# [Propuesta] Desarrollando un chatbot para refugiados: nuestra experiencia en Saturdays.AI

## Contexto y equipo

\*\*\*OJO el enfoque del artículo, revisar para evitar el “ellos”

3.Hoy en día tenemos a disposición teléfonos móviles, desde los accedemos a todo tipo de información y aplicaciones. Aun así, hay tareas que se nos siguen resistiendo... Como sería el caso de encontrar la información sobre un trámite en específico. Algunas veces acabamos perdiéndonos en los menús y subpáginas de las páginas oficiales… hasta que llegamos al trámite que buscamos. ¡Eureka!

(Bueno sí, “eureka” por el momento, porque otra cosa es que nos falte algún dato por aportar... y que cuando consigamos la documentación, seamos capaces de repetir la búsqueda sin tener que invertir casi el mismo tiempo que la primera vez…!)

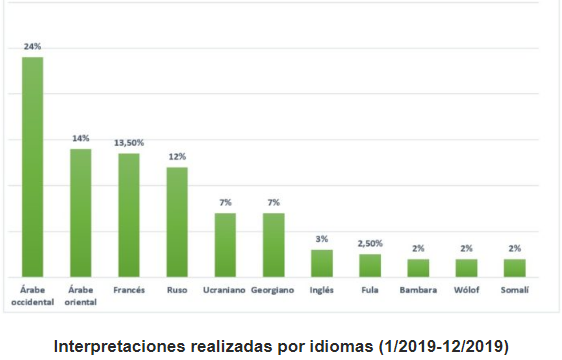
Muy bien, esto sucede a aquellos que entienden el idioma e incluso han visitado las páginas oficiales en varias ocasiones, para uno u otro tipo de trámites. ¿Nos podemos imaginar cómo sería si jamás hubiésemos visitado estas páginas? ¿Y si además, la información no estuviese en nuestro idioma nativo? Por no hablar de estar tratando de localizar información para regularizar nuestra situación en el país destino, sin poder recurrir al país origen a por orientación…

1.A continuación compartimos nuestro planteamiento: Supongamos que debes salir de un país en el que sufres por tu integridad. Llegas a España, y consideras el solicitar protección internacional. Ni siquiera hablas el idioma del país de destino… ¿Dónde puedes dirigirte? ¿Cuáles son los primeros trámites que te van permitir regularizar tu situación?

Consigues llegar a un servicio de atención… tienes muchas preguntas. Ves que están colapsados y no solo con eso, además tienes la dificultad del idioma.

Como habréis imaginado, en este caso nos enfocamos en el colectivo de refugiados; sabemos que actualmente los servicios de atención están saturando, \*\*\*Ivy, ayuda aquí plis! generar colapso sería correcto? y en muchas ocasiones no pudiendo llegar a ofrecer un servicio de calidad y en base a los tiempos prefijados. En esta situación, ¿sería útil disponer de un aplicativo que pueda responder a las primeras consultas o por lo menos las más generales y en base a las últimas actualizaciones, de manera que los servicios pudieran enfocarse en las especificidades de cada caso?

Decidimos ponernos las manos a la obra preparando un chat de atención específico para este colectivo, que pudiera proveer respuestas generales entorno a esta situación y atender en diferentes idiomas.



Fuente: <https://www.cear.es/projects/traduccion-e-interpretacion/>

2.Antes de ponernos de lleno en los detalles, nos presentamos:

Este equipo de trabajo se conformó dentro Saturdays AI:

“[Saturdays AI](https://medium.com/saturdays-ai) es una organización con la misión de formar a las personas para que aprendan Inteligencia Artificial de forma colaborativa, y basada en proyectos de impacto social.”

Siendo un equipo interdisciplinar en ciencias de la salud y desarrollo de aplicaciones, nos disponemos a preparar un chatbot para resolución de las primeras consultas que pueda tener una persona que llega a España y en la situación de refugiado.

Nuestro objetivo es brindar información específica y relevante a las consultas de los refugiados, para facilitarles el acceso a recursos y servicios que les sean útiles.

Sobre nosotros…

(\*\*\*Aquí creo que podriamos enlazar nuestros nombres a linkedin, por ejemplo)

Ivy

Antonio

Josué

Sílvia

El proyecto tal y como ha transcurrido, puede dividirse en dos fases:

