GEMINI AI CHAT WEB PART FOR SHAREPOINT ONLINE

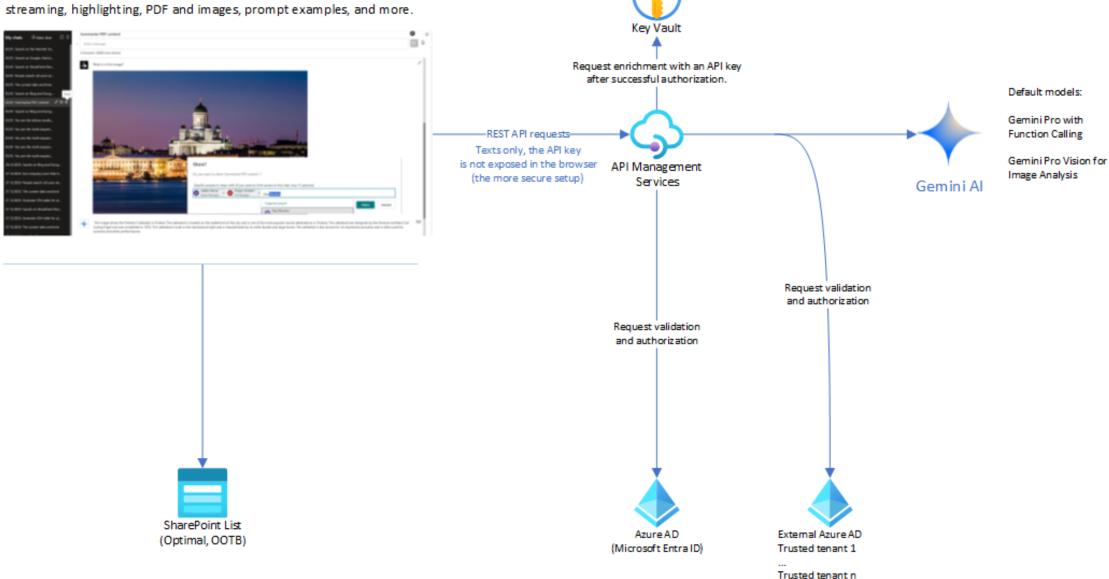
Paul Borisov, Cloud Architect

January 2024

Data Access Diagram: the following slide provides details on data privacy

Share Point Online web part with regular and Full-Screen appearance.

Optional internet and data connections, voice input and speech out, streaming, highlighting, PDF and images, prompt examples, and more.



Data Privacy

Gemini AI is the latest and most capable AI model published by Google. Any data provided to endpoints of Gemini AI goes to external service locations.

- As of January 2024, Gemini AI is not yet GDPR-compliant.
- RESTful endpoints of Gemini AI can be accessed using a free API key generated via Google AI Studio.
- Users from EU and UK locations do not have direct access to Gemini API keys. This web part provides a workaround for them.

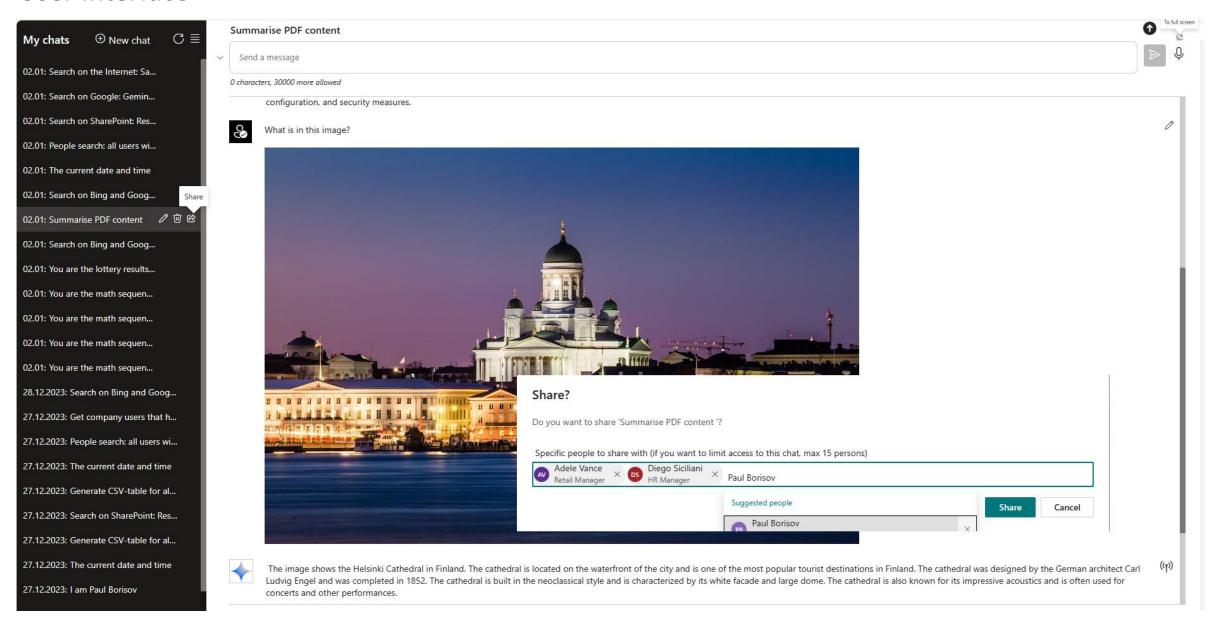
The web part interacts with **Gemini AI** endpoints published via Azure API Management service instance (APIM)

