



industriales  
etsii

Escuela Técnica  
Superior  
de Ingeniería  
Industrial

# UNIVERSIDAD POLITÉCNICA DE CARTAGENA

Escuela Técnica Superior de Ingeniería  
Industrial

## Desarrollo de un sistema automático para la generación de informes a partir de un chatbot

TRABAJO FIN DE GRADO

GRADO EN INGENIERÍA ELECTRÓNICA INDUSTRIAL Y  
AUTOMÁTICA

**Autor:** María José Martínez Bernal  
**Director:** José Manuel Cano Izquierdo  
**Codirector:** Julio José Ibarrola Lacalle

Cartagena, 02/02/2021



Universidad  
Politécnica  
de Cartagena

# Índice

1. Introducción	2
2. Objetivo	5
3. Estado del arte	6
4. Qué es un chatbot y cómo funciona	9
4.1 Caso de estudio: El chatbot en Caixabank	9
5. Metodología	18
6. Propuesta de desarrollo para la mejora de la creación de informes	21
6.1 Punto de partida y problemática concreta	21
6.2 Propuesta de desarrollo	22
7. Desarrollo	25
7.1 Tablas y extracción de datos	25
7.2 KPIs utilizados	26
7.3 Filtros de información para el informe	31
7.4 Código para la creación del informe	33
7.5 Comprobación de datos obtenidos	43
8. Resultado	44
9. Conclusiones	52
10. Bibliografía	53

## 1. Introducción

De forma anual y como segundo año consecutivo, en 2018 salió la primera convocatoria a la beca Wonnnow creada por CaixaBank y Microsoft Ibérica para poder realizar una estancia formativa profesional en las instalaciones de Caixa en Barcelona. Esta beca está dirigida a alumnas universitarias de carreras con menor presencia femenina del área de ciencias, tecnología, ingeniería y matemáticas. El galardón tiene como fin promover, poner en valor y estimular la excelencia, tanto académica como personal, de las mujeres estudiantes universitarias de las áreas antes mencionadas, dando visibilidad social al esfuerzo, tesón y capacidad femenina en estos campos.

En el tiempo que duró la beca en CaixaBank, comprendí que se trata de una entidad financiera que además de ofrecer los servicios característicos de un banco, también se ha especializado en otras áreas con las divisiones SegurCaixa y VidaCaixa. A pesar de ser independientes todas ellas trabajan en conjunto y los servicios prestados se ofrecen en las oficinas bancarias. La entidad le da mucha importancia a prestar un buen servicio al cliente, por eso su lema es “Escuchar, Hablar y Hacer”. Pero al final, la entidad se ve representada mediante las oficinas y el servicio que los empleados dan al cliente (fig. 1).

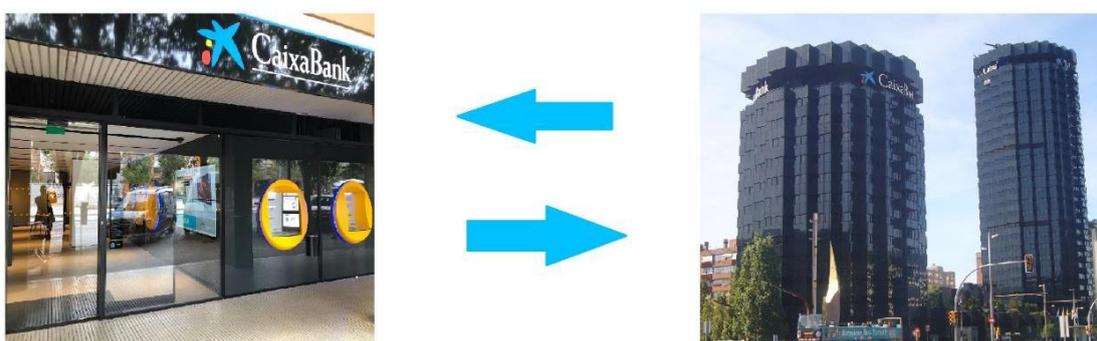


Figura 1: intercambio de información entre centros.

Entre la sede central y las oficinas debe haber un intercambio de información constante y saber cumplimentar las funciones que tiene cada una (fig. 2).

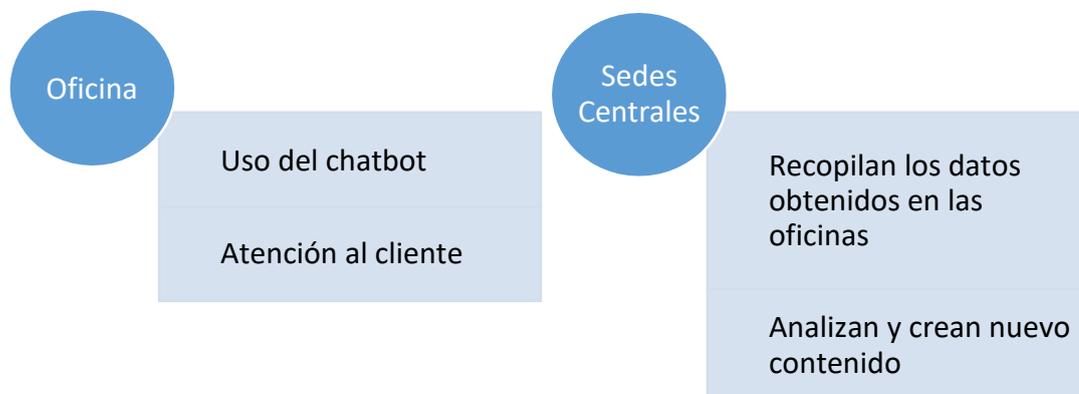


Figura 2: funcionalidades de cada sector

Tuve el honor de ser galardonada con una de las becas dirigida al asistente cognitivo de empleados y *big data* en el departamento de eficiencia y digitalización de procesos. Esta fue la rama que escogí en la solicitud de la beca, ya que dentro del programa había un ámbito de 5 disciplinas que eran las siguientes:

1. Ciberseguridad.
2. Creación de aplicaciones.
3. Agronomía.
4. Gestión de proyectos.
5. Asistente cognitivo.

A modo personal, mi elección se basó en que considero que la Inteligencia Artificial es un avance en la tecnología, el cual tiene mucho potencial ya que supone una nueva forma de que un software ponga en práctica una tarea que se le ha encargado. Además, es una tecnología que se puede aplicar en múltiples y variados campos y puedes personalizar al máximo los patrones que la forman.

Durante los primeros meses de trabajo analizando datos, uno de los aspectos en los que más había que trabajar era en los costes. Este aspecto quedaba en evidencia a final de mes con los informes que se realizan para resumir los datos obtenidos del balance mensual de cada área de la entidad. Estos informes se utilizaban para ver la evolución del *chatbot*, el cual está enfocado para uso interno de la empresa, y para ver posibles mejoras a implementar. En ellos, se pudo observar la carencia en tres aspectos básicos que afectaban fundamentalmente a la productividad:

1. Obtención de costes elevados.
2. Saturación del departamento técnico.
3. Aumento en el tiempo de producción del propio informe

La estructura del departamento de eficiencia y digitalización de procesos se estructuraba en tres grupos de trabajo:

- Asistente cognitivo.
- Comunicación.
- Logística y operaciones.

En mi caso, me asignaron al primer equipo mencionado, que previamente estaba formado por tres componentes los cuales se encargaban de supervisar el funcionamiento del *chatbot* y de las futuras implementaciones que se quisieran hacer a este. Los perfiles del equipo eran: un informático, un físico con conocimientos en múltiples lenguajes de programación y, una administradora de empresas. En el grupo cada uno tenía su función, la cual complementaba la del otro y así trabajan para obtener un objetivo común. Mi rol en el equipo era el de la automatización de funcionalidades que se pudiesen realizar dentro del *chatbot* o a raíz de datos que se generan en él. El objetivo era por tanto completar el equipo con una persona que fuera capaz de optimizar los tiempos y consecuentemente, reducirlos.

## 2. Objetivo

Según lo comentado anteriormente, esa situación supuso el punto de partida sobre el que comenzar a desarrollar la propuesta que se presenta en este trabajo como Trabajo Final de Grado. Se trata de la creación de un sistema automático, que efectúe y controle las secuencias de operaciones sin la ayuda de la actividad humana, para que generase los informes que resumen el balance del mes a partir de los datos del *chatbot*.

El objetivo principal consistirá en crear un sistema automático sólido que sea la base para la creación de todos los informes futuros. Esta propuesta será capaz de llevarse a cabo basándose en las cinco tareas siguientes:

1. Analizando el método que usaban para la obtención de datos, hay que tener un conocimiento del punto de partida para saber marcar un objetivo.
2. Organizar toda la información que se necesita para completar el informe, esto implica hacer un barrido de todos los datos que se muestran en el informe mensual, con qué tipo de datos hay que trabajar y cómo se quiere mostrar.
3. Recopilar todas las bases de datos con las que contamos para ver donde se encuentra la información que necesitamos. Es importante recopilar y analizar las bases de datos existentes para crear una relación entre todas ellas la cual nos permita generar todos los datos que se necesitan, esta parte es de las más importante ya que va a consolidar el cerebro de la aplicación.
4. Programación del código, que será la base para nuestro informe.
5. Comprobar si todo lo que se ha creado está proporcionando datos correctos o hay algún indicador, gráfica o tabla que muestra información que no debería.

### 3. Estado del arte

Actualmente los *chatbots* forman una parte importante de la tecnología, lo que podría hacer pensar que se trata de un concepto moderno, pero esta noción se remonta a setenta años atrás. Estos programas informáticos basados en mantener una conversación, chatbots, nacieron de la mano de un profesor de informática del MIT (*Massachusetts Institute of Technology*), Joseph Weizenbaum [1].

El inicio de la historia de los *chatbots* comenzó en los años 50 cuando Alan Turing, considerado el precursor de la Inteligencia Artificial, empezó a considerar la posibilidad de que las máquinas llegasen a pensar. Diseñó el Test de Turing, un postulado teórico con el que aspiraba a poder comprobar si una máquina tiene comportamientos inteligentes similares o indistinguibles del de un humano. Pero no fue hasta 1966 cuando Joseph Weizenbaum, basándose en ese mismo método, diseñó a ELIZA con el objetivo de engañar a la gente haciéndoles pensar que estaban hablando con una persona en lugar de con una máquina. Se llamaba Eliza en honor al personaje Elsa Doolittle de la obra *Pigmalión*, del dramaturgo irlandés George Bernard Shaw, que con ayuda de un profesor se transforma de una vendedora callejera a una dama de la alta sociedad.

ELIZA quería ser un *chatbot*, se convirtió en psicoterapeuta y acabó siendo un experimento social. Para llevar a cabo este programa informático, Weizenbaum identificó cinco pilares para su correcto funcionamiento que eran: el reconocimiento de palabras clave, el descubrimiento de un contexto mínimo, la elección de transformaciones apropiadas, la generación de respuestas adecuadas o la capacidad de reaccionar ante la ausencia de palabras críticas. Después de varios intentos de prueba y error, consiguió que el sistema reconociera palabras clave en las frases y respondiera con una frase asociada a la palabra clave que había detectado, por lo que no era inteligente ni había aprendizaje automático ya que era incapaz de memorizar, simplemente reconocía palabras clave y las asociaba a ideas que ya tenía definidas (fig. 3).

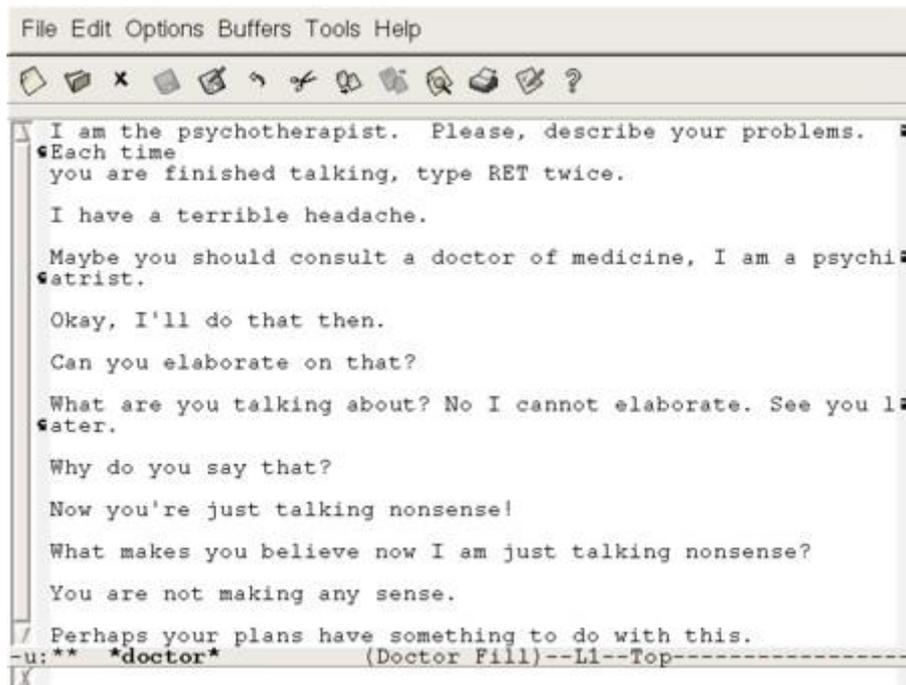


Figura 3: Conversación del *chatbot* ELIZA.

Un aspecto curioso del efecto ELIZA es que, existe un término, derivado del programa, que consiste en asociar el comportamiento de una máquina con el de un ser humano. Se refiere a la susceptibilidad de las personas a buscar una mayor comprensión en cadenas de símbolos ensartados por computadoras, tener la sensación de que estamos leyendo algo especial. El descubrimiento de este efecto ha sido muy importante en el desarrollo de la inteligencia artificial, principalmente en aquellos chatbots que intentan transmitir emociones o intentar que el usuario perciba el programa como algo más cercano.