## Fase 1

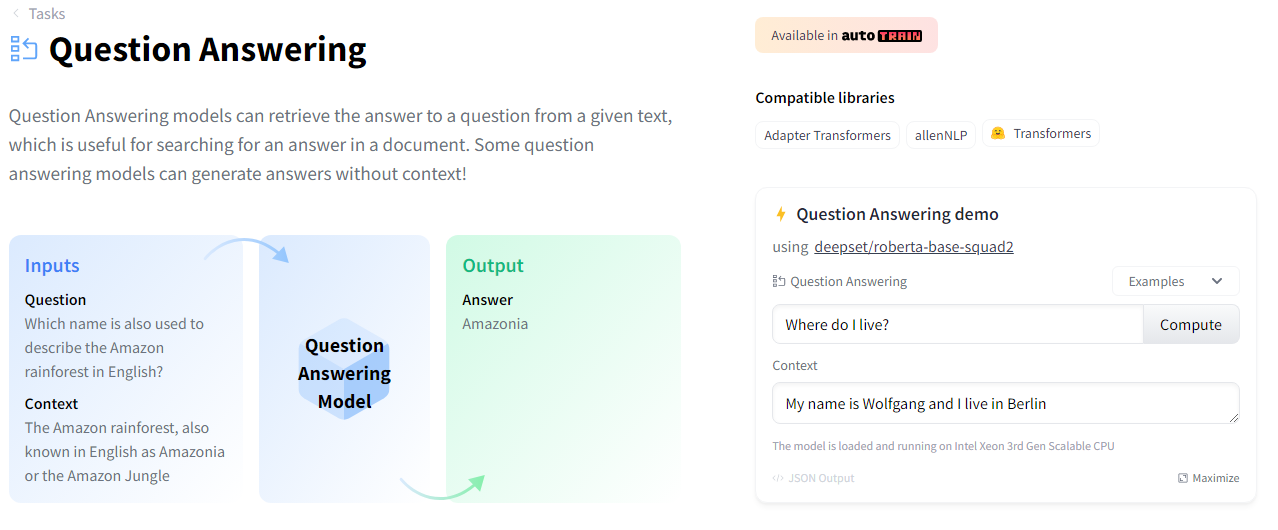
### Método 1

Nos centramos en desarrollar un chatbot al que partiendo de una pregunta que le planteemos, obtengamos una respuesta de un contexto (páginas web).

Para ello estuvimos revisando algunas opciones en [Rasa](https://rasa.com/), [HuggingFace](https://huggingface.co/) o [Dialogoflow](https://cloud.google.com/dialogflow?hl=es). Todas estas opciones parecían buenas y atractivas, no obstante decidimos buscar el dinamismo en el diálogo y el poder adaptar los pesos de los modelos, por lo que finalmente decidimos explorar opciones en HuggingFace.

Al indagar, vimos que en esta misma plataforma consta bastante documentación explicativa y una página de búsqueda de modelos preentrenados, que pueden usarse directamente o haciendo fine-tunning de los mismos. En este caso nos decantamos por modelos de [Question Answering](https://huggingface.co/tasks/question-answering). Probaríamos diferentes modelos de este tipo y veríamos cuál se ajustaba más a nuestro objetivo.

Como opciones de cara a la creación de una webapp, consideramos la subida del modelo final a HuggingFace como una posible opción, dado que en la página de información de los modelos publicados se facilita la documentación relacionada, contexto y API para probar el funcionamiento del modelo (\*\*\*Veis bien el vocabulario? Tenemos enlace finalmente sobre este paso?). De este modo podíamos contribuir a la comunidad con nuestro modelo ajustado y a su vez aprovechar la caja habilitada en la misma plataforma para el testeo de funcionamiento de cada modelo:



### Dataset1

Habíamos localizado el tipo de modelo a entrenar; y debíamos prepararnos para hacer el fine-tunning del modelo. Debíamos entender qué formato debían tener los datos para poder incluirlos en el fine-tunning.

Partimos de artículos del BOE para armar nuestro dataset;

Sobre estos artículos, nos disponíamos a formular preguntas y subrayar las respuestas dentro de cada artículo.

Descubrimos (\*\*\*Cómo lo descubrimos? Recordáis si por ejemplo fue a través de github o huggingface mismo?) que existe una plataforma, [Haystack](https://annotate.deepset.ai/) que permite preparar el dataset interactuando a nivel usuario; 1)Importando los contextos, 2)incluyendo las preguntas y 3)subrayando los contextos dentro de la misma plataforma. Tras ello, la misma plataforma ofrece la opción de descargar los datos en formato csv, excel o JSON (en este último caso, se obtiene un documento JSON con el formato equivalente al formato de datos para los modelos BERT).

### Resultado1

Llegamos a incluir los datos procedentes del BOE en diferentes modelos BERT y similares, y la mayoría de modelos funcionaron proveyendo respuestas consistentes a las preguntas formuladas. Cabe considerar que el modelo con mejor resultado y consistencia en respuestas fue el modelo [RoBERTa](https://huggingface.co/docs/transformers/model_doc/roberta), recogido en el notebook “SQUAD\_es\_v8\_GPU” desde el quinto apartdo/” ES\_Roberta\_spanish\_SimpleTransformersQA” (\*\*\*Cuál dijimos finalmente?), disponible en github:

(\*\*\*Inserir ejemplo).

### Consideraciones1

Ahora nos quedaba valorar si ampliar el dataset, subir el modelo a HuggingFace y validar que la caja para testeo funcionase; o incluso ver si añadir nuevas funcionalidades, para que el modelo fuese capaz de responder a diferentes tipos de consulta (no solo de tipo “legal”).

No obstante... hubo algo que no nos acababa de encajar. A la hora de poner el chatbot en productivo, queríamos que se tratase de un aplicativo amigable: aunque este modelo cumple la función de localizar la información requerida dentro de un contexto, no era capaz de responder de forma cordial a un "Hola" o a un "Qué tal".

Estas consideraciones nos llevaron a la segunda fase de nuestro proyecto;

## Fase 2

### Contexto2

Consideramos diferentes opciones para solucionar la problemática anterior.

