

Análisis de chats

Índice

- Punto de partida.....3**
- Solución técnica.....4**
 - Fichero de tickets
 - Unidad de análisis
 - Análisis de las conversaciones
 - Método
 - Tipos de resultados
 - Resultados detallados del proyecto
 - Valor retrospectivo de los resultados
 - Valor predictivo de los resultados
 - Nota metodológica
 - Nota sobre volumen de datos
- Presupuesto.....17**
 - Requerimientos operativos
 - Precio del proyecto inicial
 - Precio de las repeticiones periódicas
- Plan de acciones19**

Punto de partida

Rhy ofrece una plataforma de *live chats* con agentes comerciales, que dan soporte a compradores en puntos de venta virtuales. Los chats generan ficheros de conversaciones de cuyo análisis podría obtenerse información para identificar patrones de mensajes, en una y otra dirección (inbound/outbound). Tales patrones podrían ser útiles en tareas como:

- Elaboración de argumentarios comerciales.
- Procesos de resolución de incidencias.
- Construcción del chat.

En este contexto, presentamos una propuesta de colaboración con Rhy, compuesta por tres elementos:

1. Una solución técnica para el análisis de las conversaciones.
2. Un presupuesto.
3. Un plan de acciones.

Solución técnica

Fichero de *tickets*

El fichero de *tickets*, con formato .csv, tiene la siguiente cabecera o campos:

- 'id'
- 'action'
- 'data'
- 'trigger'
- 'action date'

A partir de este fichero, el primer paso de nuestra solución será la identificación de conversaciones.

Unidad de análisis

La unidad de análisis de nuestra solución es la conversación. Una conversación es una secuencia de mensajes (líneas) entre comprador y agente que comprenda, como mínimo, un línea con el valor "reply.text.create" en el campo 'action' y una línea, como máximo, con el valor "Close" en el campo 'data' cuando el valor de 'action' sea "status.change".

Éstas son las asunciones que hacemos sobre el contenido del fichero:

- ▶ Cada línea (registro) se corresponde con un mensaje cuyo emisor (comprador, agente) está identificado.
- ▶ Los interlocutores de una conversación pueden identificarse por el valor del campo 'trigger': un comprador tendrá el valor "unknown@chat.inbenta.com"; un agente, cualquier otro valor que represente una cuenta de correo electrónico.
- ▶ Las conversaciones tendrán siempre el valor "Spanish" en el campo 'data' cuando el valor de 'action' sea "ticket.create.language": no será necesario un análisis multi-idioma.
- ▶ Las conversaciones pueden estar clasificadas por el agente, que habrá consignado en aquellos registros con valor "classifications.add" en el campo 'action' uno de los siguientes valores en el campo 'data':
 - "1": consulta
 - "2": incidencia

- “3”: duda
- “4”: reclamación
- “6”: gaming
- “7”: no gaming
- “9”: entrada
- “10”: medio
- “11”: pro

Una conversación puede estar clasificada simultáneamente con un valor de 1-4, de 6-7 y de 9-11.

```
7708,status.change,Open,geraldineivg@gmail.com,1612170212
7708,reply.text.create,"BUENOS DIAS",unknown@chat.inbenta.com,1612170237
7708,reply.text.create,"QUIERO COMPRAR UN PORTATIL",unknown@chat.inbenta.com,1612170247
7708,reply.text.create,"Hola, buenos días",geraldineivg@gmail.com,1612170251
7708,reply.text.create,"en que le puedo ayudar?",geraldineivg@gmail.com,1612170258
7708,reply.text.create,"Y NO ME COGE LO DEL SUJETO PASIVO",unknown@chat.inbenta.com,1612170259
7708,reply.text.create,REVENDEDOR,unknown@chat.inbenta.com,1612170262
7708,reply.text.create,"EN MI CUENTA",unknown@chat.inbenta.com,1612170267
7708,reply.text.create,"Ese tipo de información, no la manejamos nosotros. Debe comunicarse directamente con soporte técnico
7708,reply.text.create,OK,unknown@chat.inbenta.com,1612170326
7708,reply.text.create,GRACIAS,unknown@chat.inbenta.com,1612170328
7708,reply.text.create,"https://www.pccomponentes.com/soporte/contacto
Este es el horario de soporte técnico: de 9:00 a 19:45 de lunes a viernes, y de 9:00 a 16:00 los sábados y domingos. Le dej
",geraldineivg@gmail.com,1612170333
7708,reply.text.create,"Aquí le dejo la información. Puede contactarles y consultarle.",geraldineivg@gmail.com,1612170371
7708,reply.text.create,"Tiene otra duda?",geraldineivg@gmail.com,1612170380
7708,reply.text.create,Hola,geraldineivg@gmail.com,1612170456
7708,status.change,Close,geraldineivg@gmail.com,1612170520
7709,ticket.create.title,"Chat from Unknown",unknown@chat.inbenta.com,1612170221
7709,ticket.create.description,,unknown@chat.inbenta.com,1612170221
7709,ticket.create.source,Chat,system,1612170221
7709,ticket.create.language,Spanish,system,1612170221
7709,ticket.create.status,New,system,1612170221
7709,ticket.create.priority,Medium,system,1612170221
7709,ticket.create.queue,1,system,1612170221
```

Conversación

Interlocutores

Idioma

```
8062,classifications.add,2,geraldineivg@gmail.com,1612455293
8062,classifications.add,6,geraldineivg@gmail.com,1612455293
8062,classifications.add,9,geraldineivg@gmail.com,1612455293
```