- APIM instance can be deployed to US zone to resolve location-based access restrictions for EU and UK customers.
- APIM consistently validates identities of SharePoint users for each individual request. If the request originates from authorized domains, APIM retrieves the api-key from the secure vault and injects it into the request before forwarding it to the AI endpoint. This process ensures that the api-key does not get exposed in the browser.
- Chats are private and visible only to their creators. Creators have the option to share their chats when this feature is enabled in the web part settings (disabled by default). Creators can share their chats with everyone or only with specific people in the company.

In addition to the default configuration, the web part provides the option to access to Gemini AI endpoints directly with an API key explicitly stored in the web part properties.

- This setup, while the least secure, can provide a quicker start. It is not recommended for production use, but it can be used for quick tests or in situations where you do not have access to Azure API Management.
- The stored key is encrypted in the web part properties and displayed as *** in the Property Pane. However, it will travel in browser requests and can be viewed within the DEV tools > Network > Request headers.
- If you are in Europe or UK, you should use a VPN to connect to Gemini AI endpoints in this setup.

User Interface



Key features

- Default support of Gemini Pro and Gemini Pro Vision
- Chat history stored into a SharePoint Custom List with the optional data encryption (off by default)
- Integration with Azure API Management Service to provide seamless access from European and UK locations.
- Global and Private Chat Sharing for selected Azure AD accounts
- Full-Screen Mode
- Consecutive event streaming
- Code Highlighting with Configurable Styles
- Unlimited Length of Chat History (Configurable)
- Upper and Lower side positioning for the Prompt Text Area
- Analysis of an uploaded PDF and summarization of its content
- Analysis of uploaded images and description of their content
- Optional dropdown box with examples for prompt text.
- Optional voice input to prompt text.
- Optional speech synthesis to read out AI-generated texts by default. The standard Web Speech API requires selecting the preferred language; using default page language is not always optimal.
- Optional data encryption for Chats data.
- Configurable Formats for Dates: Default is Finnish, which can be changed to "en-US" in Web Part Settings.

Data integrations

The web part supports optional integrations with external data using the Function Calling feature of Gemini Pro.

- Search on the Internet with Bing and Google
- SharePoint Search, which works under the current user identity and permissions
- Company Users
- Local Date and Time

Advanced features

- Built-in integration with API Management Service.
 - For customers that do not have it, the web part also supports direct queries to API with an api-key (configurable)
- Chat History is stored into a SharePoint list with the optional data encryption
 - SharePoint list supports up to 5000 Chats
 - The web part provides built-in "One-Click" list provisioning, Chat Privacy and Security Trimming for SharePoint Search

Optional Integrations: SharePoint Search, Company Users, Local Date & Time

How it works

The Function Calling feature is available in Gemini AI Pro model. The flow for using this feature typically looks like:

- 1. In the process of interacting with Gemini AI, you can send a request with options to call available integration functions that you have provided.
- 2. If Gemini AI determines that it requires specific data, it will request to call a function with parameters that it deems suitable.
- 3. To meet Gemini Al's demand, you can call the function with the proposed parameters on behalf of Gemini Al. This allows you to obtain raw results locally, which you can then submit to Gemini Al. You have control over what data you submit.
- 4. If Gemini AI determines that it needs more information, it may demand calling another function, and this process can continue, as necessary.
- 5. Gemini AI handles the received data and transforms it as you have requested. The final response from OpenAI will include the results based on the processed data.

```
// Several show cases are available in the web part:
// Search in SharePoint for "your text". Format the results as an HTML table.
// Get company users that have names starting with P. Format the results as an HTML table.
// Date, Time, Date and time.
export default class FunctionHelper {
 private static available: { [key: string]: any } = {
   companyUsers: this.companyUsers,
   currentDateOrTime: this.currentDateOrTime,
   currentDateAndTime: this.currentDateAndTime,
   peopleSearch: this.companyUsers,
   searchSharepoint: this.searchSharepoint,
   searchOnInternet: this.searchOnInternet,
   searchOnBing: this.searchOnInternet,
   searchOnGoogle: this.searchOnGoogle,
```

Integrations



Search on Bing and Google: Open AI vs Gemini. Show 5 most recent results with links and dates from each.