* Podíamos incluir algunas preguntas-respuestas programadas, y tras ello un botón para activar la funcionalidad de uso del modelo;
* Podíamos usar modelos conversacionales, que integran ambas funcionalidades;
* Podíamos decantarnos por un modelo de tipo de generación de texto, y entrenarlo para que supiera responder de forma adecuada a consultas relativas al colectivo de refugiados;
* Podíamos integrar diferentes modelos, uno para las preguntas-respuestas específicas y otro más "humano".

### Método2

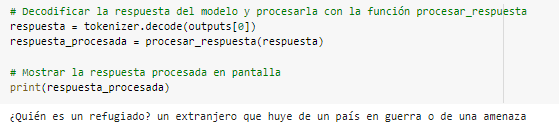
Y como os podréis imaginar, como veníamos participando de un curso de IA, decidimos decantarnos por alguna de las tres últimas opciones, considerando que la primera podría haber sido una opción en caso de haber de entregar un Producto Mínimo Viable (MVP) en cualquier otro contexto. Y de estas tres… valoramos según:

* En cuanto al uso de modelos conversacionales, al revisar en HuggingFace vimos que no constaban modelos de este tipo en idioma español; por lo que por el momento decidimos descartar esta opcion;
* Junto con el auge de los modelos de generación de texto [OpenAI](https://chat.openai.com/chat), decidimos decantarnos por esta opción y probarlo en consultas relativas orientadas a nuestro objetivo; y en caso de no funcionar de forma óptima, tratar de hacer fine-tunning del modelo con nuestros datos.
* En cuanto a la integración de diferentes modelos, cada uno con un mejor performance en una u otra área, decidimos considerarlo tras haber obtenido buenos resultados en los modelos que podrían formar parte de esta solución.

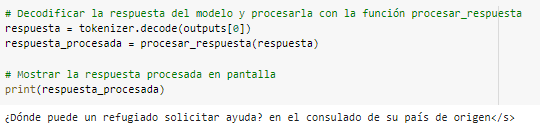
En paralelo estuvimos trabajando en la implementación del modelo resultado a [Streamlit](https://streamlit.io/), para poder proveer demo de la solución.

### Resultado2

A través de HuggingFace localizamos el modelo [Bloom](https://huggingface.co/bigscience/bloom), un modelo de generación de texto. Tras probar diferentes versiones del modelo preentrenado, vimos que la versión que podíamos llegar a implementar desde google colab era la 1b1, con XXX parámetros y XXX. \*\*\*(No encuentro la tabla en github que contenía los diferentes modelos Bloom, ¿la teneís localizada?). Y efectivamente llegamos a obtener respuestas consistentes a preguntas generales:



Al probar de realizar algunas consultas sobre el colectivo de refugiados, vimos que las repuestas no se ajustaban a una resolución real... en algunos casos podían incluso llegar a suponer un riesgo. Por ejemplo:

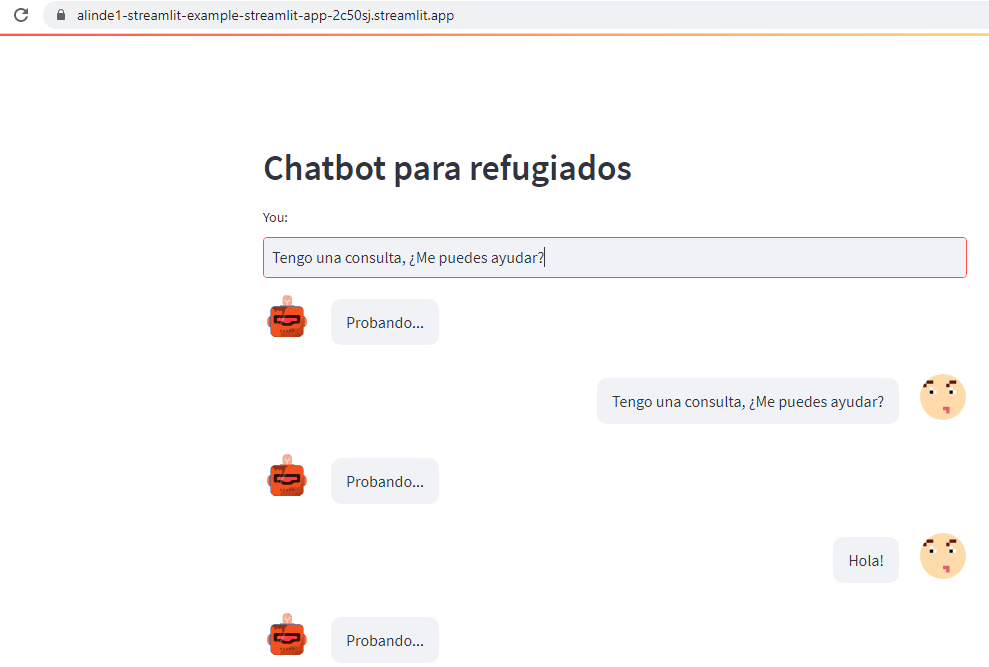


Por ello al realizar este primer testeo, consideramos imprescindible realizar fine-tunning del modelo incorporando por lo menos los datos de los que disponíamos, al modelo preexistente. Os compartimos algunos detalles sobre este proceso en el siguiente apartado.

