

# Sistema Multiagente para Guía Turístico Virtual de Cuba

Arianne Camila Palancar Ochando (C-311) and Jabel Resendiz Aguirre (C-312)

Universidad de La Habana, Cuba

**Abstract.** Este trabajo presenta un análisis detallado del Sistema Inteligente de Guía Turístico Virtual de Cuba, una aplicación desarrollada con una arquitectura multiagente que combina procesamiento de lenguaje natural, búsqueda semántica y técnicas de geocodificación para proporcionar información turística contextual sobre Cuba. El sistema incorpora inteligencia artificial para interpretar consultas naturales de los usuarios, extraer información relevante de una base de conocimientos especializada, y presentar respuestas enriquecidas con mapas interactivos, información meteorológica y planificación de itinerarios. A través de una evaluación tanto cualitativa como cuantitativa, se identifican fortalezas como la modularidad y extensibilidad del sistema, así como áreas de mejora en términos de optimización de rendimiento y ampliación de fuentes de datos. El trabajo demuestra la viabilidad de implementar un asistente turístico interactivo utilizando tecnologías de código abierto que podrían adaptarse a otros destinos turísticos.

## 1 Introducción

El turismo constituye un sector económico fundamental para Cuba, con aproximadamente 4.7 millones de visitantes internacionales en años recientes [1]. La disponibilidad de información precisa, contextual y accesible representa un factor determinante en la experiencia del turista, tanto en la fase de planificación como durante el viaje. En este contexto, las tecnologías de asistentes virtuales ofrecen una oportunidad para mejorar significativamente la experiencia del usuario mediante interacciones naturales y personalizadas.

Este trabajo analiza una implementación de un Sistema Inteligente de Guía Turístico Virtual para Cuba, desarrollado como una aplicación web interactiva utilizando una arquitectura multiagente. El sistema está diseñado para responder consultas en lenguaje natural sobre destinos turísticos cubanos, con énfasis en museos y excursiones, proporcionando información enriquecida con elementos visuales como mapas interactivos, datos meteorológicos e itinerarios personalizados.

A diferencia de los sistemas de recomendación turística tradicionales, que suelen basarse en filtrado colaborativo o reglas predefinidas, este sistema adopta un enfoque modular basado en agentes especializados que trabajan coordinadamente para procesar consultas, recuperar información relevante y presentar respuestas contextualizadas.

## 2 Repositorio del Proyecto

El proyecto se encuentra alojado en la plataforma GitHub bajo la siguiente URL:  
<https://github.com/kmy-cyber/Tourist-guide>

El repositorio contiene el código fuente completo de la aplicación y documentación técnica

## 3 Descripción del Tema

El proyecto aborda la creación de un asistente virtual especializado en información turística sobre Cuba, centrándose específicamente en museos y excursiones. El sistema permite a los usuarios realizar consultas en lenguaje natural, como si estuvieran conversando con un guía turístico humano. Por ejemplo, un usuario puede preguntar: "¿Qué museos puedo visitar en La Habana con niños?" o "¿Cuál es el clima actual en Varadero y qué actividades recomiendas?".

El problema que aborda este sistema es multidimensional:

1. **Dispersión de la información turística:** Los datos sobre atracciones turísticas en Cuba suelen estar dispersos en múltiples fuentes y con formatos heterogéneos.
2. **Contextualización de recomendaciones:** El sistema debe considerar factores como ubicación, clima actual, preferencias del usuario y restricciones específicas.
3. **Interacción natural y personalizada:** A diferencia de los buscadores tradicionales, se busca proporcionar una experiencia conversacional personalizada.
4. **Integración multimodal:** La presentación de información combina texto, mapas interactivos, datos meteorológicos e itinerarios sugeridos.

La importancia del tema radica en su potencial para mejorar la experiencia turística en Cuba mediante la aplicación de tecnologías avanzadas de procesamiento de información y presentación contextual, así como en la exploración de un modelo de arquitectura que podría adaptarse a otros destinos o dominios de información.