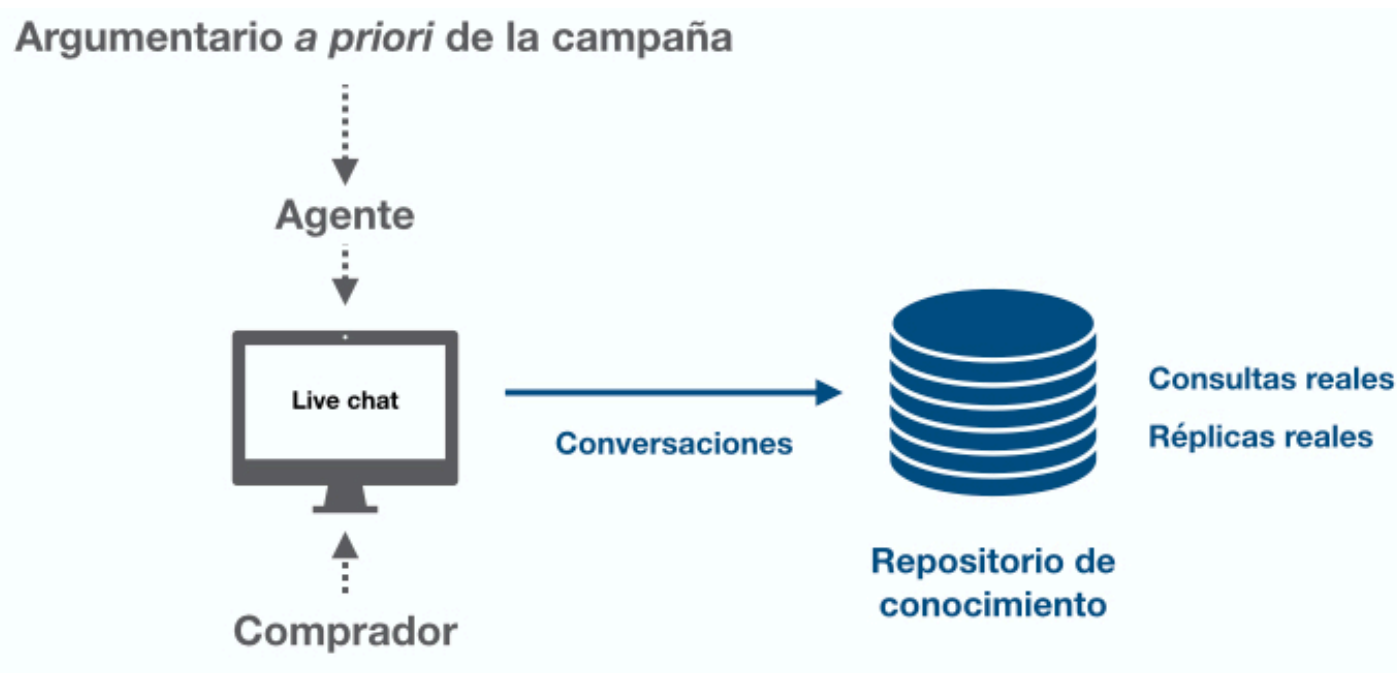
Clasificaciones

Análisis de las conversaciones

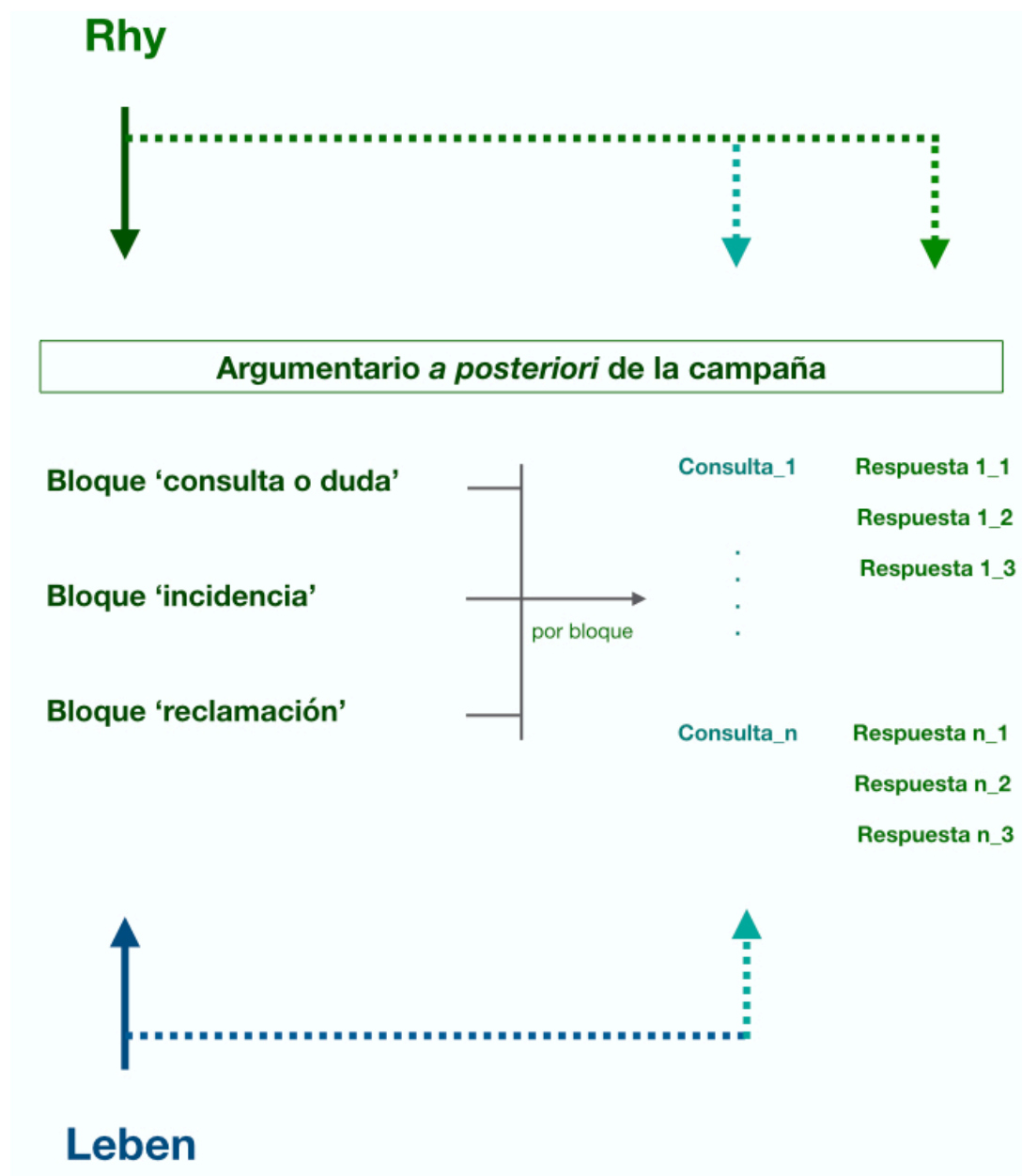
Método

La intuición detrás de nuestro modelo de análisis es la siguiente:

1. Imaginemos que con todas las respuestas dadas por los agentes, en las conversaciones de una misma campaña, compusiéramos un **repositorio de conocimiento**.
 - El repositorio lo construiríamos por tipo de conversación, atendiendo al primero de los tres criterios con los que los agentes la clasifican: consulta o duda (“1” y “3”), incidencia (“2”) y reclamación (“4”).



2. La calidad de las **respuestas reales** la valoraríamos con un **argumentario *a posteriori*** con el que simular conversaciones con el repositorio:
 - Preguntas/consultas similares a las que han hecho los compradores: 5-10 por tipo de conversación.
 - Respuestas imaginarias, consideradas idóneas para las consultas anteriores: 1-4 por consulta.

