

OPTIMIZACIÓN DE BÚSQUEDA AIRBNB

DATOS MASIVOS

RESUMEN DEL PROYECTO

La creciente cantidad de alojamientos en Airbnb dificulta a los usuarios encontrar la opción perfecta para sus necesidades.

Nuestra solución basada en datos masivos permitirá a los usuarios obtener recomendaciones altamente precisas y ajustadas a sus preferencias, maximizando así su satisfacción y brindándoles una experiencia de viaje excepcional.



RETOS

RETO 1
**DEMASIADAS
OPCIONES**



RETO 2
**ORDENAMIENTO
RESTRICTIVO**



RETO 3
**SESGOS DE
INFORMACIÓN**



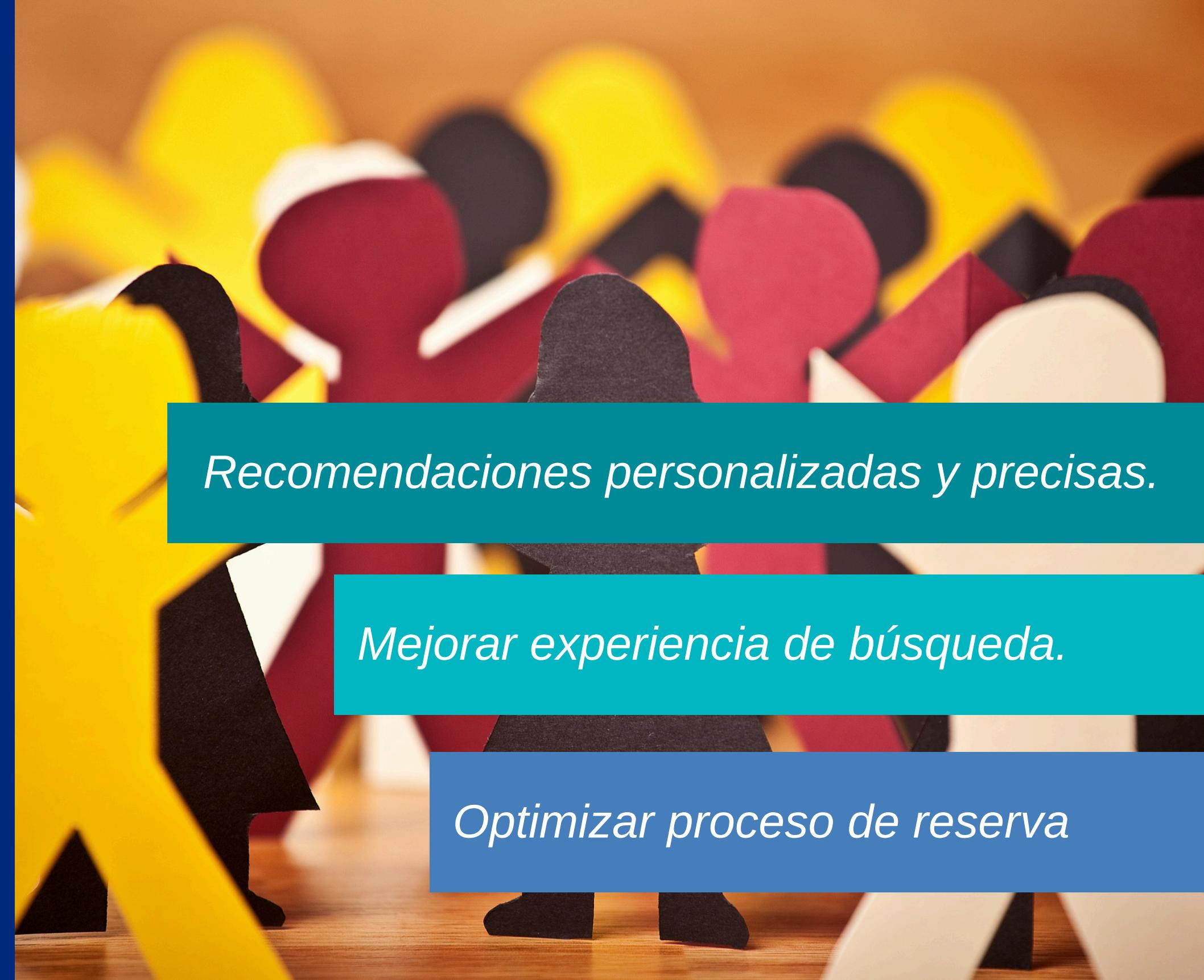
RETO 4
**DIFICULTAD DE
COMPARACIÓN**



METAS DEL PROYECTO

Simplificar y agilizar el proceso de búsqueda de alojamientos, ahorrando tiempo y esfuerzo a los usuarios.

Queremos que puedan encontrar rápidamente opciones relevantes, comparar características clave y tomar decisiones informadas de manera eficiente.



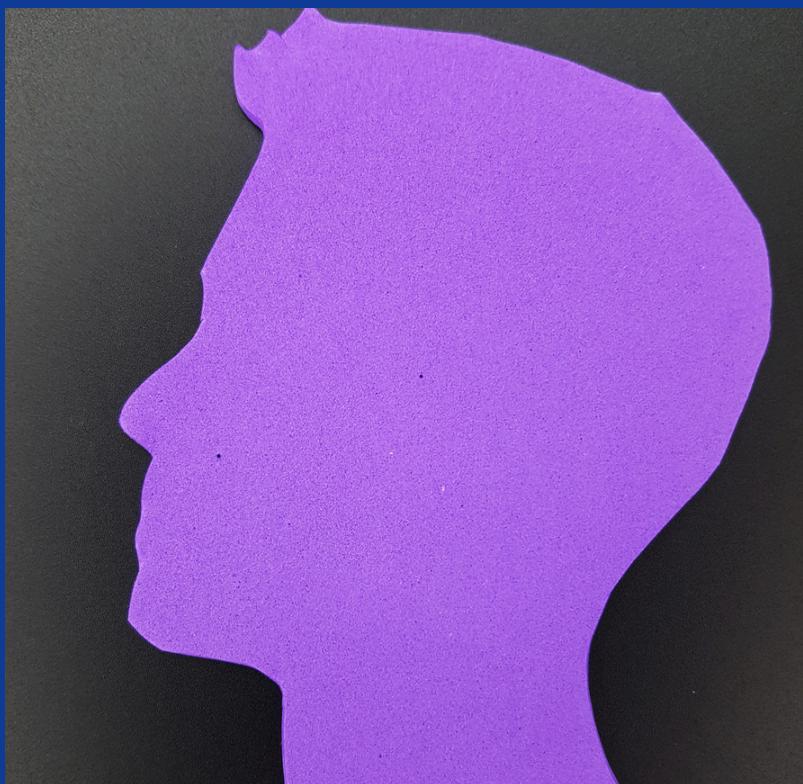
Recomendaciones personalizadas y precisas.

Mejorar experiencia de búsqueda.