A raíz de la creación del primer chatbot, ELIZA, desarrollaron nuevos *chatbots* como A.L.I.C.E. (*Artificial Linguistic Internet Computer Entity*) en el 1995 que forma parte del proyecto Pandora, Albert One en 1995 por Robby Garner, Smarterchild fundado en el 2000 desarrollado por *ActiveBuddy* Inc. O Mitsuki en 2005 creado por Steve Worswick. Este último fue el primer boceto de asistente virtual ya que no solo respondía asociando ideas, si no que podía ayudarte de una manera más personalizada y tener una mayor interacción humana.

La diferencia entre un *chatbot* y un asistente virtual reside en que el asistente, tiene una plataforma interactiva en la cual no solo entienden el idioma, sino también el significado de lo que dice el usuario. Pueden proporcionar una imprevisibilidad a su comportamiento y así tener una larga interacción humana, mientras que el *chatbot* está basado en texto y está programado para responder solo a un determinado conjunto de preguntas o declaraciones.

Actualmente, algunos de los asistentes virtuales con mejor rendimiento son [2]: Siri que fue creado en 2011, Alexa, el asistente de Amazon lanzado en 2014, *Google Assistant* del 2016 y Woebot, desarrollado en 2017 por la psicóloga Alison Darcy que ayuda a reducir los síntomas de depresión.

Como referencia para la ayuda de mejora del *chatbot* y la creación de informes me inspiré en el de la empresa *Salesforce*. Es una empresa creada por Marc Benioff que nació en el 1999 y ha ido creciendo debido a que apostó por la innovación hasta convertirse en una de las empresas líder en las soluciones CRM (Customer Relationship Management). Fue en 2016 cuando implanto su *chatbot* llamado Einstein, permitiendo así a sus usuarios acceder a una información más completa sobre sus contactos, ser más ágiles, productivos y anticiparse a sus necesidades.

## 4. Qué es un *chatbot* y cómo funciona

Un *chatbot* es una tecnología que nos permite tener una conversación por chat con un programa informático. Está construido a través de Inteligencia Artificial, la cual le permite aprender sobre las consultas que se le realizan, pudiéndole así añadir a su memoria la información deseada. A medida que se van realizando más consultas, el *chatbot* adquiere mayor cantidad de datos, lo que permite tener más experiencia, de manera que va tomando seguridad en sus respuestas, de modo que es capaz de conocer por situaciones anteriores cierto tipo de respuestas que han funcionado y consecuentemente, cuáles no. La finalidad es tener la capacidad de resolver las preguntas que el usuario pueda tener y de este modo, resolverlas de forma eficiente, fácil y rápida. Así conseguiríamos un producto final eficiente, de cómoda accesibilidad e intuitivo.

Como las interacciones con este servicio son más rápidas y sencillas generan experiencias agradables para el usuario. Esta tecnología ofrece un gran refuerzo de cara a la atención al cliente, quitándole carga de trabajo a los trabajadores del *Call Center* y al servicio técnico. Tiene como función la proporción de información, resolución de problemas buscando sobre todo la mejora en la atención a los clientes.

### 4.1 Caso de estudio: El *chatbot* en Caixabank

El funcionamiento de todos los *chatbots* parten una base con el mismo funcionamiento, pero luego cada empresa amplía el suyo propio proporcionándole una serie de prestaciones diferentes tales como ayuda psicológica, automatización de procesos o atención al cliente. En este apartado se explicará en detalle el *chatbot* en Caixabank, ya que es en el que pude trabajar intensamente durante seis meses.

Como punto de partida, es importante comentar que la mayoría de las gestiones de Caixabank se producen en las oficinas bancarias. Estas son un lugar donde la institución financiera ofrece asistencia personal y automatizada a los clientes, por lo que es importante que cuenten con una buena fuente de información para consolidar un servicio de calidad. Este servicio presenta unas gestiones que se basa en responder y gestionar posibles incertidumbres que pueda tener un cliente. Las dudas que pueden surgir suelen estar relaciones con nuevos planes que ha lanzado la entidad, operaciones que no sabe cómo realizar o contrataciones de algunos de los planes que ofrecen (fig. 4). Como resulta ser una información muy concreta y en muchos de los casos son planes que acaban de lanzarse, puede también generarle dudas al propio empleado que tiene que ofrecer el servicio. Cuando se da esta situación, el empleado intenta ponerse en contacto con el departamento técnico para ver si este puede resolver las dudas generadas.

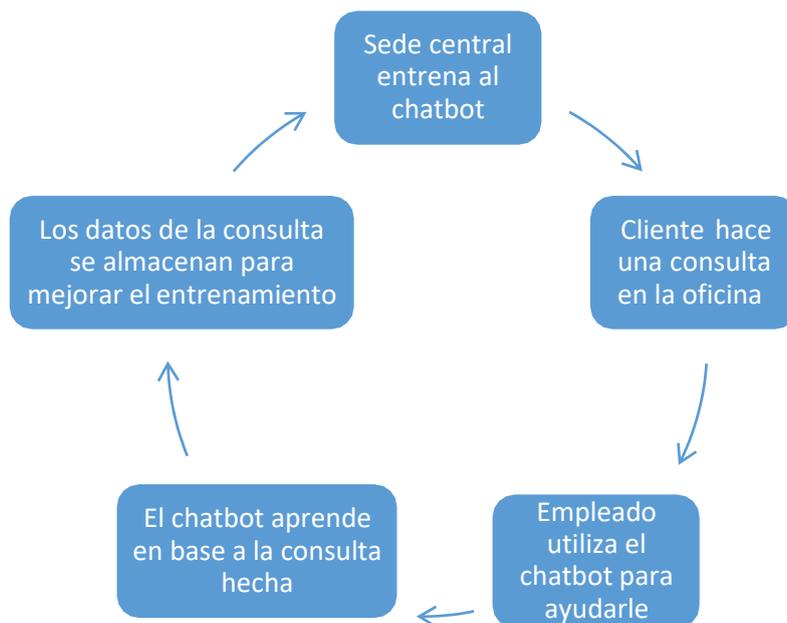


Figura 4: Ciclo de la información del *chatbot*

Para precisar la focalización del objeto de estudio, se creó una estructura buscando la eficacia en la operativa de resolución de consultas lo más satisfactorio posible. Un esquema en cuanto a la configuración podría ser el siguiente:

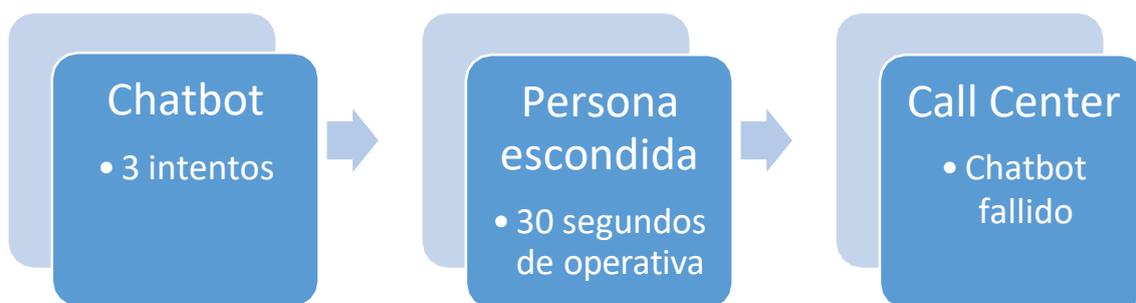


Figura 5: Estructura de respuesta a una consulta

Esta disciplina científica, la inteligencia artificial del *chatbot*, está alimentada con alrededor de 7000 *FAQs* (*Frequently Asked Questions*) para ser capaz de resolver las posibles dudas que generan los servicios ofrecidos a través de la entidad y de sus áreas. Pero para tener la posibilidad de dar un servicio de lo más óptimo posible, el *chatbot* debe saber relacionar bien toda la información que tiene almacenada y esto se consigue mediante entrenamiento. En el caso particular de Caixa, al disponer con tantos datos recopilados la tasa de entendimiento para captar lo que se le pide tiene que ser mucho mayor. Cada consulta tiene una intención, *intent* en lenguaje informático, de lo que el usuario quiere obtener o recibir. Por ello es fundamental diseñar una lógica que permita interpretar aquello que el usuario podría pedir y definir por tanto la posible respuesta que el *chatbot* tendrá ante esa interacción.

La base del entrenamiento será exponer al *chatbot* al mayor número de asociaciones posibles. Esto se hace estableciendo respuestas fijas que queremos que el *chatbot* ofrezca, utilizando sinónimos para poder ampliar la red de preguntas a entender

y etiquetar textos para que pueda detectar temas directamente. La asociación de sinónimos se programa para que una misma pregunta formulada de diferentes formas con otras palabras sea capaz de entenderla. También hay que entrenarle [3] para los casos en los que no entienda lo que se le pregunta. Entra aquí en juego el procesamiento del lenguaje natural, idioma hablado o escrito por humanos para propósitos generales de comunicación, y la capacidad del *chatbot* de pedir a la persona que repita una pregunta con otro lenguaje, antes de guiarle a acciones erróneas debido a una mala interpretación. Dentro del entrenamiento hay que tener en cuenta el idioma en el que se le realiza la consulta. Este está entrenado para entender el castellano y, además, al ser una entidad bancaria catalana cuenta también con el aspecto bilingüe y puede responder también catalán.

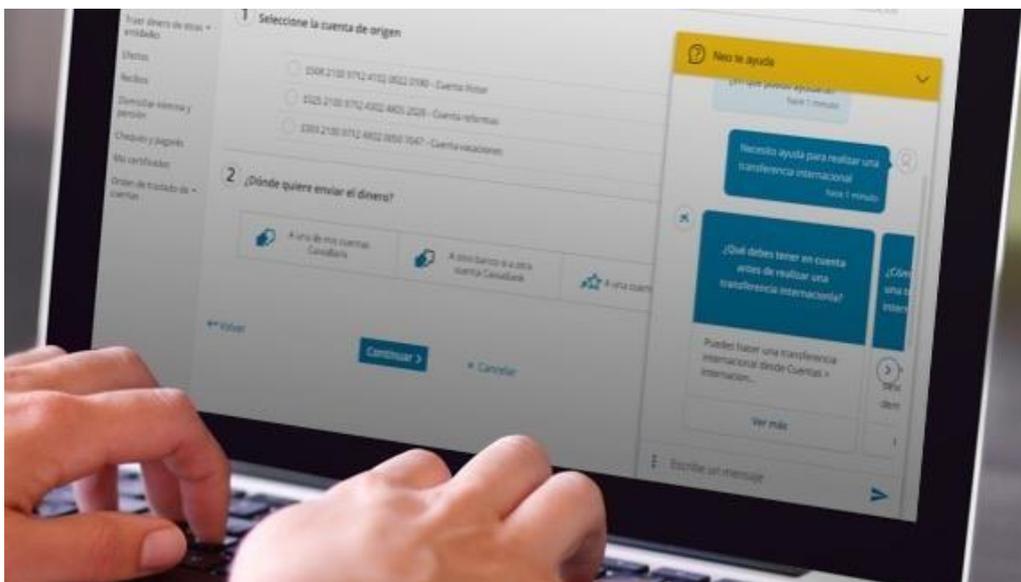


Figura 6: Chat a modo de pop-up

El funcionamiento ideal de un *chatbot* se produce a través de una consulta del usuario hacia el sistema, este entiende la pregunta al primer intento y le proporciona la información deseada sin tener ninguna duda. La consulta se realiza a través de un chat que aparece como una ventana de diálogo situada en la esquina inferior derecha a modo de *pop-up* (fig. 6). Una vez comienzas la conversación, las respuestas pueden ser



expresiones que ha aprendido o corrigiendo respuesta que ha podido dar. De esta manera, se sigue ampliando la información y, por tanto, aprendiendo para poder ofrecer respuestas más precisas



Figura 8: Boceto de la pregunta del *chatbot* para retroalimentarse

Una vez visto un caso ideal, pasamos a las situaciones en las que el *chatbot* no entiende la consulta en un primer momento, lo que nos lleva a la estructura de la fig. 2. En esta imagen, cada uno de los tres bloques representan una estructura de organización al que la consulta puede derivarse. Entendiéndose como nivel cero el *chatbot*, el nivel uno la persona escondida y nivel dos el *Call Center*. A continuación, se presentan las situaciones que pueden darse en cada uno de los niveles:

### Nivel cero

Es la que consideraremos como fase inicial. El empleado, el que consideramos el usuario para el *chatbot*, desde la oficina realiza una consulta al *chatbot*, y una vez la pregunta se ha realizado se pueden dar dos casos.

- El primero, es el caso ideal, el *chatbot* entiende la pregunta a la primera y le genera la solución a su problema. Esta respuesta puede ser una *FAQ* corta,

simple o guiada. Siendo una *FAQ* corta la que se muestra directamente en el flujo del chat, la guiada necesita seguir unos pasos que le va indicando el *chatbot* durante la conversación y la simple son las que se muestran en el *frame* de la derecha (fig. 4).

- El siguiente caso sucede cuando el *chatbot* no entiende bien el primer intento de pregunta, entonces el empleado la reformula de nuevo. De nuevo trabajaría la Inteligencia Artificial que está programada para que el usuario pueda explicarle su consulta en tres intentos. Probará a resolverla durante este número de intentos y si no puede, bien porque la pregunta no está bien formulada o porque no sabe asociarlo a ninguno de los contenidos que tiene, procederá a escalar el caso al siguiente nivel.

