

La importancia de las palabras en el entrenamiento del Chatbot Marketeer

- **Empezar:** Lo primero a tener en consideración para entender si un producto está bien elaborado o no es pensar en lo que el usuario querrá encontrar. La única forma de saberlo es con el propio uso. Si el usuario pregunta algo que no tiene respuesta, no quiere decir que esté mal hecho o que tu sistema funcione mal. Solo quiere decir que le falta algo de contenido. Eso es lo bueno de Marketeer: el contenido no es estático, es totalmente dinámico y evoluciona con el cliente. No es como una web que tú creas, de la que el usuario se va y no sabes por qué. Aquí el usuario te preguntará clara y directamente y podrás ver qué es lo que más le interesa.
- **Crear la base de conocimiento:** No te preocupes si al inicio parezca que falle o que no conteste a nada, es bastante normal porque estás creando algo nuevo, que es una base de conocimientos. Y si nunca la has hecho con anterioridad, es de las cosas más importantes que tiene la empresa: el conocimiento de lo que a mis clientes les preocupa, todo en un mismo sitio para poder consultarlo.
- **Qué le interesa a mi cliente:** Te darás cuenta de qué es lo que le interesa al usuario y verás que es bastante redundante, así que pronto podrás hacer un gran avance, simplemente dejando funcionar a la herramienta para que dé respuestas mientras tú editas las que no tengan contenido.
- **Preguntas genéricas vs. específicas:** Si tengo que responder a la pregunta “¿tienes agua de 500ml.”, mi respuesta no puede ser “sí tengo”, porque si un usuario pregunta “¿tienes agua de 200ml.?” es posible que la respuesta también sea “sí tengo”. La razón es que el lenguaje natural funciona por proximidad, algo que puedes configurar en *tu widget* y no hay tanta distancia entre ambas preguntas como para tomarla como errónea. Por consiguiente, te recomendamos que tu respuesta sea “Tenemos botellas de 1l., 500ml. y 200ml.”. Así, si un usuario pregunta: “¿tenéis botellas de 2 litros?”, la respuesta que se obtenga tenga, por defecto, el valor correcto.
- **Respuestas con sentido:** Es absolutamente indispensable que entiendas que la pregunta es igual de importante que su respuesta. Marketeer evalúa el grupo de información, no solo las preguntas o las respuestas, sino su contexto. Para definir este grupo de preguntas, analiza el sentido de sus respuestas (que es lo que interesa), más que incluir preguntas similares, por lo que las respuestas deben ser claras.
- **Uso y autoaprendizaje:** Es realmente importante el uso que le das a tu bot para mejorar sus respuestas. Cada vez que activas el *Machine Learning*, tu bot aprende de su funcionamiento. Si envías una pregunta con una respuesta, el bot aprende que ambas son sinónimos.
 - Si la editas, el bot la creará como variante, lo cual después discriminaría cuál es la importante, si la nueva o la anterior.
 - Si agregas nuevas preguntas desde el *live chat*, evitas pensar qué debe tener tu base de conocimiento; simplemente dejas que guarde todo para luego validarlo.
 - Cada pregunta se guarda como borrador, lo que verás que se visualiza desde la interfaz del cliente, pero no desde la interfaz del usuario. Así, después puedes entrar en la carpeta “chat” de tu base de conocimiento y activar, mover y quitar una pregunta guardada como borrador, tras revisarla y hacerla pública.