En cuanto a la demo de la solución, estuvimos preparándola en Streamlit; compartimos a continuación una demo de la solución actualizada, en el enlace siguiente. Y en el caso de cualquier cambio prevemos informarlo en github:

\*\*\*Modificar o incluir enlace a nuestra solución, ¿sería el siguiente? A fecha 22/01 por la mañana visualizo “Probando”:

<https://alinde1-streamlit-example-streamlit-app-2c50sj.streamlit.app/>



\*\*\*En paralelo trabajamos en desarrollar un chat programado con algunas preguntas-respuestas, con el objetivo de poder entregar un MVP en caso que la solución eje no quedase operativa. Disponible en:

### Dataset2

Esta fase del proyecto se desarrolló en Diciembre/2023;

Bloom es un modelo de publicación reciente; se subió a HuggingFace a fecha XXX(\*\*\*¿Se os ocurre cómo consultar la fecha en que quedó disponible? Si os parece podemos incorporar captura de pantalla para validación de los mentores).

Creemos que ello nos lleva a que, al tratar de localizar información en web, nos hemos encontrado con dificultades de cara a localizar documentación explicativa y formato de los datos.

Entonces estuvimos trabajando en la adaptación del dataset para la ejecución de trainer.train() en base a la documentación disponible y más adelante, incluso partiendo de los errores recibidos al ejecutar el notebook. No obstante por el momento no hemos alcanzado a aplicar ajustes en el modelo.

### Consideraciones2

En cuanto a la adaptación de los datos para esta opción, posteriormente consultamos al chatbot de OpenAI y nos indicó las siguientes consideraciones; hecho que nos lleva a considerar que esta podría ser una próxima accion a realiar para avanzar en el fine-tunning del modelo.

Cada línea del archivo debe ser un diccionario JSON que contenga al menos dos campos: "prompt" y "response". El campo "prompt" es el texto que se usará como entrada para generar una respuesta, mientras que el campo "response" es el texto esperado como respuesta. Los datos deben estar limpios y procesados previamente para evitar problemas en el entrenamiento del modelo.

Ejemplo:

Copy code

{"prompt":"¿Qué es un refugiado?", "response":"Un refugiado es una persona que ha huido de su país debido a la persecución, el conflicto armado, la violencia generalizada o violaciones graves de los derechos humanos, y que no se puede o no quiere regresar a ese país debido al temor a represalias."}

Una vez que tengas tu conjunto de datos en el formato correcto, puedes utilizar la herramienta de fine-tuning de OpenAI para entrenar tu modelo con tus datos específicos. Es importante mencionar que para hacer un fine tuning de calidad es necesario contar con un gran volumen de datos, y que estos sean relevantes y de calidad.

No obstante, cabría valorar otras opciones en cuanto a la solución final:

## Consideraciones finales

A lo largo del proceso, nos enfrentamos a varios desafíos, como la necesidad de encontrar una forma de brindar respuestas relevantes y precisas a las consultas de los refugiados. También tuvimos que aprender a trabajar con herramientas y tecnologías nuevas. Sin embargo, estos desafíos también nos permitieron aprender mucho y desarrollar habilidades valiosas.

Consideramos que la filosofía (explícita) de HuggingFace es una práctica realmente interesante, que nos lleva a optimizar los recursos considerando también el impacto en el medio ambiente, así como conectarnos a una comunidad global. Esta plataforma nos ha sido de gran ayuda de cara a localizar modelos y comprender los avances en este ámbito.

Aun así, hemos visto que hacer el fine-tunning de los datos no siempre es tan sencillo! En caso que por ejemplo tú, como lector de esta documento, dispusieses de alguna información sobre este tema... estaríamos encantados de leer tus comentarios!

Puedes contactarnos a través de: XXX\*\*\*(github? medium?)

Creemos que el próximo paso del proyecto podría basarse en o bien explorar las opciones de Rasa (modelos conversacionales), Dialogoflow o bien partir de los modelos de los que disponemos para implementar un enfoque mixto:

- Saludos y preguntas-respuestas generales a través de un modelo de generación de texto, y

- Preguntas-respuestas específicas a través de otro modelo. En este punto, valorar si sería de interés ampliar o no el corpus.

El switch entre un modelo u otro podría realizarse a través de un botón, o aún mejor aplicando un modelo de clasificación que permita detectar de qué tipo de input se trata en cada momento. Y en este caso incluso valorar la inclusión de un !disclaimer.

(\*\*\*Qué os parecen estas propuestas y que queden plasmadas así en el artículo?)

Aunque todavía queda trabajo por hacer, estamos emocionados de haber llegado hasta aquí en el desarrollo de nuestro chatbot para refugiados. Creemos que este proyecto tiene el potencial de marcar una gran diferencia en la vida de las personas que huyen de sus hogares y buscan una nueva oportunidad. Continuaremos trabajando en mejorar nuestro chatbot, haciendo fine tuning específico para el área de refugiados y probando el chatbot en diferentes idiomas. Esperamos poder compartir nuestros resultados y aprendizajes con la comunidad, con la esperanza de inspirar a otros a unirse a nosotros en esta misión.

