

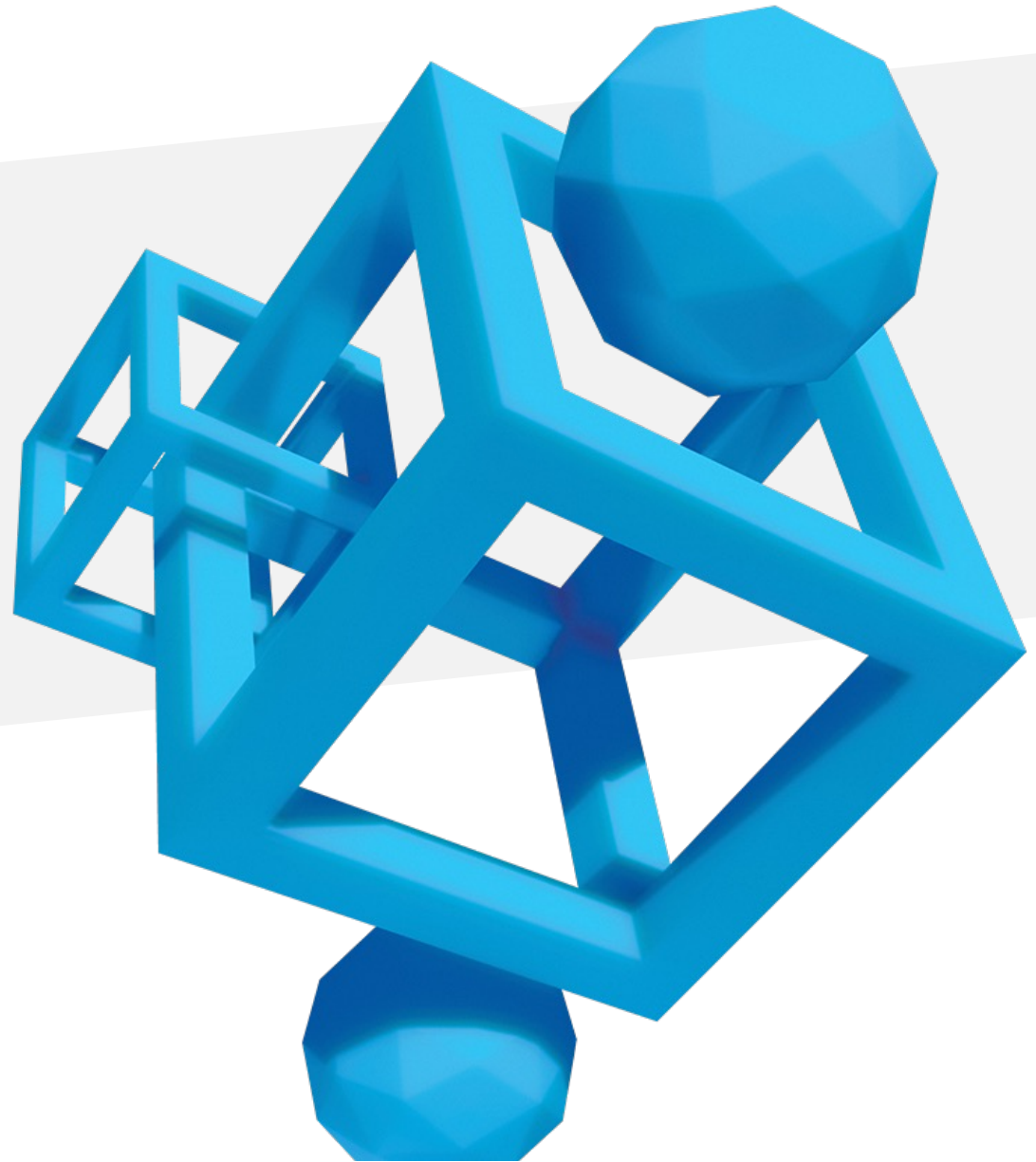
The VAST AI Assistant

ESUG 2026

Johan Brichau and Kris Gybels

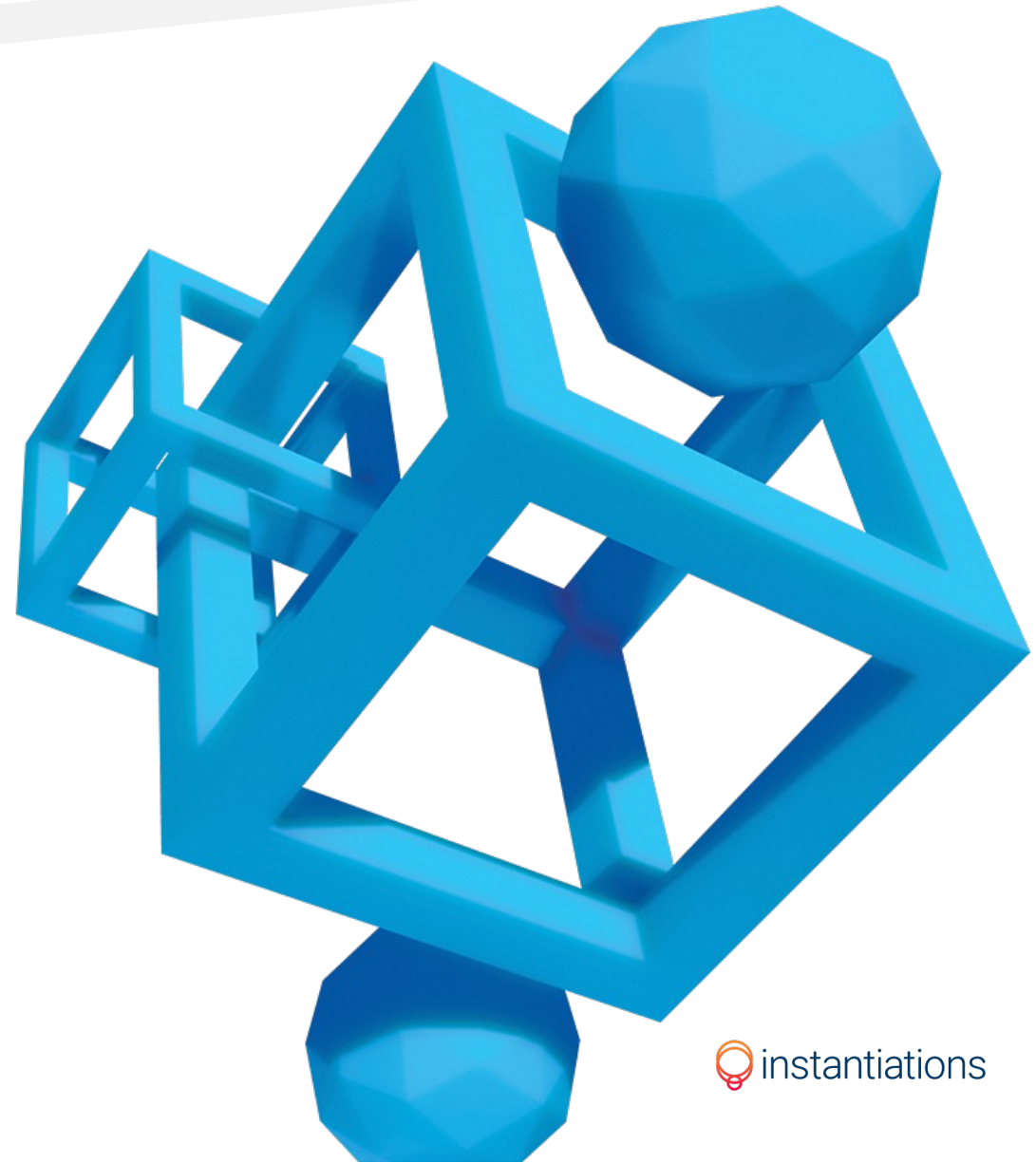
Senior Software Engineers

✉ jbrichau@instantiations.com | kgybels@instantiations.com



Agenda

- VAST AI Assistant goals
- How to get started
- The Assistant
- IDE integration
 - Code changes browser
 - Debugger
 - Envy semantic diff
- Applications