## **Nivel uno**

El nivel uno o “persona escondida”, siendo este un empleado del Call Center, intentará resolver el problema para evitar escalar la consulta al siguiente nivel. Este concepto implica tener a una persona, en horario de oficina, detrás de una pantalla proporcionándole soporte al *chatbot*. Se dedica a supervisar las preguntas en las que el asistente duda. Este nivel entra en acción cuando se dan los siguientes casos.

- El *chatbot*, después de los tres intentos, no ha dado con la respuesta a lo que se le está preguntando. La consulta la redirecciona al siguiente nivel y este tiene un límite de tiempo para responder, 30 segundos. Esto sucede así para que el usuario no demore mucho tiempo y para que no se note el salto al siguiente nivel. La persona pasará a responder su consulta y si se ha excedido de tiempo pasaría a ser un caso para el segundo nivel.
- El *chatbot* sabe que le está respondiendo bien a la pregunta, pero el empleado le responde que no es la información que busca. Estos casos suelen darse cuando el empleado quiere hablar directamente con un humano y no quiere hacerle caso a la información que este le proporciona. Así que, se deriva a la “persona escondida” y esta verifica si lo que le está mostrando es de verdad erróneo o no.

Si la persona le confirma que es correcta el *chatbot* vuelve a mostrarle la misma respuesta, en caso contrario, la persona del primer nivel seguirá la conversación con el usuario por chat indicándole que ahora está hablando con un humano. Es preferible tener este nivel ya que ayuda a mantener el objetivo de ahorrar dinero, una llamada supone generar un coste mayor que un chat.

## **Nivel dos**

Se llega a este punto cuando el *chatbot* después de tres intentos no ha podido solucionar la consulta y cuando la fase de la “persona secreta” tampoco ha sido satisfactoria. Una vez en esta fase, se genera un código que se le muestra al usuario por el chat y así podrá realizar la llamada desde su Smartphone corporativo y hablar con un agente del Call Center. Este proceso de generación de código se ha bautizado como *Click to Call*. Funciona de manera que el código que se genera, se le muestra al usuario, este lo introduce en la aplicación del Smartphone y recibirá la llamada. Esta llamada será más rápida y eficaz debido a que el operador tiene en la pantalla del ordenador la traza de la consulta, por lo tanto, el usuario no tendrá que explicarle su problema desde el principio. Una vez se ha llegado a este punto ya la conversación en el *chatbot* se pierde y ya estaría listo para otra consulta.

En todo momento se busca priorizar el uso del *chatbot* ya que el objetivo es conseguir que se reduzcan el número de llamadas a *Call Center*. Con la implantación y la utilización de la herramienta de Inteligencia Artificial, los usuarios serán autosuficientes y podrán responder las consultas que les realizan los clientes sin necesidad de comunicarse con otros departamentos. De esta forma se conseguirá una atención más personalizada y eficaz.

Una breve conclusión inicial a la que llegaríamos hasta el momento, es que considerando el soporte de 24 horas que ofrece el *chatbot*, gracias a la automatización de procesos y adquisición de información en tiempo real, se tratará de una herramienta de lo más útil para empresarios como para clientes.

## 5. Metodología

Como ya se ha mencionado en la introducción, el rol que ejercía en el proyecto estaba focalizado en la automatización de los posibles procesos del asistente cognitivo.

Teniendo en cuenta que un proceso es una sucesión de instrucciones que pretenden llegar a un estado final o que persiguen realizar una tarea concreta, dentro del *chatbot* se podrían automatizar procesos tales como operaciones bancarias, consultas de citas con clientes o consultas de saldos. Por ello, un proceso correctamente formulado se caracterizaría por la capacidad de generar ahorros en los costos de funcionamiento, mejorar la productividad, agilizar procedimientos de gestión y mejorar en la calidad de los servicios.

La referencia principal a la hora de investigar y emplear como base de inspiración para poder proponer mejoras o nuevas funcionalidades al *chatbot* de CaixaBank fue el propio *chatbot* desarrollado por la empresa Salesforce (fig. 9) [3]. Es una plataforma de gestión de las relaciones con los clientes (CRM, Customer Relationship Management) basada en la nube que proporciona a todos los departamentos de su organización, incluidos los de marketing, ventas, servicio al cliente y comercio electrónico, una visión unificada de sus clientes en una plataforma integrada.



Figura 9: Chatbot de Salesforce

Me pareció una referencia muy interesante sobre la que basarme porque es una empresa puntera desarrollando esta herramienta para la gestión óptima de ventas, además dispone de un software muy robusto y ofrece un excelente servicio al cliente. Además, hacen uso de la Inteligencia Artificial para clasificaciones predictivas, información procesable y para la automatización inteligente, por lo que también me basé en esta empresa para la creación de informes y la visualización de datos. Ofrece una serie de herramientas de creación de informes de gran capacidad que funcionan conjuntamente para ayudar a comprender y realizar acciones en sus datos, con el añadido de que ellos funcionan con datos en tiempo real.

Teniendo en cuenta esta referencia, para poder desarrollar y presentar las que se iban a proponer, me apoyé en la metodología de las ocho fases (8D) [4]. Esta consiste en el perfeccionamiento continuo de la aplicación debido a que busca obtener una serie de mejoras al respecto que sean incrementales y sostenibles. La propia metodología del 8D se estructura como era de suponer en ocho fases y consiste en:

1. Formación del grupo de mejora: al trabajar en un equipo multidisciplinar se obtienen diversas opiniones sobre las posibles mejoras a desarrollar. Cada integrante proporciona una visión más enfocada en los conocimientos de su área de conocimiento.
2. Definición del problema: para poder definirlo primero hay que ordenar y clasificar los datos con los que contamos, de esta manera se obtiene el proceso, variables y elementos afectados. Para comprobar si todo estaba bien definido seguí la técnica del interrogatorio, respondiendo *al what, why, when, how, where and who*.
3. Implementación de soluciones de contención: implica la acotación del problema, evitando que nos alejemos de la raíz de la solución.
4. Identificación de la raíz de la causa: consiste en la recolección de datos para los indicadores, para poder llevar a cabo este punto me hice preguntas tales como ¿cuántos datos se necesitan? o ¿qué requisitos tenemos?

Una vez estas preguntas han sido resueltas, se puede decir que tenemos la base del análisis causal de la situación del problema. De esto modo, se da comienzo a la lluvia de ideas en base a los indicadores para obtener el motivo del problema.

5. Análisis de soluciones para las causas raíces: en esta fase se realiza a generación de ideas, donde en nuestro caso se utilizó el método del juicio diferido, listé todas las alternativas que se me habían ocurrido para abordar el problema.
6. Comprobación de las causas raíces: se elige entre las diversas alternativas cual deberá implementarse. Se decidió comenzar con una prueba piloto ya que nos ayuda a probar tanto la solución como el proceso de implementación. Así, se puede realizar una comprobación de los datos y comprobar cómo reacciona el empleado ante este nuevo informe. De esta manera, si los resultados son positivos, se puede comenzar con la creación del informe automatizado.
7. Prevención de reocurrencias del problema: el objetivo es que la solución se sostenga en el tiempo y que los ya antiguos problemas no se vuelvan a presentar. Para ello se realiza una comprobación de la herramienta para ver si reporta bien los datos y se comprueba que los empleados la están consultando.
8. Cierre y reconocimiento del equipo mejora de proceso: consiste en formalizar el cierre del proyecto y reconocer que el trabajo ha sido valorado.

## 6. Propuesta de desarrollo para la mejora de la creación de informes

Mensualmente las áreas SegurCaixa y VidaCaixa, que están integradas en Caixabank, tienen que mandar un informe donde se recojan los resultados de cómo está funcionando el chatbot. La primera es una compañía de seguros de salud y VidaCaixa es una entidad de previsión social. El informe relacionado con estas áreas, sirve para ver la evolución que está teniendo, analizar procesos que se puedan automatizar y comprobar si los objetivos propuestos se están alcanzando.

El proceso inicial que seguía la entidad junto con las áreas para poder montar el informe era lento y poco eficaz. La finalidad que tenía era recopilar la información relativa a cómo ha funcionado el *chatbot* a lo largo de cada mes. Este desarrollo consistía en que para reunir toda la información tenían que enviar varios correos para solicitar los datos a distintos departamentos. Al ser una rutina periódica produce una pérdida constante del tiempo de los empleados y una reducción en la eficiencia y optimización del proceso.

### 6.1 Punto de partida y problemática concreta

Fundamentalmente el problema parte de una gestión individual por departamento de su propia información por medio de informes mensuales que no están conectados con un núcleo central. Esta información no está conectada entre ellos, por lo que reunirla y organizarla se convierte en un proceso lento e ineficaz. Considerando esta situación se llegó a la conclusión de crear un *dashboard* automatizado, documento en el que se reflejan las principales métricas o KPI (*Key Performance Indicator*) que intervienen en la consecución de los objetivos de una estrategia. En él, se podrían acortar estos procesos y tendrían todos los datos necesarios, evitando así, miles de correos entre departamentos solicitando datos y filtros. Gracias a este *dashboard*, sólo tendrían que

darle a un botón dentro de la aplicación para tener los datos actualizados.

Automatizando este proceso se conseguirían las siguientes ventajas:

- Optimización de la jornada laboral gracias a la automatización de los procesos
- Reducción del tiempo en la obtención de datos
- Departamentos más autosuficientes
- Trabajadores más resolutivos en sus tareas
- Información más precisa y detallada de los servicios ofertados
- Representaciones más dinámicas de la información para facilitar su comprensión
- Reducción de los tiempos de análisis de los datos. La información se podrá obtener por franjas horarias, días, semanas, meses o años.
- Mejora de la atención al cliente gracias a un análisis más exhaustivo de las consultas

De tal manera que, los datos del informe se obtendrían extrayéndolos directamente del *dashboard*. Cada área podría obtener los datos que necesiten aplicando ellos mismos los filtros, fechas y categorías que necesiten. Por lo que así se conseguiría un informe donde solo tengan que actualizar los datos sin necesidad de comunicarse con otros departamentos.

## 6.2 Propuesta de desarrollo

Para obtener todas estas mejoras se propuso una automatización del proceso basado en una *query* que abarcaría todos los datos con los que trabajan las áreas. Además, este proceso permite visualizar los indicadores de manera directa y descargarse además los datos obtenidos en el momento.

El sistema propuesto a automatizar, se basó en el software *QlikSense* [5] que es el que utiliza la empresa, para mí era la primera experiencia, pero es una herramienta que se caracteriza por un fácil manejo, por ser muy intuitiva y por ofrecer una experiencia

analítica muy completa. Me pareció una buena opción ya que es una plataforma de análisis de datos que tiene un motor asociativo de análisis único. Además, brinda la oportunidad de hacer mejoras diarias, obteniendo así una empresa verdaderamente basada en datos.

Como método de investigación utilicé el de la observación. Este está basado en describir, explicar y tener la secuencia perfectamente estudiada. Mediante La observación sistemática intenté agrupar la información a partir de ciertos criterios fijados previamente. Requiere categorizar hechos, conductas y eventos que se han de observar.

Se tuvo que aprender sobre las técnicas de Inteligencia Artificial existentes [6] y catalogar el *chatbot* en el que yo trabajaba en cuáles se enfocaba. Llegué a la conclusión de que estaba formado por dos:

- *Machine learning*: es la rama que busca el desarrollo de técnicas que permitan a los ordenadores aprender por sí mismos. Para ello se crean programas que pueden generalizar ciertas respuestas a partir de información sin estructurar, que se suministra como ejemplos. Con ello, se induce al conocimiento por parte del ordenador.
- Sistemas expertos: sistema de información que se basa en el conocimiento de un área de aplicación de gran complejidad y muy específica, en este caso el tema de la banca. Sirve como asistente consultor y experto para los usuarios de su interfaz, es el *chatbot* quien ayudará al usuario a revolver la duda sobre un concepto o transacción que no tenga clara. Son entornos que proporcionan respuestas sobre problemáticas muy específicas. Es una fusión entre aprendizaje automático y establecer algunas reglas fijas. Al final ambos tienen el mismo objetivo: construir una máquina inteligente que pueda realizar tareas inteligentes y dar respuestas inteligentes.

La aplicación podrá contar con implementaciones o posibles mejoras que vayan surgiendo con el avance del tiempo ya que las necesidades de una empresa suelen ir variando conforme a la época o etapa en la que se encuentren. Pero la proyección de esta aplicación se estima que tenga un uso indefinido. Al final se trata de una herramienta que resulta útil para la extracción de datos, por lo que se cuenta con que tenga una duración permanente.

## 7. Desarrollo

Una vez se conoce el objetivo del informe y las necesidades de cada área a las que va dirigido este proyecto se procede a la organización de toda la información. Para comenzar con la creación del informe, es necesario establecer los KPIs que necesitamos, qué filtros se van a utilizar, cómo representar la información y con qué tablas de datos contamos para dicha extracción.

### 7.1 Tablas y extracción de datos

Este sería el primer paso para la obtención y estructuración de la información que necesitamos. Para una extracción de datos adecuada es necesario conocer todas las bases de datos que existen y qué datos contienen cada una. Es fundamental hacer una recopilación de todos los datos con los que contamos en cada tabla para ver cuales nos resultarían necesarios y serían convenientes para poder relacionarlas entre sí. También se analizan qué datos tienen en común cada una de las tablas para poder enlazar unas con otras. Estas tablas ya existían en la base de datos con la que ellos cuentan, así que solo tuve que acceder a esta y estudiarlas.