\*\*\*Valorar si incluir en artículo o reservarlo para nuestra revisión más adelante:

*No obstante... no os vamos a engañar. No pudimos resistir la tentación de consultar al chatbot de OpenAI cómo nos recomendaría seguir... y aquí os compartimos sus respuestas:*

Ah, y por cierto. Habréis visto que algunos de los párrafos de esta publicación se presentan en azul. Se trata de párrafos generados por el chat de OpenAI, al consultarle cómo nos proponía exponer las conclusiones de nuestro proyecto ☺

Gracias AI Saturdays por esta gran oportunidad!

## Dejamos a disposición...

\*\*\* Valorar si usar mismo github o "cerrar" este y abrir otro para pruebas

Compartimos el [Repositorio GitHub](https://github.com/josuemzx/Chatbot-para-refugiados) del proyecto, donde podréis encontrar información actualizada sobre el proyecto.

Respecto al MVP con una solucion programada, podéis encontrarlo disponible en:\*\*\*Valorar si incluir

## Recursos/Webgrafía:

\*\*\*Incluir aquí webs contexto, web BOE, XXX

# *[Borrador] Desarrollando un chatbot para refugiados: nuestra experiencia en Saturdays.AI*

## *Contexto y equipo*

*Llegados a este punto, podemos estar de acuerdo en el gran potencial que reside en el ámbito de la Inteligencia artificial (seguro que no lo habíais escuchado nunca ;)). En nuestro caso, incluso hemos podido experimentar cuán rápido está evolucionando hoy en el ámbito de NLP. Bajo nuestro punto de vista, en IA se están desarrollando múltiples aplicaciones que nos permitirán optimizar múltiples tareas y cada vez mejor, de forma que podamos centrarnos en nuestro día a día en aportar mayor valor.*

*Pero, centrémonos…*

## *Sobre el proyecto*

### *Introducción*

### *Método y Resultados*

*4.Nos planteamos presentar un chatbot (un chat interactivo) donde la persona que se encuentre en esta situación, pueda plantear consultas y este chat pueda proveer información relacionada, a la vez proporcionando la fuente.*

*4.Para ello, vimos que existen plataformas que recogen diferentes modelos preentrenados en NLP (Natural Language Processing).*

* *HuggingFace.*
* *Rasa.*

*4.La documentación de proyecto está disponible en github:\*\*\**

*4.Robot OpenAI*

*Supimos de la existencia de HuggingFace, una plataforma que XXX .*

*\*\*\*Uso de HuggingFace: poner la filosofía y el enlace (filosofía de usar modelos ya entrenados por volumen de computación, etc)*

*Y lo logramos. Llegamos a implementar un modelo de Question Answering (HuggingFace, más detalle más adelante en este artículo), que nos respondía de forma coherente a las consultas que le podíamos formular.*

*Estuvimos trabajando en la inclusión de datos específicos del colectivo refugiados procedentes del BOE. El modelo requería \*\*\*(fine tuning?) un documento en formato json y con determinadas entradas. Para generar este formato de docuemnto, encontramos una plataforma, Haystack con la que tratando los datos a nivel usuario (cargando los contextos), en la misma página se pueden incluir preguntas y respuestas. Y estos datos pueden exportarse tanto en formato Excel, csv o incluso json para los modelos BERT.*

*No obstante, había un problema… y es que si bien el modelo respondía bien, era muy poco humano. No era capaz de responder un “Hola” o un “Buenos días”. De forma que se nos hacía poco amigable…*

*\*\*\*Redactar una posible solución más adelante.*

*Y dado que el área de la IA está actualmente en auge y en especial en NLP a través del chat de OpenAI, decidimos buscar otro tipo de solución que englobase que el modelo fuese más “humano” y entrenarlo para que pudiese responder a casuisticas más concretas sobre el colectivo de refugiados.*

*\*\*\*Redactar otras opciones, aquí o más adelante.*

*\*\*\*Explicar qué es un modelo QA, un modelo Conversacional, y un modelo de Generación de texto.*

*El modelo ideal para ello hubiese sido un modelo de tipo Conversacional. Pero… problema (ojo al “dato”, guiño-guiño ;)): En el momento de implementar y trabajar el modelo, en HuggingFace no aparecían modelos entrenados a Español. Por lo tanto, decidimos seguir el modelo de OpenAI y optar por un modelo de Generación de texto.*

*Aunque el modelo de OpenAI está disponible para probarse, no se dispone del código en abierto. Y decidimos buscar alternativas que nos permitieran implementar un proyecto similar y profundizar o buscar en fuentes oficiales, información para el colectivo de refugiados. E incluso llegar a probar si funciona en diferentes idiomas.*

*Llegamos a localizar un modelo llamado bloom, entrenado por el \*\*\*pdte contrastar gobierno francés, de código disponible y subido a HuggingFace; que parecía ser potente en cuanto a generación de texto se refiere. Y nos pusimos a la obra para ajustar este modelo y poder usarlo para el fin que nos habíamos propuesto.*

*\*\*\*Redactar alternativa: Chat en gravio programando algunas preguntas y respuestas, para cumplir con el propósito. Aunque no incluía IA!*

