Use of KIT's KI-Toolbox to Support Programming in Visual Studio Code

Version 1; 15.01.2026

Important Disclaimer

This manual only describes the **technical integration of the API** and the criteria for model selection within the plugin. It does not serve as a guide to high-quality programming, software architecture, or the general use of AI in software development.

For guidance on best practices in vibe coding (i.e., a conversational, AI-powered programming approach in which developers work with an assistant like Cline to design, review, and improve code step by step), check out the linked documentation at the end of the document.

1. Technical requirements

- **KIT VPN** or a direct connection to the KIT network.
- Visual Studio Code with installed **Cline Plugin**.
- A personal **API key**, generated in the KIT KI-Toolbox.
  Further information on the generation of the API key and API use at KIT – e.g. on data protection and general conditions – can be found in the **central instructions for API use at KIT (PDF).**

Installing the Cline plugin in Visual Studio Code
If the Cline plugin is not already installed, do the following:

1. In Visual Studio Code, open the **Extensions pane** (icon in the left sidebar or Ctrl+Shift+X / Cmd+Shift+X).
2. Search for **Cline: AI Assistant**.
3. Select the listing of publisher **Saoud Rizwan** or go directly to the marketplace page: https://marketplace.visualstudio.com/items?itemName=saoudrizwan.claude-dev.
4. Click **Install** and restart Visual Studio Code if necessary.

You can then proceed with the configuration described in Section **2. Setup in Visual Studio Code**.

2. Setup in Visual Studio Code

1. Open Cline via the icon in the sidebar.
2. Click on the **gear icon (⚙)** to open the settings.
3. Configure the provider:
    - **API Provider:** Select **"OpenAI Compatible".**
    - **Base URL:** Enter https://ki-toolbox.scc.kit.edu/api.
    - **API Key:** Insert your KIT API key.

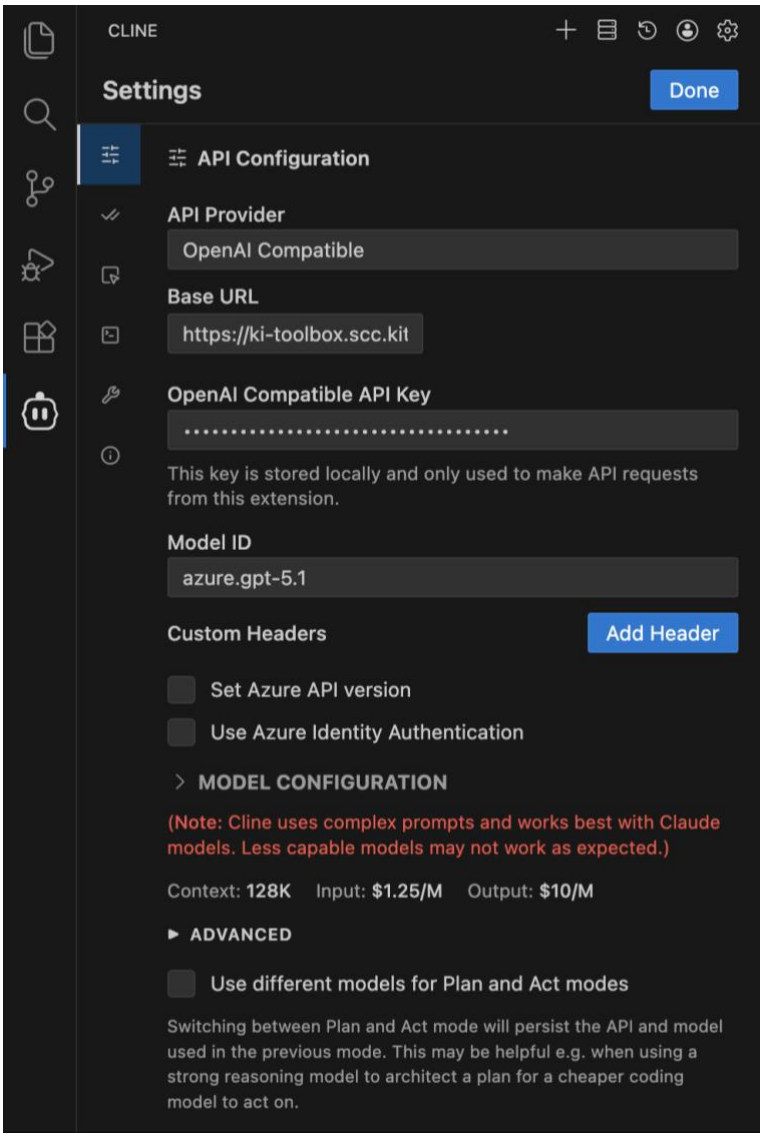

Figure 1: Example for the settings in the plugin. The costs of the respective model can also be stored for better traceability.

Functional test after setup

To check if Cline is correctly connected to the KI-Toolbox, you can perform the following short test:

1. Open any project or folder in Visual Studio Code.
2. Start Cline via the icon in the sidebar.
3. Enter a simple request, for example: *"Analyze the structure of this project and describe the most important files to me."*

If Cline can access the project files and provide a comprehensible description, setting up the API is usually correct.