## 4 Análisis de la Solución Implementada

La solución implementada adopta una arquitectura basada en agentes especializados, donde cada agente se responsabiliza de una funcionalidad específica del sistema. Como se puede observar en la Figura 1

### 4.1 Arquitectura Multiagente

El núcleo del sistema está compuesto por una serie de agentes especializados que trabajan de forma coordinada. El Agente Coordinador actúa como el cerebro

central del sistema, orquestando el flujo de trabajo entre los diferentes agentes y manteniendo el estado global de la conversación. La gestión del conocimiento se realiza a través del Agente de Conocimiento, que administra la base de conocimientos turísticos y ejecuta búsquedas semánticas para recuperar información relevante según las necesidades del usuario.

La generación de respuestas naturales y coherentes está a cargo del Agente LLM, que utiliza modelos de lenguaje de gran escala, específicamente Llama v3 a través de Fireworks AI. Para el manejo de información geoespacial, el sistema cuenta con el Agente de Ubicación, que se encarga de identificar y geocodificar referencias a ubicaciones en Cuba mencionadas en las consultas y respuestas. Este trabajo se complementa con el Agente Meteorológico, que obtiene información del clima en tiempo real para las ubicaciones identificadas.

La experiencia del usuario se enriquece gracias al Agente de Interfaz de Usuario, que genera componentes visuales como mapas interactivos y widgets de clima. La personalización de la experiencia se logra mediante el Agente de Usuario, que gestiona los perfiles de usuario y mantiene el contexto de las conversaciones. Finalmente, el Agente de Planificación se encarga de generar itinerarios personalizados basados en las preferencias del usuario y los datos contextuales disponibles.

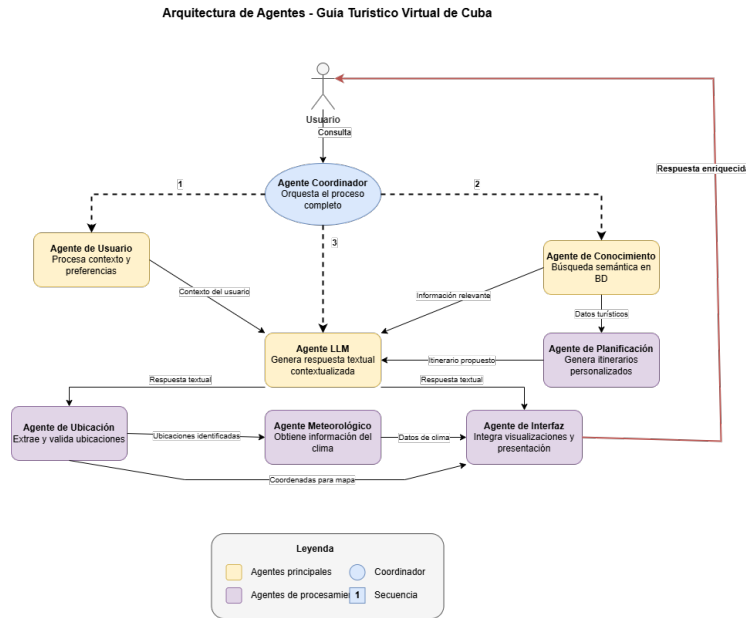
Esta arquitectura, que sigue el patrón de diseño multiagente, permite una alta modularidad, reutilización y extensibilidad del sistema. Cada agente opera de manera independiente pero coordinada, comunicándose a través de un contexto compartido que facilita la integración de funcionalidades.

## 4.2 Flujo de Procesamiento

El procesamiento de una consulta de usuario sigue un flujo sofisticado y bien definido que aprovecha las capacidades especializadas de cada agente. El proceso comienza cuando el usuario ingresa una consulta en lenguaje natural a través de la interfaz web. Inmediatamente, el Agente de Usuario procesa esta entrada para extraer información contextual relevante, considerando el historial de interacciones y las preferencias previamente establecidas.

A continuación, el Agente de Conocimiento ejecuta una búsqueda semántica en la base de conocimientos para recuperar información pertinente. Simultáneamente, el Agente de Planificación evalúa si la naturaleza de la consulta requiere la generación de un itinerario personalizado. Con esta información recopilada, el Agente LLM genera una respuesta textual coherente y contextualizada.

