This code sets up a backend system for a chatbot that answers questions based on a user's résumé, leveraging Google Generative AI models. Here's a concise overview:

1. **Environment Initialization:**
    - Libraries like ‘express’, ‘body-parser’, ‘cors’, ‘fs’, and path are imported.
    - Google AI models for embeddings and text generation are initialized.
2. **Express Server Setup:**
    - An Express server is created, configured with body-parser for JSON parsing and cors for cross-origin request handling.
3. **Embedding Functions:**
    - Functions are defined to generate and utilize embeddings for texts. The Euclidean distance between these embeddings is calculated to determine relevance.
4. **Query Processing:**
    - A function performs relevance searches by comparing the query embedding with precomputed document embeddings and sorts results based on relevance.
5. **Answer Generation:**
    - A second AI model generates responses using the query and the most relevant document contexts.
6. **Document Handling:**
    - Résumé content is loaded and preprocessed from a file.
7. **Precomputed Embeddings:**
    - Embeddings for documents are computed and stored for efficient query handling.
8. **POST Endpoint:**
    - A /ask endpoint processes incoming questions, identifies relevant document sections, and returns generated responses.

This setup provides a RESTful API that processes user questions and generates responses based on résumé content using AI models.