

# Prospectiva de ChatGPT para Reporting de Análisis Financiero

---

Tras probar varias tecnologías y desarrollar tres POCs de diferentes opciones para el proyecto me permito comentar cuales son los caminos –opciones– de posible desarrollo que se tienen por delante.

## 1. Front Local

---

Esta fue la primera opción que se probó dado que al inicio yo desconocía que las API de OpenAI para sus modelos GPT no incluyen una conexión a la búsqueda web.

Esta opción permite generar información baasado en los prompts ( $\approx 15$ ) proveídos por Alberto de manera autónoma. Se puede alternar entre diferentes modelos de OpenAI (o1, GPT4, GPT4o) dependiendo del prompt.

El resultado con el modelo GPT-4 fue un poco corto y desactualizado; sin embargo, con el modelo "o1" que tiene "capacidades de razonamiento" y fue entrenado más recientemente; las respuestas mejoraron y la longitud del output se duplicó sin perder calidad de contenido. El modelo "GPT-4o" también presenta información más actualizada pero carece de "razonamiento".

A pesar de producir contenido de calidad, esta configuración no produjo información completamente reciente.

## Pros

- Posibilidad de alternar entre modelos.
- Flexibilidad en el tamaño de la respuesta generada.
- Posibilidad de asegurar privacidad y control sobre información enviada o preguntada.
- Generación de reporte automatizada con un bucle para ejecutar todos los prompts.
- Posibilidad de registrar todas las interacciones por parte del usuario a nivel local. (Mensajes, hora, respuestas).
- Apertura para restricción a preguntas o comportamiento pre-determinado. Es decir; se puede bloquear entrada de texto y en su lugar proveer de texto pre-determinado en el caso de que se quiere realizar más preguntas aparte de los prompts.

## Cons

- Información atrasada de meses o aveces unos pocos años.
- Limitación en generación de información numérica.
- El modelo se desenfoca un poco y puede producir respuestas muy diferentes bajo la misma pregunta en diferentes ocasiones por falta de un contexto de la empresa fuerte.
- Las actualizaciones y mejoras se deberían realizar a nivel local.
- No hay gráficos interactivos.

- Alto costo operativo.

## Tiempo de Desarrollo

- Medio (2-3 semana)

## 2. ChatGPT de OpenAI

---

Usando el ya existente FrontEnd integrado con su propia API se podría ahorrar mucho trabajo. Se construyó y probó una configuración para este ChatBot en su propia página web que lo pueden revisar [aquí](#) (Se necesita cuenta ChatGPT+).

Los prompts se han cargado a la configuración y se requiere interacción por parte del usuario para que se proceda al siguiente prompt. Es decir, este proceso no está automatizado completamente.

Se logran generar datos financieros y existen búsqueda web integrada que intenta extraer la información más actualizada posible.

### Pros

- Lista para el uso.
- Desarrollo externo que viene con actualizaciones y mejoras periódicas.
- Basado en la nube.
- Coste continuo y predecible bajo modelo de suscripción.
- Integrada con búsqueda web.

### Cons

- Poco control sobre la interfaz para limitar las opciones de interacción.
- Dificultad para interceptar información confidencial a nivel local.
- Limitaciones de personalización e integración con fuentes de información externas a OpenAI y dependientes de otras APIs pagadas como Bloomberg o YFinance.
- Búsqueda web poco personalizable y no fiable.
- Calidad de respuesta varía según la demanda global.
- Tiene el costo operativo más alto en comparación a las otras opciones.

## Tiempo de Desarrollo

- Bajo (1 semana)

## 3. Front Local + Internet

---

Esta fue el último POC que desarrollé este pasado fin de semana (Jan 25 - 26). Esta configuración une las fortalezas de ambas opciones previamente propuestas pero a un más grande costo de programación.

A través de APIs como "Yahoo Finance", "Google Custom Search API", y un LLM local —O Llama— logré producir un reporte completamente actualizado con información de hasta el mismo día. También logré graficar cotizaciones de la empresa requerida al igual que informar el histórico de sus índices al LLM para que este lo analice; Demostré igual la posibilidad de realizar cualquier gráfico interactivo en base a tablas o data extraída online de manera dinámica.

Con una pequeña configuración de extracción de información logré interceptar información tipo noticias dentro de páginas web confiables como Bloomberg News o el sitio web de la SEC.

Esta configuración tiene la capacidad de asociar noticias y artículos sobre la empresa junto con su comportamiento de índices internos como los EBITDA; porque al leer fuentes numéricas brutas extraídas como las de YFinance y mezclarlas con noticias como de NY Times tiene mayor contexto y conocimiento para consumir y procesar.

## Pros

- Respuesta profunda y detallada.
- Información actualizada.
- Posibilidad de poner la aplicación detrás de un Firewall.
- Integración de diferentes fuentes de información.
- Reducción de costos en llamadas API a OpenAI tras apoyarse en información tercera como llamadas de Google Search API.
- Producción de gráficos estadísticos interactivos configurables.
- Tiene el costo operativo de proveedores más bajo de todas las opciones.

## Cons

- Alta complejidad de desarrollo.
- Requiere acceso y configuración de varias APIs pagadas.
- Mayor necesidad de instalación de programas y librerías —herramientas de software— al servidor local —la computadora que prestaría el servicio— que podría verse impactado por la baja eficiencia del departamento de infraestructura interno sobre tareas similares.

## Tiempo de Desarrollo

- Alto (3-5 semanas)

## Conclusión

---

A pesar de que estas opciones se presentaron de manera aislada y contrastada, en realidad se podrían también considerar fases o hitos de desarrollo. Ordenando las opciones por tiempo de desarrollo tendríamos 2->1->3.

Considerando que la opción escogida sea **3** o **1**; se puede probar la usabilidad y versatilidad de una herramienta similar como la opción **2**.

Bajo los requerimientos descritos; la opción **1** queda fuera de juego ya que no tiene información actualizada sin embargo es importante conocerla y saber por qué no funciona. Por ende, queda decidir de las opciones **2** y **3**.

\*El tiempo de desarrollo es un estimado que solo considera el trabajo de la fase de desarrollo y no de el despliegue.