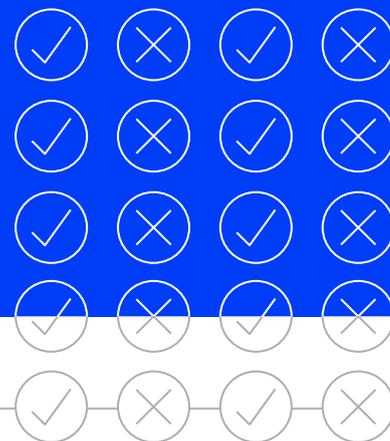


¿Son claras las respuestas del Ayuntamiento de Madrid?



El Ayuntamiento de Madrid quiere que su comunicación con la ciudadanía sea clara y transparente. El primer paso para conseguirlo ha sido evaluar la calidad de sus respuestas y, así, en un segundo paso, mejorarlas. Con este fin, se ha embarcado en un proyecto ambicioso: **usar la tecnología como medio para analizar la claridad de sus comunicaciones.**

Objetivo del proyecto

Analizar la calidad de las respuestas del Ayuntamiento de Madrid a las solicitudes de la ciudadanía. Dos ejes de análisis:

- Claridad.
- Personalización.

Cómo se hizo: Con un modelo de aprendizaje automático

Es un **software** que predice la claridad y personalización de las comunicaciones mediante técnicas de aprendizaje automático y Procesamiento del Lenguaje Natural.

5 pasos del proceso:



1 Recoger, limpiar y ordenar un corpus de textos:

Casi 200.000 pares de solicitudes y respuestas sirven de ejemplo al modelo.

2 Definir las métricas de claridad y personalización:

Indicadores que ayudan al modelo a evaluar los textos.

a. 9 métricas lingüísticas para medir la claridad:

- Uso de palabras fuera del diccionario.
- Uso de conectores discursivos.
- Número de signos de puntuación por palabra.
- Referencias y citas literales a leyes, decretos, etc.
- Número de estructuras pasivas por número de palabras.
- Número de nexos subordinantes por número de palabras.
- Detección de tecnicismos.
- 1000 palabras más comunes en español.
- Número medio de palabras por frase.

b. 5 métricas lingüísticas para medir la personalización:

- Coincidencia de lemas.
- Uso de títulos indeterminados.
- Distancia de solicitud-respuesta.
- Detección de plantillas predefinidas.
- Coincidencia de entidades nombradas.

3 Entrenar el mejor modelo de aprendizaje automático:

Encontrar la mejor configuración del modelo e implementar las métricas y el corpus para que los analice.

4 Analizar los resultados del entrenamiento:

Evaluar si el modelo ha “aprendido” a analizar los textos de forma correcta.

5 Instalar el modelo en una API Rest:

Instalar el modelo en una API rest o servicio **online** para que el Ayuntamiento pueda hacer pruebas y analizar más de sus comunicaciones.

Conclusiones

Las respuestas del Ayuntamiento son:

- **Claros** porque no contienen muchas palabras fuera del diccionario y no incluyen demasiadas referencias a leyes.
- **Claros, pero necesitan mejorar** en el uso de signos de puntuación, las estructuras pasivas, el número medio de palabras por frase, el número de tecnicismos, las frases subordinadas y el uso de las palabras más comunes en español.
- **No son claros** en cuanto al uso de conectores discursivos.
- **Personalizadas** porque limitan el uso de títulos indeterminados y plantillas predefinidas. Además, las respuestas están relacionadas con los temas de las soluciones.
- **No son personalizadas** porque las respuestas no comparten el vocabulario general ni las entidades mencionadas con el texto de la solicitud.