Enfocándonos en el objetivo de crear un sistema automático para la generación de informes, encontramos la base para un correcto funcionamiento en el buen uso de los datos. Para conseguir esto, se necesitaba hacer un estudio de dónde se encontraban todos los que necesitábamos. Los datos se almacenan en tablas y cada tabla contiene unos datos diferentes, salvo en algunos casos que comparten alguno de ellos y es esto lo que nos permitirá relacionar unas tablas con otras. Se realizó un estudio de qué datos estaban en qué tabla para así poder tenerlo todo organizado y fuese más sencillo la extracción de los datos. A continuación, explico la organización que seguí, las tablas que contenía y su respectivo nombre:

- Preguntas que realizan los empleados al *chatbot*:  
ELE\_CCC.V\_VV\_COGNI\_FORMU\_OFICINA  
ELE\_CCC.V\_CC\_TM\_INTENTS\_OFICINA
- *FAQs* mostradas durante la consulta:  
ELE\_CCC.V\_CC\_TM\_FAQS\_OFICINA
- Espacios de trabajo y áreas en las que se organiza la empresa:  
ELE\_CCC.V\_CC\_TM\_ESPACIOS\_OFICINA  
ELE\_CCC.V\_CC\_TM\_AREAS\_OFICINA
- Categoría y subcategoría de las *FAQs*:  
ELE\_CCC.V\_CC\_TM\_TOPICS\_OFICINA  
ELE\_CCC.V\_CC\_TM\_SUBTOPICS\_OFICINA
- Organización de centros por zona geográfica:  
ELE\_CTM.CTM\_TM\_PH  
ELE\_CCC.V\_CC\_COGNI\_AGENTES\_OFICINA

Hay ciertas variables que van a ser comunes en casi todas las tablas y eso va a permitir que se puedan relacionar todas entre sí. En este caso resultan ser:

- Código de chat -> PK\_CHAT
- Código de la *FAQ* -> PK\_FAQ

## 7.2 *KPIs* utilizados

Este sería el paso natural tras la extracción de datos realizada con el código anterior, puesto que podríamos obtener una gran variedad de indicadores, *KPIs*, que nos van a ayudar a la hora de facilitar los datos en el informe. Algunos de estos indicadores se pueden obtener directamente del código, pero otros se consiguen mediante una fórmula (fig. 10). Esta fórmula se crea mediante variables sacadas a través del código o

simplemente mediante cálculos, como puede ser la obtención de un dato en formato porcentaje o la media de formulaciones realizada.

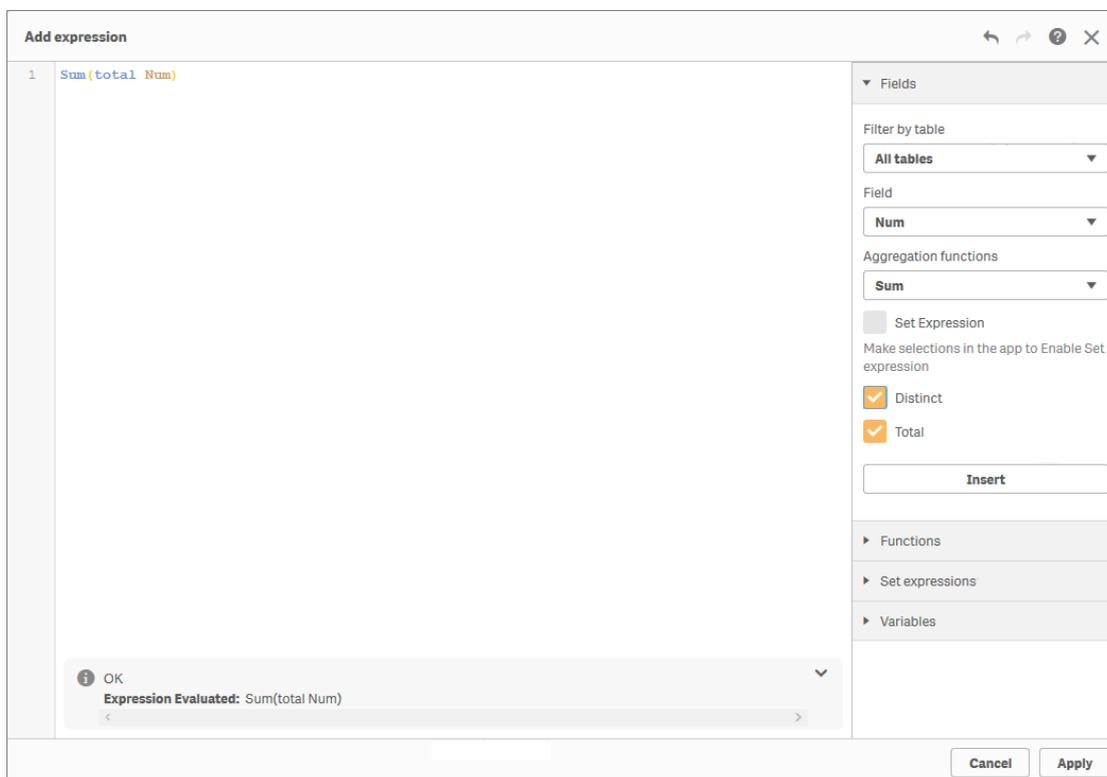


Fig 10: Ejemplo de la creación de indicadores mediante fórmula.

Para crear un indicador o fórmula es importante saber diferenciar entre parámetro y variable ya que cada uno funciona de una manera diferente. No se puede mostrar de igual manera y hay que seguir un procedimiento especial para cada una. La definición de cada una es la siguiente:

- **Parámetro:** permite intercambiar los datos existentes entre procedimientos o funciones guardadas, es decir, pasar el valor de los datos a un procedimiento o función ya guardado. Una de sus características principales es que, si no se establece un valor para el código de retorno, este devolverá como consecuencia un 0.

- **Variable:** mantiene un valor único respecto a datos específicos. Este valor se devuelve como un valor de retorno de una función. Cuando se usa un variable es sustituida por su valor o la definición de la variable. Por ejemplo, puede usarse como contador en una función bucle.

Tanto para las variables como para los parámetros una vez dentro de la creación del informe se les puede cambiar el nombre para que sea más intuitivo (fig. 11). El nombre que se utiliza para la programación a veces no resulta igual de entendible para el usuario final, por lo que se puede poner cualquier otro.



Figura 11: Información de una variable

Una vez conocida la diferencia entre estos dos conceptos, voy a enumerar y explicar los *KPIs* que se han utilizado y el motivo de su necesidad:

- **Vol\_FAQs:** indica la cantidad total de *FAQs* que hay. Conforme se aplica algún filtro, este indicador muestra solo la cifra de *FAQs* que hay correspondiente a los datos específicos que se solicitan.

El estado de las *FAQs* se fija en su software, ahí es donde se crean, modifican y se eliminan. Este corpus está conectado con el cerebro del *chatbot*, por lo tanto,

cuando se cambian estados en el software por consecuencia se cambian también en el *chatbot*.

- **FAQ\_produccion:** indica cuántas son las que contienen la información completa de lo que se quiere mostrar.
- **FAQ\_inactiva:** indica cuántas se encuentran dentro del cerebro del *chatbot* pero no se van a mostrar debido a que en el software las han deshabilitado por información antigua o en desuso.
- **FAQ\_nueva:** indica cuántas se han añadido recientemente. El título de nuevas les dura un mes.
- **FAQ\_corta:** se muestra la información directamente en el flujo de la conversación.
- **FAQ\_guiada:** la respuesta a la consulta sigue unos pasos que le va indicando el *chatbot* durante la conversación.
- **FAQ\_simple:** indica cuántas se muestran en el *frame* de la derecha de la conversación.
- **FAQ\_sinmostrar:** número de *FAQs* que el *chatbot* nunca ha llegado a mostrar. Esto sirve para poder hacer limpieza del corpus de *FAQs*.
- **Absorción global:** porcentaje de consultas satisfactorias del *chatbot*.
- **Absorción intent:** porcentaje de consultas resueltas satisfactoriamente cuando se acaba de iniciar un servicio.
- **Absorción solo cognitive:** porcentaje de consultas resueltas satisfactoriamente a través solo del *chatbot*.
- **Hist\_Absorción:** muestra el histórico de la absorción. Es independiente a cualquier filtro porque muestra una información fija.
- **Hist\_EscaladoFacturable:** muestra el histórico de la absorción que genera un coste. Los escalados a llamada o chat se suman y se muestra su histórico. Es independiente de cualquier filtro ya que es un valor fijo.
- **Visualización:** número de veces que una *FAQ* ha sido mostrada por el *chatbot* o que el empleado a clicado sobre ella.
- **C2C:** número de veces que se ha ofrecido el *click to call* al empleado que ha realizado una consulta.

- **Suma Chat y C2C:** suma total de las conversaciones que se escalan a llamada y las que pasan a *chat* con un agente. Esta suma son los escalados que tienen un coste para la empresa.
- **Chat con agente:** número de consultas que no se han resuelto con el *chatbot* y han pasado a una conversación con un humano.

Cuando el *chatbot* escala al siguiente nivel las consultas se pueden dividir en dos casos. Las solicitadas, significan que, aunque se haya lanzado la petición, no ha sido recogida por el siguiente nivel. Esto puede darse debido a un tiempo de espera demasiado largo, a que no había ningún operador libre o porque quien ha hecho la consulta ha cerrado el *chat*. Los atendidos son los que una vez ha sido escalada, la petición ha sido recogida y pasa a ser resuelta por ese nivel.

- **Vol\_Hidden\_Atendido:** cantidad de consultas que han pasado a la “persona oculta” y están siendo atendidas.
- **Vol\_Hidden\_Solicitado:** cantidad de consultas que se han escalado a la “persona oculta” pero que no han llegado a ser atendidas por esta.
- **Vol\_Chat\_Aten:** cantidad de consultas que han pasado a un *chat* con un humano y han sido atendidas.
- **Vol\_Chat\_Sol:** cantidad de consultas que se han escalado a un chat con un humano pero que no han llegado a ser atendidas.

El SLA (Service Level Agreement) como sus siglas indican, es un contrato en el que se fija el servicio concreto que se quiere dar. En este caso el contrato iba relacionado con el tiempo, la empresa de que gestiona el *Call Center* tiene que dar una respuesta en menos de treinta segundos.

- **SLA\_Hidden:** muestra el tiempo medio en el que la “persona oculta” responde a las consultas.

- **SLA\_Chat:** muestra el tiempo medio en el que el humano con el que se va a iniciar la conversación inicia el *chat*.
- **Num\_Conversaciones:** indica cuántas conversaciones se tienen en total. Cada formulación no es una conversación, por lo que una conversación puede contener más de una formulación.
- **Escalado:** número de veces que el *chatbot* no ha podido resolver la consulta y ha tenido que escalarla a otro nivel.
- **Preguntas por empleado:** media de consultas que realiza un empleado en el *chatbot*.
- **Numero formulaciones:** cantidad de preguntas que realiza un empleado al *chatbot*. Muestra la suma total de todas las formulaciones que realizan los empleados, conversaciones y temas. Se puede filtrar para que se muestren los datos de días, empleados o categorías en concreto.

### 7.3 Filtros de información para el informe

El objetivo que tienen los filtros es poder obtener la información específica que se necesita. Estos filtros ayudan a profundizar en la búsqueda concreta de lo que se pide, ya que de primeras en el informe sale toda la información general con la que cuenta. Lo primero que se ve son los datos más recientes en referencia a la fecha en la que tenemos cargados los datos y por eso era necesario la creación de filtros concretos en función a las necesidades.



Figura 12: Distribución global de los filtros

Para un uso más cómodo y práctico de la aplicación, se optó por tener los mismos filtros y en el mismo orden en cada una de las páginas (fig. 12), así la experiencia de usuario sería mucho más fácil. El orden y su correspondiente explicación es la siguiente:

- Los dos primeros filtros van relacionados con la fecha. El que aparece primero es un **calendario** (fig. 13) en el cual puedes seleccionar un día concreto, uno o varios días sueltos, semanas completas, etc. Mientras que en **PK\_ANYOMES** (fig. 14) se te permite filtrar la información de un mes completo:

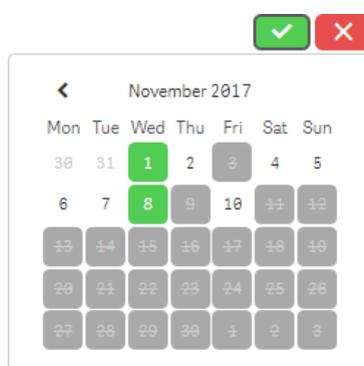


Figura 13: Calendario general



Figura 14: Filtro por meses

- **DES\_AREA** te permite seleccionar el área o áreas específicas de las que necesitas obtener la información.
- **ET** son los espacios de trabajo.
- **Categoría y subcategoría** de las *FAQs*. Están relacionadas entre sí, porque al seleccionar una categoría te va a indicar que subcategorías se encuentran dentro

de ese filtro. Se puede seleccionar directamente solo la subcategoría que se necesita.

- **TITULO\_FAQ** para poder ir directo a una en concreto.
- **ORIGEN\_LINK** por si se quiere comprobar, por ejemplo, qué *FAQs* vienen de una dirección concreta. Aparecen todos los orígenes de donde se encuentran las *FAQs* y, por ende, desde donde puede proceder ese *click*.

Todos los filtros son compatibles entre sí. No hay ninguno que no pueda estar junto a otro ya que al final son diferentes capas de información. No habrá condicionantes en los filtros por lo que esto permite encontrar un dato de manera más precisa

## 7.4 Código para la creación del informe

Para facilitar la lectura del código para futuras modificaciones que se quieran hacer, hice comentarios de cada bloque del código ya que la manera de programar de cada persona es diferente y así poder ahorrarle tiempo a la siguiente persona que quiera comprender su funcionamiento.

Se trata de relacionar las tablas para conseguir llevar a cabo la extracción de datos deseada. Se van relacionando tablas y conectando datos que pertenecen a las distintas tablas para así poder obtener la traza completa. Todas las tablas no tienen la misma información por lo que es importante relacionarlas bien e ir obteniendo los datos necesarios de cada una sin que se pisen unos datos a otros.