- **Entidades:** En el lenguaje natural es importante el tema de entidades: palabras únicas. Por ejemplo, yo creo la pregunta “¿qué es un *showroom*?” y tengo otra pregunta, “¿qué precio tiene la instalación?”. Si pregunto “¿qué precio tiene el *showroom*?”, *showroom* es una entidad, algo único. Es decir, no figura en ninguna otra pregunta e incluso tiene más peso si no figura en el diccionario de la lengua. Por lo tanto, es posible que la respuesta que ofrezca el bot cuando se le pregunta utilizando la palabra *showroom* tenga el peso suficiente como para que sea altamente probable que la respuesta obtenida de casi cualquier pregunta sea la de “¿qué es un *showroom*?”.
- **Verbos:** Es mejor utilizar los mismos verbos para definir una acción, por ejemplo, si creo en la base de conocimiento la pregunta “cómo añadir preguntas” y otra que sea “cómo agregar *links*”, cuando el usuario pregunte “cómo agregar preguntas”, la acción “agregar” tendrá más sentido en “*links*”, por lo que el bot la escogerá como primaria por ser la acción equivalente, aunque “añadir” sea un posible sinónimo. Esto es debido a que la acción tiene un peso mayor.
- **Diferentes maneras de formular una pregunta:** Para entender qué es un sinónimo, debes entender que Marketeer no solo considera como sinónimos palabra sueltas, sino que trabaja por proximidad. Así que los **sinónimos** son palabras o expresiones equivalentes de frases. Si se introduce en la base de conocimiento la pregunta “método para añadir el código en su página”, no se deben crear variantes de esta pregunta agregando solamente sinónimos dispares, tales como: “método para agregar el código en su página”, o “método para añadir el código en su web”. Lo más productivo es ir primero al entrenador y testar.
 - Sabiendo los resultados obtenidos, éstos se pueden ir ajustando.
 - Recomendamos agregar formas distintas de decir lo mismo, pero no demasiadas: una o dos son suficientes. Lo importante es que la respuesta esté bien escrita. Por ejemplo: “copiar el código en tu web”. Así, ya de por sí, estás entrenando sinonimia de frases, no solo de palabras.
- **Diferencias entre SEO y procesamiento de lenguaje natural (PLN):** Para el posicionamiento SEO de un contenido se debe indicar cuáles son las palabras de peso de la web. El SEO funciona muy mal cuando usas palabras que NO ESTÁN en tu web. En cambio, el PLN funciona mejor. Como analogía, la página web sería la “respuesta” y el título de la página sería la “pregunta”. Si quieres que el bot ofrezca una respuesta de la mejor manera posible, debes indicar primero que la página (lo que sería “tu respuesta”) tenga el contenido apropiado bien explicado: intenta que el contenido de las respuestas sea claro y conciso, que no sea excesivo, y que contenga diferentes formas de preguntar para acceder a dicho contenido. Por ejemplo, con palabras cortas:
 - Pregunta: Tabla de precios
 - Respuesta: Tenemos planes de 100, 200 y 300 euros.
 - Pregunta relacionada / Sinónimos
 - ¿Qué planes tenéis?
 - ¿Cuánto vale la suscripción?
 - ¿Cuánto debo pagar al mes?
 - Esto hará que si preguntas “planes de 300 euros”, “cuánto vale por mes” o cualquier contenido cruzado que puedas imaginar, incluido “valor de la mensualidad”, el bot encuentre esta respuesta.