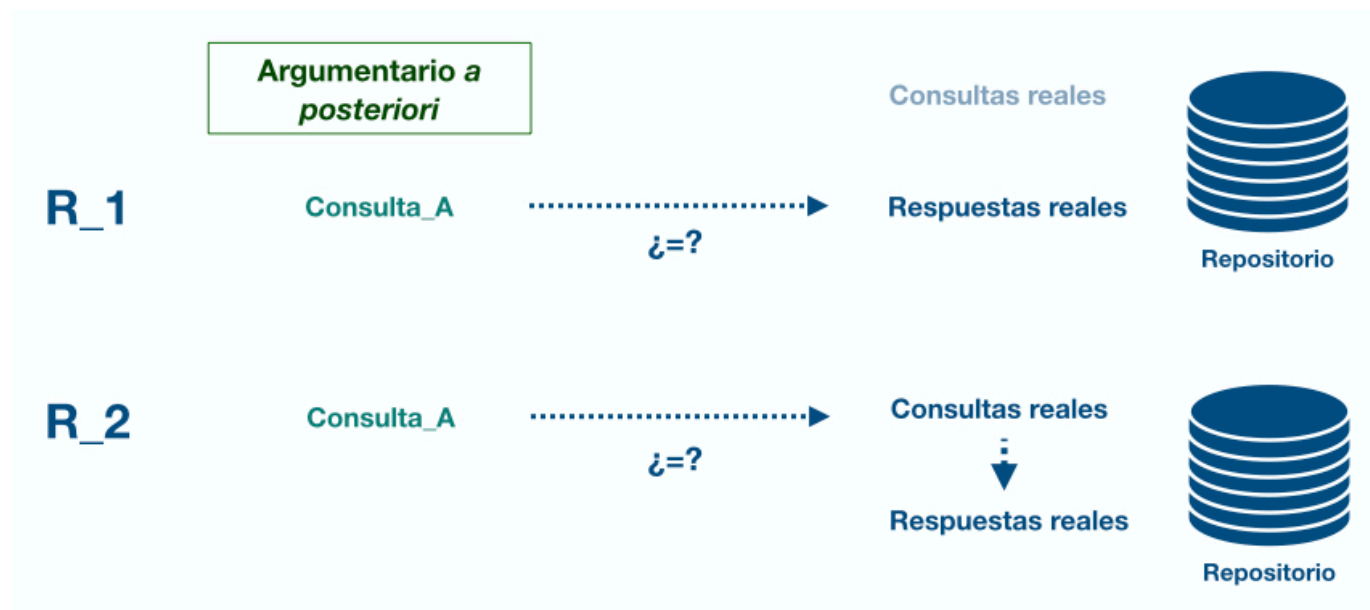


3. El repositorio respondería las consultas de dos formas, R_1 y R_2:

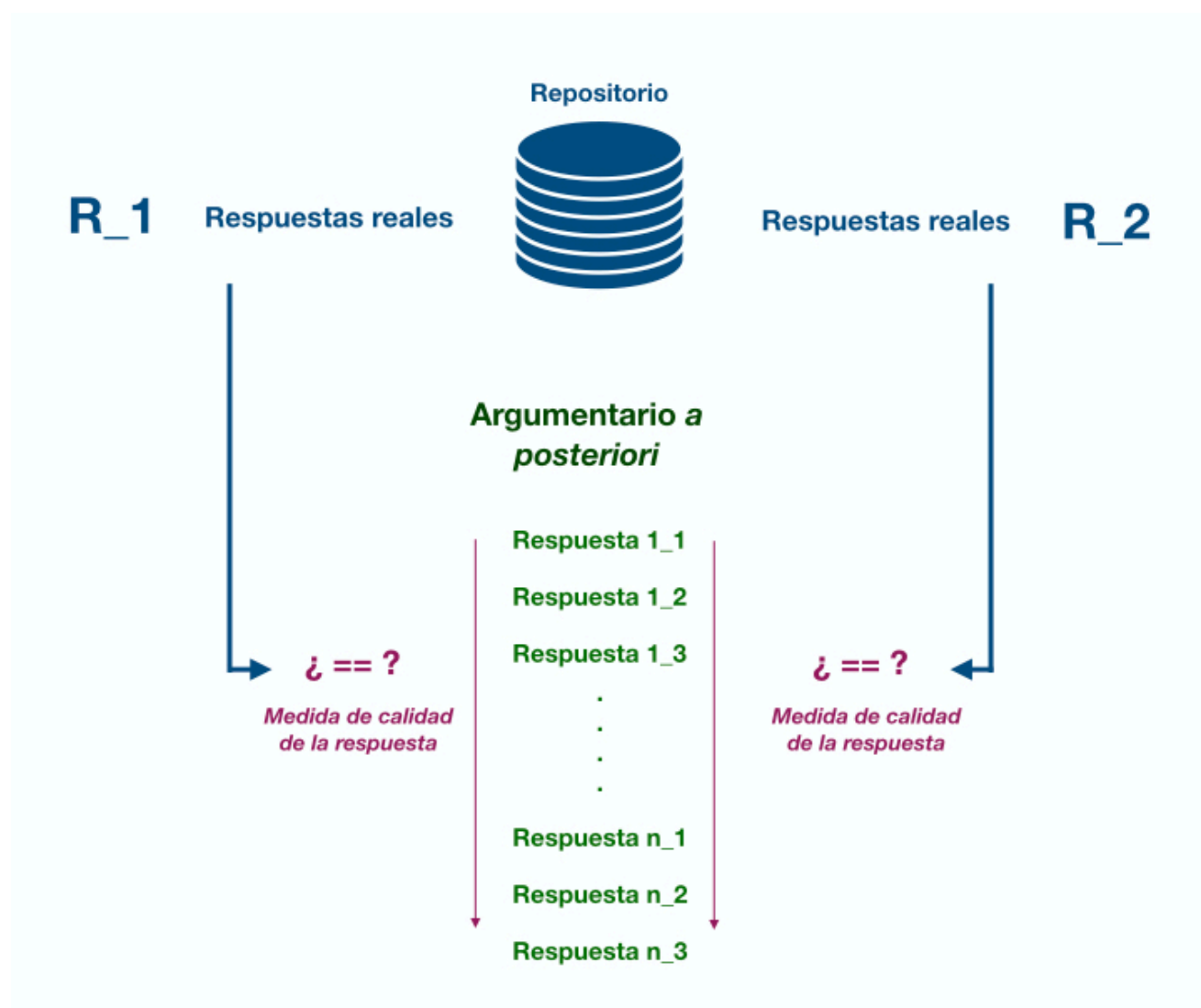
- **R_1:** se considerarán mejores respuestas reales aquellas **más similares**¹ a cada una de las **consultas** del argumentario *a posteriori*.
- **R_2:** primero buscaríamos las consultas reales **más similares**² a cada consulta del argumentario y, a continuación, obtendríamos las respuestas reales dadas a esas consultas.

¹ Por 'más similares' entendemos las situadas en último o penúltimo percentil del score de similitud que definiremos: el 10%-20% con el score más alto.

² El 10%-20% con el score de similitud más alto.



4. Las diferencias entre las respuestas de R_1 y R_2 podríamos interpretarla como la aportación del *live chat* frente a una consulta sin intermediación, donde el comprador interroga directamente un motor de búsqueda a la medida de la campaña.



Ejemplo¹

Consulta_A: “Estoy buscando una tarjeta de video para mi equipo”

R_1:

- Respuesta real similar a Consulta_A:

“Tenemos varias opciones de tarjetas de video, con diferentes precios”.

R_2:

- Consulta real similar a Consulta_A:

“Necesito una tarjeta de video no muy cara”.

- Respuesta real:

“Esta semana está de oferta la serie RTX. ¿La conoces?”

Respuesta en el **argumentario a posteriori**:

“Actualmente tenemos una promoción de las tarjetas de las series RTX y GTX de NVIDIA.”

En R_1, la respuesta real elegida utiliza un vocabulario similar al de la Consulta_A.

En R_2, primero se buscaría una consulta **real** similar a la consulta_A y luego se buscaría la respuesta **real** que el agente escogió para esa consulta. En nuestro ejemplo, en la respuesta real no aparecen ni el término ‘tarjeta’, ni el término ‘vídeo’, que sí aparecen en la consulta **real** similar a la Consulta_A; sin embargo, la respuesta es más informativa que la dada en R_1: ha sido el agente quien ha asociado la serie RTX a ‘tarjeta de vídeo’.

En este ejemplo, nuestro modelo podría arrojar unos scores de similitud de 0,6 y 0,85, respectivamente, para las respuestas en R_1 y R_2.

¹ Los ejemplos presentan una versión simple de lo que podrían ser tanto las entradas en el argumentario a posteriori, como las secciones de diálogo denominadas como respuestas R_1 y R_2. En la versión real, podrían ser, más bien, *bolsas de palabras*.

Tipos de resultados

¿En qué se diferencian las respuestas **reales** en R_1 de las de R_2?