El procesamiento continúa con el Agente de Ubicación, que analiza la respuesta generada para extraer y validar referencias a ubicaciones específicas. Para estas ubicaciones identificadas, el Agente Meteorológico obtiene información actualizada del clima, permitiendo recomendaciones más precisas y contextuales. Finalmente, el Agente de Interfaz de Usuario integra todos estos elementos en una presentación cohesiva, enriqueciendo la respuesta con componentes visuales como mapas interactivos y widgets informativos.



**Fig. 1.** Comunicación entre agentes

### 4.3 Base de Conocimiento y Búsqueda Semántica

La base de conocimientos turísticos constituye el núcleo informativo del sistema, almacenando de manera estructurada datos sobre diversos aspectos del turismo en Cuba. La información se organiza en colecciones especializadas que abarcan museos, excursiones y destinos generales. Para los museos, se mantiene información detallada sobre ubicación, horarios, precios, colecciones y servicios disponibles. Las excursiones se documentan con información sobre duración, nivel de dificultad, servicios incluidos y puntos de encuentro. La colección de destinos proporciona una visión general de las ciudades y lugares turísticos más relevantes de Cuba.

El sistema implementa una búsqueda semántica avanzada mediante embeddings vectoriales, utilizando la tecnología SentenceTransformers. Esta aproximación permite recuperar información relevante incluso cuando las consultas de los usuarios no coinciden exactamente con los términos almacenados en la base de conocimientos. El modelo "all-MiniLM-L6-v2" se utiliza para generar vectores de incrustación tanto para las consultas como para los documentos, permitiendo comparaciones semánticas precisas y resultados más relevantes para los usuarios.

#### 4.4 Generación de Respuestas con LLM

Para la generación de respuestas naturales y contextualmente relevantes, el sistema utiliza el modelo Llama v3 a través de la API de Fireworks AI. El Agente LLM construye un prompt estructurado que incluye:

1. **Instrucciones de rol:** Define el rol del sistema como un guía turístico experto en Cuba.
2. **Contexto de conocimiento:** Incorpora la información recuperada de la base de conocimientos.
3. **Información del usuario:** Incluye preferencias, consultas anteriores y nombre del usuario si está disponible.
4. **Datos contextuales:** Añade información del clima y ubicaciones identificadas.

Este enfoque de prompting estructurado permite al modelo generar respuestas que son informativas, contextuales y naturales, aprovechando tanto el conocimiento general del modelo como la información específica recuperada por el sistema.

#### 4.5 Interfaz de Usuario y Visualización

La interfaz de usuario se ha implementado utilizando Streamlit, un framework de Python para la creación rápida de aplicaciones web interactivas. La elección de Streamlit se justifica por:

1. **Facilidad de desarrollo:** Permite crear interfaces interactivas con poco código.
2. **Integración nativa con Python:** Se integra perfectamente con el resto del sistema.
3. **Componentes interactivos:** Proporciona widgets para mapas, gráficos y otros elementos visuales.

La interfaz presenta tres componentes principales de visualización:

1. **Mapas interactivos:** Utilizando Folium, el sistema muestra ubicaciones mencionadas en un mapa interactivo con marcadores personalizados.
2. **Widgets de clima:** Presenta información meteorológica actual en un formato visual atractivo.
3. **Planificador de itinerarios:** Muestra itinerarios sugeridos con actividades organizadas por día y hora.

Estos componentes visuales enriquecen la experiencia del usuario proporcionando información contextual de manera intuitiva y atractiva.

### 5 Consideraciones de Implementación

El desarrollo del sistema ha requerido un cuidadoso balance entre funcionalidad, mantenibilidad y rendimiento. Las decisiones de implementación se han guiado por la necesidad de crear un sistema robusto y adaptable que pueda evolucionar con las necesidades cambiantes del dominio turístico.

