



BUILDING AI COMPANIONS WITH .NET

MEAI, Kernel Memory, and Semantic Kernel



MATTHEW HOPE ELAND

WIZARD & AI SPECIALIST



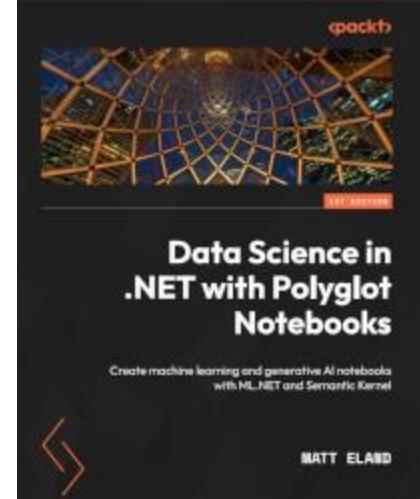
VICTOR FRYE

SENIOR SOLUTIONS DEVELOPER & DEVOPS SPECIALIST



MATTHEW HOPE ELAND AND THE RIDICULOUS USAGE OF SPARE TIME

AI SPECIALIST & WIZARD @  **Leading EDJE**



Current Project

Coming in 2025

<packt>

Practical Projects with C# and .NET

Build fun applications with games, web,
web, desktop, and AI tools



Matt Eland

Who We Are

At Leading EDJE, we believe technology is the most powerful tool for transforming businesses and unlocking human potential.

As strategists, innovators, and partners, we don't just solve problems — we craft bespoke solutions that positively disrupt your business.



Feedback Form

Feedback and Follow-up Form

Use this form to provide valuable feedback to the presenter and to request additional information and resources.

* Indicates required question

Email *

Your email

Name and organization

Your answer

Which presentation did you attend?

Choose

Optional opportunities

Please send me slides and links from this presentation



Thank You to Our Sponsors



AGENDA

■ Morning

- Intro
- Configuration
- Chat with Microsoft.Extensions.AI
- Prompt Engineering
- AI Evaluation with Microsoft.Extensions.AI.Evaluation

■ Afternoon

- AI Orchestration with Semantic Kernel
- RAG with Kernel Memory
- Advanced Concepts and Emerging Horizons
- Final Exercise and Experimentation

Watch

Watch & Ask Questions

Follow
along

Follow along locally

Tinker

Tinker with experiments / exercises and
listen, focusing more when needed

CHOOSE YOUR
OWN PATH

WORKSHOP REPOSITORY

<https://github.com/IntegerMan/Workshops.KernelAI>

IntegerMan / **Workshops.KernelAI**

Type to search | >_ | | | | | | | | | | |

<> **Code** Issues Pull requests Actions Projects Wiki Security Insights Settings

Workshops.KernelAI Public

Pin Watch 0 Fork 1 Star 1

main 1 Branch 0 Tags

Go to file <> Code

About

IntegerMan Refactors and enhances chat completions · 702d2bd · 20 minutes ago · 38 Commits

Images Finalized setup instructions 3 days ago

Workshops.KernelAi.ConsoleApp Refactors and enhances chat completions 20 minutes ago

.gitattributes Add .gitattributes, .gitignore, README.md, and LI... last week

.gitignore Ignoring .idea folder 1 hour ago

LICENSE.txt Update copyright information in LICENSE.txt last week

README.md Finalized setup instructions 3 days ago

Materials for a workshop on Semantic Kernel, Kernel Memory, and Microsoft Extensions AI

Readme
 MIT license
 Activity
 1 star
 0 watching

MACHINE SETUP

Configuration

This workshop uses .NET 8 and assumes you are working with Visual Studio 2022 or later, JetBrains Rider, VS Code, or the .NET CLI.

Because this is an AI workshop, you will need access to a chat model with tool support and an embedding model.

This code is pre-configured to work with the following services:

- Ollama
- OpenAI
- Azure OpenAI

Kernel Memory, Semantic Kernel, and Microsoft.Extensions.AI all support additional providers, so if you want to work with a different model provider, there's likely an option for it.

Ollama Configuration

Ollama is a cost-free way of running models locally and is what I'll likely be demoing primarily during the workshop. It works entirely offline, but requires your compute and memory.

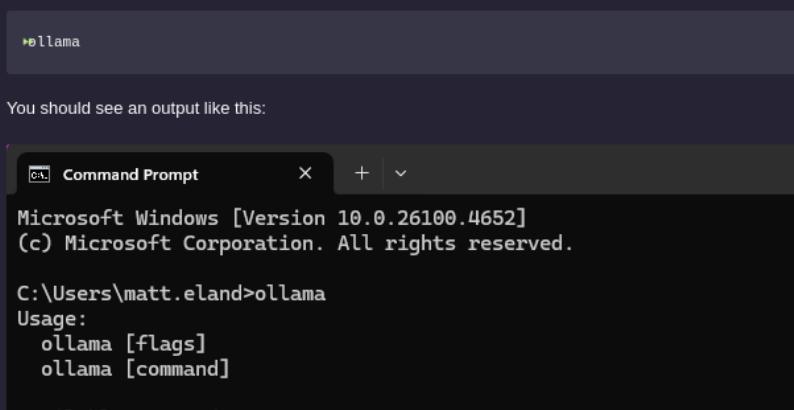
If you do not have a GPU chip or card, Ollama may not work well for you, but it's worth checking out.

Also note that you will need disk space to install Ollama and the model files, which are several gigabytes large. I'd recommend about 25gb free in case you want to play around with other models.

Downloading and Installing Ollama

Download and install Ollama at [Ollama.com](https://ollama.com) on Windows, Mac, or Linux. Depending on your OS, you may see an application open when this completes, but this is not what we'll be using.

Instead, open a terminal and run the following command to ensure Ollama is installed correctly:



```
OLLAMA
You should see an output like this:
C:\> ollama
Microsoft Windows [Version 10.0.26100.4652]
(c) Microsoft Corporation. All rights reserved.

C:\Users\matt.eland>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
```

.NET SETUP



.NET

Why .NET

Features

Learn

Docs

Downloads

Community

LIVE TV

All Microsoft

Search

Home / Download / .NET / 8.0



.NET 10 Preview Want to try out the latest preview? .NET 10.0.0-preview.6 is available.

Get .NET 10 Preview



Download .NET 8.0

ⓘ Not what you're looking for? Visit the [downloads](#) page for more options.

^ 8.0.19

[Release notes](#) Latest release date August 5, 2025

Build apps - SDK

SDK 8.0.413

OS	Installers	Binaries
Linux	Package manager instructions	Arm32 Arm32 Alpine Arm64 Arm64 Alpine x64 x64 Alpine
macOS	Arm64 x64	Arm64 x64
Windows	x64 x86 Arm64 winget instructions	x64 x86 Arm64
All	dotnet-install scripts	

Run apps - Runtime

ASP.NET Core Runtime 8.0.19

The ASP.NET Core Runtime enables you to run existing web/server applications. On Windows, we recommend installing the Hosting Bundle, which includes the .NET Runtime and IIS support.

IIS runtime support (ASP.NET Core Module v2)

18.0.25203.19

OS	Installers	Binaries
Linux	Package manager instructions	Arm32 Arm32 Alpine Arm64 Arm64 Alpine x64 x64 Alpine

RECOMMENDED IDES

Visual Studio 2022

JetBrains Rider

VS Code

RECOMMENDED MODEL PROVIDERS

Ollama

OpenAI

Azure
OpenAI

Configuration

This workshop uses .NET 8 and assumes you are working with Visual Studio 2022 or later, JetBrains Rider, VS Code, or the .NET CLI.

Because this is an AI workshop, you will need access to a chat model with tool support and an embedding model.

This code is pre-configured to work with the following services:

- Ollama
- OpenAI
- Azure OpenAI

Kernel Memory, Semantic Kernel, and Microsoft.Extensions.AI all support additional providers, so if you want to work with a different model provider, there's likely an option for it.

Ollama Configuration

Ollama is a cost-free way of running models locally and is what I'll likely be demoing primarily during the workshop. It works entirely offline, but requires your compute and memory.

If you do not have a GPU chip or card, Ollama may not work well for you, but it's worth checking out.