Comenzando a explicar el código, la parte mostrada en la fig. 15 es para fijar la fecha. Está en formato año y mes y en esta línea de código es desde donde se nos permitirá seleccionar el mes del cual queremos los datos.

```

1  /*Cambiar el 202003 segun La fecha hasta La que se quieran cambiar Los datos
2  si se quiere febrero seria 202002*/
3  set MES = 202003;

```

Figura 15: Selección de la fecha

La fig. 16 realiza la creación a los links. Se cargan las distintas procedencias de donde puede venir el *click* de la *FAQ* y luego la creación de la URL para el *chatbot*.

```

5  //Se cargan Los distintos origenes de Los Links por Los que Luego se filtran
6  LOAD * INLINE [
7    ORIGEN_LINK
8    link_alertas
9    link_circular
10   link_actualidad
11   link_myplanning
12   link_NOMBREPORTAL
13 ];
14
15 //URL para La creacion del Link para La visualizacion de FAQs a través de NEO
16 LOAD * INLINE [
17   URL_DEFECTO
18   https://gppini.pro.absisccloud.lacaixa.es/util/redirector/redirectorTF7.html?tf7_url=/CA.OFI/CCcognitivo/inicio.tf7?_flowId=inicio
19 ];
20

```

Figura 16: Origen y carga del link

En el siguiente bloque (fig. 17), se hace referencia a una de las visualizaciones que se muestran en la página 1 del informe (fig. 24), el *Wordcloud*, y el apartado de *map2* sirve para su creación. Consiste en la agrupación de las preguntas que tienen mayor número de búsquedas e ir agrupándolas en bloques. Cuantas más consultas tenga una pregunta, más grande aparecerá en la nube. La clave de este cuadro es que preguntas que se realizan con el mismo contenido, pero con maneras distintas de formularlo, se asocien a la misma nube y así se pueda ver de manera general los temas específicos que se consultan.

```

21 map2:
22 mapping LOAD *
23 Inline [
24 AttCode, Attribute
25 %, ''
26 ?, ''
27 ¿, ''
28 !, ''
29 ñ, 'n'
30 á, 'a'
31 à, 'a'
32 é, 'e'
33 è, 'e'
34 í, 'i'
35 ì, 'i'
36 ò, 'o'
37 ó, 'o'
38 ù, 'u'
39 ú, 'u'
40 ] ;

```

Figura 17: Creación del *Wordcloud*

La fig. 18 corresponde a la carga de datos para poder crear el histórico de un año por espacio de trabajo. En el mismo código se han creado variables que luego se utilizarán para la visualización del informe, como son:

- VOLUMEN\_HISTÓRICO
- ESCALADO\_HISTORICO\_F: hace referencia al que se factura
- ESCALADO\_HISTORICO
- ESCALADO\_CHAT

El histórico en las variables anteriores, hace referencia a todo un año, es independiente a la fecha en la que se carguen los datos porque no funciona por mes. Se muestra una evolución anual, por lo que no le afectará el filtro de la fecha.

---

```

/*Apartado donde se cargan Los datos para crear el historico de un año por ET
se cargan Los historicos del volumen y del escalado general y facturable*/
Historicos:
LOAD PK_ANYOMES,
COD_ET as ET,
VOLUMEN_HISTORICO,
ESCALADO_HISTORICO_F,
ESCALADO_HISTORICO,
ESCALADO_CC,
ESCALADO_CHAT;

Select PK_ANYOMES,
COD_ET,
count(distinct PK_CHAT) as VOLUMEN_HISTORICO,
sum(CASE WHEN (IND_ESC_CHAT_AGENT+IND_ESC_CC) > 0 THEN 1 ELSE 0 END) AS ESCALADO_HISTORICO_F,
sum(CASE WHEN (IND_ESC_CHAT_AGENT+IND_ESC_CC+IND_ESC_HIDDEN) > 0 THEN 1 ELSE 0 END) as ESCALADO_HISTORICO,
sum(CASE WHEN (IND_ESC_CHAT_EFEC)='1' THEN 1 ELSE 0 END) as ESCALADO_CHAT,
sum(CASE WHEN (IND_ESC_CC)='1' THEN 1 ELSE 0 END) as ESCALADO_CC
from ELE_CCC.V_CC_COGNI_FORMU_OFICINA where PK_ANYOMES >= $(MES)-100
group by PK_ANYOMES,COD_ET
order by PK_ANYOMES;

```

Figura 18: Bloque de los históricos

En el segundo bloque que termina en la parte 5 del bloque (fig. 19, 20, 21 y 22), es donde se encuentra la matriz de la aplicación. Reside el cuerpo de cómo va a funcionar la aplicación y como está organizada. La aplicación se podía plantear de varias maneras, pero al final se ha enfocado de modo que todo está organizado según la formulación de cada pregunta (en una conversación puede haber más de una formulación). Los datos se pueden obtener individuales desde el momento que el empleado realiza la consulta hasta que esta termina. Se puede seguir la traza completa que sigue la operación para así no perder detalle y poder ser conscientes de donde residen la mayoría de los problemas.

```

/*En este apartado esta todo el cuerpo de La app
SE ORGANIZA POR FORMULACION*/
Cognitive:
LOAD PK_ANYOMES,
PK_CHAT as ID_CHAT,
DES_INTENT as Intent,
PK_FAQ as ID_FAQ,
DES_CASTELLANO as Titulo_FAQ,
DES_CATALAN as Titulo_FAQ_Cat,
date(FECHA_PREGUNTA) as FECHA_PREG,
DES_ESPACIO as ET,
FAQ_WC_1 as Respuesta,
PREGUNTA,
ESCALADO,
COD_RESULTADO as Resultado_conv,
IND_FAQ_MOSTRADA as FAQ_mostrada,
FK_USUARIO as ID_Usuario,
FAQ_TIEMPO_LECTURA,
COD_RESULTADO_FINAL,
ESCALADO_FACTURABLE,
ESTADO,