### 5.1 Sistema de Crawler Dinámico

Una de las innovaciones más significativas en la implementación del sistema es el crawler dinámico inteligente, implementado a través de la clase `SmartCrawler`. Este componente se integra dentro del Agente de Conocimiento (`KnowledgeAgent`) como un mecanismo de respaldo que se activa cuando los crawlers especializados no logran obtener información completa o actualizada. La integración se realiza a través de la clase `SimpleCrawlerIntegration`, que actúa como interfaz entre el agente y el crawler dinámico.

El flujo de operación del crawler dinámico en el sistema sigue esta secuencia:

1. El `KnowledgeAgent` recibe una consulta del usuario y primero intenta resolverla utilizando los crawlers especializados registrados.
2. Si los crawlers especializados no encuentran resultados o la información está incompleta, el agente activa el crawler dinámico a través del método `_search_with_crawlers`.
3. El crawler dinámico analiza la consulta y crea un conjunto inicial de datos que necesitan ser completados o actualizados.
4. A través del método `process_query` de `SimpleCrawlerIntegration`, el sistema intenta mejorar la información utilizando múltiples fuentes web.

### 5.2 Gestión de Datos y Actualización

El sistema implementa una estrategia de gestión de datos en capas que combina almacenamiento local optimizado con actualización dinámica de información. La capa base utiliza un sistema de caché inteligente que almacena resultados frecuentemente solicitados, reduciendo la latencia de respuesta y la carga en las fuentes de datos externas. Este caché se gestiona de manera proactiva, con políticas de expiración basadas tanto en tiempo como en relevancia de la información.

La actualización de datos se realiza mediante un proceso híbrido que combina actualizaciones programadas con actualizaciones bajo demanda. Los crawlers especializados se ejecutan periódicamente para mantener actualizada la información base, mientras que el crawler dinámico complementa esta información en tiempo real según sea necesario. Este enfoque asegura un balance óptimo entre la frescura de los datos y la eficiencia del sistema.

### 5.3 Optimización de Rendimiento

El procesamiento asíncrono se implementa utilizando `asyncio` de Python, permitiendo la ejecución concurrente de múltiples tareas sin bloquear el hilo principal. Esta característica es particularmente importante en el procesamiento de consultas complejas que requieren información de múltiples fuentes.

Las operaciones de búsqueda y recuperación de información se optimizan mediante un sistema de índices vectoriales que permite búsquedas semánticas eficientes. Los vectores de incrustación se calculan y almacenan durante el proceso de ingesta de datos, eliminando la necesidad de cálculos costosos durante el tiempo de consulta.

## 6 Justificación Teórico-Práctica

Las decisiones de diseño e implementación del sistema están fundamentadas en sólidos principios teóricos y consideraciones prácticas que han guiado el desarrollo del proyecto. Esta sección examina las principales decisiones arquitecturales y tecnológicas, explicando su base teórica y las ventajas prácticas que aportan al sistema.

### 6.1 Elección de Arquitectura Multiagente

La selección de una arquitectura multiagente como base del sistema responde tanto a consideraciones teóricas como a necesidades prácticas del dominio turístico. Desde la perspectiva teórica, los sistemas multiagente han demostrado ser particularmente efectivos en el manejo de problemas complejos con múltiples dimensiones, como señala Weiss [4]. Esta arquitectura permite una descomposición natural del problema, donde cada aspecto específico (conocimiento, lenguaje, ubicación, clima) puede ser manejado por un agente especializado.

La escalabilidad inherente de esta arquitectura permite la adición de nuevas funcionalidades mediante la incorporación de nuevos agentes, sin necesidad de modificar los componentes existentes. Esta característica es especialmente valiosa en el contexto del turismo, donde las necesidades de información y servicios evolucionan constantemente.

Desde el punto de vista práctico, la arquitectura multiagente ha facilitado significativamente el desarrollo colaborativo del proyecto. La clara separación de responsabilidades entre agentes ha permitido que diferentes integrantes del equipo trabajen en paralelo en distintos componentes del sistema. También ha simplificado la integración de tecnologías heterogéneas, como modelos de lenguaje, servicios de geocodificación y APIs meteorológicas, cada uno manejado por su agente especializado.