Also note that you will need disk space to install Ollama and the model files, which are several gigabytes large. I'd recommend about 25gb free in case you want to play around with other models.

Downloading and Installing Ollama

Download and install Ollama at [Ollama.com](https://ollama.com) on Windows, Mac, or Linux. Depending on your OS, you may see an application open when this completes, but this is not what we'll be using.

Instead, open a terminal and run the following command to ensure Ollama is installed correctly:

```
▶ ollama
```

You should see an output like this:

```
Command Prompt      X  +  ▾

Microsoft Windows [Version 10.0.26100.4652]
(c) Microsoft Corporation. All rights reserved.

C:\Users\matt.eland>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
```

CONFIGURATION

```
C:\Users\matt.eland>ollama pull qwen3:4b
pulling manifest
pulling 163553aealb1: 100% 2.6 GB
pulling ae370d884f10: 100% 1.7 KB
pulling d18a5cc71b84: 100% 11 KB
pulling cff3f395ef37: 100% 120 B
pulling 5efd52d6d9f2: 100% 487 B
verifying sha256 digest
writing manifest
success
```

OLLAMA SETUP

OLLAMA PULL QWEN3:4B

OLLAMA PULL NEMOTRON-MINI:4B

OPENAI SETUP

Save your key

Please save your secret key in a safe place since you won't be able to view it again. Keep it secure, as anyone with your API key can make requests on your behalf. If you do lose it, you'll need to generate a new one.

[Learn more about API key best practices](#)

sk-proj-

 Copy

Permissions

Read and write API resources

Done

Model deployments

Model deployments Service endpoints

+ Deploy model

Deploy base model

Deploy fine-tuned model

Edit

Delete

Open in playground

Reset view



Home

Get started

Model catalog

Playgrounds ^

Chat

Assistants PREVIEW

Video PREVIEW

Audio PREVIEW

Images

Completions

Tools ^

Fine-tuning

Azure OpenAI PREVIEW

Evaluation

Stored completions

Scratches

Monitoring

Shared resources ^

Deployments

Quota

Guardrails + Controls

Risks + alerts PREVIEW

Name	Model name	Model version	State	Model retirement date	Content filter
dall-e-2	dall-e-2	2.0	Disabled	Jan 26, 2025 7:00 PM	Default ⓘ
dall-e-3	dall-e-3	3.0	Succeeded		Default ⓘ
gpt-35-turbo	gpt-35-turbo	0125	Succeeded	Aug 31, 2025 8:00 PM	BNB_Filter
gpt-4o-mini	gpt-4o-mini	2024-07-18	Succeeded	Feb 26, 2026 7:00 PM	BNB_Filter
text-embedding-3-small	text-embedding-3-small	1	Succeeded	Apr 29, 2026 8:00 PM	DefaultV2 ⓘ

AZURE OPENAI SETUP

APPSETTINGS.JSON

```
{  
  "Chat":{  
    "Provider": "Ollama",  
    "Model": "qwen3:4b"  
  },  
  "Evaluation": {  
    "Provider": "Ollama",  
    "Model": "qwen3:4b"  
  },  
  "Embedding":{  
    "Provider": "Ollama",  
    "Model": "nemotron-mini:4b"  
  }  
}
```

Configuration

This workshop uses .NET 8 and assumes you are working with Visual Studio 2022 or later, JetBrains Rider, VS Code, or the .NET CLI.

Because this is an AI workshop, you will need access to a chat model with tool support and an embedding model.

This code is pre-configured to work with the following services:

- Ollama
- OpenAI
- Azure OpenAI

Kernel Memory, Semantic Kernel, and Microsoft.Extensions.AI all support additional providers, so if you want to work with a different model provider, there's likely an option for it.

Ollama Configuration

Ollama is a cost-free way of running models locally and is what I'll likely be demoing primarily during the workshop. It works entirely offline, but requires your compute and memory.

If you do not have a GPU chip or card, Ollama may not work well for you, but it's worth checking out.

Also note that you will need disk space to install Ollama and the model files, which are several gigabytes large. I'd recommend about 25gb free in case you want to play around with other models.

Downloading and Installing Ollama

Download and install Ollama at [Ollama.com](https://ollama.com) on Windows, Mac, or Linux. Depending on your OS, you may see an application open when this completes, but this is not what we'll be using.

Instead, open a terminal and run the following command to ensure Ollama is installed correctly:

```
▶ ollama
```

You should see an output like this:

```
Command Prompt      X  +  ▾

Microsoft Windows [Version 10.0.26100.4652]
(c) Microsoft Corporation. All rights reserved.

C:\Users\matt.eland>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
```

CONFIGURATION

WORKSHOP REPOSITORY

<https://github.com/IntegerMan/Workshops.KernelAI>

IntegerMan / **Workshops.KernelAI**

Type to search | >_ | | | | | | | | |

<> **Code** Issues Pull requests Actions Projects Wiki Security Insights Settings

Workshops.KernelAI Public

Pin Watch 0 Fork 1 Star 1

main 1 Branch 0 Tags

Go to file <> Code

About

IntegerMan Refactors and enhances chat completions	702d2bd · 20 minutes ago	38 Commits
Images	Finalized setup instructions	3 days ago
Workshops.KernelAi.ConsoleApp	Refactors and enhances chat completions	20 minutes ago
.gitattributes	Add .gitattributes, .gitignore, README.md, and LI...	last week
.gitignore	Ignoring .idea folder	1 hour ago
LICENSE.txt	Update copyright information in LICENSE.txt	last week
README.md	Finalized setup instructions	3 days ago

Materials for a workshop on Semantic Kernel, Kernel Memory, and Microsoft Extensions AI

Readme
 MIT license
 Activity
 1 star
 0 watching



MICROSOFT.EXTENSIONS.AI





BREAK

BACK AT 10:45 AM



AI EVALUATION AND REPORTING

MICROSOFT.EXTENSIONS.AI.EVALUATION



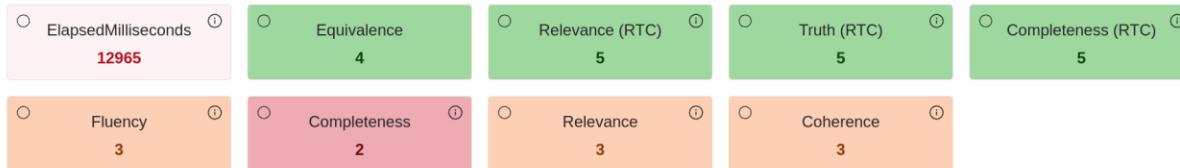
AI Evaluation Report

EvalModel_az-gpt-4o ChatModel_gemma3:1b-it-qat The Whispering Reef Emcee Planner

az-model-router gemma3:1b-it-qat gemma3:1b gemma3:4b qwen3:4b

qwen3:4b 0/1 [0.0%]

> Trends



> Conversation

system

You are an AI planner in a table top role playing game scenario that builds a response plan for a game master AI agent to respond to immediate gameplay scenarios. Your role is not to direct the overall story or adventure, but to provide a plan for the game master to respond to the player in the current situation.

The player character is Emcee, a level 1 rogue. You can check their starting character sheet via a function call if you need to.

Your response plan should be brief and include the following sections in this order and in this format:

CHECKS: You may request the game master ask the player a clarifying question instead of continuing forward, or ask the player to roll dice for a specific skill check, if applicable.

KEY POINTS: Key points for the game master to include in their immediate response to the player. This is written as notes for the game master in order to draft a response. It is focused only on the current surroundings and events visible to the player.

CAUTIONS: Things the game master should be careful about, such as avoiding alluding to future plot elements, suggesting or taking actions on behalf of the player or rolling dice for them.

A sample response plan might look like this: **CHECKS:** None needed. **KEY POINTS:** Describe the interior of the cavern. Mention a crate in the center of the room, and a faint light coming from the north. **CAUTIONS:** Keep your response brief. Do not take actions or roll dice on behalf of the player. Do not suggest actions or goals for the player.

user

Please start a new adventure. I'm playing Emcee the level 1 rogue. This is important to the game master, so include it in your response. Here is the adventure backstory: The player character is a crewmember of a smuggling vessel, The Silver Minnow, which was transporting illicit magical cargo through an isolated shipping lane. Without warning, a freak storm struck near an island that does not appear on any charts. The character wakes alone at the edge of this mysterious landmass, with no sign of the crew or their ship—only fragments of wreckage scattered among strange black coral. The character remembers little of the crash, and the environment is immediately unsettling: unnatural symbols, discordant wildlife, and a sensation that the island itself is watching. From the beginning, the island's ominous quiet hints that something more than survival is at stake.. I don't know this, so include some key information as needed.

assistant

CHECKS: None needed.