```

Figura 19: Comienzo del corpus del código

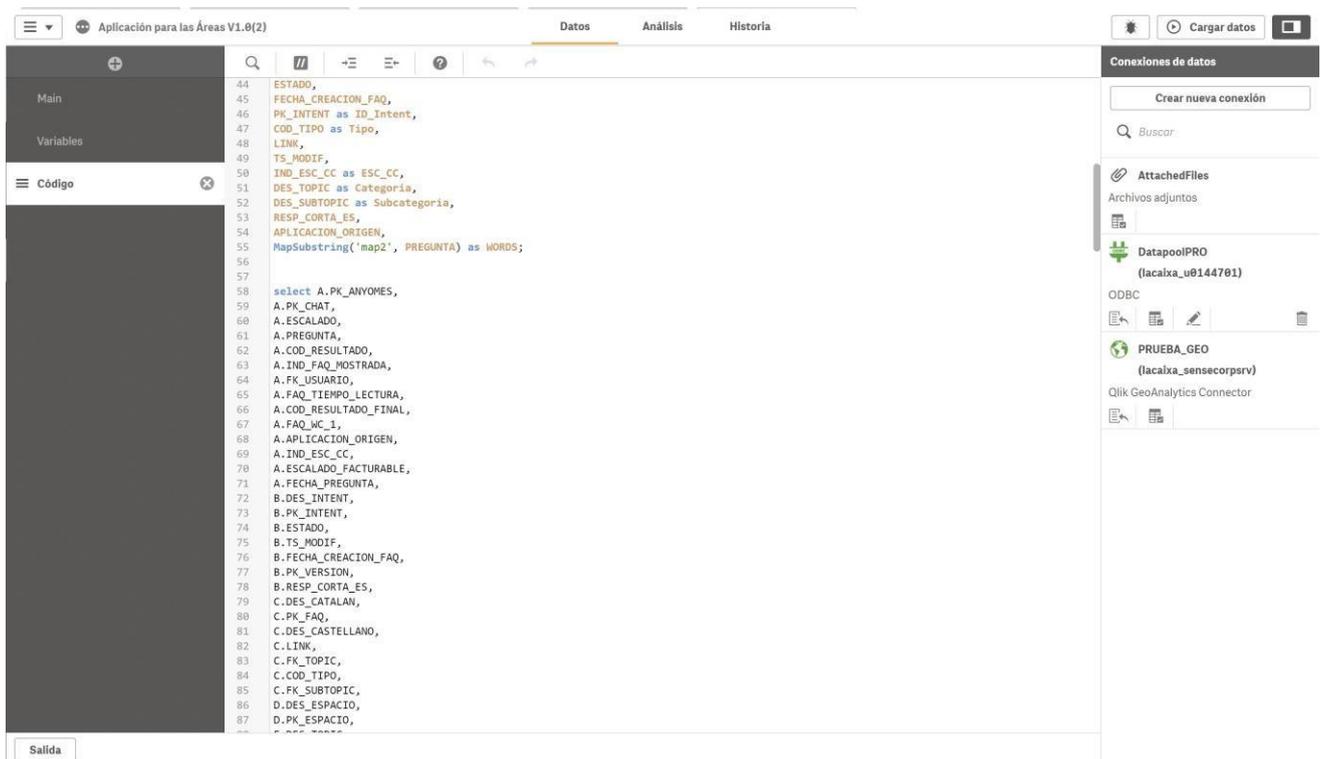


Figura 20: Continuación del código base

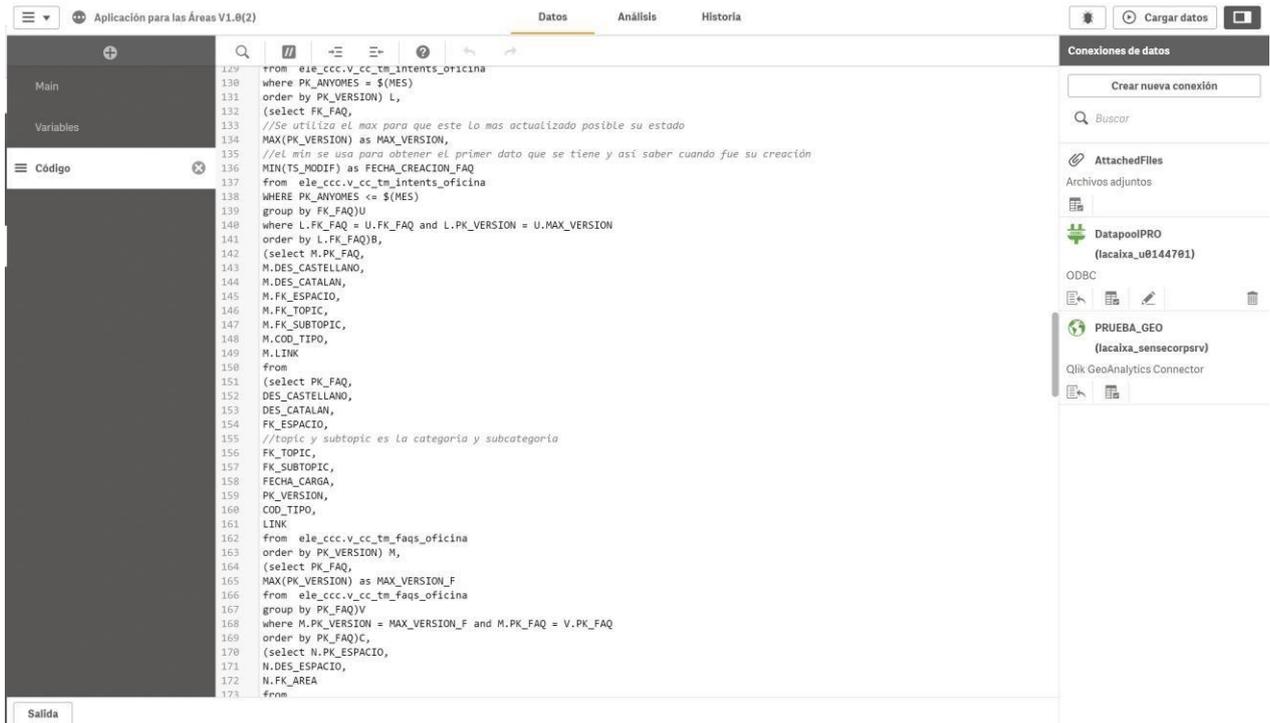


Figura 21: Continuación II del código base

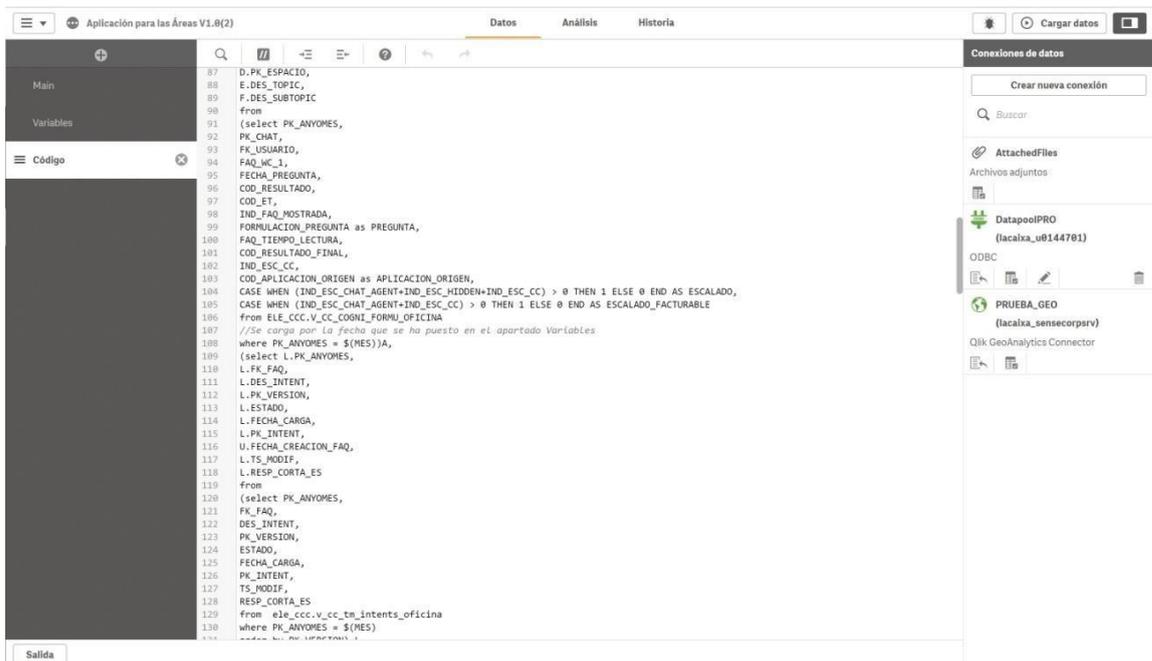


Figura 22: Continuación III del código base

En la fig. 23 se encuentra la secuencia donde se consigue obtener el *topic* y *subtopic* de una FAQ. Con este código se consigue la posibilidad de agruparlas y así poder filtrarlas en el informe. Al final del bloque, es donde se realiza la relación entre tablas para poder conseguir su conexión.

```

172 N.FK_AREA
173 from
174 (select PK_ESPACIO,
175 DES_ESPACIO,
176 FK_AREA,
177 PK_VERSION
178 from ele_ccc_v_cc_tm_espacios_oficina ) N,
179 (select PK_ESPACIO,
180 MAX(PK_VERSION) as MAX_VERSION_E
181 from ele_ccc_v_cc_tm_espacios_oficina
182 group by PK_ESPACIO) S
183 where N.PK_ESPACIO= S.PK_ESPACIO AND N.PK_VERSION = MAX_VERSION_E),
184 (select Z.DES_TOPIC,
185 Z.PK_TOPIC,
186 Z.FK_ESPACIO
187 from
188 (select DES_TOPIC,
189 PK_TOPIC,
190 PK_VERSION,
191 FK_ESPACIO
192 from ELE_CCC_V_CC_TM_TOPICS_OFICINA) Z,
193 (select PK_TOPIC,
194 MAX(PK_VERSION) as MAX_VERSION_C
195 from ELE_CCC_V_CC_TM_TOPICS_OFICINA
196 group by PK_TOPIC) T
197 where Z.PK_TOPIC= T.PK_TOPIC AND Z.PK_VERSION = MAX_VERSION_C),
198 (select Q.DES_SUBTOPIC,
199 Q.DES_TOPIC,
200 Q.FK_TOPIC
201 from
202 (select DES_SUBTOPIC,
203 PK_SUBTOPIC,
204 PK_VERSION,
205 FK_TOPIC
206 from ELE_CCC_V_CC_TM_SUBTOPICS_OFICINA) Q,
207 (select PK_SUBTOPIC,
208 MAX(PK_VERSION) as MAX_VERSION_S
209 from ELE_CCC_V_CC_TM_SUBTOPICS_OFICINA
210 group by PK_SUBTOPIC)
211 where Q.PK_VERSION = MAX_VERSION_S and Q.PK_SUBTOPIC=0.PK_SUBTOPIC)F
212 //Las conexiones entre tablas
213 where A.FAQ_MC_I(+) = B.DES_INTENT and C.PK_FAQ = B.FK_FAQ and D.PK_ESPACIO(+) = C.FK_ESPACIO
214 and E.FK_ESPACIO (+) = D.PK_ESPACIO and E.PK_TOPIC (+) = C.FK_TOPIC and F.PK_SUBTOPIC(+) = C.FK_SUBTOPIC;
215

```

Figura 23: Creación del *topic* y *subtopic*

En la siguiente parte del código (fig. 25), se busca obtener la información geográfica. Esta documentación es necesaria para luego plasmarla en el informe como una visualización. Se busca poder formar una cadena (fig. 24), es decir, asociar a cada empleado a una oficina la cual está en una ciudad que reside en una autonomía que pertenece a un territorio concreto creado en Caixabank. Entonces, una vez se tiene toda la traza, se podrá organizar el mapa según la zona geográfica que se prefiera. Las zonas geográficas se organizan en este orden:



Figura 24: Organización de zonas geográficas

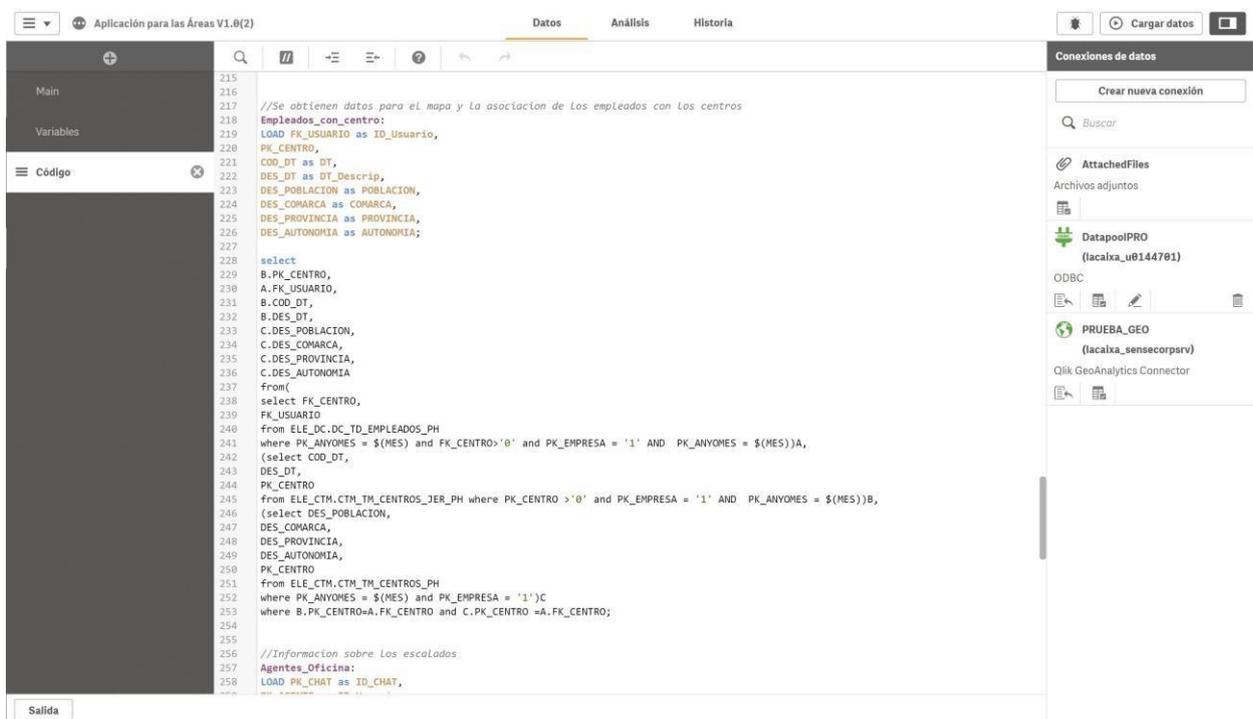


Figura 25: Organización geográfica

Siguiendo todo el código, ahora tiene lugar la parte (fig. 26) donde se encuentra la relación de los escalados. Se analiza la traza de la llamada, desde que se inicia la conversación en el *chat* hasta que o termina satisfactoriamente en el o se escala a otro nivel (fig. 5). Una vez se escala a otro nivel ya no cuenta esa información para el *chatbot*, se genera una consulta nueva a la que se le seguirá la traza desde el nivel en el que se encuentra. Cuando en un *chat* no se resuelve la consulta del usuario, se escala. En este caso se buscaba obtener el análisis de cuántas resultan no satisfactorias.

```

LOAD PK_CHAT as ID_CHAT,
PK_AGENTE as ID_Usuario,
FK_TIPO_AGENTE,
FK_TIPO_AGENTE_NEXT,
FECHA_INI_PETICION,
date(FECHA_INI_ATENCION) as FECHA_INICIO_AT,
date(FECHA_FIN_ATENCION) as FECHA_FIN_AT,
NUM_SEG_ESPERA,
NUM_SEG_ATENCION,
DES_ACCIONES_TIP,
DES_ET as Categoria_Chat,
date(FECHA_DATOS) as FECHA_DATOS;

select PK_CHAT,
PK_AGENTE,
//este paso se hizo para limpiar las tablas de datapool y poder obtener bien los datos
CASE WHEN FK_TIPO_AGENTE='0' THEN FK_TIPO_AGENTE_NEXT ELSE FK_TIPO_AGENTE END as FK_TIPO_AGENTE,
CASE WHEN (FK_TIPO_AGENTE='0' and FK_TIPO_AGENTE_NEXT is not null) then null ELSE FK_TIPO_AGENTE_NEXT END as FK_TIPO_AGENTE_NEXT,
CASE WHEN FECHA_INI_PETICION is null THEN FECHA_INI_ATENCION -30 ELSE FECHA_INI_PETICION END AS FECHA_INI_PETICION ,
FECHA_INI_ATENCION,
FECHA_FIN_ATENCION,
NUM_SEG_ESPERA,NUM_SEG_ATENCION,
DES_ACCIONES_TIP,
DES_ET,
FECHA_DATOS
from ELE_CCC.V_CC_COGNI_AGENTES_OFICINA
where PK_ANYOMES = $(MES) and
(FK_TIPO_AGENTE_NEXT <> '0' OR FK_TIPO_AGENTE <> '0' ) and PK_AGENTE <> 'system';

```

Figura 26: Obtención de la traza de la llamada

Para concluir con la *query* creada, organizamos las áreas y espacios con los que están relacionadas las *FAQ* (fig. 27 y 28). Se relacionan varias tablas desde donde se encuentra la matriz de la información y se relaciona para obtener ese dato y poder filtrar también con esa información. La parte del código que aparece en la fig. 27 sirve para mantener la clasificación actualizada y que no se confunda con conceptos antiguos.

```

//Obtener La descripcion de espacio y area de manera actualizada
AREAYESPACIO:
LOAD DES_ESPACIO,
DES_AREA;

select DES_ESPACIO,
DES_AREA
from
(select PK_ESPACIO,
DES_ESPACIO,
FK_AREA
from
(select PK_ESPACIO,
DES_ESPACIO,

```

Figura 27: Actualización de datos para esos campos

```

//Obtener La descripcion de espacio y area de manera actualizada
AREAYESPACIO:
LOAD DES_ESPACIO,
DES_AREA;

select DES_ESPACIO,
DES_AREA
from
(select PK_ESPACIO,
DES_ESPACIO,
FK_AREA
from
(select PK_ESPACIO,
DES_ESPACIO,
FK_AREA,
PK_VERSION,
FECHA_CARGA
from ele_ccc.v_cc_tm_espacios_oficina
order by PK_VERSION) N,
(select MAX(PK_VERSION) as MAX_VERSION_E
from ele_ccc.v_cc_tm_espacios_oficina)
where N.PK_VERSION = MAX_VERSION_E)D,
(select PK_AREA,
DES_AREA
from
(select PK_AREA,
DES_AREA,
PK_VERSION,
FECHA_CARGA
from ele_ccc.v_cc_tm_areas_oficina
order by PK_VERSION) P,
(select MAX(PK_VERSION) as MAX_VERSION_A
from ele_ccc.v_cc_tm_areas_oficina)
where P.PK_VERSION = MAX_VERSION_A)E
where E.PK_AREA(+)=D.FK_AREA;

```

Figura 28: Parte 9 del código

## 7.5 Comprobación de datos obtenidos

Una vez que ya se tiene todo el código, se han creado los indicadores, filtros y visualizaciones, se tiene que comprobar que los resultados obtenidos sean los correctos. La comprobación de datos se realizó de dos maneras.

La primera, se basa en la comprobación de que todos los datos que aparecen en el informe son coherentes, es decir, que en ningún momento obtenemos números negativos y que se muestran cifras dentro de los rangos estudiados. Se van aplicando los diferentes filtros y a la vez se van observando como varían los datos mostrados. Cuando se ha observado que no hay ningún error notable, pasamos al siguiente método.

La segunda, es una capa más profunda de verificación de datos. Se introducen los filtros necesarios para poder comprobar con otros informes anteriores los resultados que muestra el automatizado. Además, contamos con la posibilidad de comprobar algunos datos con softwares anteriores donde se encuentra esa información, por lo que se pueden verificar los resultados de este informe con los del corpus original.

Cuando ya se tienen los datos comprobados y, por tanto, el código es correcto, aún queda el último paso. Para que todo el mundo pueda utilizar este informe, tiene que pasar por unos controles de calidad de la empresa. El código tiene que seguir una serie de patrones estéticos para que la aplicación pase a ser oficial. Los patrones que se tuvieron que seguir para pulir el código fueron:

- No podía haber espacios, los nombres de los indicadores tenían que estar separados mediante barras baja, por ejemplo, TITULO\_FAQ.
- Tenían que ser abreviaturas que todo el mundo pudiera entender, por ejemplo, C2C.
- En el código, cada *select* debía ir en una columna distinta.
- Los nombres de los indicadores debían estar en mayúsculas.

## 8. Resultado

Cuando ya se tienen las *queries*, lenguaje estándar que permite traer datos de una o más tablas, actualizar contenidos o eliminarlos, creadas, tablas relacionadas e indicadores necesarios para la obtención de datos, se debía pasar al siguiente paso: la visualización. La correcta representación es un paso fundamental ya que todo el trabajo queda plasmado en un informe con el que los empleados van a tener que trabajar. Una buena visualización permitirá tener una mejor experiencia de usuario y conseguirá fácilmente la autosuficiencia, la cual es una de los objetivos del proyecto.