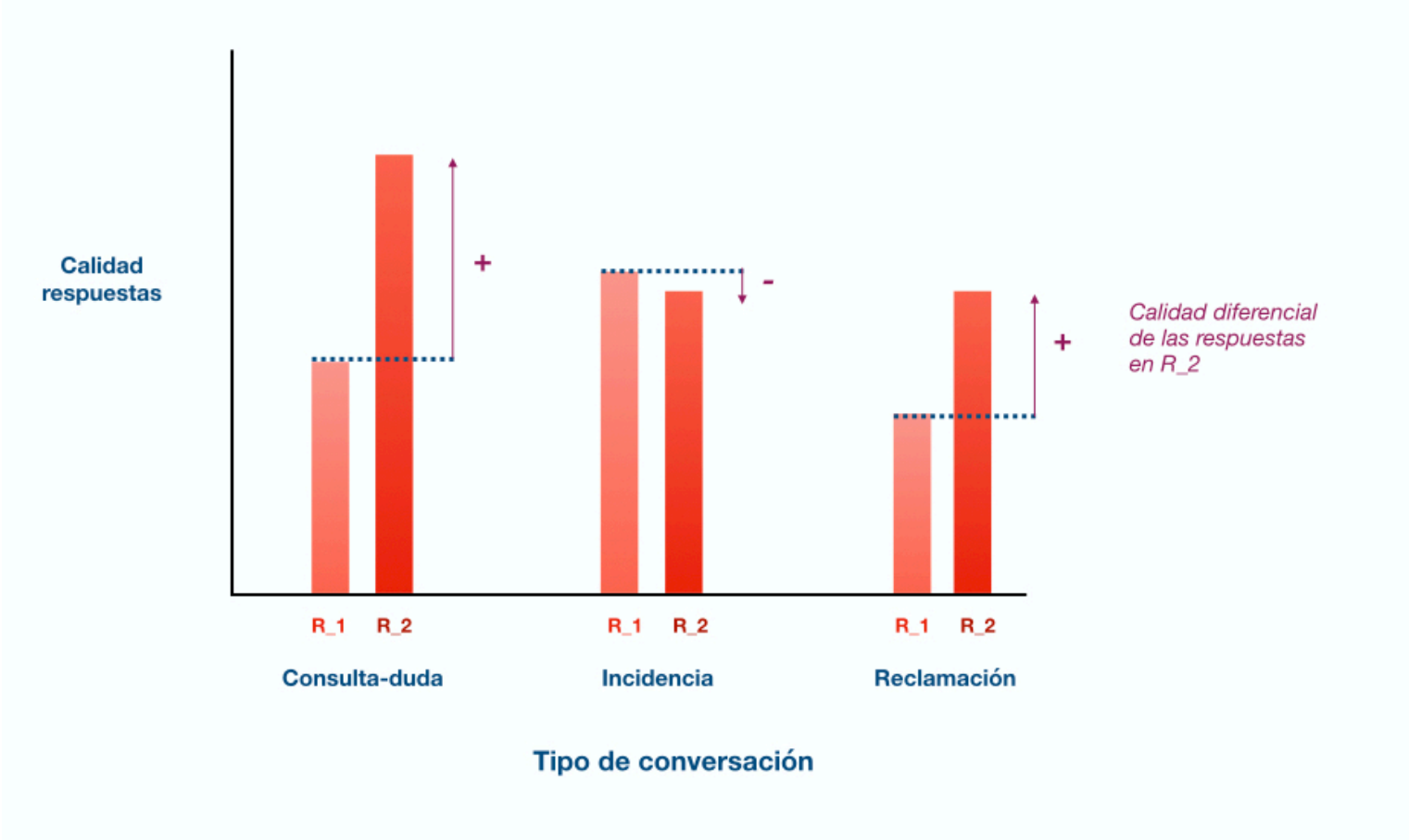
Las respuestas en **R_1** son las que un comprador obtendría en una *conversación* sin la asistencia de un agente, formulando consultas directamente contra un motor de búsqueda muy simple, alimentado con las respuestas históricas reales de campañas anteriores.



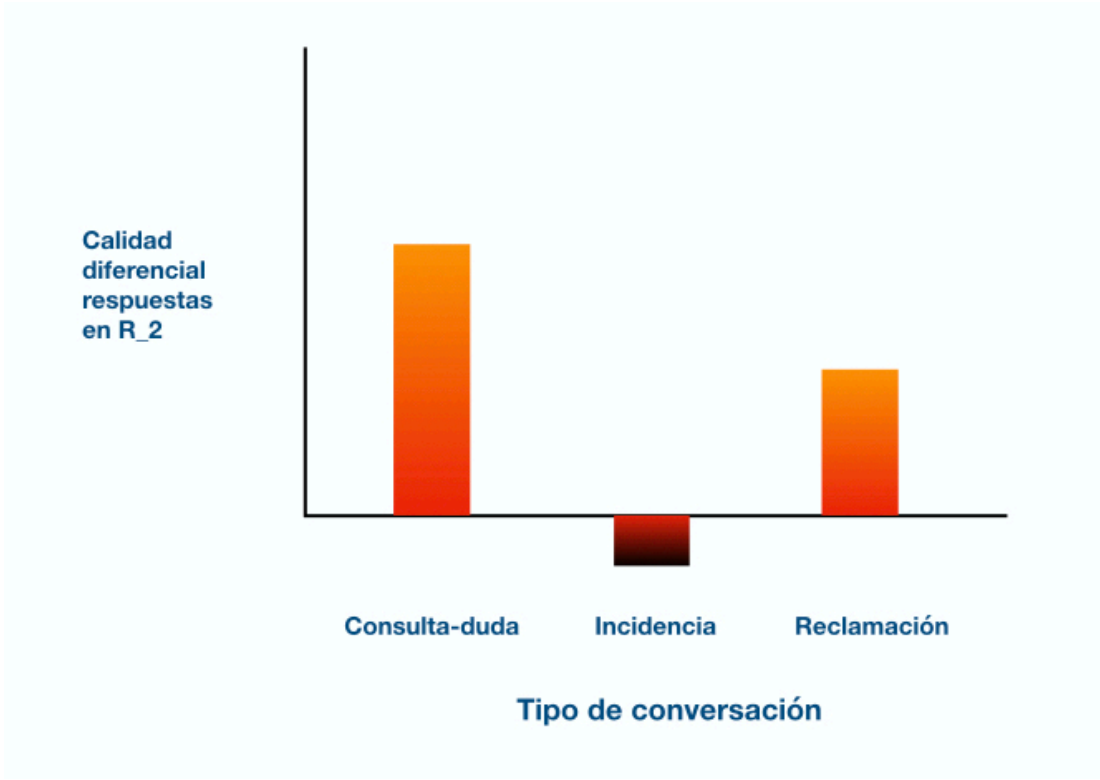
Las respuestas en **R_2** son las que un comprador obtendría en una conversación con un agente que responde las consultas de un comprador basándose en las respuestas históricas que se dieron anteriormente a esas mismas consultas o a otras muy parecidas; es decir, basándose en argumentarios de venta anteriores.