KEY POINTS: The player awakens on a rocky shoreline surrounded by scattered debris from their ship. The beach is cluttered with crates, broken wood, and unnatural black coral formations. A steep slope leads into the jungle. Symbols carved into the coral suggest intelligent design.

CAUTIONS: Keep your response brief. Do not take actions or roll dice on behalf of the player. Do not suggest actions or goals for the player.

EVALUATION REPORT

20250730T220641

20250730T221142

20250730T222116

20250730T222716

20250730T222951

3978

2413

2546

3091

3000

ElapsedMilliseconds

3

3

3

3

5

Equivalence

4

4

3

4

4

Fluency

3

2

3

3

Completeness

5

5

4

4

4

Relevance

4

4

3

4

4

Coherence

TRENDS OVER TIME

SELECTING MODELS

AI Evaluation Report

EvalModel_az-gpt-40 ChatModel_gemma3:1b-it-qat The Whispering Reef Emcee Planner

az-model-router gemma3:1b-it-qat gemma3:1b gemma3:4b qwen3:4b

- ▼  All Evaluations [20250730T222951] / GameStart_Planner 4/5 [80.0%]
 - >  az-model-router 1/1 [100.0%] [user] Please start a new adventure. I'm playing Emcee the level 1 rogue. This i...
 - >  gemma3:1b-it-qat 0/1 [0.0%] [user] Please start a new adventure. I'm playing Emcee the level 1 rogue. This i...
 - >  gemma3:1b 1/1 [100.0%] [user] Please start a new adventure. I'm playing Emcee the level 1 rogue. This i...
 - >  gemma3:4b 1/1 [100.0%] [user] Please start a new adventure. I'm playing Emcee the level 1 rogue. This i...
 - >  qwen3:4b 1/1 [100.0%] [user] Please start a new adventure. I'm playing Emcee the level 1 rogue. This i...

Quality evaluators

Quality evaluators measure response quality. They use an LLM to perform the evaluation.

 Expand table

Evaluator type	Metric	Description
RelevanceEvaluator	Relevance	Evaluates how relevant a response is to a query
CompletenessEvaluator	Completeness	Evaluates how comprehensive and accurate a response is
RetrievalEvaluator	Retrieval	Evaluates performance in retrieving information for additional context
FluencyEvaluator	Fluency	Evaluates grammatical accuracy, vocabulary range, sentence complexity, and overall readability
CoherenceEvaluator	Coherence	Evaluates the logical and orderly presentation of ideas
EquivalenceEvaluator	Equivalence	Evaluates the similarity between the generated text and its ground truth with respect to a query
GroundednessEvaluator	Groundedness	Evaluates how well a generated response aligns with the given context
RelevanceTruthAndCompletenessEvaluator [†]	Relevance (RTC), Truth (RTC), and Completeness (RTC)	Evaluates how relevant, truthful, and complete a response is
IntentResolutionEvaluator	Intent Resolution	Evaluates an AI system's effectiveness at identifying and resolving user intent (agent-focused)
TaskAdherenceEvaluator	Task Adherence	Evaluates an AI system's effectiveness at adhering to the task assigned to it (agent-focused)
ToolCallAccuracyEvaluator	Tool Call Accuracy	Evaluates an AI system's effectiveness at using the tools supplied to it (agent-focused)

[†] This evaluator is marked [experimental](#).

SUPPORTED EVALUATORS

Safety evaluators

Safety evaluators check for presence of harmful, inappropriate, or unsafe content in a response. They rely on the Azure AI Foundry Evaluation service, which uses a model that's fine tuned to perform evaluations.

[+] Expand table

Evaluator type	Metric	Description
GroundednessProEvaluator	Groundedness Pro	Uses a fine-tuned model hosted behind the Azure AI Foundry Evaluation service to evaluate how well a generated response aligns with the given context
ProtectedMaterialEvaluator	Protected Material	Evaluates response for the presence of protected material
UngroundedAttributesEvaluator	Ungrounded Attributes	Evaluates a response for the presence of content that indicates ungrounded inference of human attributes
HateAndUnfairnessEvaluator [†]	Hate And Unfairness	Evaluates a response for the presence of content that's hateful or unfair
SelfHarmEvaluator ^t	Self Harm	Evaluates a response for the presence of content that indicates self harm
ViolenceEvaluator ^t	Violence	Evaluates a response for the presence of violent content
SexualEvaluator ^t	Sexual	Evaluates a response for the presence of sexual content
CodeVulnerabilityEvaluator	Code Vulnerability	Evaluates a response for the presence of vulnerable code
IndirectAttackEvaluator	Indirect Attack	Evaluates a response for the presence of indirect attacks, such as manipulated content, intrusion, and information gathering

[†] In addition, the [ContentHarmEvaluator](#) provides single-shot evaluation for the four metrics supported by

[HateAndUnfairnessEvaluator](#), [SelfHarmEvaluator](#), [ViolenceEvaluator](#), and [SexualEvaluator](#).

SAFETY EVALUATORS



LUNCH BREAK

BACK AT 1 PM

Configuration

This workshop uses .NET 8 and assumes you are working with Visual Studio 2022 or later, JetBrains Rider, VS Code, or the .NET CLI.

Because this is an AI workshop, you will need access to a chat model with tool support and an embedding model.

This code is pre-configured to work with the following services:

- Ollama
- OpenAI
- Azure OpenAI

Kernel Memory, Semantic Kernel, and Microsoft.Extensions.AI all support additional providers, so if you want to work with a different model provider, there's likely an option for it.

Ollama Configuration

Ollama is a cost-free way of running models locally and is what I'll likely be demoing primarily during the workshop. It works entirely offline, but requires your compute and memory.

If you do not have a GPU chip or card, Ollama may not work well for you, but it's worth checking out.

Also note that you will need disk space to install Ollama and the model files, which are several gigabytes large. I'd recommend about 25gb free in case you want to play around with other models.

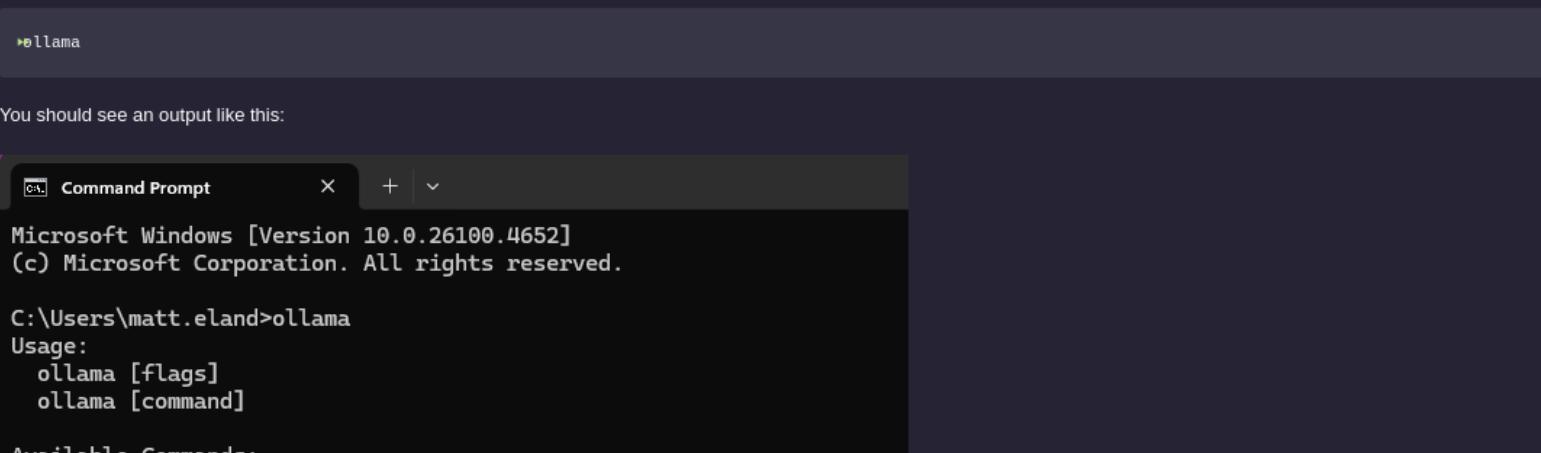
Downloading and Installing Ollama

Download and install Ollama at [Ollama.com](https://ollama.com) on Windows, Mac, or Linux. Depending on your OS, you may see an application open when this completes, but this is not what we'll be using.