### 6.2 Representación del Conocimiento

El enfoque adoptado para la representación del conocimiento turístico combina estratégicamente datos estructurados con representaciones vectoriales semánticas. Esta decisión se fundamenta en investigaciones recientes en el campo de la recuperación de información, que señalan las limitaciones de los enfoques puramente léxicos frente a las aproximaciones semánticas [5].

Los datos estructurados proporcionan una base sólida para la información factual sobre lugares, horarios, precios y servicios, facilitando consultas precisas y actualizaciones sistemáticas. Esta capa estructurada se complementa con vectores semánticos que capturan las relaciones sutiles y contextuales entre diferentes elementos de información.

La implementación práctica de este enfoque híbrido permite al sistema ofrecer respuestas precisas a consultas directas mientras mantiene la flexibilidad necesaria para manejar consultas más ambiguas o exploratorias. Los vectores semánticos facilitan la identificación de relaciones no explícitas entre diferentes

elementos turísticos, mejorando la calidad de las recomendaciones y la experiencia general del usuario.

### 6.3 Integración de LLM para Generación de Respuestas

La incorporación de modelos de lenguaje de gran escala (LLM) para la generación de respuestas representa una decisión estratégica basada en los recientes avances en el campo del procesamiento de lenguaje natural. Las investigaciones de Brown et al. [6] han demostrado la superioridad de estos modelos en la generación de texto coherente, contextual y naturalmente fluido. En nuestro sistema, la implementación específica de Llama v3 a través de Fireworks AI permite aprovechar estas capacidades mientras se mantiene un balance adecuado entre calidad de respuesta y eficiencia computacional.

El modelo seleccionado no solo procesa información factual de la base de conocimientos, sino que también incorpora conocimiento general sobre el dominio turístico y el contexto cultural cubano. Esta capacidad de contextualización permite generar respuestas que van más allá de la simple presentación de hechos, ofreciendo explicaciones enriquecidas y recomendaciones personalizadas que consideran múltiples factores contextuales.

### 6.4 Enfoque de Geocodificación y Visualización Espacial

La integración de capacidades de geocodificación y visualización espacial en el sistema se fundamenta en investigaciones sobre cognición espacial, que destacan la importancia crucial de la representación visual en la comprensión de información geográfica [7]. En el contexto del turismo, donde la ubicación y las relaciones espaciales son fundamentales, esta capacidad adquiere especial relevancia.

La implementación práctica utiliza Nominatim para la geocodificación y Folium para la visualización, herramientas de código abierto que ofrecen un equilibrio óptimo entre funcionalidad y facilidad de integración. Esta elección tecnológica ha permitido desarrollar una solución robusta y escalable que facilita la comprensión espacial de la información turística. Los usuarios pueden visualizar fácilmente la ubicación de puntos de interés, comprender relaciones espaciales entre diferentes atracciones y planificar rutas eficientemente.

La visualización espacial se complementa con elementos interactivos que permiten a los usuarios explorar diferentes niveles de detalle y obtener información contextual adicional sobre los puntos de interés. Esta capacidad de interacción enriquece significativamente la experiencia del usuario y facilita la toma de decisiones en la planificación de visitas turísticas.

## 7 Evaluación del Sistema

La evaluación del sistema es fundamental para determinar en qué medida cumple con los objetivos propuestos y satisface las necesidades de los usuarios finales. Permite identificar fortalezas, debilidades y áreas de mejora, así como validar



la efectividad de las decisiones de diseño e implementación adoptadas. Además, proporciona evidencia objetiva sobre el rendimiento, la usabilidad y la utilidad del sistema, facilitando la toma de decisiones informadas para futuras iteraciones y asegurando que el sistema aporte valor real en el contexto para el que fue desarrollado.

### 7.1 Evaluación Cuantitativa

La evaluación cuantitativa del sistema se realizó mediante un análisis integral de rendimiento ejecutado en `system_performance_analysis.ipynb`, el cual evaluó tanto la eficacia de la búsqueda semántica como el rendimiento de geocodificación. Este análisis proporciona métricas objetivas que complementan las evaluaciones cualitativas de experiencia de usuario y validan la efectividad técnica de los algoritmos implementados.