Una forma de valorar la idoneidad de las conversaciones del *live chat*, desde el punto de vista de su adecuación a un argumentario de ventas, es medir la calidad de las respuestas obtenidas en R_2 y compararlas con la calidad de las respuestas obtenidas en R_1. Ése es el método que proponemos en nuestra solución técnica: comparar la *calidad* de una conversación contra un buscador a medida con la calidad de una simulación del *live chat* con un agente.



O de forma más sintética:

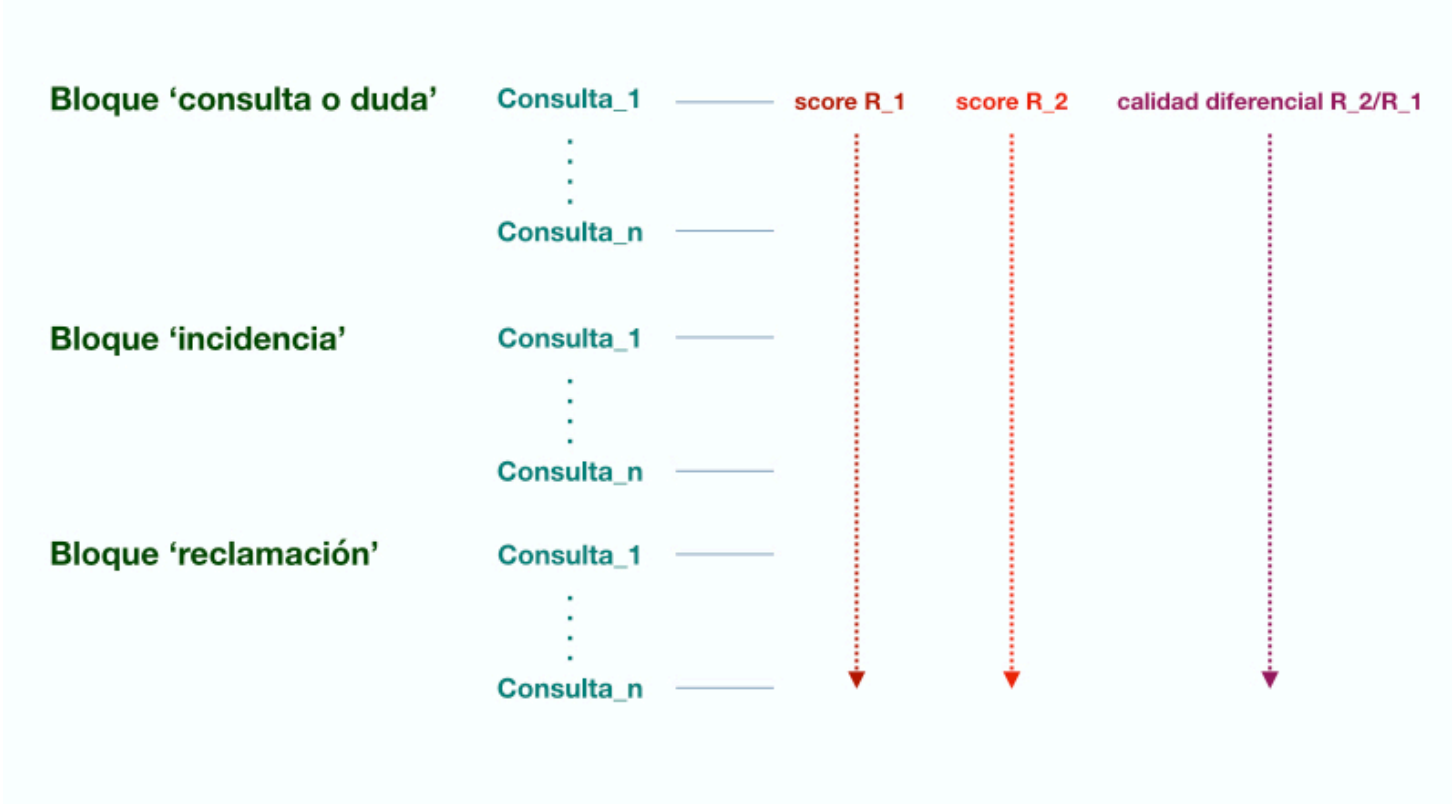


Los gráficos anteriores muestran un ejemplo donde, agregando las medidas de calidad de las respuestas reales por tipo de conversación, nuestra simulación del *live chat* arroja mejores resultados en ‘consulta-duda’ y ‘reclamación’, y peores en ‘incidencia’, respecto a las medidas de calidad de las conversaciones contra nuestro simple buscador.

Resultados detallados del proyecto

- Para **cada consulta-pregunta** incluida en el argumentario *a posteriori*, obtendremos los siguientes resultados¹:
 1. Media aritmética y mediana del *score*² de similitud de las respuestas en R_1.
 2. Media aritmética y mediana del *score* de similitud de las respuestas en R_2.
 3. Calidad diferencial en valor absoluto y porcentual.
- Para cada tipo de conversación, los mismos resultados anteriores agregados por tipo.
- Los resultados anteriores agregados para el conjunto de todas las preguntas-consultas del argumentario *a posteriori*.

Argumentario *a posteriori* de la campaña



¹ También podrán consultarse, una a una, las respuestas reales consideradas tanto en R_1 como en R_2.

² Recordamos que el *score* de similitud lo calculamos respecto a las respuestas en el argumentario *a posteriori* que corresponden a la consulta-pregunta cuyos resultados calculamos.