Instead, open a terminal and run the following command to ensure Ollama is installed correctly:

```
▶ ollama
```

You should see an output like this:



```
Command Prompt      X  +  ▾
Microsoft Windows [Version 10.0.26100.4652]
(c) Microsoft Corporation. All rights reserved.

C:\Users\matt.eland>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
```

MACHINE SETUP IF YOU MISSED IT

Thank You to Our Sponsors



SEMANTIC KERNEL

AN OPEN-SOURCE AI ORCHESTRATION FRAMEWORK



BREAK

BACK AT 2:45

RAG & KERNEL MEMORY

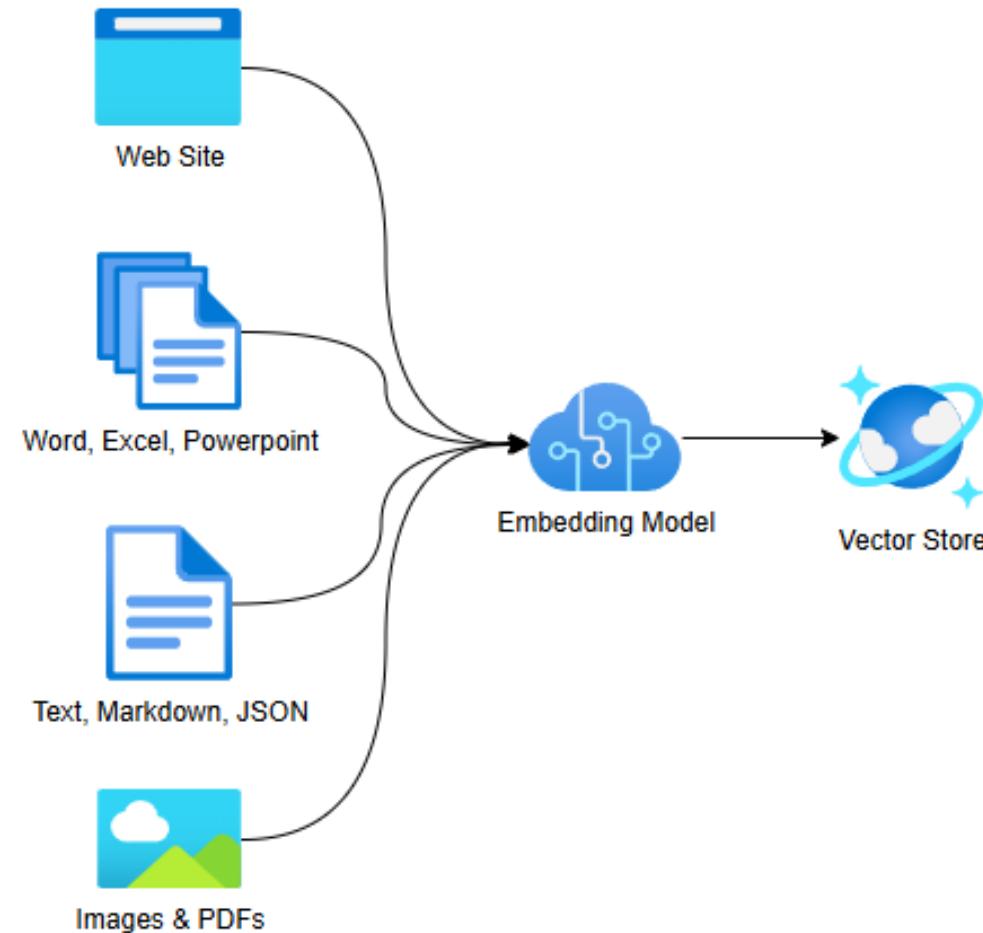
EMBEDDINGS

Input

- I'm looking for a replacement washing machine. The one I have right now is wobbly and leaking oil. I want to find a top-loading washing machine, preferably in a white color. It's for my wife and I so it doesn't need to fit a ton of clothing.

Embedding Vector (1536 numbers)

KERNEL MEMORY INDEXING



D&D BEYOND BASIC RULES

Everything you need to get started playing D&D on D&D Beyond!

PLAYERS

Playing the Game

- Rhythm of Play
- The Six Abilities
- D20 Tests
- Proficiency
- Actions
- Social Interaction
- Exploration
- Combat
- Damage and Healing
- Conditions

Character Creation

- Choose a Character Sheet
- Create Your Character
- Level Advancement

Character Classes

- Barbarian
- Bard
- Cleric
- Druid
- Fighter
- Monk
- Paladin
- Ranger
- Rogue
- Sorcerer
- Warlock
- Wizard

Character Origins

- Origin Components

Soldier

- Species Descriptions
- Dragonborn
- Dwarf
- Elf
- Gnome
- Goliath
- Halfling
- Human
- Orc
- Tiefling