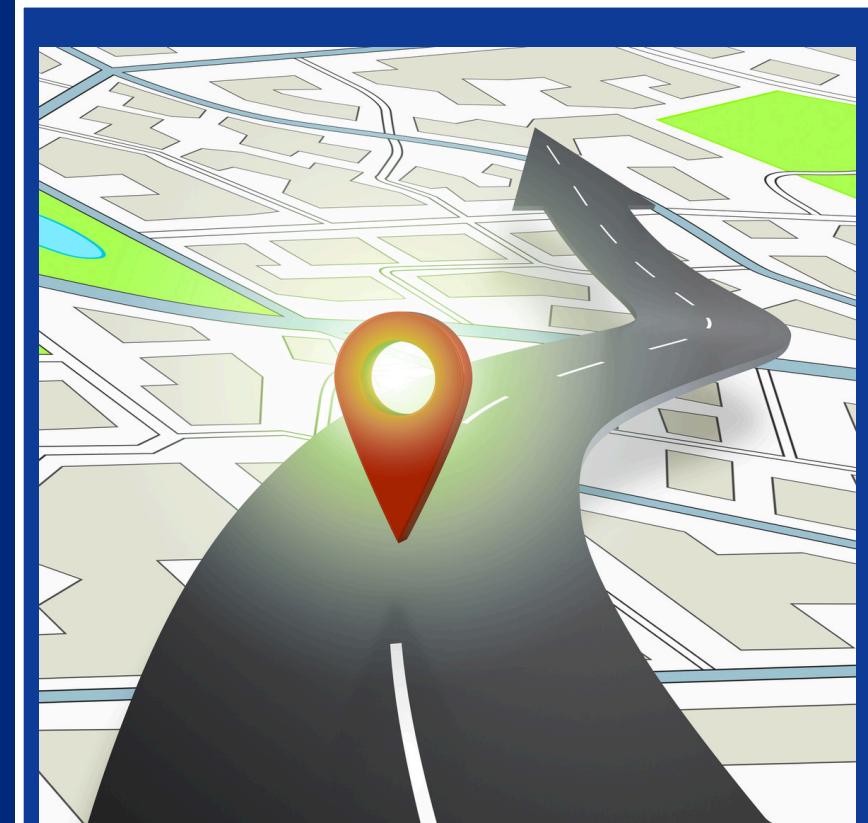
Optimizar proceso de reserva

FILTRADO AIRBNB

ASPECTOS QUE TOMA AIRBNB PARA
RECOMENDACIÓN.



PERFIL



CERCANÍA



FECHA

PROCESAMIENTO DE LENGUAJE NATURAL

- El crecimiento exponencial de los datos disponibles en forma de texto, tanto en línea como fuera de línea, ha generado la necesidad de desarrollar tecnologías capaces de comprender, interpretar y generar texto de manera efectiva.
- El NLP permite a las organizaciones extraer información valiosa, realizar análisis de sentimientos, automatizar tareas basadas en texto y mejorar la interacción humano-computadora a través de la comprensión y generación de lenguaje natural.



PROYECTO

Emplear alternativas de procesamiento de texto para encontrar maneras más eficientes de comparar información en las búsquedas de airbnb.

Al encontrar mejores resultados en las búsquedas podemos emplear técnicas más complejas para el procesamiento de lenguaje natural.



EXTRACCIÓN

- Trabajaremos con datos extraídos de airbnb, específicamente con las descripciones de todos los alojamientos en CDMX

PROCESAMIENTO

- Al obtener los datos con web scrapping debemos limpiarlos para poder trabajar con ellos

IMPLEMENTACIÓN

- Ocuparemos el algoritmo de índice inverso para poder comparar grandes cantidades de texto