*4.Explicad de forma ordenada como resolvéis el problema: tipo de problema, datos, exploración, insights, modelo, etc... Lo que habéis aprendido.*

*5.Exponed vuestros resultados.*

*6.Volved a hacer preguntas de implicación y terminar con los pasos futuros del proyecto.*

*7.Añadir link a la ppt, el código en github, la webapp o demo si tenéis...*

*Y efectivamente, tras revisar el notebook publicado en HuggingFace, conseguimos que el chat respondiese de forma coherente a las consultas propuestas. Y actualmente seguimos trabajando en nutrir el modelo con consultas específicas del colectivo de refugiados y ampliación de idiomas. O incluso considerar la integración de diferentes modelos, que de un lado permitan que un chatbot responda de forma más o menos humana y a su vez fuese capaz de proveer un contexto y respuestas concretas…*

### *Conclusiones y próximos pasos*

*Nuestro objetivo es poder proveer de información al colectivo de refugiados, de forma que las consultas generales puedan resolverse a través de un chatbot capaz de encontrar las respuestas disponibles en web, compatible en diferentes idiomas. Y de este modo, reservar los recursos disponibles (generalmente voluntariado) a tareas o casos que requieran una revisión más específica.*

*Pretendemos contribuir a una mejor destinación de los recursos disponibles para poder brindar un servicio más especializado en los casos que se requiera.*

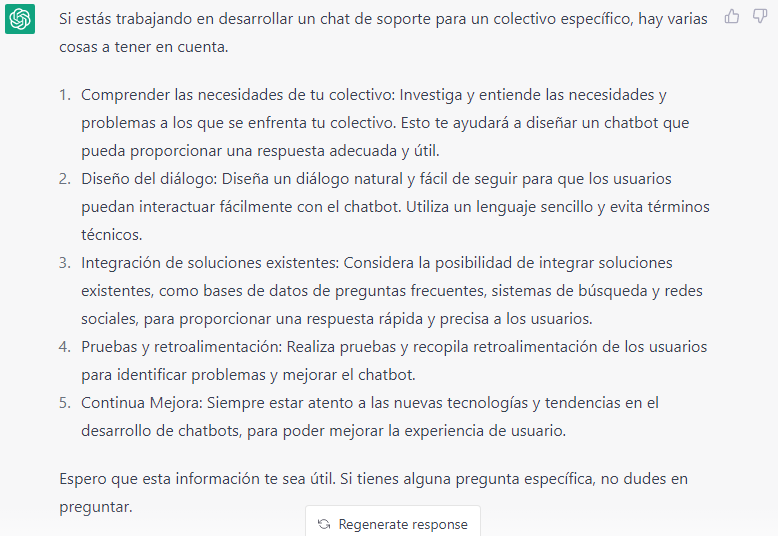
# Recursos a tener en cuenta

## Eduflow

[Eduflow] Para escribir el borrador os recomiendo seguir varias pautas:

1. Empezad conectando con el lector. Haced preguntas de implicación. Y exponed el tema en un párrafo. ([ejemplo](https://medium.com/saturdays-ai/de-cero-a-chatbot-4204612c5c72?source=collection_category---6------0-----------------------))
2. Presentad brevemente quienes sois y por qué os motiva este proyecto.
3. Enlazar lo anterior con la técnica a través de una historia. Puede ser vuestra misma historia desde que entráis en AI Saturdays, algo personal o cualquier otro ejemplo.
4. Explicad de forma ordenada como resolvéis el problema: tipo de problema, datos, exploración, insights, modelo, etc... Lo que habéis aprendido.
5. Exponed vuestros resultados.
6. Volved a hacer preguntas de implicación y terminar con los pasos futuros del proyecto.
7. Añadir link a la ppt, el código en github, la webapp o demo si tenéis...

OpenAI:



## Webgrafía

-Bibliografía o información de partida-

**Chatbot para la resolución de consultas, trámites y soporte a colectivos de riesgo (título pendient ajustar\*\*\*)**

Equipo AI Saturdays formado por:

Antonio Linde, Ivy Chen, Josué Huaman, Silvia Miró

**Problemática (pendiente documentar\*\*\*)**

Nuestro objetivo es poder dar soporte o facilitar el acceso a información por parte de colectivos que suelan trasladarse a servicios asistenciales, que puedan estar actualmente gestionados por voluntarios y/o saturados. De forma que el acceso a estos servicios pueda optimizarse y poder destinar los recursos de una forma más específica y proveer atención más especializada.

**Solución propuesta (hipótesis, pendiente acordar equipo\*\*\*)**

0. [MVP] Creación de un chatbot que al recibir una consulta en formato de enunciado o pregunta:

1. Localice información relacionada en la página oficial,
2. Devuelva:
   1. Explicación de la situación en lenguaje “de persona no especialista”
   2. Fuente: fragmento de la página origen de la información, y enlace. Y fragmento/s de importancia subrayados.

En siguientes fases, valorar si incluir:

* Identificación de idioma de entrada, así como traducción de la respuesta al idioma en cuestión
* Nuevas funcionalidades (detección necesidades y nivel escolarización, servicios sociales…)
* Análisis de sentimiento y sugerir-derivar-alertar a otros servicios

**Proceso**

Se revisan posibles soluciones, modelos previamente entrenados que puedan servir como base para la construcción de un chatbot con estas características. Se localizan soluciones en:

* HuggingFace.
* Rasa.