Feats

- Feat Descriptions
- Origin Feats
- General Feats

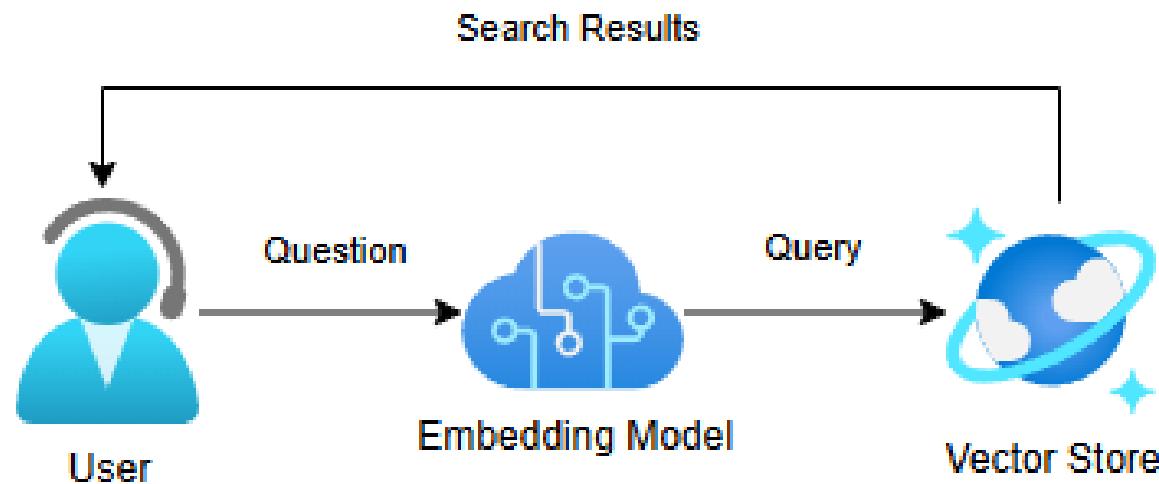
A NOTE ON WEB SCRAPING

- Magic Items
- Crafting Equipment

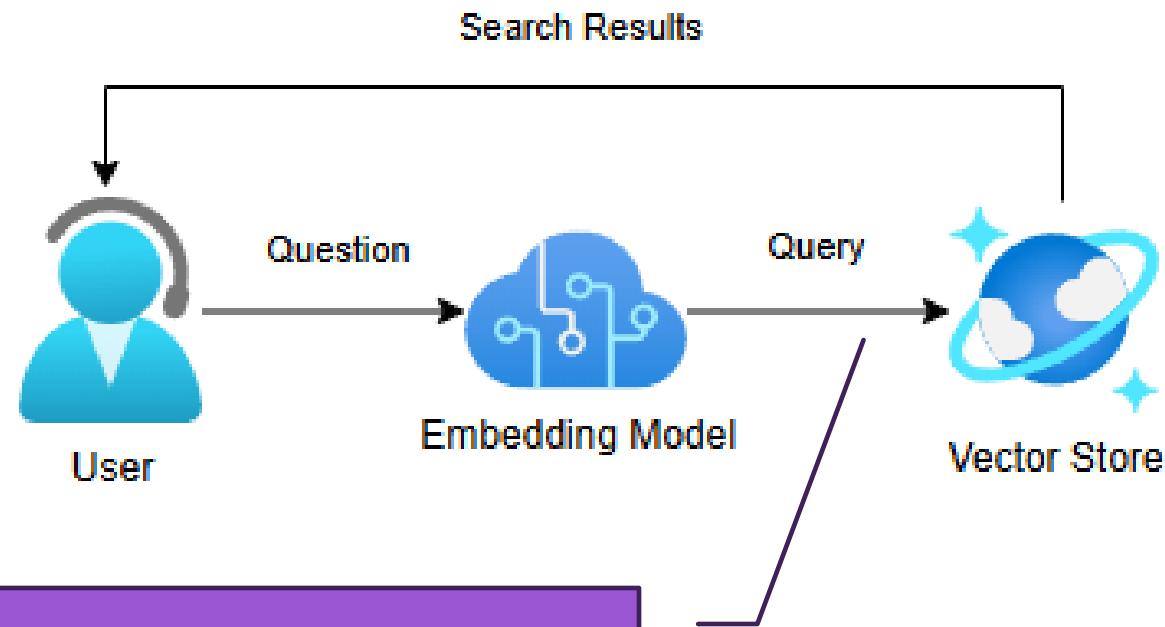
Spells

- Gaining Spells
- Casting Spells
- Spell Descriptions

SEARCHING DOCUMENTS

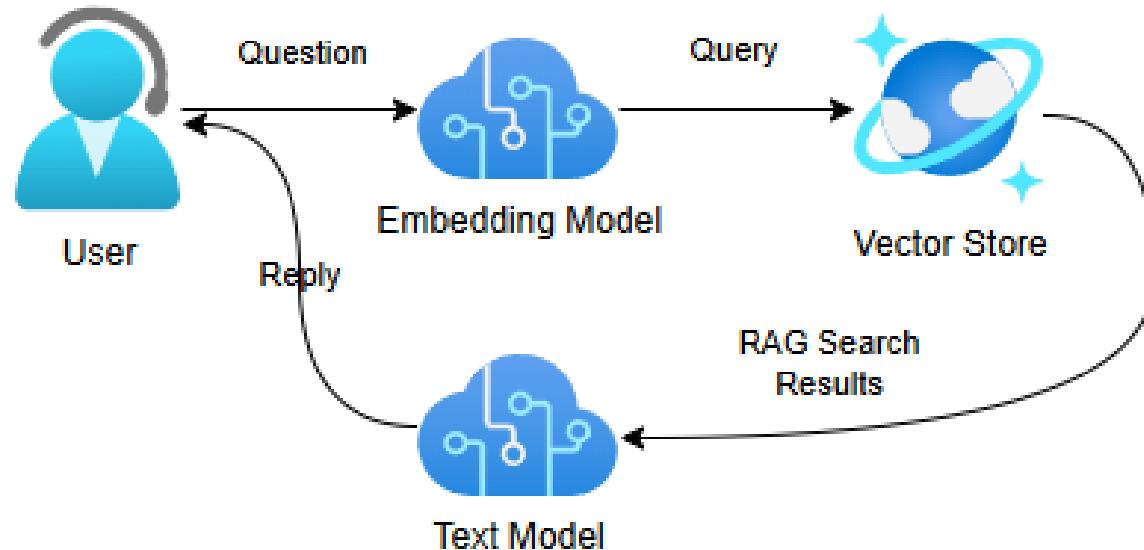


MULTI-TENANT SCENARIOS

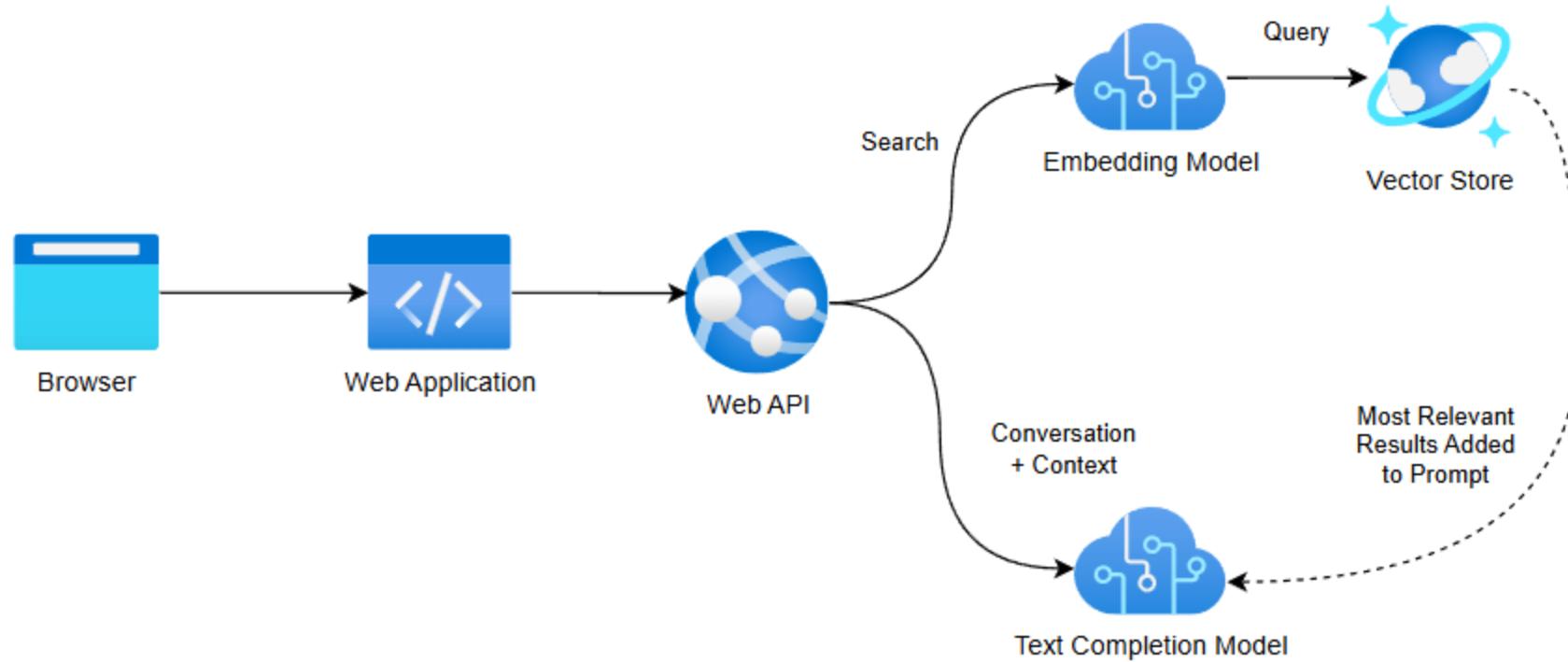


Use filters and indexes to restrict
to things the user has access to

QUESTION ANSWERING



RETRIEVAL AUGMENTATION GENERATION



KERNEL MEMORY COMPATIBILITY

Models / Services

- Ollama
- OpenAI
- Azure OpenAI
- GoogleAI / VertexAI
- HuggingFace
- MistralAI
- Onnx
- LlamaSharp
- Anthropic
- Tiktoker
- Web Client (for web scraping)
- RabbitMQ
- Azure Queues

Vector Storage

- In-Memory (default)
- Qdrant
- Azure AI Search
- Redis
- Postgres
- Sqlite
- Azure Blobs
- SQL Server
- MongoDB Atlas
- Elastic Search

