicaciones Ltda. (2017). *¿Qué es un internet dedicado y cuando se necesita?* . Recuperado el 17 de Noviembre de 2018, de ProntoIP: <http://www.prontoip.com/internet-dedicado-cuando-se-necesita/>

Quiñonez, F. (2016). *Estudio de la red telefónica IP basada en Elastix instalada en la comunidad salesiana María Auxiliadora*. Esmeraldas: Pontificia Universidad Católica del Ecuador. Obtenido de <https://repositorio.pucese.edu.ec/bitstream/123456789/632/1/QUI%C3%91ONEZ%20ANGULO%20FRANCISCO%20XAVIER.pdf>

Ransbotham, S., & Kiron, D. (2018). *Using Analytics to Improve Customer Engagement*. MIT Sloan Management Review.

Redpuntos. (2016). *Las entrañas de IBM Watson: así funciona la plataforma de IA más avanzada del mundo*. Obtenido de <http://www.unblogenred.es/las-entranas-de-ibm-watson-asi-funciona-la-plataforma-de-inteligencia-artificial-mas-avanzada-del-mundo/>

Ries, E. (2009). *Minimum Viable Product: a guide*. Obtenido de <http://www.startuplessonslearned.com/2009/08/minimum-viable-product-guide.html>

Rouse, M. (2017). *What is Cloud Object Storage?* Recuperado el Noviembre de 2018, de SearchStorage: <https://searchstorage.techtarget.com/definition/cloud-object-storage>

Sánchez, I. R. (2003). *Enfoque económico social de evaluación de proyectos de inversión*. .

(Doctoral dissertation, Tesis doctoral en Ciencias Económicas, CDICT, Universidad de La Habana).

Santos, T. (2008). Estudio de factibilidad de un proyecto de inversión: etapas en su estudio.

Contribuciones a la Economía. Obtenido de <http://www.eumed.net/ce/2008b/>

Schoeller, A. (2018). *Future-Proof Your Customer Service: Build An AI-Infused Cognitive*

Contact Center. Forrester. Recuperado el Septiembre de 2018, de <https://reprints.forrester.com/#/assets/2/1358/RES137693/reports>

Schwartz , R., Mooney , K., & Heller, C. (2018). *The AI-enhanced customer experience. A sea change for CX strategy, design and development*. IBM Institute for Business Value .

Sucely, G. (2008). *Elaboración de estudios de factibilidad en la adopción de TIC's*. Guatemala: Universidad de San Carlos de Guatemala.

Suff, P., Reilly, P., & Mercer, M. (2005). *Your call: Managing reward and performance in call centres*. Institute for Employment Studies.

Suhm, B., & Peterson, P. (2001). Evaluating commercial touch-tone and speech-enabled telephone voice user interfaces using a single measure. *CHI Extended Abstracts*.

Tarokh, M. J., & Ghahremanloo, H. (2007). Intelligence CRM: a contact center model. *Service Operations and Logistics, and Informatics*, 1-6.



Torres Pluas, P. R., & Lopez Jaramillo, N. E. (2013). *Estudio de Factibilidad para mejorar la atención al cliente con la implementación de un Call Center en la CNEL Milagro.*

Milagro: Universidad Estatal de Milagro.

Uribe-Echevarría, V., & Morales, G. (2010). *Atendiendo a los clientes de los clientes: La industria del call center y sus condiciones laborales.* Santiago, Chile: Dirección del Trabajo.

Ward, J., Daniel, E., & Peppard, J. (2008). Building Better Business Case for IT Investments. *MIS Quarterly Executive*, 7(1).