Partimos del modelo XXX de HuggingFace, previamente preentrenado.

Elaboramos nueva bdd para hacer fine-tunning del modelo, planteando consultas de entrada para el modelo y fuente para su resolución.

**Modelos Question answering en ESP**

Modelos encontrados

BERT en ESP:

<https://jaimesendraberenguer.medium.com/bert-para-responder-preguntas-sobre-squad-2-0-en-espa%C3%B1ol-5842748f051a>

Modelos encontrados y validados

**Fuentes de datos:**

* A considerar para presentación:

<https://www.un.org/es/global-issues/migration>

<https://www.un.org/es/global-issues/refugees>

<https://peacekeeping.un.org/es> --> Misiones y su estado actual

<https://www.cear.es/projects/traduccion-e-interpretacion/> 🡪 Justificación en caso de incluir idiomas en el bot. Apoyo a servicio de traducción e interpretaciones

* A cosiderar para artículo y presentación, funcionamiento BLOOM:

<https://huggingface.co/docs/transformers/model_doc/bloom>

<https://huggingface.co/blog/bloom-megatron-deepspeed>

<https://openreview.net/forum?id=UoEw6KigkUn>

<https://openreview.net/pdf?id=UoEw6KigkUn>

<https://github.com/bigscience-workshop/Megatron-DeepSpeed>

* Idea de estructura para ReadMe github:

<https://github.com/bigscience-workshop/petals>

* De interés para el equipo de trabajo, contextualización:

<https://www.acnur.org/pacto-mundial-sobre-refugiados.html> --> Contexto. El [Pacto Mundial sobre los Refugiados](http://acnur.org/5c782d124) es un marco para una distribución previsible y equitativa de la carga y la responsabilidad, reconociendo que una solución sostenible a las situaciones de refugiados no se puede lograr sin la cooperación internacional.

<https://www.acnur.org/noticias/noticia/2016/7/5b9008e74/refugiado-o-migrante-cual-es-el-termino-correcto.html> --> Refugiado o Migrante? Pregunta: quién/cómo se define?\*\*\*

<https://elpais.com/internacional/2015/08/28/actualidad/1440781136_652160.html> --> Refugiado o Migrante

* Generar base de datos – Preguntas:

<https://www.acnur.org/fileadmin/Documentos/BDL/2011/7575.pdf?view#page=12>

<https://www.acnur.org/5c782d124#_ga=2.191202474.1084684248.1668149931-1264021629.1668149931> 🡪 Desde la página 22 (1.2). Este documento nos puede servir de respaldo o de esquema para generar ideas de preguntas o ámbitos de atención. Sobretodo desde el punto B. Ambitos que requieren apoyo > 2.Atención de las necesidades y apoyo a las comunidades.

<https://www.icj-cij.org/en> --> Corte Internacional de Justicia. Resoluciones actualizadas, textos jurídicos. Valorar si puede ser de interés, echando un primer vistazo entiendo que se trata de resoluciones a nivel de instituciones y habría que ver si se puede bajar o si puede tener impacto a nivel de consultas individuales. Por ejemplo: <https://www.icj-cij.org/public/files/publications/handbook-of-the-court-en.pdf> --> 8. Cases Brought Before the Court. Otra opción sería que pudiera funcionar como titulares o incluso generar notificaciones en app. A considerar que este documento es de 2018

<https://www.iom.int/es> --> **La OIM trabaja en estrecha colaboración con sus asociados gubernamentales, intergubernamentales y no gubernamentales.**

* **Migración y desarrollo**
* **Migración facilitada**
* **Migración reglamentada**
* **Migración forzada.**
* Esta pagina la veo como que se pudiera desarrollar notificaciones al publicarse nuevo contenido.

<https://www.acnur.org/noticias/noticia/2016/3/5b9007344/preguntas-frecuentes-sobre-los-terminos-refugiados-y-migrantes.html> --> Preguntas frecuentes sobre los términos ‘refugiados’ y ‘migrantes’

* <https://www.acnur.org/5b0766944.pdf> --> Convención sobre el Estatuto de los Refugiados, 1951. Este documento creo que puede ser interesante tomarlo como glosario.
* <https://www.acnur.org/5b076e994#_ga=2.124149898.1084684248.1668149931-1264021629.1668149931> 🡪 Convención de la oua por la que se regulan los aspectos específicos de problemas de los refugiados en África

<https://www.cear.es/> 🡪 Comisión Española de Ayuda al Refugiado

* <https://www.cear.es/sections-post/area-juridica-cear/> 🡪 Principales actuaciones (para orientar consultas a chatbot, para introducción en presentación) e Información Jurídica (alimentar modelo).

<https://www.cear.es/informes-pais-origen/> 🡪 Se podría considerar este enlace para generar filtro en la consulta de entrada. Valorar si contactar a [**centro.documentacion@cear.es**](http://centro.documentaci%C3%B3n@cear.es/) para recibir información de años anteriores o incluso poder generar modelo predictivo?