Sure, here are the 5 most recent results for Open AI vs Gemini with links and dates from Bing and Google:

Bing

- * Google Unveils Gemini, Claiming It's More Powerful Than OpenAl's GPT-4 December 20, 2023
- * Gemini vs OpenAl: Which Al Model Is Better? December 19, 2023
- * OpenAl vs. Google Al: Who Will Win the Al Race? December 18, 2023
- * Google's Gemini Al: A Potential Challenger to OpenAl's Dominance December 17, 2023
- * Google Al vs. OpenAl: The Battle for Al Supremacy December 16, 2023

Google

- * Google Reveals Gemini: Meet the Al Language Model Rivaling OpenAl's GPT-4 December 19, 2023
- * Google's Gemini Al Model Takes On OpenAl's GPT-4 in the Al Language Model Race December 18, 2023
- * The Race Between Google Al and OpenAl: Who Will Dominate the Al Landscape? December 17, 2023
- * Google Al and OpenAl: Comparing the Two Al Powerhouses December 16, 2023
- * OpenAl vs. Google Al: The Future of Al Language Models December 15, 2023

I hope this is helpful!

Upload and Summarize PDF content



Summarise PDF content



OK. Here's a summary of the document provided:

- Azure OpenAl Chat is a web part that allows for the utilization of both GPT-3.5 and GPT-4 models in creating chat responses, allowing for the understanding of context and generation of relevant responses.
- The web part offers many features, including chat history management, with the option to store history in Azure SQL or locally, and file sharing, allowing multiple users to collaborate on documents.
- Chat responses can be customized by adjusting frequency and relevancy of results, changing dialogue tone or style, and defining the length of messages. Additionally, users can utilize premade or create custom code snippets for automated completion.
- Built-in security features include Azure API Management to safely access Azure OpenAl APIs and protect API keys from exposure, as well as frontend authentication to block unauthorized API access.

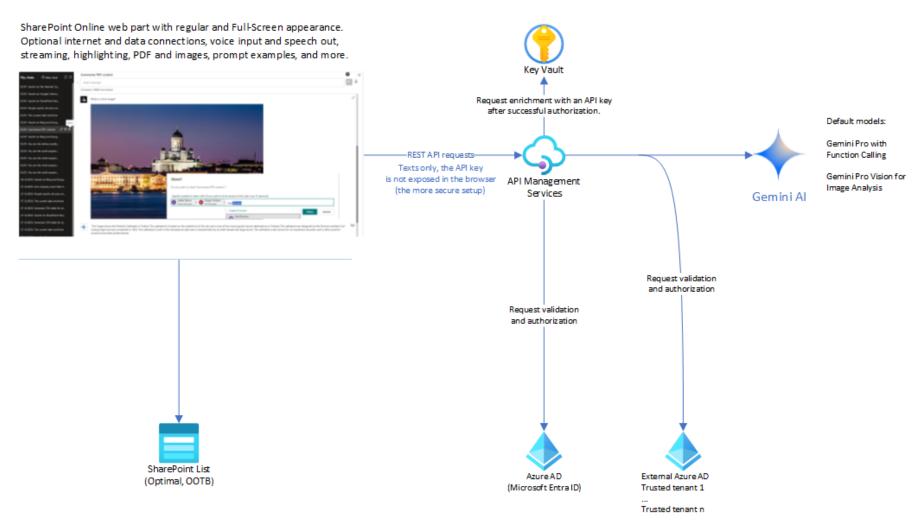
The following are the key technical specifications:

- Azure OpenAl service offers a variety of language models, with GPT-3.5 being the primary choice. Both GPT-3.5 and GPT-4 models can be enabled and configured; however, only one model's response can be displayed at a time.
- Azure API Management is the recommended method for accessing Azure OpenAI APIs securely.

Using Gemini Pro Vision for Image Analysis: a Technical Diagram



What is in this image?





This is an architecture diagram that shows how to use Azure Key Vault to protect API keys in a SharePoint Online web part. The diagram shows how the web part makes a request to the API Management service, which in turn makes a request to Azure Key Vault to retrieve the API key. The API key is then used to make a request to the SharePoint Online API.

Using Gemini Pro Vision for Image Analysis: a Photo



What is in this image?





The image shows the Helsinki Cathedral in Finland. The cathedral is located on the waterfront of the city and is one of the most popular tourist destinations in Finland. The cathedral was designed by the German architect Carl Ludvig Engel and was completed in 1852. The cathedral is built in the neoclassical style and is characterized by its white facade and large dome. The cathedral is also known for its impressive acoustics and is often used for concerts and other performances.

Gemini AI Chat web part for SharePoint Online, configurable options