ÍNDECE INVERSO

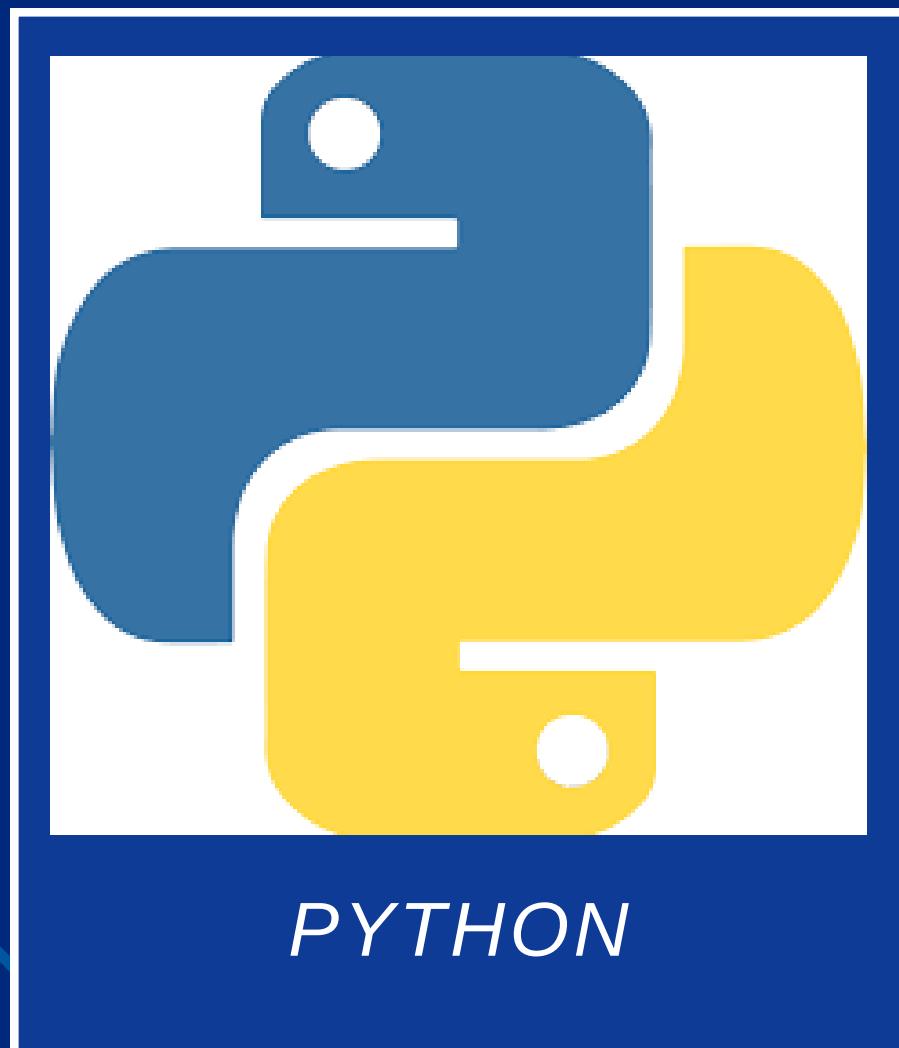
Método para altos grados de similitud que consiste en crear un índice que mapea cada palabra o término del texto con los documentos en los que aparece.

La presencia de un índice invertido acelera el proceso de búsqueda, ya que no es necesario examinar cada documento individualmente.



TECNOLOGÍAS

Herramientas usadas



MÉTRICAS

Comparación de similitudes con diferentes métodos

	<i>Test 1</i>	<i>Test 2</i>	<i>Test 3</i>
<i>Distancia del coseno</i>	.44	.51	.64
<i>Hamming</i>	.86	.89	.86
<i>Jaccard</i>	.13	.31	.13