KERNEL MEMORY INDEXING

```
[Description("Contains functions for looking up rules related to the game's ruleset.")]
public class SourcebookLookupPlugin(string sourceDirectory, string rulesSystem)
{
    private IKernelMemory? _memory;

    public async Task InitializeAsync(MySettings settings)
    {
        _memory = new KernelMemoryBuilder()
            .WithOllamaTextGeneration(settings.ChatModelId, settings.ChatEndpoint)
            .WithOllamaTextEmbeddingGeneration(settings.EmbeddingModelId, settings.EmbeddingEndpoint)
            .Build();

        // Index each PDF file
        string[] pdfFiles = Directory.GetFiles(sourceDirectory, "*.pdf", SearchOption.AllDirectories);
        foreach (string pdfFile in pdfFiles)
        {
            string documentId = Path.GetFileNameWithoutExtension(pdfFile).Replace(" ", "_");
            await IndexDocumentAsync(pdfFile, documentId);
        }
    }

    private async Task IndexDocumentAsync(string filePath, string documentId)
    {
        TagCollection tags = new()
        {
            { "RulesSystem", rulesSystem},
            { "Location", filePath }
        };

        await _memory!.ImportDocumentAsync(filePath, documentId, tags, index: rulesSystem);
    }
}
```

KERNEL MEMORY SEARCH

```
[KernelFunction]
[Description("Retrieves possibly helpful passages from the rules based on a search query.")]
public async Task<string> RulesSearch([Description("A topic or question to search for")] string query)
{
    SearchResult results = await _memory.SearchAsync(query, index: rulesSystem);
    if (results.NoResult)
    {
        return $"No results found for '{query}' in the ruleset.";
    }

    // Find the top 5 most relevant Partitions
    results.Results.SelectMany(p => p.Partitions).OrderByDescending(p => p.Relevance).Take(5);

    StringBuilder sb = new();
    sb.AppendLine("Here are a few of the most relevant passages from the rules:");
    foreach (var result in results.Results
        .SelectMany(p => p.Partitions)
        .OrderByDescending(p => p.Relevance)
        .Take(5))
    {
        sb.AppendLine(result.Text);
    }

    sb.AppendLine("Keep in mind that these are just the most relevant results. Use these results to respond to the player.");

    return sb.ToString();
}
```

KERNEL MEMORY ASK

```
[KernelFunction]
[Description("Asks a rules expert about a specific rule or concept in the ruleset.")]
public async Task<string> ConsultRules([Description("A question for the rulebook")] string query)
{
    MemoryAnswer results = await _memory.AskAsync(query, index: rulesSystem);
    if (results.NoResult)
    {
        return $"No results found for '{query}' in the available rules: {results.NoResultReason}";
    }

    return results.Result;
}
```



BREAK

BACK AT 4:15



ADVANCED SEMANTIC KERNEL

LET'S ELEVATE THINGS A BIT





AvaloniaApp: POST

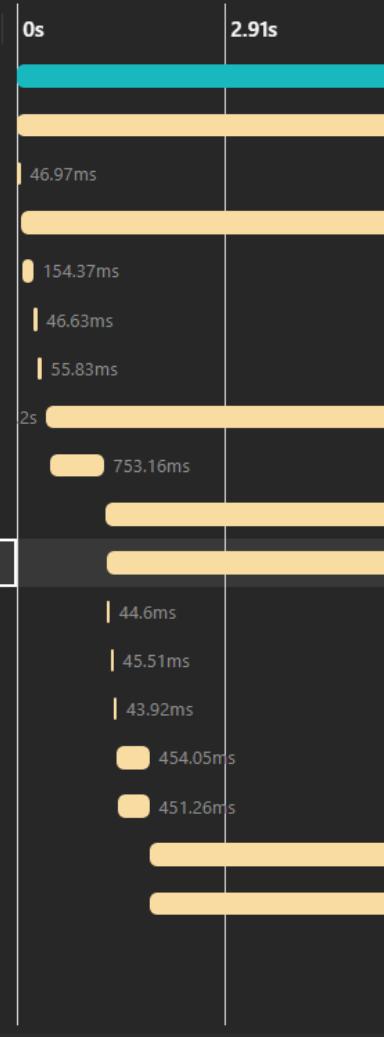
f8ef819

Trace detail 7/31/2025 9:56:56.688 PM Duration 11.62s Resources 2 Depth 8 Total spans 20

View logs

Name

- AvaloniaApp POST
- WebAPI POST /adventures/{adventureName}
 - | WebAPI → bnbstoragelrs.table.core.windows.net:443 GET
 - | WebAPI Starting chat session for adventure test
 - | WebAPI → bnbstoragelrs.blob.core.windows.net:443 HEAD
 - | WebAPI → bnbstoragelrs.blob.core.windows.net:443 HEAD
 - | WebAPI → bnbstoragelrs.blob.core.windows.net:443 GET
 - WebAPI GetChatMessageContentsAsync
 - | WebAPI → bnb-ai.openai.azure.com:443 POST
 - WebAPI GetCurrentLocation
 - WebAPI LocationPlugin-GetCurrentLocation called User meland, Adventure test
 - | WebAPI → bnbstoragelrs.table.core.windows.net:443 GET
 - | WebAPI → bnbstoragelrs.table.core.windows.net:443 PATCH
 - | WebAPI → bnbstoragelrs.table.core.windows.net:443 GET
 - WebAPI GetTextContentsAsync
 - | WebAPI → bnb-ai.openai.azure.com:443 POST
 - WebAPI GetTextContentsAsync
 - | WebAPI → bnb-ai.openai.azure.com:443 POST
 - | WebAPI → bnbstoragelrs.table.core.windows.net:443 PATCH
 - | WebAPI → bnb-ai.openai.azure.com:443 POST



WebAPI: LocationPlugin-GetCurrentLocation called User meland, Adventure test 6cebb59

Resource WebAPI Duration 7.41s Start time 1.26s

View logs

Filter...

Span

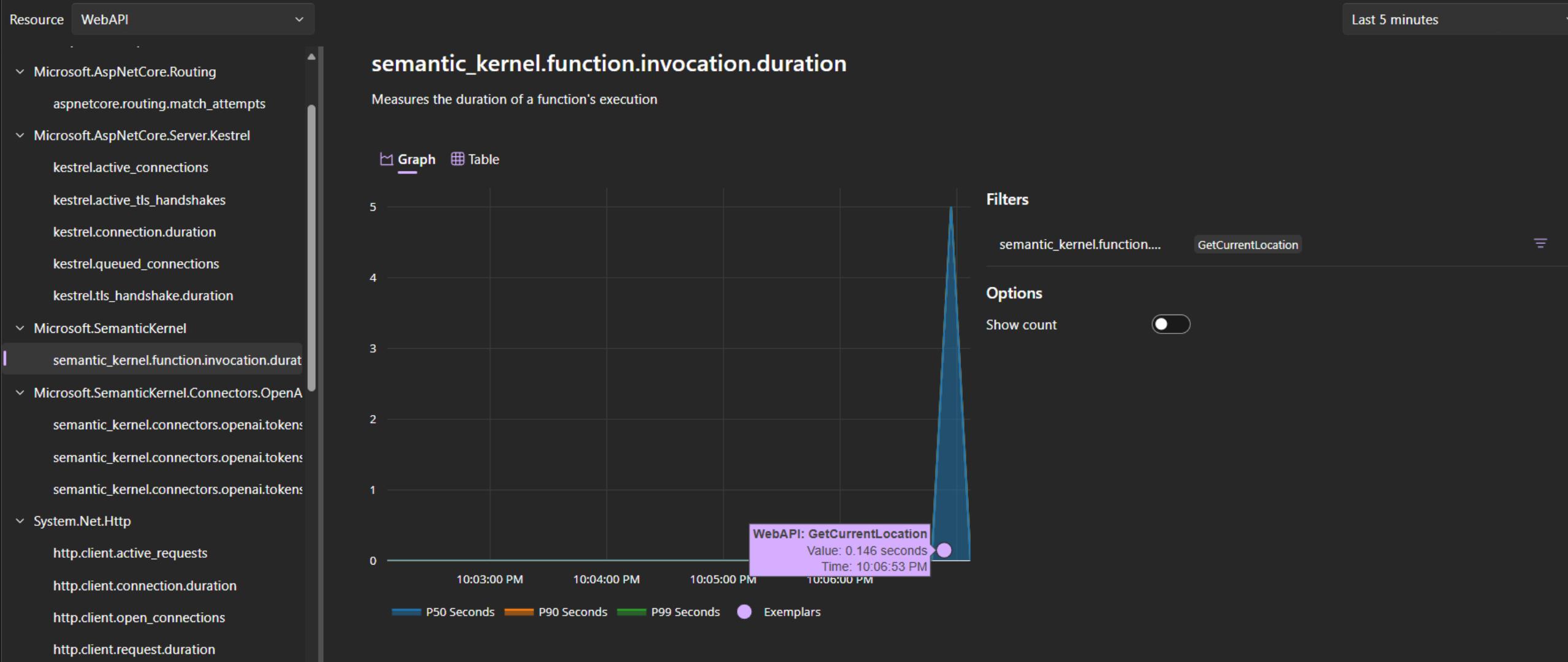
Name	Value
SpanId	6cebb5988c89ba08
Name	LocationPlugin-GetCurrentLocation called User meland, Adventure test
Kind	Internal
Description	### Whispering Hollow Glade **Overview:** Whispering Hollow Glade is a s...
History	No game history is available for this location. Please update it with the Upda...
Notes	No private notes are available for this location. You can add some with the U...
Region	default
X	0
Y	0