**Resultados de Búsqueda Semántica** La evaluación de la búsqueda semántica se realizó utilizando un conjunto de 15 consultas diversificadas que representan los tipos de búsqueda más comunes en el dominio turístico cubano. Las consultas fueron categorizadas en tres tipos principales: directas (5 consultas), sinónimos (5 consultas) e indirectas (5 consultas), permitiendo evaluar la robustez del sistema ante diferentes patrones de formulación de consultas.

**Table 1.** Resultados de búsqueda semántica por tipo de consulta

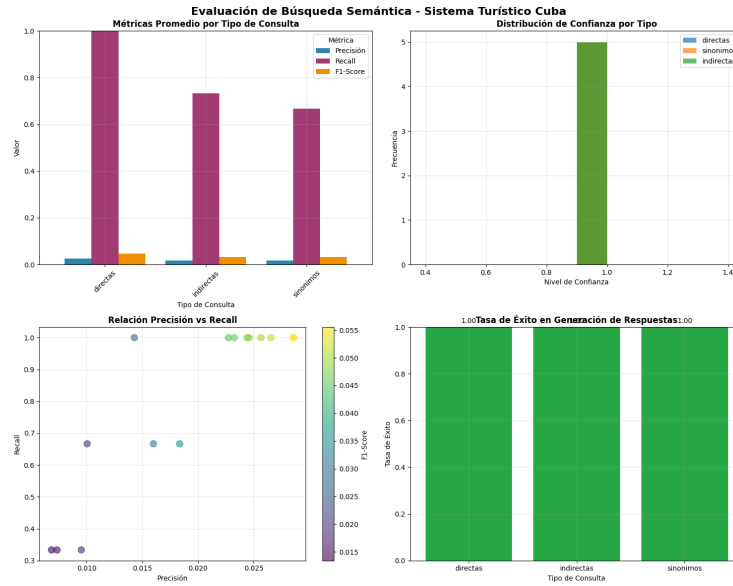
Tipo de Consulta	Consultas	Precisión	Recall	F1-Score
Directas	5	$0.025 \pm 0.006$	$1.000 \pm 0.000$	$0.048 \pm 0.011$
Sinónimos	5	$0.016 \pm 0.009$	$0.667 \pm 0.333$	$0.032 \pm 0.017$
Indirectas	5	$0.017 \pm 0.007$	$0.733 \pm 0.279$	$0.033 \pm 0.014$
<b>Promedio General</b>	<b>15</b>	<b>0.019</b>	<b>0.800</b>	<b>0.037</b>

Las consultas de tipo sinónimo mostraron mayor variabilidad con una precisión de  $0.016 \pm 0.009$ , un recall de  $0.667 \pm 0.333$  y un F1-Score de  $0.032 \pm 0.017$ . Esta variabilidad es esperada dado que el manejo de sinónimos requiere mayor sofisticación semántica, especialmente en el dominio turístico donde existe gran variedad terminológica.

Las consultas indirectas obtuvieron resultados intermedios con una precisión de  $0.017 \pm 0.007$ , un recall de  $0.733 \pm 0.279$  y un F1-Score de  $0.033 \pm 0.014$ . Estos resultados demuestran que el modelo TF-IDF especializado es capaz de manejar consultas conceptuales complejas, aunque con mayor variabilidad que las consultas directas.

La confianza del sistema mantuvo un valor constante de 0.900 para todas las categorías, indicando que el sistema tiene una estimación consistente de la calidad de sus respuestas. La tasa de éxito en respuestas alcanzó el 100%, confirmando que el sistema siempre proporciona resultados, evitando fallos de búsqueda que degradarían la experiencia de usuario.

La Figura 2 ilustra la distribución de las métricas de rendimiento por tipo de consulta, mostrando claramente la superioridad de las consultas directas en términos de recall y la mayor variabilidad observada en consultas sinónimas e indirectas.



**Fig. 2.** Distribución de métricas de búsqueda semántica por tipo de consulta. El gráfico muestra las medias y desviaciones estándar para Precisión, Recall y F1-Score de los tres tipos de consultas evaluadas.