VAST AI Assistant Goals

AI Development at Instantiations

- Embedding AI into real systems, workflows, and augmenting how people actually work
- Balancing innovation with stability, sustainability, and trust

The Approach: Responsible Evolution™

VAST AI Assistant goals

- Improve developer productivity and daily workflows
- Seamless and deep integration of natural language and AI-assisted development in the VAST IDE
 - Autonomous exploration of the system architecture to gather context before attempting to explain code or suggest modifications
 - Fully integrated with the live environment and live coding workflow
- Focus on "Human in the loop" AI-assisted software development
- Allow agentic workflows

How to get started

Open Source

 Platform ▾ Solutions ▾ Resources ▾ Open Source ▾ Enterprise ▾ Pricing

[Sign in](#) [Sign up](#) 

 [instantiations](#) / **VAST-Platform-AI-Assistant** Public

[Notifications](#) [Fork 3](#) [Star 7](#)

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Security and quality](#) [Insights](#)

 [main](#) ▾ **VAST-Platform-AI-Assistant** / **README.md** 



 **eMaringolo** Add instructions to upgrade from previous installations 8b01b16 · yesterday [History](#)

[Preview](#) [Code](#) [Blame](#) 180 lines (125 loc) · 11.7 KB [Raw](#)    

VAST Platform AI Assistant

This repository provides the integration tools for AI-assisted development within the [VAST platform](#) development environment. It enables developers to use natural language to query codebases, analyze runtime errors, run tests, and automate development tasks directly from the IDE.

Supported Model Providers

The VAST Platform AI Assistant can be used with different Large Language Models (LLMs). It supports [Google's Gemini models](#) as well as other models through any API compatible with the OpenAI Responses API, as offered for example by [Ollama](#), [LMStudio](#), and [LiteLLM](#). Note that not all models may work; models (such as those from [Ollama](#) or [LMStudio](#)) need to support tool calling while thinking or reasoning.

The Assistant in practice

Demo time

Dev Image: AI Assistant

ChatEditOptionsDebugHelp

Can you explain how things need to be set up using "Settings" in the "Options" menu in the chat window?

Thoughts

To set up the AI Assistant via the **Settings** dialog (accessible from the **Options** menu in the chat window), you'll need to configure two main areas: **Providers** and **Models**.

Here is the step-by-step workflow to get things running:

1. Set Up a Provider

A Provider represents the service hosting the AI models (like OpenAI or Google Gemini) and contains your connection credentials.

1. Select **Providers** from the list on the left side of the Settings window.
2. Click the **+** button to create a new provider.
3. Select a **Provider Type** from the dropdown (for example, **OpenAI** or **Gemini**).
4. Enter a **Label** to easily identify this provider (e.g., "My OpenAI").
5. The **URL** field will typically auto-populate based on the provider type, but you can adjust it if you are using a custom endpoint.
6. Enter your **API Key** for the service.
7. Click **Save** (at the bottom).
8. *Optional but recommended:* Click the **Refresh** button next to "Provider Models:". This tests your connection and API key by fetching the list of available models from the provider.

2. Set Up a Model

Once your Provider is registered, you need to register a Model configuration that the chat window can actually use.

