High-Level Design:

The system is structured into three core **functional agents** and a final **report generation** module. Each agent performs a specific role in generating actionable insights and relevant datasets for AI/ML use cases.

Core Components:

1. **Industry Research Agent**:
    - Gathers industry-specific data using the Serper API.
    - Forms the basis for generating AI/ML use cases.
2. **Use Case Generation Agent**:
    - Uses OpenAI's GPT models to generate AI/ML use cases tailored to the industry.
    - Focuses on Generative AI solutions, such as document search, automated reporting, and chat systems.
3. **Dataset Search Agent**:
    - Searches for datasets on platforms like GitHub, Kaggle, and Hugging Face based on the generated use cases.
    - Uses the Serper API for efficient dataset discovery.
4. **Markdown Report Generator**:
    - Compiles industry insights, use cases, and datasets into a structured markdown file for easy sharing and implementation.

Low-Level Design:

1. Industry Research Agent:

- **Input**: A string representing the target industry or company.
- **Process**:
    - Query the Serper API with "industry overview" as the context.
    - Parse the API response to extract titles, links, and snippets.
- **Output**: A list of industry insights

2. Use Case Generation Agent:

- **Input**: Industry insights (list of dictionaries).
- **Process**:
    - Combine insights into a formatted prompt.
    - Use OpenAI's GPT-4 to generate use cases tailored to the industry.
    - Include predefined suggestions (e.g., document search, chat systems).
- **Output**: A list of use cases

3. Dataset Search Agent:

- **Input**: Use cases (list of dictionaries).
- **Process**:
    - For each use case, query the Serper API with dataset-specific queries (e.g., "dataset for AI-powered customer feedback analysis").
    - Collect datasets from GitHub, Kaggle, and Hugging Face.
- **Output**: A list mapping use cases to relevant datasets

4. Markdown Report Generator:

- **Input**: Industry insights, use cases, and datasets.
- **Process**:
    - Format the data into markdown sections (#, ##, - for bullets).
    - Save the report as a .md file with a timestamped filename.
- **Output**: A markdown file containing:
    - Industry research insights.
    - AI/ML use cases.
    - Relevant datasets

Input Flow:

1. **Input**: User specifies an industry or company.
2. **Industry Research**: Fetch insights to form the basis of Generative AI solutions.
3. **Use Case Generation**: GPT-4 creates innovative solutions.
4. **Dataset Search**: Queries datasets for implementation.
5. **Report Compilation**: Outputs a well-structured markdown document.

This modular approach ensures scalability, reusability, and ease of integration.