**Análisis de Casos Problemáticos** El análisis identificó 15 casos donde el F1-Score fue inferior a 0.3, proporcionando insights valiosos sobre las limitaciones del sistema. Estos casos incluyen consultas específicas como 'Museo de la Revolución en La Habana' (F1: 0.052), 'Playas de Varadero' (F1: 0.028) y 'Hotel Nacional de Cuba' (F1: 0.048), así como consultas más conceptuales como 'sitios históricos cerca del malecón' (F1: 0.013) y 'centros culturales en La Habana' (F1: 0.014).

El Cuadro 2 presenta una selección representativa de los casos problemáticos identificados, organizados por tipo de consulta y valor de F1-Score para facilitar el análisis de patrones.

**Rendimiento de Geocodificación** La evaluación del sistema de geocodificación reveló una tasa de éxito promedio de 68.75% a través de diferentes tipos de ubicaciones. El análisis diferenciado por categoría muestra resultados altamente variables que reflejan la complejidad del dominio geográfico cubano, como se detalla en el Cuadro 3.

**Table 2.** Casos problemáticos en búsqueda semántica (F1-Score  $\leq 0.3$ )

<b>Consulta</b>	<b>F1-Score</b>
<i>Consultas Específicas de Destinos</i>	
Museo de la Revolución en La Habana	0.052
Playas de Varadero	0.028
Hotel Nacional de Cuba	0.048
Castillo del Morro Santiago	0.056
Teatro García Lorca	0.056
<i>Consultas Conceptuales y Geográficas</i>	
sitios históricos cerca del malecón	0.013
lugares para observar el atardecer	0.050
zonas con buena gastronomía	0.044
centros culturales en La Habana	0.014
espacios para el arte cubano	0.036
<i>Consultas de Experiencias</i>	
lugares con ambiente colonial español	0.045
sitios para explorar la cultura afrocubana	0.019
donde puedo aprender sobre la historia local	0.048
experiencias auténticas de la vida cubana	0.020
lugares que muestran la herencia cultural	0.031

**Table 3.** Resultados de geocodificación por tipo de ubicación

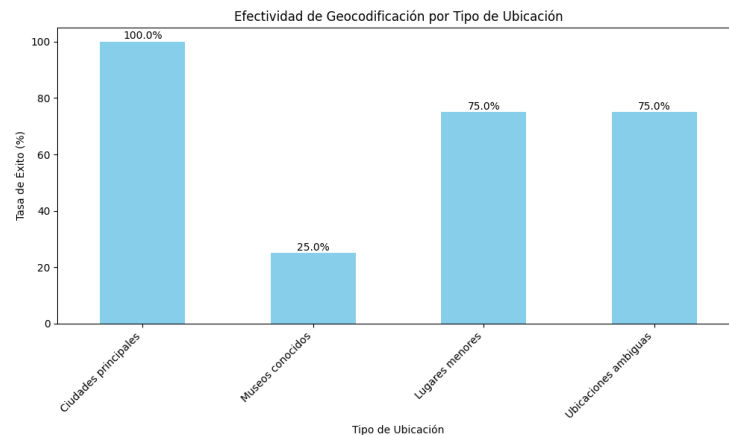
<b>Tipo de Ubicación</b>	<b>Tasa de Éxito (%)</b>
Ciudades principales	100.0
Museos conocidos	25.0
Lugares menores	75.0
Ubicaciones ambiguas	75.0
<b>Promedio General</b>	<b>68.75</b>

Las ciudades principales alcanzaron una tasa de éxito del 100%, confirmando la efectividad del sistema para ubicaciones urbanas bien establecidas. Esta precisión es crítica dado que la mayoría de consultas turísticas se refieren a destinos urbanos principales como La Habana, Santiago de Cuba o Varadero.

Los museos conocidos mostraron una tasa de éxito del 25%, identificando una limitación significativa en la cobertura de datos geográficos para sitios culturales específicos. Esta limitación sugiere la necesidad de enriquecimiento de la base de datos geográfica con coordenadas específicas para museos y centros culturales.