1. Select **Models** from the list on the left side of the Settings window.
2. Click the **+** button to create a new model configuration.
3. Select the **Provider** you just created from the dropdown.
4. Select the specific **Model** you want to use from the second dropdown (e.g., **gpt-4o** or **gemini-1.5-pro**). This list is populated from the models fetched during the setup.
5. Enter a **Label** (e.g., "GPT-4o standard"). This label is what will appear in your chat window when you are selecting which model to talk to.
6. Under **Model-specific Parameters**, you can optionally tweak values like:
 - **Temperature:** Controls randomness/creativity (usually between 0.0 and 2.0).
 - **Top-p:** An alternative to temperature for controlling diversity.
 - **Reasoning Effort/Summary** (if supported by your model type).
7. Click **Save**.

3. Use the Model

Once you have saved your model, you can close the Settings window. In the main Chat Window, go back to the **Options** menu, find the **Model** submenu, and select the model created for your model. You are now ready to chat!

Send

Tokens used (Prompt: 26351 Total: 27821) / 1048576

16.0.0x64-b604-dev: AI Assistant Settings

ProvidersModelsFunctionsMCP Servers

Sample MCP Server

Name: Sample MCP Server

URL: http://localhost:3000/mcp

Dev Image: AI Assistant

ChatEditOptionsDebugHelp

How can I open a chat with you from a workspace?

Thoughts

Exploring Chat Implementation

I'm currently exploring the VAST Platform Smalltalk environment to understand how to initiate a chat session, likely with an AI assistant. I'm focusing on identifying any relevant classes or methods related to "chat," "assistant," "openai," or "llm". I'm trying OpenAIChat and Chat to see if there is a match.

Function calls:

Exploring Chat Initiation

I'm currently focused on how to best initiate the chat window. I'm considering different approaches, such as `AIChatWindow new open` or `AIChatWindow chat: AIChat new`. I'll analyze existing globals to determine the simplest method for the user; ultimately, they could just evaluate `AIChatWindow` in their Workspace.

To open a chat window with me from a Workspace, you can evaluate the following Smalltalk code:

```
AIChatWindow new open
```

Make sure you select the code and execute it (e.g., using **Ctrl+E** or **Right-Click -> Execute**).

If you need to configure the AI provider or select a different model, you can do so from the **Options -> Settings...** and **Options -> Model** menus in the chat window once it opens.