Teniendo en cuenta los objetivos y características buscadas, se utilizó la herramienta analítica llamada *QlikSense*. Es una plataforma de análisis y visualización de datos, basada en una arquitectura de múltiples nubes. Cuenta con API abiertas, inteligencia aumentada y un motor asociativo de análisis que la hace única. Los *highlights* de la experiencia usuario que puede aportar son los siguientes:

- Indexa y comprende las relaciones en sus datos para una exploración mucho más libre.
- Genera visualizaciones inteligentes, informes personalizados y cuadros dinámicos.
- Permite volúmenes de datos de todos los tamaños sin comprometer los estándares de rendimiento, confiabilidad y gobierno.
- Pensada para el autoservicio ya que tiene una interfaz muy intuitiva.
- Se obtiene un análisis de datos potente, flexible y personalizado que se puede visualizar en cualquier dispositivo (ordenador, móvil...)
- Exploración de datos de una manera libre, con un simple *click*.

De manera global, cada página cuenta con los mismos filtros y en el mismo orden para que así resulte más práctico. Además, aparece en la esquina superior izquierda las fechas entre las que se ha realizado la carga de datos y un menú desde el

cual puedes ir directo a la página que necesitas. Cada pestaña tiene un nombre que hace referencia a la descripción de lo que aparece en la página.

A continuación, voy a mostrar la visualización final de cada página del informe y explicar más detalladamente el motivo de su representación.

En la primera pestaña (fig. 29), se muestran 3 gráficos los cuales nos ayudan a tener una idea general acerca del uso de *chatbot*. La explicación de los gráficos es la siguiente:

- El gráfico principal (número 1 de fig. 29), indica el volumen total de veces que se ha usado el *chatbot*. Esta información es útil para poder observar la evolución del uso que se le está dando, el cual debería ir aumentando con el paso de los meses.
- El *wordcloud* (número 2 de fig. 29) muestra qué es lo que más tienden a preguntar los empleados para así poder ver donde hay que centrarse en ayudarles.
- Desde el mapa (número 3 de fig. 29) se puede observar qué regiones realizan más consultas en proporción a la población correspondiente de cada una.

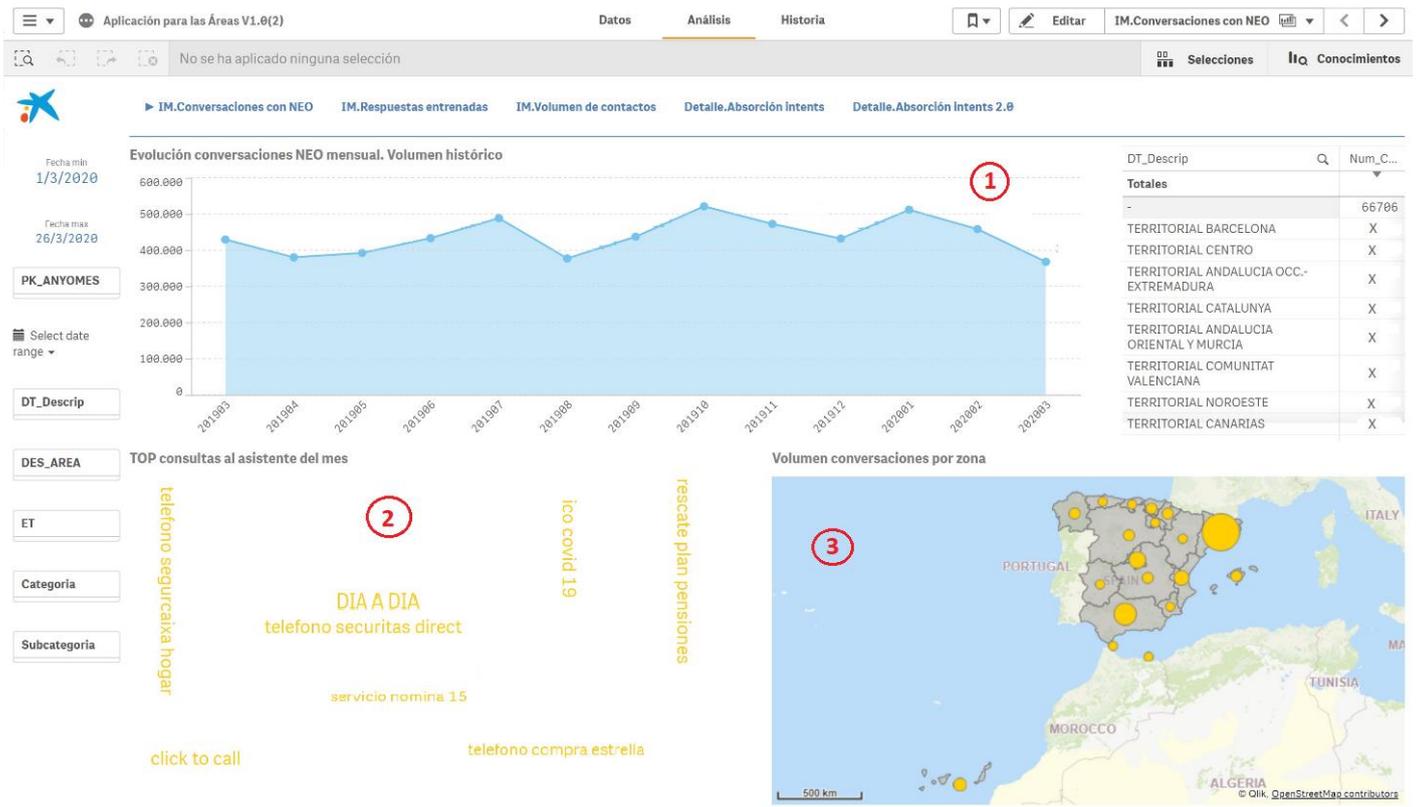


Figura 29: Página 1 del informe

La siguiente pestaña (fig. 30), muestra varios *KPIs* (número 1 de fig. 30) relacionados con la cantidad de *FAQs* que se tienen, se muestran en formato dato y en conjunto en un gráfico de queso (número 4 de fig. 30). Esto es importante debido a que una *FAQ* tiene distintos estados:

- Inactiva: sigue existiendo en el corpus, pero ya no la mostrará el *chatbot*. Esto puede darse debido a que hay que actualizarla, contiene información que ya no está vigente o que nadie la ha consultado y entonces se opta por retirarla.
- Producción: son las que se muestran en el *chatbot*, las que contienen la información completa de lo que se quiere mostrar.
- Nueva: son las que están recién validadas y sirven para poder llevar una estimación de cuántas *FAQs* se van creando a lo largo del mes.

Otra variante de las *FAQs* es que se clasifican en tres tipos:

- Simple: se muestran en un *frame* a la derecha de la conversación. Aparece la visualización de un texto con toda la información solicitada explicada. Un ejemplo puede ser una norma, una pregunta general sobre un *topic* y que entonces te muestre todo el contenido.
- Corta: la información solicitada por el empleado se va a mostrar en el flujo de la conversación como una respuesta breve y concisa. Cuando se hace una pregunta muy concreta, sabe cómo responder a ese contenido específico.
- Guiada: implica que hay una serie de pasos a seguir, por lo que el *chatbot* te irá indicando el camino e irá asegurándose de que no te pierdes. Por ejemplo, podría ser como configurar el teléfono de la oficina o si hay algún problema con el material informático.

Por último, aparecen dos gráficas que van en relación a los *KPIs* mostrados. La primera (número 2 de fig. 30) muestra la absorción que ha tenido la *FAQ*. Esto ayuda para saber cómo están funcionando y si habría que hacer modificaciones en alguna. La segunda (número 3 de fig. 30) sirve para saber dónde residen las dudas, ya que muestran las *FAQs* más consultadas y se compara con cómo de útil es la información ofrecida en esta.

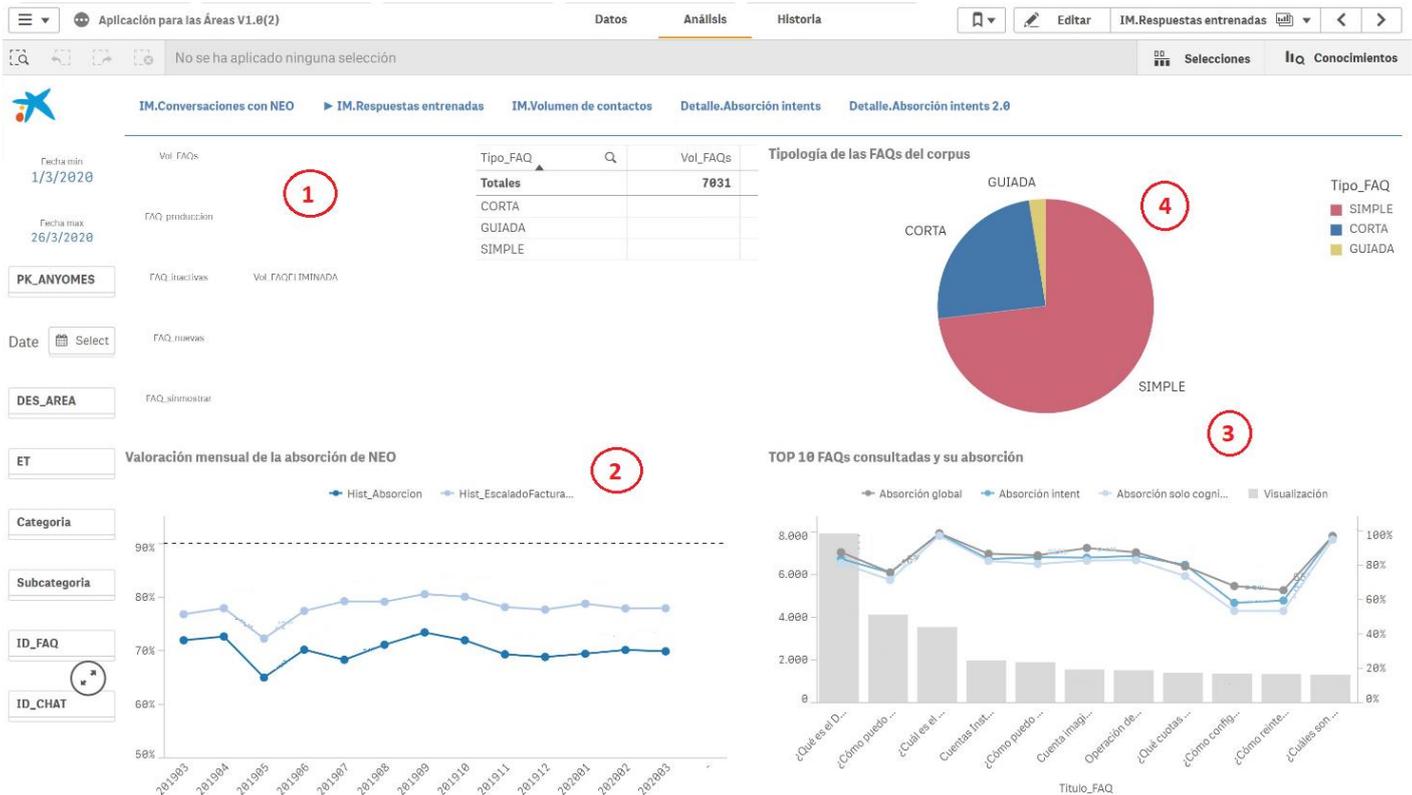


Figura 30: Página 2 del informe

En la pestaña 3, nos enfocamos en la comparativa de llamadas y *chat* (número 1 de fig. 31). Es una página importante ya que nuestro objetivo es reducir el número de llamadas al máximo para así poder reducir mucho en costes.

Una llamada se realiza cuando el empleado desde la consulta del chat no le ha quedado clara la información y entonces se le proporciona un número donde poder dirigirse. De este modo, con la gráfica mensual se puede ir comprobando cómo el uso del chat va mejorando ya que la proporción de llamadas ya no es tan alta.

Otro enfoque que se da en la página es el análisis de las zonas geográficas (número 2 de fig. 31). Es interesante analizar en qué territorios y autonomías lo utilizan más y el motivo que esto conlleva. Se hace una comparativa de las consultas que pasan a llamada por autonomías y del número total de conversaciones por distrito territorial.