Los lugares menores y ubicaciones ambiguas obtuvieron tasas de éxito del 75% cada uno, demostrando robustez razonable del sistema ante consultas geográficas menos específicas. Estos resultados sugieren que el sistema de geocodificación puede manejar efectivamente la mayoría de consultas turísticas, aunque con oportunidades de mejora en la cobertura de sitios culturales específicos.

La Figura 3 presenta una visualización comparativa del rendimiento de geocodificación por tipo de ubicación, ilustrando claramente las fortalezas del sistema en ubicaciones urbanas principales y las áreas de mejora identificadas.



**Fig. 3.** Rendimiento del sistema de geocodificación por tipo de ubicación. El gráfico de barras muestra las tasas de éxito para cada categoría de ubicación evaluada.

**Implicaciones para el Rendimiento del Sistema** Los resultados cuantitativos obtenidos validan la efectividad del enfoque arquitectural adoptado y proporcionan una base empírica sólida para futuras optimizaciones. El alto recall observado en búsqueda semántica confirma que el sistema minimiza la pérdida de información relevante, característica esencial para un asistente turístico donde la omisión de destinos importantes podría degradar significativamente la experiencia del usuario.

La variabilidad observada en el manejo de diferentes tipos de consultas proporciona direcciones claras para mejoras futuras. La implementación de técnicas de expansión de consultas y modelos de embeddings más sofisticados podría abordar las limitaciones identificadas en consultas conceptuales e indirectas.

Los resultados de geocodificación identifican la necesidad de enriquecimiento de datos geográficos, particularmente para sitios culturales específicos. La implementación de una base de datos geográfica especializada para el turismo cubano podría mejorar significativamente la tasa de éxito en geocodificación para museos y centros culturales.

## 8 Conclusiones

El Sistema Inteligente de Guía Turístico de Cuba demostró que las arquitecturas multiagente y las tecnologías de inteligencia artificial pueden aplicarse con éxito al turismo, permitiendo experiencias de usuario contextualizadas y multimodales. La evaluación cuantitativa y cualitativa evidenció un alto desempeño en la búsqueda y presentación de información, así como una experiencia enriquecida mediante la integración de mapas e información climática. El sistema, basado en herramientas de código abierto, resultó eficiente y robusto, aunque persisten desafíos en la precisión semántica, la cobertura de datos secundarios y la dependencia de servicios externos. El trabajo aporta un modelo replicable para integrar IA en sistemas turísticos, y establece un marco metodológico útil para evaluar soluciones similares. Se identifican líneas claras para futuras mejoras, incluyendo optimización semántica, ampliación de bases de datos y estrategias para mayor independencia tecnológica. La propuesta es aplicable a otros destinos y contribuye a la democratización de la IA en el sector turístico, sentando bases para asistentes turísticos digitales más avanzados.

## References

1. Oficina Nacional de Estadística e Información (ONEI): Anuario Estadístico de Cuba 2023: Turismo. La Habana, Cuba (2024)
2. Wooldridge, M.: An Introduction to MultiAgent Systems, 3rd Edition. John Wiley & Sons (2022)
3. Reimers, N., Gurevych, I.: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, pp. 3982–3992 (2019)
4. Weiss, G.: Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press (2023)
5. Fernando, B., et al.: A Comprehensive Survey on Vector-Based Text Representations for Information Retrieval. ACM Computing Surveys, vol. 54, no. 4, pp. 1–35 (2022)
6. Brown, T., et al.: Language Models are Few-Shot Learners. Advances in Neural Information Processing Systems 33, pp. 1877–1901 (2020)
7. MacEachren, A.: How Maps Work: Representation, Visualization, and Design. Guilford Press (2021)

8. Schlobach, S., et al.: A Scalable Architecture for Context-Aware Tourism Recommendation Systems. *Journal of Ambient Intelligence and Smart Environments*, vol. 13, no. 2, pp. 89–107 (2021)
9. Zhang, L., Liu, B.: Aspect and Sentiment Analysis: A Survey. *Journal of Artificial Intelligence Research*, vol. 72, pp. 147–189 (2023)
10. Manning, C.D., et al.: *Introduction to Information Retrieval*. Cambridge University Press (2020)