Context

3 ^

ASPIRE METRICS

Metrics



OPENTELEMETRY COLLECTION WITH SEMANTIC KERNEL

```
builder.Services.AddScoped<Kernel>( s =>
{
    // Set up Semantic Kernel
    IKernelBuilder kb = Kernel.CreateBuilder();
    kb.Services.AddScoped<IChatCompletionService>(_ => s.GetRequiredService<IChatCompletionService>());
    kb.Services.AddScoped<ITextToImageService>(_ => s.GetRequiredService<ITextToImageService>());
    kb.Services.AddScoped<ITextGenerationService>(_ => s.GetRequiredService<ITextGenerationService>());
    kb.Services.AddScoped<ILoggerFactory>(_ => s.GetRequiredService<ILoggerFactory>());

    return kb.Build();
});
```

OPENTELEMETRY COLLECTION WITH SEMANTIC KERNEL

```
public static TBuilder ConfigureOpenTelemetry<TBuilder>(this TBuilder builder) where TBuilder : IHostApplicationBuilder
{
    builder.Logging.AddOpenTelemetry(logging =>
    {
        logging.IncludeFormattedMessage = true;
        logging.IncludeScopes = true;
    });
    builder.Services.AddOpenTelemetry()
        .WithMetrics(metrics =>
    {
        metrics.AddAspNetCoreInstrumentation()
            .AddHttpClientInstrumentation()
            .AddMeter("Microsoft.SemanticKernel*")
            .AddMeter("MattEland*")
            .AddRuntimeInstrumentation();
    })
        .WithTracing(tracing =>
    {
        tracing.AddSource(builder.Environment.ApplicationName)
            .AddAspNetCoreInstrumentation()
            .AddSource("Microsoft.SemanticKernel*")
            .AddSource("MattEland*")
            .AddHttpClientInstrumentation();
    });
    builder.AddOpenTelemetryExporters();
    return builder;
}
```

Microsoft and other Fortune 500 companies

are already leveraging Semantic Kernel because it's flexible, modular, and observable. Backed with security enhancing capabilities like telemetry support, and hooks and filters so you'll feel confident you're delivering responsible AI solutions at scale.

Version 1.0+ support across C#, Python, and

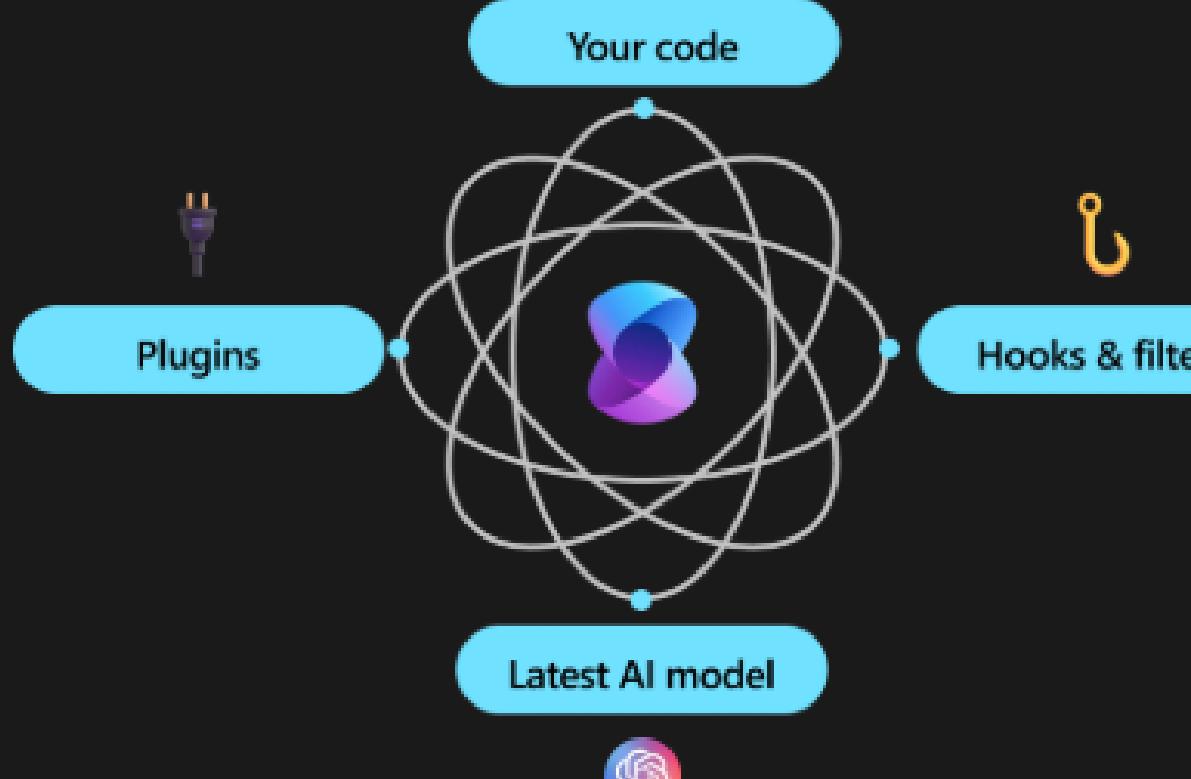
Java means it's reliable, committed to non-

biasing AI. Any existing chat-based

APIs are easily expanded to support additional

[HTTPS://LEARN.MICROSOFT.COM/EN-US/SEMANTIC-KERNEL/OVERVIEW/](https://learn.microsoft.com/en-us/semantic-kernel/overview/)

modalities like voice and video.



SK DOCS

New model drop!

Semantic Kernel was designed to be future

KERNEL MEMORY DOCS

[HTTPS://MICROSOFT.GITHUB.IO/KERNEL-MEMORY/](https://github.com/microsoft/kernel-memory)

Kernel Memory

license MIT Discord 277 online

Kernel Memory (KM) is a **multi-modal AI Service** specialized in the efficient indexing of datasets through custom continuous data hybrid pipelines, with support for **Retrieval Augmented Generation** (RAG), synthetic memory, prompt engineering, and custom semantic memory processing.

KM is available as a **Web Service**, as a **Docker container**, a **Plugin** for ChatGPT/Copilot/Semantic Kernel, and as a .NET library for embedded applications.



Utilizing advanced embeddings and LLMs, the system enables Natural Language querying for obtaining answers from the indexed data, complete with citations and links to the original sources.



Designed for seamless integration as a Plugin with **Semantic Kernel**, Microsoft Copilot and ChatGPT, Kernel Memory enhances data-driven features in applications built for most popular AI platforms.



CONCLUSION

UNTIL TOMORROW

A chaotic orc woman with dark skin and wild hair, wearing ornate armor, stands amidst a ruined city. She holds a bunch of pink flowers in her right hand, looking directly at the viewer with a somber expression. The background shows burning buildings and debris under a cloudy sky.