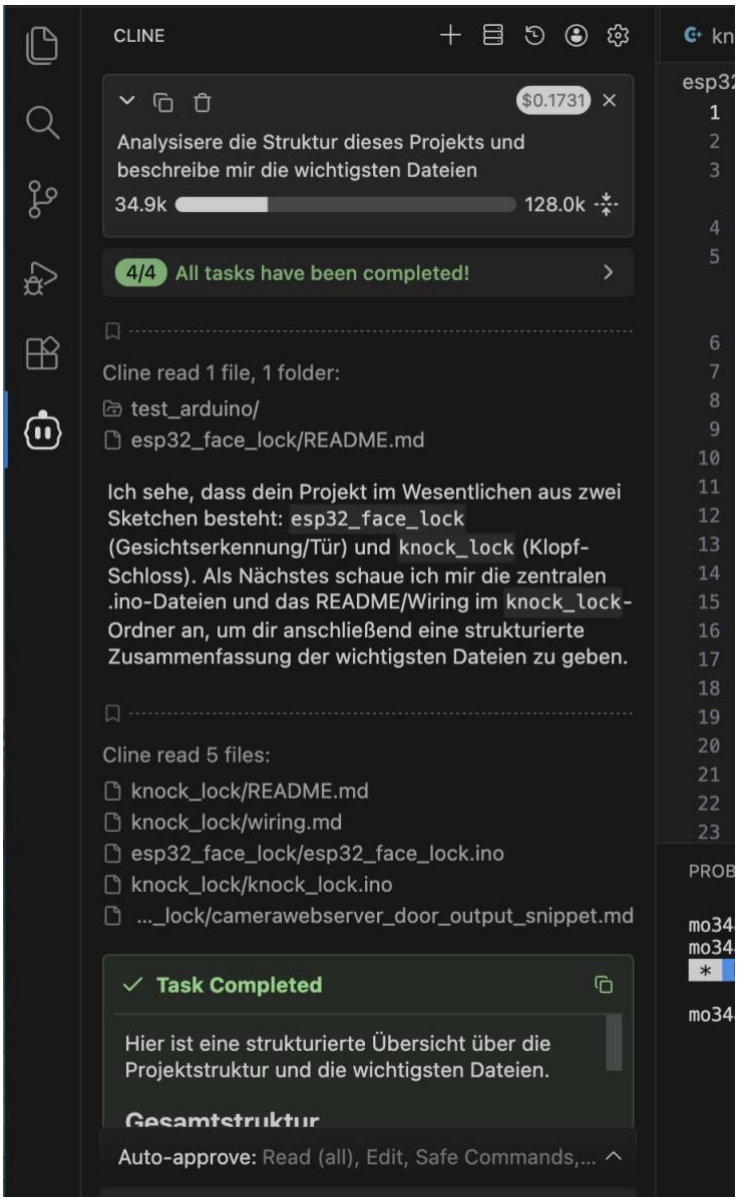

Figure 2: Starting Cline in the sidebar (icon) and performing the functional test

3. Model selection and tool usage

In Cline, the choice of model is crucial because the model acts as a "brain" that understands commands and controls tools (such as creating files or terminal commands). The KI-Toolbox also stores the costs of the models. This aspect should also be taken into account when choosing. For example, azure.gpt-5.1 has a cost of $1.25 in and $10 out per 1 million tokens, while azure.gpt-5-mini has a cost of $0.25 in and $2 out. In the advanced settings of the plugin, you can also store these costs for better transparency and then see the associated costs for each task.

A. Powerful Cloud Models
For complex tasks that require high reliability in tool usage, the models provided through Azure are recommended.

- **Recommendation:** azure.gpt-5.1 or azure.gpt-5-mini.
- **Features:** These models feature excellent reasoning and a very stable design of Cline tools.
- **Note:** The data will be transferred to Microsoft Azure (Germany).

B. Local Models at KIT
If you want to work with highly confidential data or only use the SCC's infrastructure, choose local models.

- **Recommendation:** kit.qwen3-vl-235b-a22b-instruct.
- **Features:** Local models such as variants of Qwen3 are now considered powerful enough to effectively apply Cline to tasks.
- **Advantage:** Your data remains completely within the infrastructure of KIT.

C. Non-Recommended Models for API Usage
It is **not recommended** to use the standard-external and standard-local models via the API in Cline.

- **Reason:** These models have a specially adapted system prompt for the web interface. When used via the API, this system prompt can negatively affect your instructions in Cline and lead to unwanted or erroneous results.

4. Data classification

Please always choose your model according to the sensitivity of your data:

- **Public / less sensitive data:** Cloud models (azure....) or on-premises models (kit....) are possible.

- **Confidential data (without strict personal requirements):** Cloud models (azure...) or on-premises models can be partially used – in compliance with the central instructions for API usage.
- **Personal or highly confidential data: Only local models** (models in the "Local" tab in the KI-Toolbox) may be used.

5. Related Resources

Documentation at KIT:

- Central Guide to API Use at KIT (PDF) – Basic information on data protection and API details.
- Interactive API documentation (Swagger UI) – Overview of all available endpoints.

Cline & AI Development:

- Cline Model Selection Guide – Help with selection based on context windows and cost.
- Prompt Engineering Guide – Strategies for precise instructions to the AI.
- Cline Learning Path – Comprehensive tutorials on LLM basics and best practices for tool integration.
- Custom Instructions & Rules – How to customize Cline to fit your personal coding style.

Info & Contact

License Notice

Imprint

Publisher: Karlsruhe Institute of Technology (KIT) Kaiserstraße 12 76131 Karlsruhe

Contact: InformatiKOM Adenauer Ring 12 76131 Karlsruhe Germany Phone: +49 721 608-48200 E-mail: info@zml.kit.edu

Questions about the KI-Toolbox should be directed to: ki-toolbox@scc.kit.edu