Argumentario *a posteriori* de la campaña

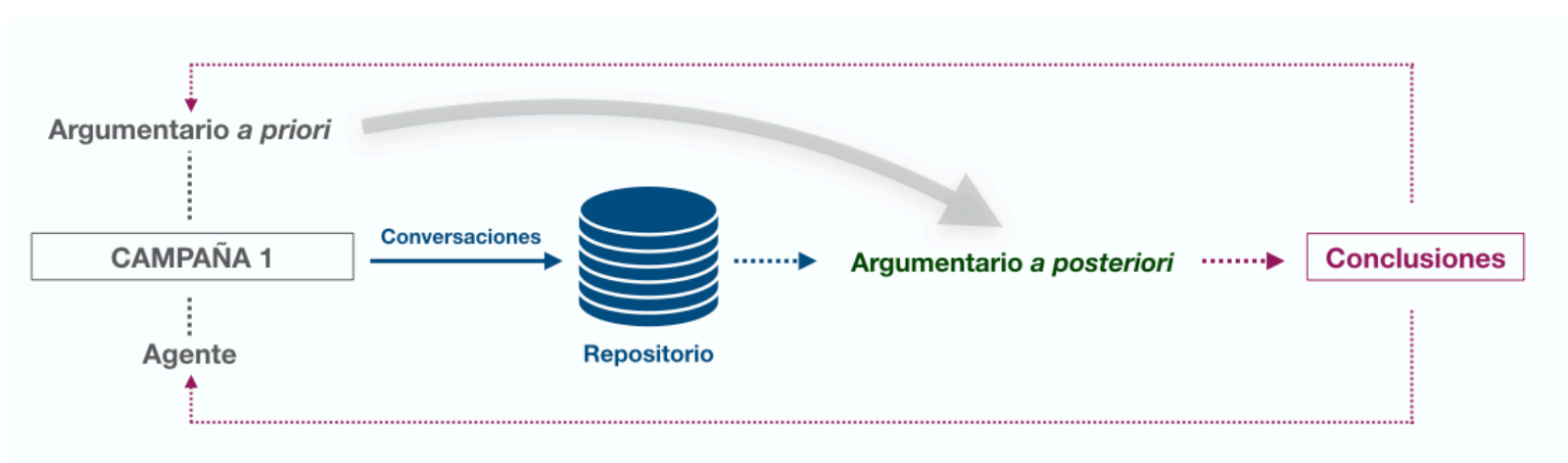


Argumentario *a posteriori* de la campaña ——— score R_1 score R_2 calidad diferencial R_2/R_1

Valor *retrospectivo* de los resultados

Si utilizamos un argumentario *a posteriori* que reproduzca, con las necesarias variaciones¹, el argumentario *a priori* de la campaña, podríamos establecer el siguiente encadenamiento de conclusiones:

1. Ha habido consultas que podrían haberse respondido mejor² en algunas conversaciones con mensajes utilizados en otras conversaciones de la misma campaña.
2. Ello puede haber sido así por dos razones: el argumentario no incluía en esa consulta los mensajes más idóneos, o el agente no ha seguido fielmente el argumentario.
3. En el primer caso, tendríamos un indicador de cuán mejorables eran ciertos bloques o consultas del argumentario. En el segundo, una indicador de en qué casos podría haberse mejorado la formación de los agentes.



¹ No conviene que sea exactamente igual. Por ejemplo, para tratar el problema de la polisemia, convendría que una misma consulta y las respuestas ideales que a ella se sugieren estuvieran redactadas de más de una manera.

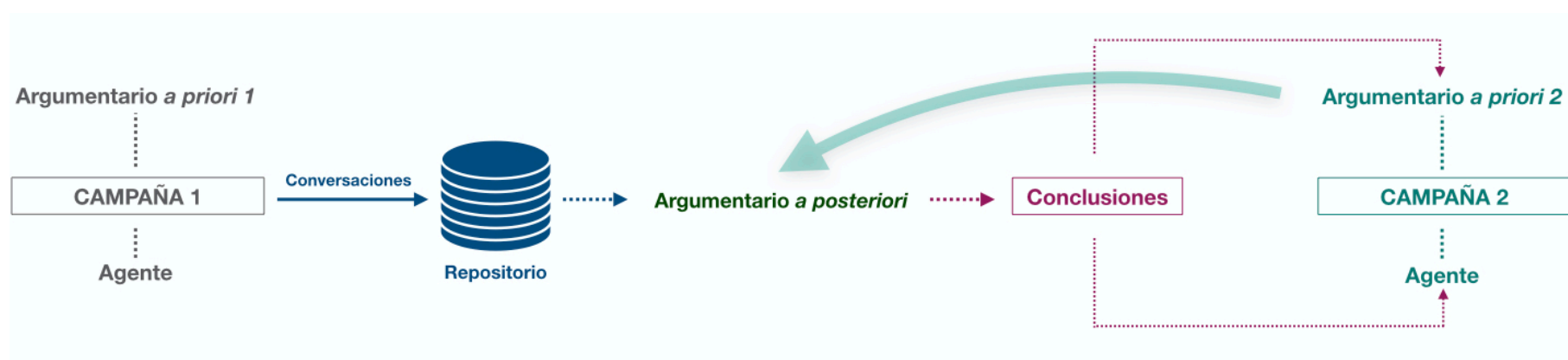
² Con un diferencial de calidad de respuesta positivo, en lugar de negativo, por ejemplo.

Valor predictivo de los resultados

Imaginemos un escenario diferente al anterior: vamos a iniciar otra campaña con un contenido comercial similar a la campaña cuyas conversaciones vamos a analizar. En ese escenario, el argumentario *a posteriori* podría ser una variante del argumentario de la **nueva** campaña.