Ideas de preguntas (valorar si poner desde país origen, y cómo garantizar confidencialidad. Quizá dejando el servicio a disposición de XX organización?):

* Cómo puedo ser reconocido como refugiado?
* A qué países me puedo dirigir?
* Qué trámites debo realizar para ser reconocido como refugiado?
* Qué trámites debo realizar para tener derechos como refugiado?
* Qué documentación debo aportar y cómo voy a estar protegido?
* Cómo puedo saber si estoy realmente protegido por el país de acogida?
* En qué consiste el derecho de asilo / derecho de acogida? Cómo puedo prolongarlo en el tiempo?
* Cómo puedo conseguir Asesoramiento legal?
* Tendré problema en la acogida por el idioma? Hay servicios de traducción, en qué centro y cómo solicitarlos?

## OpenAI

Hola!

Hola, ¿en qué puedo ayudarte hoy?

Me podrias ayudar a redactar un artículo?

¡Por supuesto! ¿De qué tema te gustaría que sea el artículo? Por favor proporciona más detalles como el título, el objetivo o el público objetivo. Asimismo, si tienes algún formato específico en mente para el artículo, házmelo saber.

Somos un equipo de trabajo dentro de la organización Saturdays.AI, una organización que brinda formaciones basadas en la práctica y centrada en desarrollar proyectos de impacto social. Hemos estado trabajando en desarrollar un chatbot para refugiados; inicialmente vimos varias opciones para desarrollar el chatbot (rasa y huggingface), finalmente tiramos por huggingface. y dentro de huggingface iniciamos usando un modelo de question answering, con la idea de dotar de información específica de páginas web a consultas específicas. No obstante, luego vimos que el chatbot era "poco humano" y empezamos a trabajar en un modelo de generación de texto, en este caso bloom. aunque el modelo nos da respuestas generales, nos gustaría poder acabar de hacer el fine tuning para la area de refugiados en específico y probarlo en diferentes idiomas, o incluso valorar nuevas opciones

Me parece un proyecto muy interesante y enriquecedor. Con base en la información que me proporcionaste, el título del artículo podría ser algo así como "Desarrollando un chatbot para refugiados: nuestra experiencia en Saturdays.AI" y el objetivo sería documentar el proceso de desarrollo del chatbot, incluyendo las diferentes opciones que consideraron, las razones por las que optaron por una en particular y los desafíos y aprendizajes que tuvieron. También podrías incluir la idea de futuros pasos, como la posibilidad de hacer fine tuning específico para el área de refugiados y probar el chatbot en diferentes idiomas.

Aquí te dejo un borrador:

Título: Desarrollando un chatbot para refugiados: nuestra experiencia en Saturdays.AI

Introducción: En Saturdays.AI, una organización que brinda formaciones basadas en la práctica y centrada en desarrollar proyectos de impacto social, hemos estado trabajando en desarrollar un chatbot para refugiados. Nuestro objetivo es brindar información específica y relevante a las consultas de los refugiados, para facilitarles el acceso a recursos y servicios que les sean útiles.

Desarrollo del chatbot: Inicialmente, consideramos varias opciones para desarrollar el chatbot, como Rasa y Hugging Face. Finalmente, optamos por Hugging Face y comenzamos a trabajar con un modelo de pregunta-respuesta. Sin embargo, pronto nos dimos cuenta de que el chatbot no era "lo suficientemente humano" y decidimos trabajar con un modelo de generación de texto. En este caso, utilizamos Bloom, un modelo de generación de texto desarrollado por OpenAI. Aunque el modelo nos brinda respuestas generales, estamos trabajando en hacer fine tuning específico para el área de refugiados y probar el chatbot en diferentes idiomas.

Desafíos y aprendizajes: A lo largo del proceso, nos enfrentamos a varios desafíos, como la necesidad de encontrar una forma de brindar respuestas relevantes y precisas a las consultas de los refugiados. También tuvimos que aprender a trabajar con herramientas y tecnologías nuevas. Sin embargo, estos desafíos también nos permitieron aprender mucho y desarrollar habilidades valiosas.

Conclusión: Aunque todavía queda trabajo por hacer, estamos emocionados de haber llegado hasta aquí en el des

(Suena interesante lo que propones! No obstante creo que te has quedado a media respuesta.. podrías seguir por favor? Gracias!)

Claro, disculpa si me quedé a medio camino. Me complace saber que te parezca interesante el artículo propuesto. Aquí continúo:

Conclusión: Aunque todavía queda trabajo por hacer, estamos emocionados de haber llegado hasta aquí en el desarrollo de nuestro chatbot para refugiados. Creemos que este proyecto tiene el potencial de marcar una gran diferencia en la vida de las personas que huyen de sus hogares y buscan una nueva oportunidad. Continuaremos trabajando en mejorar nuestro chatbot, haciendo fine tuning específico para el área de refugiados y probando el chatbot en diferentes idiomas. Esperamos poder compartir nuestros resultados y aprendizajes con la comunidad, con la esperanza de inspirar a otros a unirse a nosotros en esta misión.

Uau muchas gracias!

¡De nada! Me alegra haber podido ayudarte. Si tienes alguna otra pregunta o necesitas más información, no dudes en preguntar. Estoy aquí para ayudarte.