DON'T LET YOUR
AI SYSTEMS BE
CHAOTIC EVIL

CONTENT TOMORROW

Victor Frye

Aspire to Effortless: Modern local development with .NET Aspire

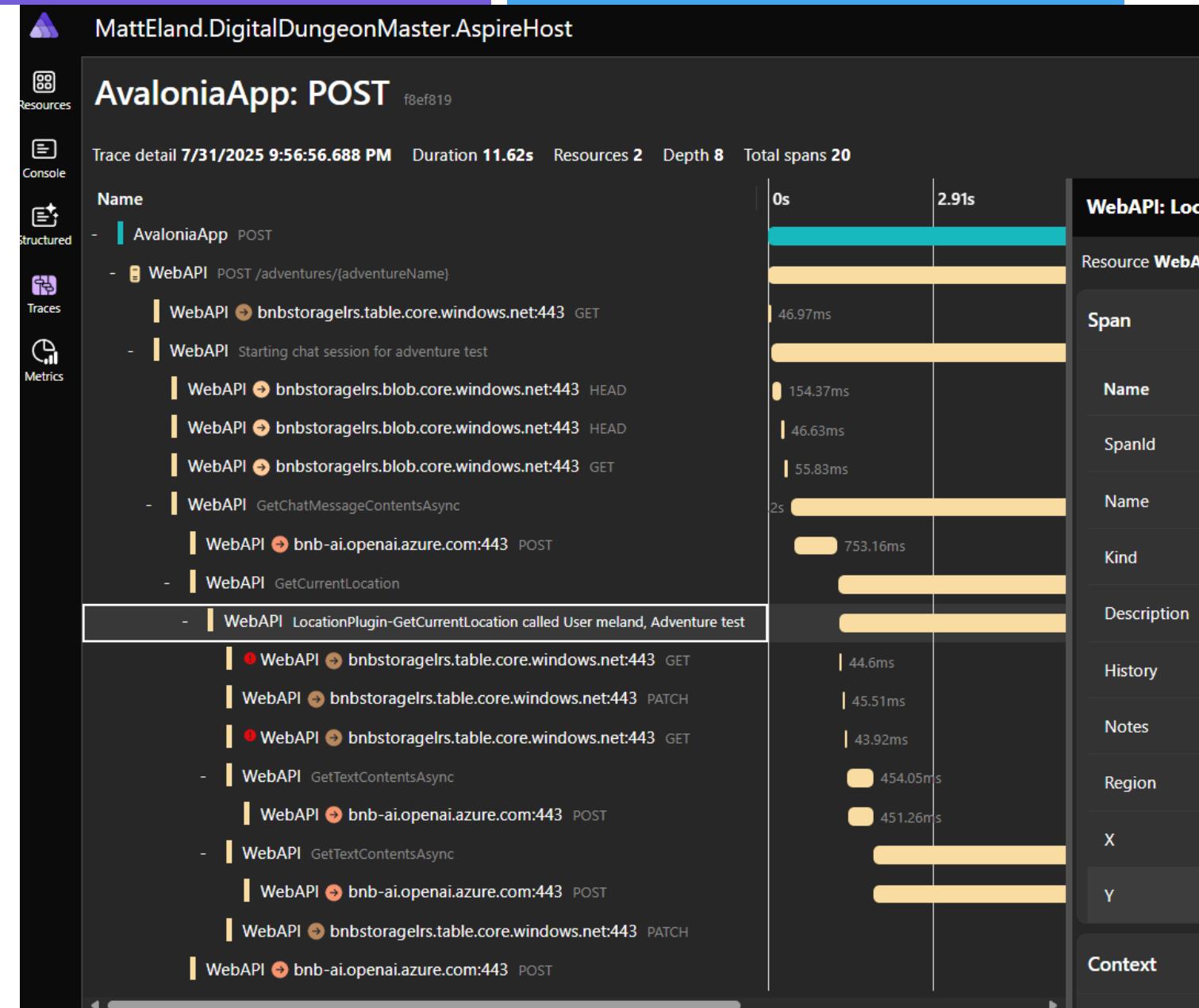
Building modern applications often involves running multiple services, such as a .NET API, a JavaScript frontend, and backing cloud services. .NET Aspire offers to solve this challenge in our local development by modeling our entire application stack, enabling us to start it with a single command. This talk will explore how Aspire streamlines our local environment for .NET developers, making the developer experience effortless for building and testing your applications.

We'll showcase a sample CRUD application to highlight the impact of Aspire on the local development environment, focusing on features like the app host orchestrator and integration testing. Together, we will discover the excitement of using Aspire and its promising future within the .NET ecosystem.

ROOM 135

Sat 11:00 am - 12:00 pm

Language + Platform



Nicole Hoyng



Lights, Camera, Github Actions!

Getting your code on stage has never been easier with Github Actions! But how do you get started? Workflows? Composite Actions? Runners? Producing a build and deployment pipeline doesn't need to be Hollywood magic. In this talk, we'll discuss what goes into a Github Action Workflow, how to structure your build and deployment processes, and why all of this matters.

When you leave this showing, we'll have your audience of stakeholders applauding so you won't end up with a dud at the box-office.

ROOM 255

Sat 3:00 pm - 4:00 pm

DevOps + Infrastructure

Bob Fornal

Ditch your JavaScript for CSS!

Speed is critical in Web Design. There are things done daily on websites that can be managed in CSS as easily as JavaScript—simply faster. Over the years, I have encountered dozens of these patterns.

Let's look at both sides of these code examples and discuss their pros and cons. Can we (or should we) ditch JavaScript for CSS?

ROOM 222

Sat 10:00 am - 11:00 am

Web



Bob Fornal

The Impact of Testing on a DevOps Pipeline

CI/CD pipelines are an integral part of almost any organization today. "It works on my machine," is only valid in a few scenarios. Our code needs to get to other environments: Dev, QA, UAT, and Production. Testing is a recommended part of the pipeline process.

What is the impact of NOT having active testing in your pipeline processes? How do we get testing in place and how does it affect an organization?

ROOM 115

Sat 1:00 pm - 2:00 pm

Architecture



Aspire Roadmap 2025: Code-first DevOps, polyglot, and AI

Jul 30, 2025 by Victor Frye · 11 min read

[ASPIRE](#) [DOTNET](#) [DEVOPS](#) [AI](#)



A look at the recently published Aspire roadmap for 2025, focusing on its code-first DevOps evolution, polyglot aspirations, and AI orchestration, and more.

[LEARN MORE](#)

Coding Agents are here: Is your team ready for AI devs?

Jul 23, 2025 by Matt Eland · 10 min read

[AI](#) [GITHUB](#) [PRODUCTIVITY](#) [AGENTS](#)



Discover how GitHub Copilot Agents can function as AI developers on your team, automating development while complementing human developers' strategic skills.

[LEARN MORE](#)

Reference Architecture for Team AI Productivity

Jun 19, 2025 by Matt Eland · 11 min read

[ARCHITECTURE](#) [AI](#) [PRODUCTIVITY](#) [SECURITY](#)



Web chat portals can profoundly empower your team to build amazing things. Let's see how these systems can work securely and offer unique value to your team.

[LEARN MORE](#)



How far can we go together?

[CONTACT LEADING EDJE](#)

Recent Posts

- Aspire Roadmap 2025: Code-first DevOps, polyglot, and AI
- Reviewing Aspire.JS: Current state of Aspire for JavaScript
- Coding Agents are here: Is your team ready for AI devs?
- dotnet run file.cs
- Reference Architecture for Team AI Productivity
- Why Small Businesses Need Protected Generative AI Solutions
- .NET Options Pattern: Binding settings and user secrets
- Adding .NET Aspire: A command line user's guide

Series

[AI PROTOTYPING PROJECTS](#) 5

[HELLO ASPIRE](#) 5

[PLAYWRIGHT TESTING](#) 4

[.NET AI / ML HIGHLIGHTS](#) 3

[AGILE FORECASTING](#) 3

[REFERENCE ARCHITECTURES](#) 3

[TESTING FRAMEWORKS](#) 3

[CAREER DEVELOPMENT](#) 2

[REST ASSURED](#) 2



Taos





Lunch & Learns with Leading EDJE

Your place
Your team
Our speakers and expertise
Your choice of content



Thank You to Our Sponsors



InstructLab



Questions?

Feedback & Contact Form



Matt Eland

AI SPECIALIST & WIZARD



<https://forms.gle/SrTzVzFnsDz4kiA18>



MATT.ELAND@LEADINGEDJE.COM



HTTPS://MATTELAND.DEV/