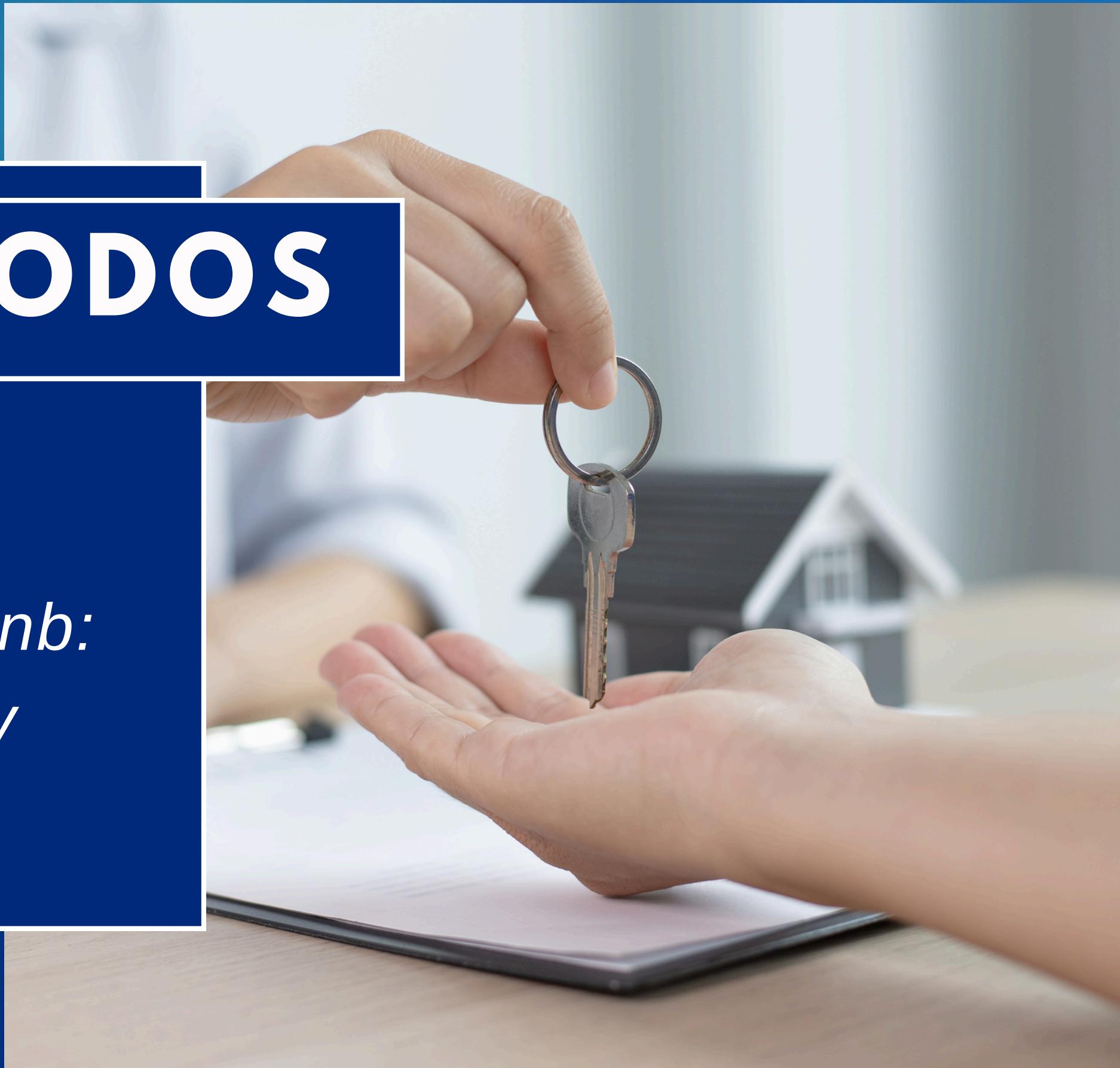
BENEFICIOS

1. Búsqueda eficiente: El método del índice invertido permite una búsqueda rápida y eficiente de alojamientos en Airbnb.
2. Personalización y relevancia: El índice invertido facilita la personalización y la entrega de resultados relevantes a los usuarios.
3. Filtrado avanzado: Con el método del índice invertido, se pueden aplicar filtros y criterios de búsqueda avanzados para refinar los resultados.
4. Actualización en tiempo real: El índice invertido se puede mantener actualizado en tiempo real.



GRACIAS A TODOS

*El futuro de la
búsqueda en Airbnb:
personalizada y
eficiente.*



PÁGINA DE REFERENCIAS

Lu, Z., & Peng, H. (2020). Airbnb Experience Description Analysis Based on Natural Language Processing. *IEEE Access*, 8, 104758-104766. DOI: 10.1109/ACCESS.2020.3003580

Gerendas, J., & Schlegl, T. (2020). Machine Learning Approach to Predicting Airbnb Rental Prices. In *The World in a Matrix* (pp. 89-105). Springer, Cham. DOI: 10.1007/978-3-030-67462-4_7

. Leskovec, J., Rajaraman, A. and Ullman, J. (2016). *Mining of massive datasets*. Delhi: Cambridge University Press.

1. Bendersky, M., Metzler, D., & Croft, W. B. (2011). Effective Query Formulations for Finding Implicit Information in Email. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 75-84). DOI: 10.1145/2009916.2009930

Pechenizkiy, M., & Wojciechowski, M. (2013). *New trends in databases and information systems*. Berlin: Springer.