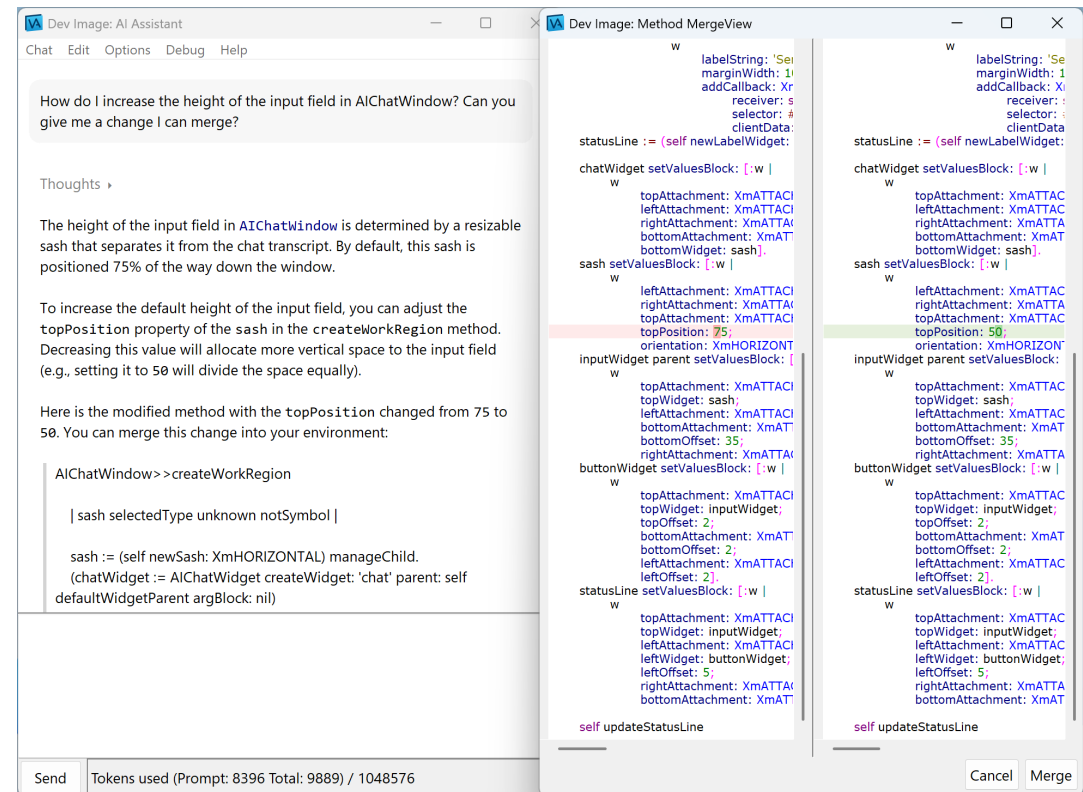
Send

Tokens used (Prompt: 8089 Total: 8358) / 1048576

IDE Integration

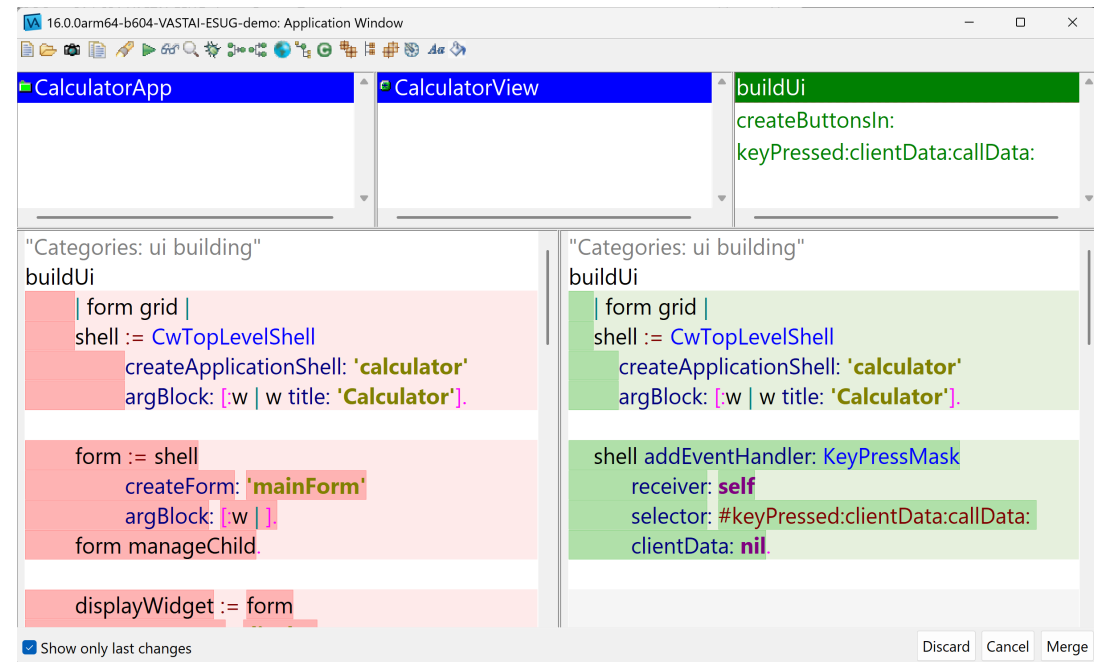
AI Code Generation – first approach

- Show **code snippets** in conversation
- Open diff view to apply code change
- Disadvantages:
 - Often produced invalid code (i.e. LLM did not validate the code first)
 - The syntax for method definitions was often not followed by the LLM
 - Not all changes are possible (e.g. remove and rename)