Asumamos que tanto el nuevo argumentario como el perfil de los agentes son muy similares a los de la campaña anterior. Al realizar el análisis de sus conversaciones, las conclusiones tendrían un valor predictivo en lugar de retrospectivo:

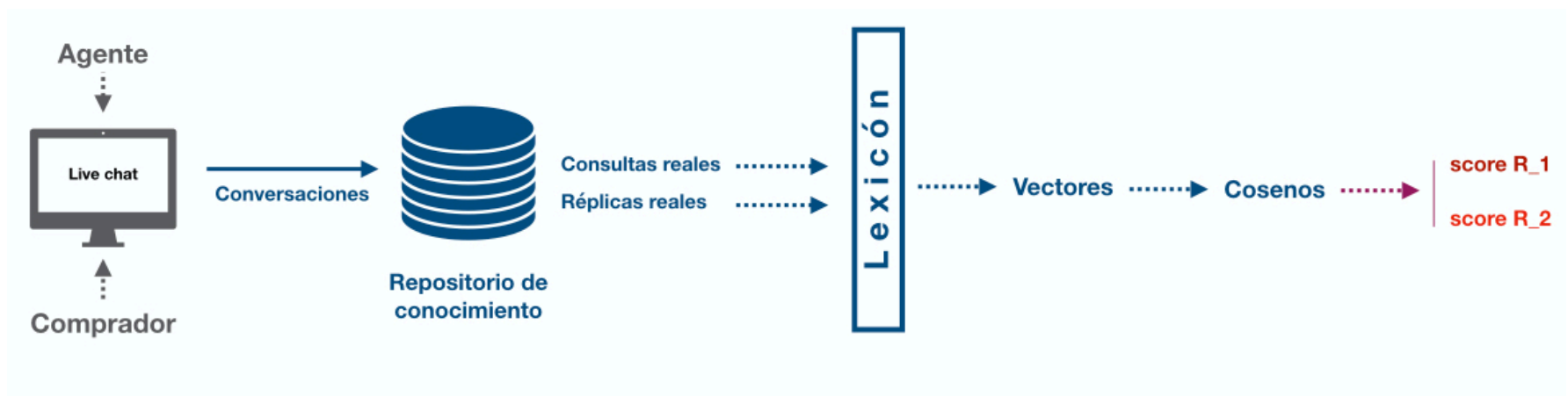
1. Cómo podemos mejorar los mensajes del nuevo argumentario, antes de ser utilizado.
2. Qué consultas conviene destacar en la formación de los agentes de la nueva campaña.



Nota metodológica

Nuestro análisis está basado en definir un ‘espacio vectorial’ con las conversaciones:

1. Representamos los mensajes que intercambian comprador y agente, dentro de una conversación, como vectores de números reales ¹.
2. Cada vector se cuantifica de acuerdo al método tf-idf ² (*term frequency - inverse document frequency*):
 - Para cada término del vocabulario empleado en todas las conversaciones ³, en el vector se consigna el producto (tf x idf) : se multiplica el número de veces que aparece el término en la conversación por una medida (invertida) del número total de conversaciones en que aparece ese mismo término.
3. Con la representación vectorial, la similitud entre dos mensajes o dos bloques de una conversación la estimamos con el coseno del ángulo que forman los dos vectores.



¹ Que definamos un vector con cada mensaje de un comprador o de un agente, o que definamos el vector a partir del conjunto de la conversación, dependerá de los datos concretos de cada fichero de conversaciones. Con conversaciones breves, centradas en un solo objetivo, un vector corresponderá la mitad de conversación respectiva de un agente o de un comprador.

² También podríamos evaluar en el proyecto la utilización del método Okapi BM25 para la misma tarea.

³ Si el vocabulario empleado en las conversaciones fuera muy extenso, podríamos plantearnos utilizar algún mecanismo de filtrado de vocablos.

Nota sobre volumen de datos

Si el número de conversaciones de un mismo tipo superara el millar, realizaríamos un muestreo estratificado por tamaño de la conversación, y circunscribiríamos el análisis a la muestra resultante de mil conversaciones.

Es importante tener en cuenta, también, la granularidad del argumentario *a posteriori*: consultas muy específicas tendrán muy pocas respuestas asociadas, y los resultados serán poco significativos, tanto desde un punto de vista estadístico como operativo.

Presupuesto

Requerimientos operativos

Los datos de una campaña han de entregarse en un fichero con formato .csv que sólo contenga la información siguiente:

1. Un registro único (cabecera) con el nombre de los campos del fichero: 'id', 'action', 'data', 'trigger' y 'action date'.
2. El resto de los registros del fichero, cuyos campos describe la cabecera.

En cuanto al volumen de datos, asumimos en este presupuesto que el fichero no contendrá más de un millón de registros.

Precio del proyecto inicial

El precio del proyecto es de 4.950 euros.

Este precio incluye:

- Nuestra participación, junto con el equipo de Rhy, en la composición del argumentario *a posteriori* de la campaña.
- La programación informática de todas las tareas relacionadas con el análisis. El desarrollo conllevará las pruebas necesarias para que pueda repetirse el análisis en campañas sucesivas, con un número mínimo de incidencias.
- La ejecución de un ciclo completo de análisis, tal como se describe en el apartado 'Análisis de conversaciones', con los datos de una campaña real.

Precio de las repeticiones periódicas

Siempre que el formato del fichero de *tickets* se mantenga sin variaciones, respecto al proyecto inicial, y el argumentario *a posteriori* sólo requiera variaciones en algunas preguntas, el precio de las repeticiones periódicas de los análisis descritos en esta propuesta sería el siguiente:

- Proyectos de 120 horas: 12,5 euros/h.
- Proyectos de 240 horas: 8,25 euros/h.
- Proyectos de 360 horas: 6 euros/h.

- Proyectos de 480 horas: 5 euros/h.
- Proyectos de 600 horas: 4,25 euros/hora

Plan de acciones

Presentamos a continuación un cronograma del proyecto asumiendo una fecha de inicio del 15 de julio y el mes de agosto como plenamente operativo. Ambas asunciones, obviamente, podrían estar sujetas a modificaciones. La duración total, en tiempo de calendario, es de unas nueve semanas.

Hitos