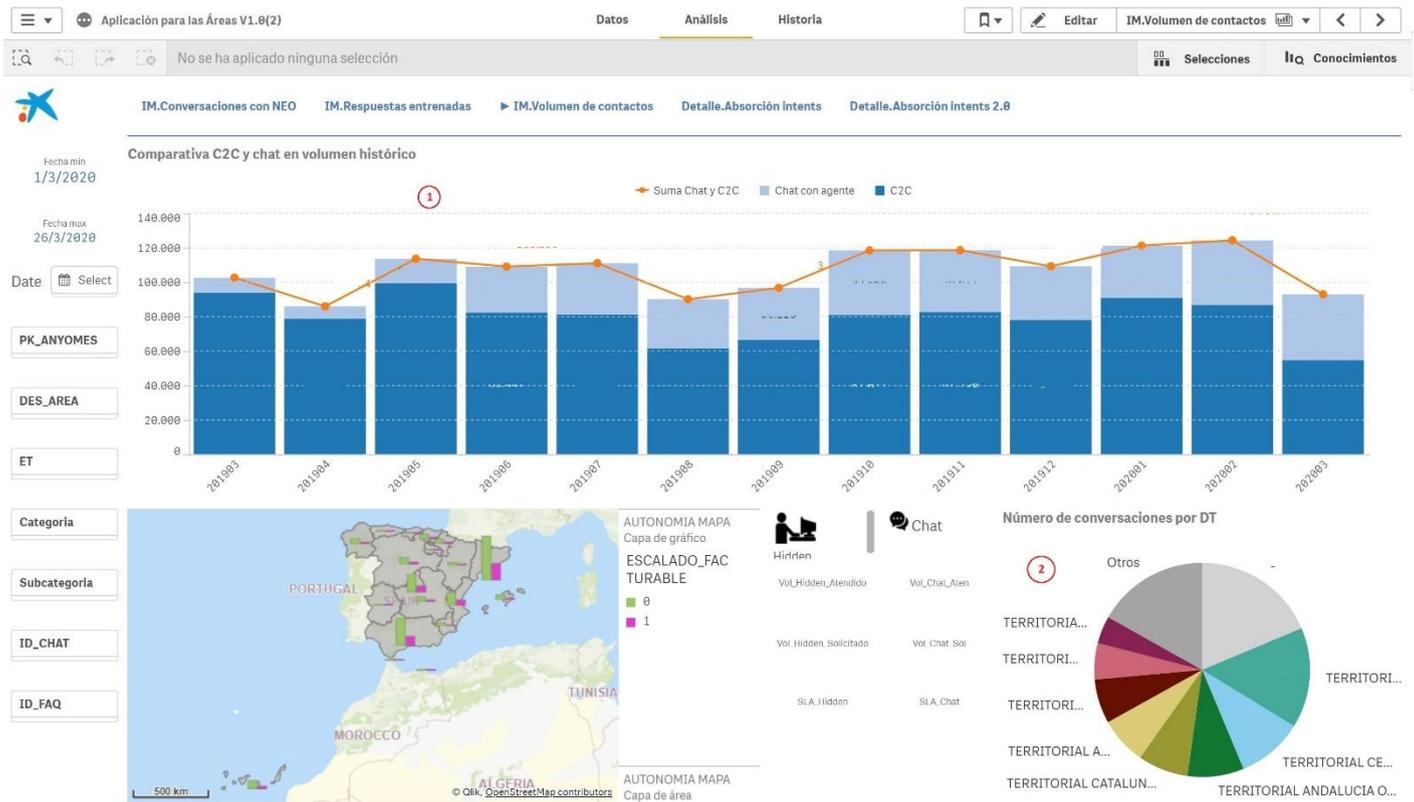


Figura 31: Página 3 del informe

En la cuarta pestaña (fig. 32), se hace un análisis más específico para cada *FAQ*. Se te permite navegar a través de todas las que existen y poder analizar cómo ha funcionado una en concreto (número 1 de fig. 32). Esto sirve para poder hacer limpieza de *FAQs* y saber cuáles son las que mejor funcionan y analizar el por qué para así poder aplicarlo en las que tienen menor tasa de éxito. Además, como todas ellas están categorizadas se puede observar cual se defiende mejor y donde residen el foco de las dudas. Permite tener a simple vista toda la información necesaria de su propio funcionamiento.

La gráfica (número 2 de fig. 32) hace la comparativa de en un día cuántas conversaciones se tienen y cuál ha sido la tasa de éxito en el chat. Cabe también destacar, que cuando aparece un número muy elevado de conversaciones suele darse el motivo de que ha habido alguna implantación nueva para la gente de oficina (nuevas ofertas, nuevas formas de proceder...).

El valor que aparecen en los KPIs (número 3 de fig. 32) son el resultado de los rangos de fechas que se han solicitado.

Además, aparecen los indicadores generales de chat y del paso intermedio entre el *chatbot* y el agente del *Call Center*.

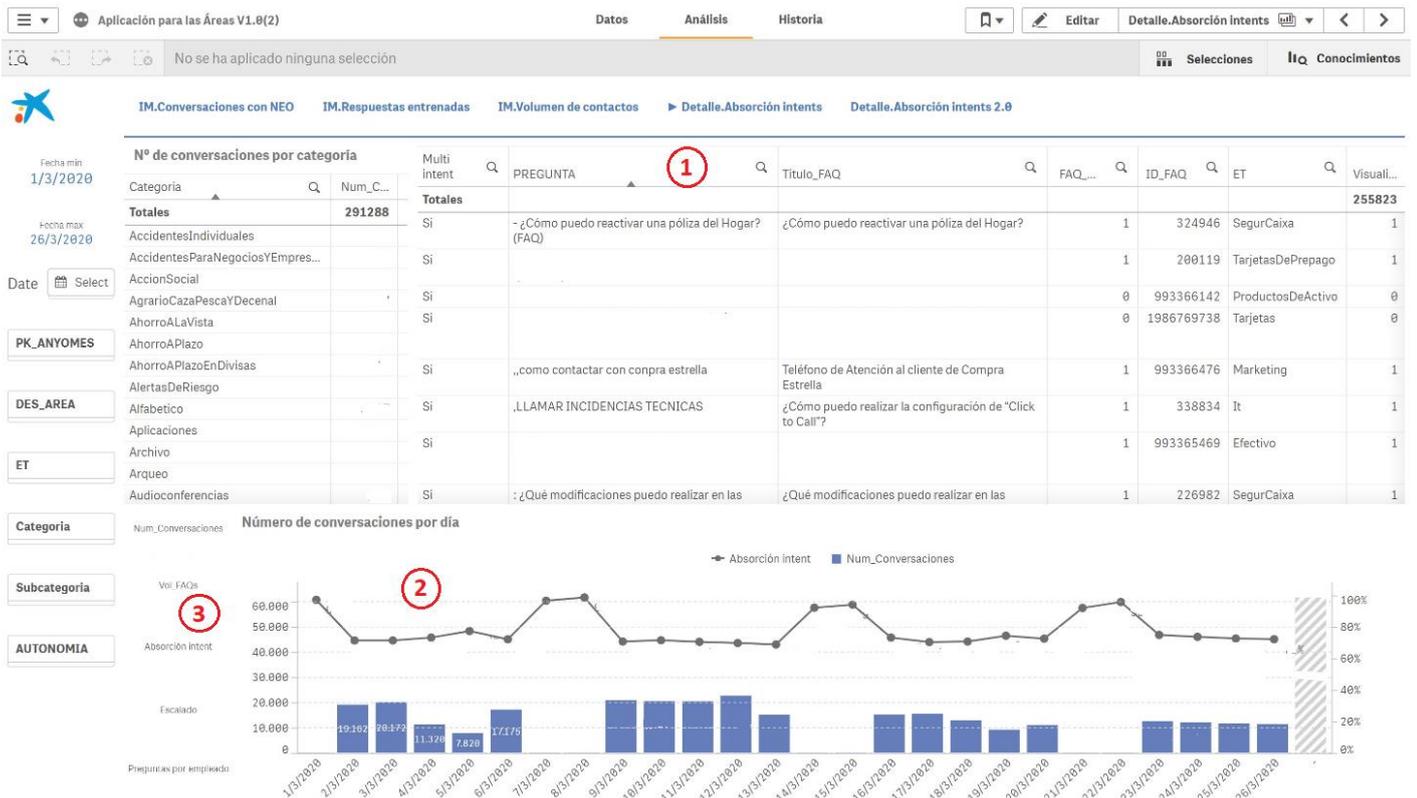


Figura 32: Página 4 del informe

En la quinta y última pestaña (fig. 33), se muestran tres gráficas. La primera (número 1 de fig. 33), hace un análisis diario de las consultas que se hacen por teléfono. Esta gráfica va mostrando la cifra exacta de los 9 últimos días según la fecha en la que hayas filtrado, y luego en la barra de otros, muestra el sumatorio de los días restantes en los que tengas el filtro de fecha. La siguiente gráfica (número 2 de fig. 33), se muestran las visualizaciones que han tenido las *FAQs* según su origen. Esto es importante para saber desde dónde consultan las dudas los empleados y desde dónde leen los artículos para así poder enfatizar en ese sitio concreto la información y de este modo poder ahorrar consultas ya que estaría toda la información bien detallada en el lugar correcto.

La última gráfica (número 3 de fig. 33), muestra el volumen de llamadas respecto a la categoría. Sirve para conocer cuáles son las que suelen generar más preguntas.

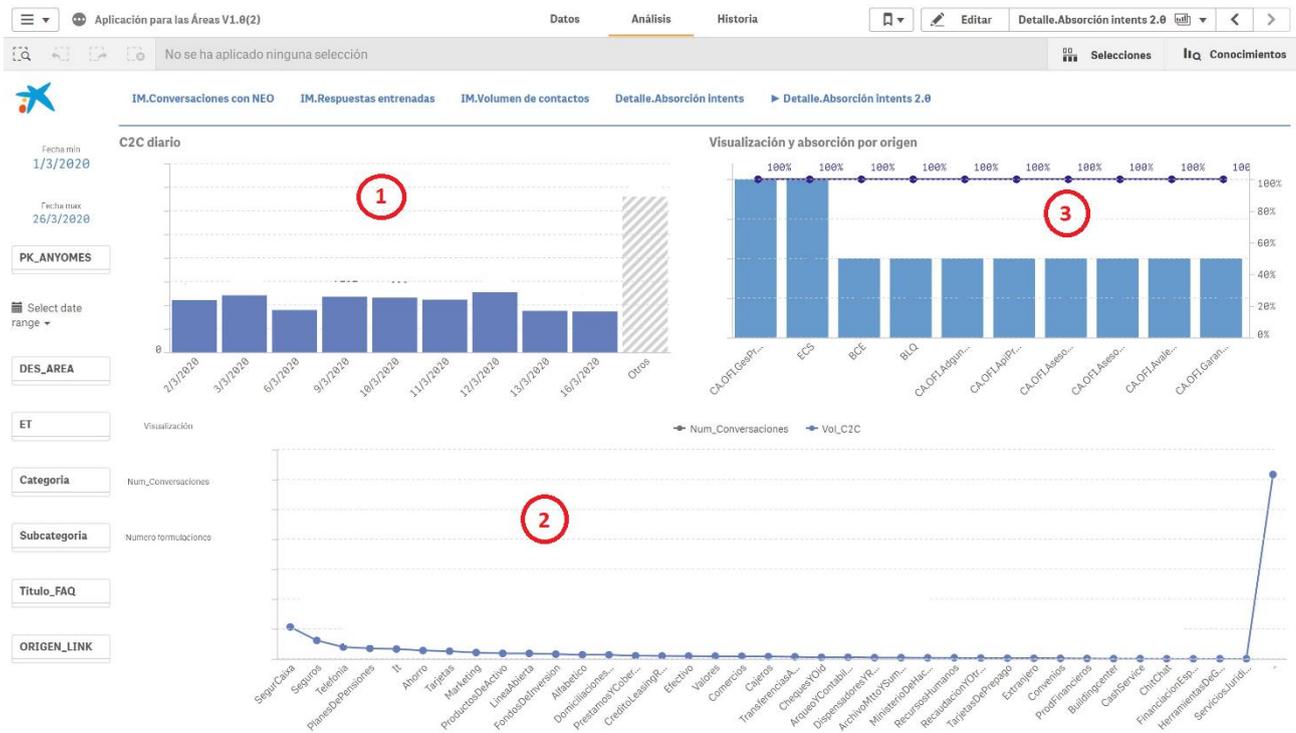


Figura 33: Página 5 del informe

## 9. Conclusiones

El objetivo final alcanzado en este proyecto de automatizar gestiones fue superado con éxito ya que actualmente se continúa usando y se lograron conseguir todos los objetivos y tareas que el proyecto conllevaba.

Una vez visto todo lo desarrollado, el balance final que se obtuvo del proyecto se basó en la obtención de los siguientes puntos:

- Un producto industrialmente útil.
- Una mejora en la eficiencia.
- La cumplimentación del desarrollo de un sistema automático.
- Proyección de uso futuro.

Mi conclusión personal sobre la incorporación al equipo del proyecto del asistente cognitivo fue de lo más interesante, ya que me permitió desarrollarme tanto a nivel personal como profesional. Pude aprender a cómo se trabaja en una empresa de ese tamaño y a cómo gestionar todos los conocimientos teóricos que se aprenden a lo largo del grado. Además, supuso una inmersión en el mercado laboral y un orgullo el hecho de que el proyecto realizado haya sido una propuesta propia que partía desde cero.

## 10. Bibliografía

[1] Página web de:

<https://www.nextpit.es/tbt-eliza-uno-de-los-primeros-chatbots-de-la-historia-1966>

[2] Página web de:

<https://www.userlike.com/es/blog/los-mejores-chatbots>

[3] Página web de:

<https://centribal.com/es/como-entrenar-un-chatbot-para-resolver-dudas/>

<https://blog.inconcertcc.com/como-entrenar-a-tu-bot/>

[3] Página web de:

<https://www.salesforce.com/es/blog/2020/05/Chatbots.html>

<https://www.muycomputerpro.com/2019/04/25/empresas-con-historia-salesforce>

[4] Página web de:

<https://www.ingenieriaindustrialonline.com/gestion-y-control-de-calidad/mejora-de-procesos-metodo-de-las-ocho-fases/>

[5] Página web de:

<https://help.qlik.com/es-ES/>

[6] Página web de:

<https://www.apd.es/tecnicas-de-la-inteligencia-artificial-cuales-son-y-para-que-se-utilizan/>

Páginas webs para documentación:

<https://svcmindustry.org/es/dictionary/what-is-the-difference-between-an-expert-system-and-artificial-intelligence/>

<https://gus.chat/machine-learning-vs-sistema-experto/>

Libros leídos para el auto-aprendizaje:

Superinteligencia: Caminos, peligros, estrategias, Nick Bostrom, TEELL EDITORIAL, S.L.; N.º 1 edición (25 febrero 2016)

Cognitive processes in planning: A report ([Report] - Rand Corporation ; R-2366-ONR)