AI Code Generation – new approach

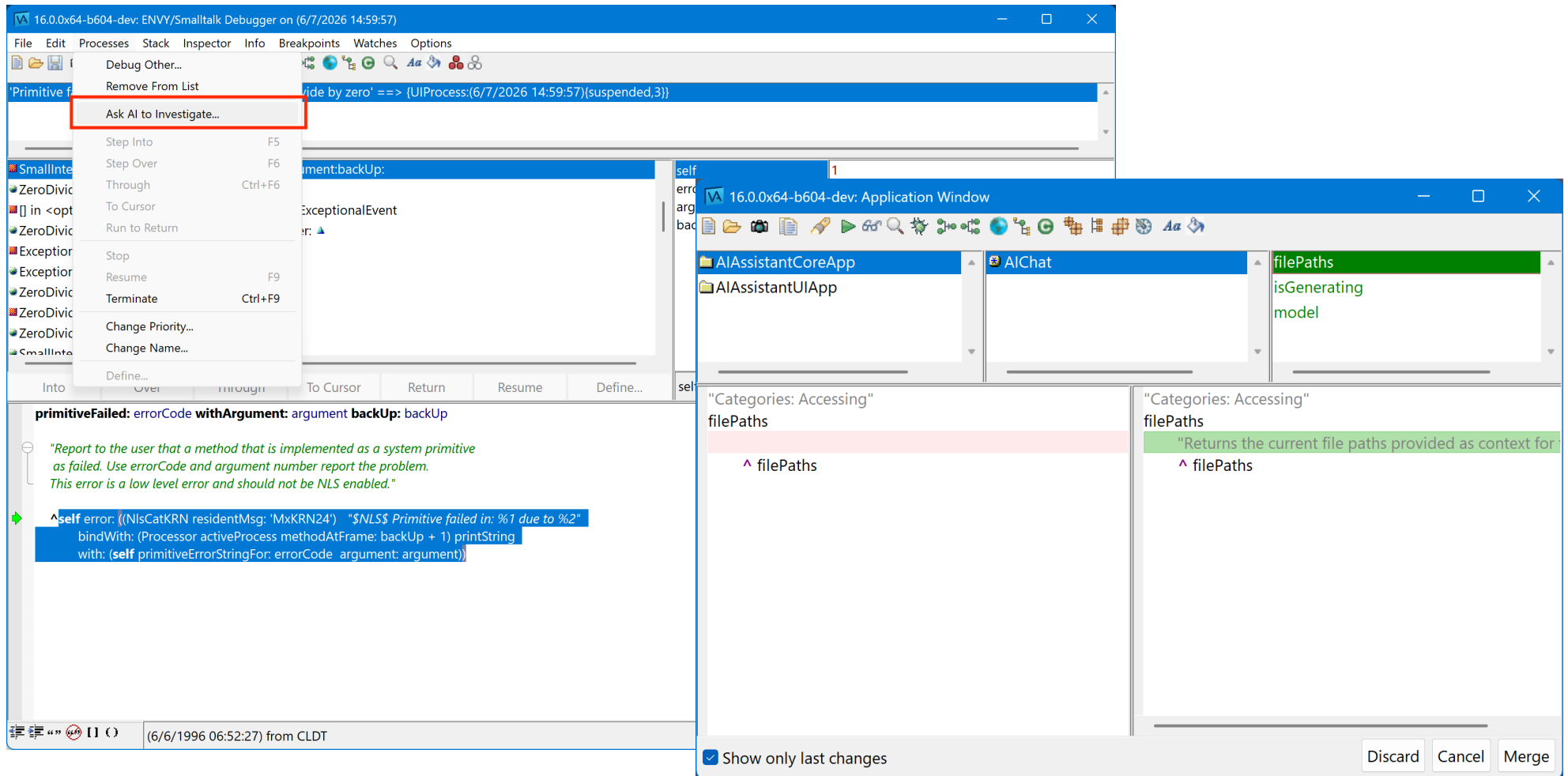
- Offer functions/tools for code modification
- Open a merge browser on the suggested changes
- Advantages:
 - Code is always validated by the functions
 - No formatting rules required, AI will more likely follow function descriptions than the system instructions
 - Dedicated Merge Browser visualization
 - Allows agentic coding