- **Forma de las preguntas:** Por suerte no buscamos *keywords* (palabras clave). Si tuvieras que poner todas las variantes para encontrar algo sería una locura. Al no usar *keywords*, no lo “entrenes como tal”. Es decir, no hace falta poner todas las formas verbales. Por ejemplo, si dices “corro mucho”, no añadas relacionadas como “yo he corrido...”, “correr”, “corriendo”, etc. Todo significa lo mismo. Si agregas relacionadas con diferentes formas verbales de la misma pregunta, estás obligando a que la acción de esa pregunta sea “única”, es decir, haces que cualquier pregunta sobre “correr” vaya a dar con esa respuesta, e intentar solucionarlo es muy complejo, ya que tendrías que agregar el resto de formas a las demás preguntas y no responderá nada exceptuando las frases escritas tal cual las has introducido.
- **Sobre-entrenamiento:** No debes agregar sinónimos a libre albedrío intentando entender las formas de expresarse de un usuario, porque al hacer un test real, posiblemente el bot no dé los resultados esperados. Lo mejor es empezar con preguntas simples y respuestas concisas, explicativas y claras. No le des vueltas ni a la pregunta ni a la respuesta, no estás vendiendo con tu propia voz, ni es una llamada, ni es un cara a cara. El usuario está **leyendo** y suele evitar leer textos muy largos. Pero tampoco vayas al extremo de contestar simplemente “sí” o “no”. Sé claro: “¿tenéis botella de 200 ml.?” con una respuesta corta, clara y concisa. “Tenemos botellas de 1l, 500 ml. y 200 ml. [Aquí](#) puedes verlas”, y puedes incluir un enlace a una página con más información, por ejemplo.
 - Adicionalmente, puedes indicarle al usuario los beneficios del agua, si crees que no va a haber otras respuestas de 200 ml. y que el usuario pudiese preguntar “¿cuáles son los beneficios del agua de 200 ml.?”. Piensa en qué te gustaría recibir como respuesta, no es tu *pitch* comercial. Cuanto más directas y claras sean las respuestas, más satisfecho estará el usuario, ya que es su tiempo y debes valorarlo.
- **Repetición de palabras:** En PLN debes saber que una palabra muy repetida es una palabra de poca importancia, es decir, todas aquellas palabras utilizadas en muchas preguntas (p. ej.: qué, cuánto, cómo, por qué...). Si introduces muchas preguntas que empiecen por “qué es...” al final el “qué es” tendrá poca importancia. Pero si creas unas siete preguntas con “qué es” y una con “definición de” y el usuario pregunta: “definición de *casa*”, probablemente, si tienes una pregunta tipo “qué es *casa*”, tienes que considerar que solo estás escribiendo 3 palabras a evaluar y 2 de ellas ya no tienen peso (“qué es”). **Así que en la suma de la proximidad**, posiblemente “definición de” y algún sinónimo de “*casa*”, aporten suficiente peso a otra pregunta como para ser la mejor posicionada. Por eso te recomendamos construir las frases de igual forma, para evitar una mala comprensión.
 - Si escribes “precio del televisor”, utilízalo en todo y no escribas otras preguntas con “cuánto vale el microondas”, o “qué pagarías por la nevera”, porque estás confundiendo al sistema y dándole importancia a la forma de expresarte, más que al contenido. Dicha importancia de la forma de expresarte puede ser apenas un 10% de variación, pero sería lo suficiente como para dar un resultado u otro.
- **Mejorando detalles:** No le des relevancia a lo que no es importante. Estás construyendo un sistema de autoservicio, claro y 24x7, por lo tanto, ese debe ser tu objetivo como solución de autoservicio.
 - Es poco constructivo crear preguntas tipo: “Me gustaría instalar un router en mi casa que no tiene balcón y tengo una ventana muy pequeña para pasar el cable”. El usuario tardaría

más en leer el enunciado que la respuesta. Con el sistema de Marketeer pasa lo mismo. Sé más claro: "Instalación correcta de nuestros routers" y su respuesta.

- De lo contrario, no solo haces que el usuario tenga que leer más, sino que además haces que al sistema no le resulte sencillo identificar qué es lo importante en la pregunta.
 - No obstante, el usuario puede hacer preguntas más largas e identificar la respuesta correcta, aunque tu pregunta sea corta. Es muy probable que si el usuario haga la pregunta de "me gustaría instalar..." y tú tengas "instalación de router", el bot le ofrezca esta última como respuesta, porque interpretaría que, de todo el contenido de tu base de conocimiento, lo importante de la pregunta que hizo el usuario es el router y su instalación.
-
- **Test para entrenar:** La forma correcta de saber si tu contenido está creado de la mejor manera en la base del conocimiento es entrenarlo con frases muy largas. Para ello, escribe textos largos de tu web y pregúntale al bot desde el entrenador para ver qué resultados se obtienen. Si entre ellos está la respuesta correcta, has entrenado correctamente la herramienta. De lo contrario, posiblemente debas mejorar tus respuestas, no solo las preguntas, ya que Marketeer tiene como objetivo ofrecer respuestas. Adicionalmente, prioriza darle sentido a las respuestas, pero lo primordial ante preguntas muy largas de los usuarios es que las respuestas estén bien redactadas.