AI-assisted debugging

- Smalltalk is a live coding environment
 - Development in the debugger
- Trigger the AI from the debugger
 - New chat: investigate a stand-alone error
 - Select existing chat: report error in its development context
- Run and debug test cases
- More 'live' features in the future:
 - Inspect objects
 - Navigate the object memory

Demo time 2



ENVY version history tools

Compare the versions 'b600 + 74792d' and 'b604 +74872a' for the 'ENVY/Image SCI' config map and describe what has changed.

16.0.0arm64-b604-VASTAI-ESUG-demo: AI Assistant - ENVY/Image SCI Config Changes

Chat Edit Options Debug Help

Compare the versions 'b600 + 74792d' and 'b604 +74872a' for the 'ENVY/Image SCI config' map and describe what has changed.

Thoughts ▶

Comparing the configuration map ENVY/Image SCI editions b600 + 74792d and b604 + 74872a reveals changes focused around adding support for the MSGPEEK socket flag on SSL sockets, allowing you to peek into the socket's receive buffer without actually consuming the data.

Here is the detailed breakdown of the changes in the `SciSslOpenSSLInterface` application:

- 1. `SciSslOpenSSLInterface` class >> `#_PRAGMA_SciSslFunctions`**
 - Added the FFI definition for the underlying OpenSSL `SSL_peek` C function to the `SciSslFunctions` pool dictionary.
- 2. `SciSslOpenSSLLibraryDispatcher` >> `#callSSL_peekWith:with:with:`**
 - (New Method)** This method was added to wrap the newly declared `SSL_peek` platform function. It invokes the C function using the standard FFI mechanism and includes standard error checking block for OpenSSL calls (`value <= 0`).
- 3. `SciSslBlockingDispatcher` >> `#callRecvWithSocket:with:with:with:`**
 - (Changed Method)** Previously, the `flags` argument was ignored and it unconditionally performed an `SSL_read`.
 - It was updated to check the `flags` integer argument for `SciSocketConstants::MSGPEEK`. If the `MSGPEEK` flag is present, it now delegates to the new `callSSL_peekWith:with:with:` method. Otherwise, it behaves as before and calls `callSSL_readWith:with:with:`.
 - The method comment was also updated to specify that the `flags` argument handles `MSGPEEK`.

VAST team uses the VAST AI

- Rapid prototyping
 - Code merge view; diff algorithm; user interface;...
- Test case generation
 - Repetitive work, sometimes a lot of boilerplate
- Assist in exploring bugs
- Fixed a couple of edge case bugs in the VM
- Generate platform function definitions for external libraries
 - Provide C header files and documentation to product a library wrapper
- Development of the AI Assistant itself
 - Rapid prototyping, exploration, ...



Thank you for attending!

VAST AI Assistant

Kris Gybels and Johan Brichau

Senior Software Engineers

✉ kgybels@instantiations.com | jbrichau@instantiations.com

Questions?

Contact

General Inquiry
info@instantiations.com

Sales
sales@instantiations.com

VAST Support Portal
vast-support.instantiations.com

North America, Toll Free
855 476 2558

International
+1 503 263 0058

© Instantiations, Inc. All rights reserved. 'Instantiations' and the 'intersecting circle design' are registered trademarks of Instantiations, Inc. in the United States. All product names, trademarks, and registered trademarks are property of their respective owners. Company, product, and service names not owned by Instantiations are used for identification purposes only. Use of these names, trademarks, and brands does not imply endorsement.