

Integrate Multimodal Language Models into .NET Backends with OpenAI .NET API and .NET Aspire

think
tecture

Generation AI Meetup
Karlsruhe, 2024-10-17

Kenny Pflug

[@feO2x](#)

Software Architect Consultant



Today's menu

- Use Case: Information Extraction from text, speech, and images
- OpenAI API for .NET for accessing GPT-4o and Whisper
- Endpoint design and prompt management
- Integration into .NET Aspire-based backend infrastructure
- A small excursion into authentication security

Kenny Pflug

Consultant Software Architect @ Thinktecture AG

- Distributed Systems with ASP.NET Core
- .NET internals
- Cloud-native



kenny.pflug@thinktecture.com

[@fe02x](https://www.linkedin.com/in/fe02x)

<https://www.thinktecture.com>

The Use Case: filling out a complex form

Schadenmeldeformular

ADAC Versicherung AG - 81362 München

ADAC

▼ Mitgliedsnummer

Schadenfall

in Deutschland ☐ im Ausland ☐

▼ Name / Vorname des Mitglieds

▼ Telefon tagsüber

▼ Mail / Fax

▼ PLZ / Ort

▼ Straße / Nr.

Angaben zum Schaden:

▼ Datum des Schadens

▼ Uhrzeit

Panne

☐

Unfall

☐

Kfz-Diebstahl

☐

Sonst.

☐

▼ Anz.d.Reisenden inkl.Fahrer

▼ Schadenland

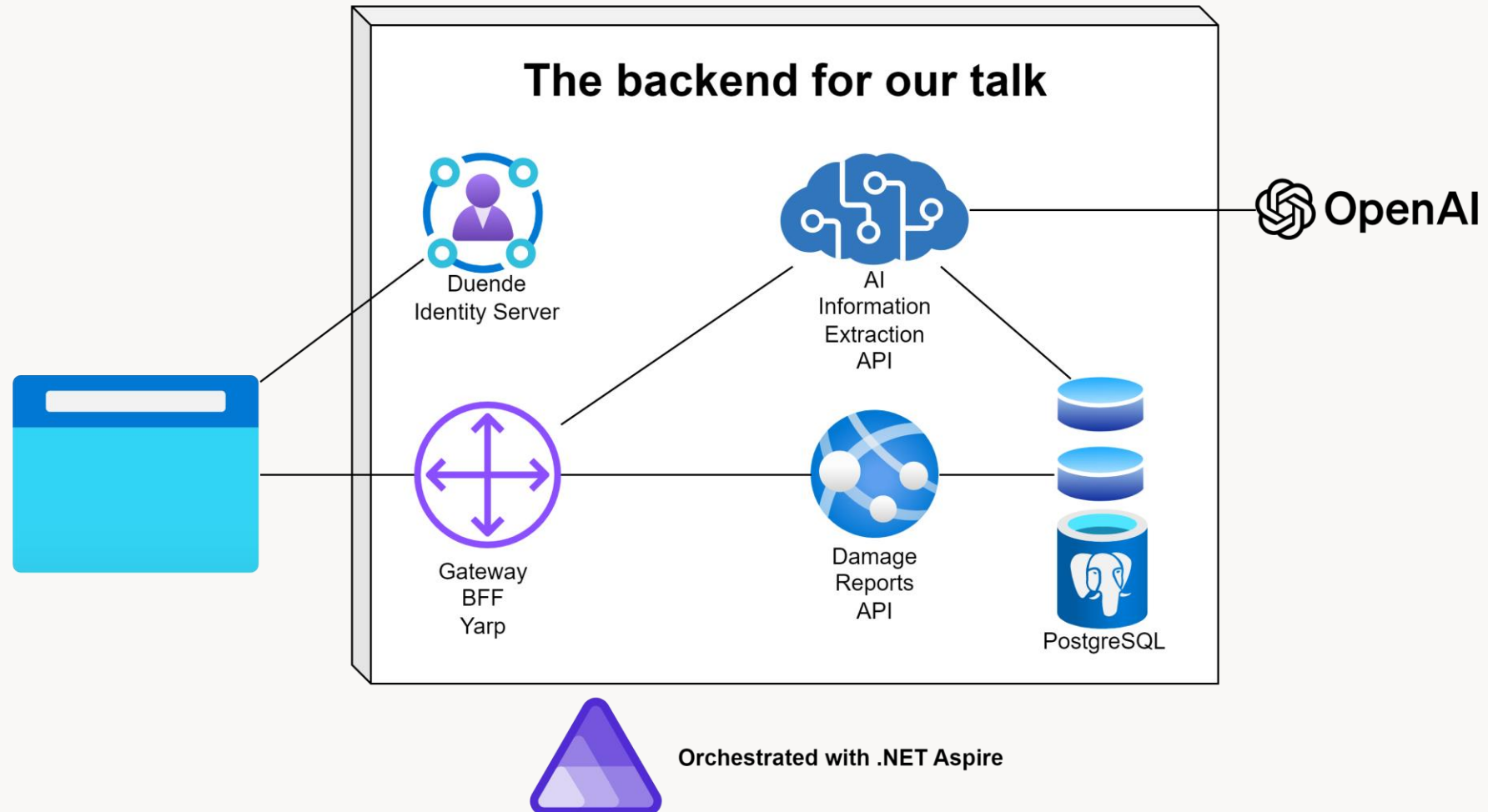
▼ PLZ / Ort des Schadens

▼ ca. Entfernung zum Wohnort

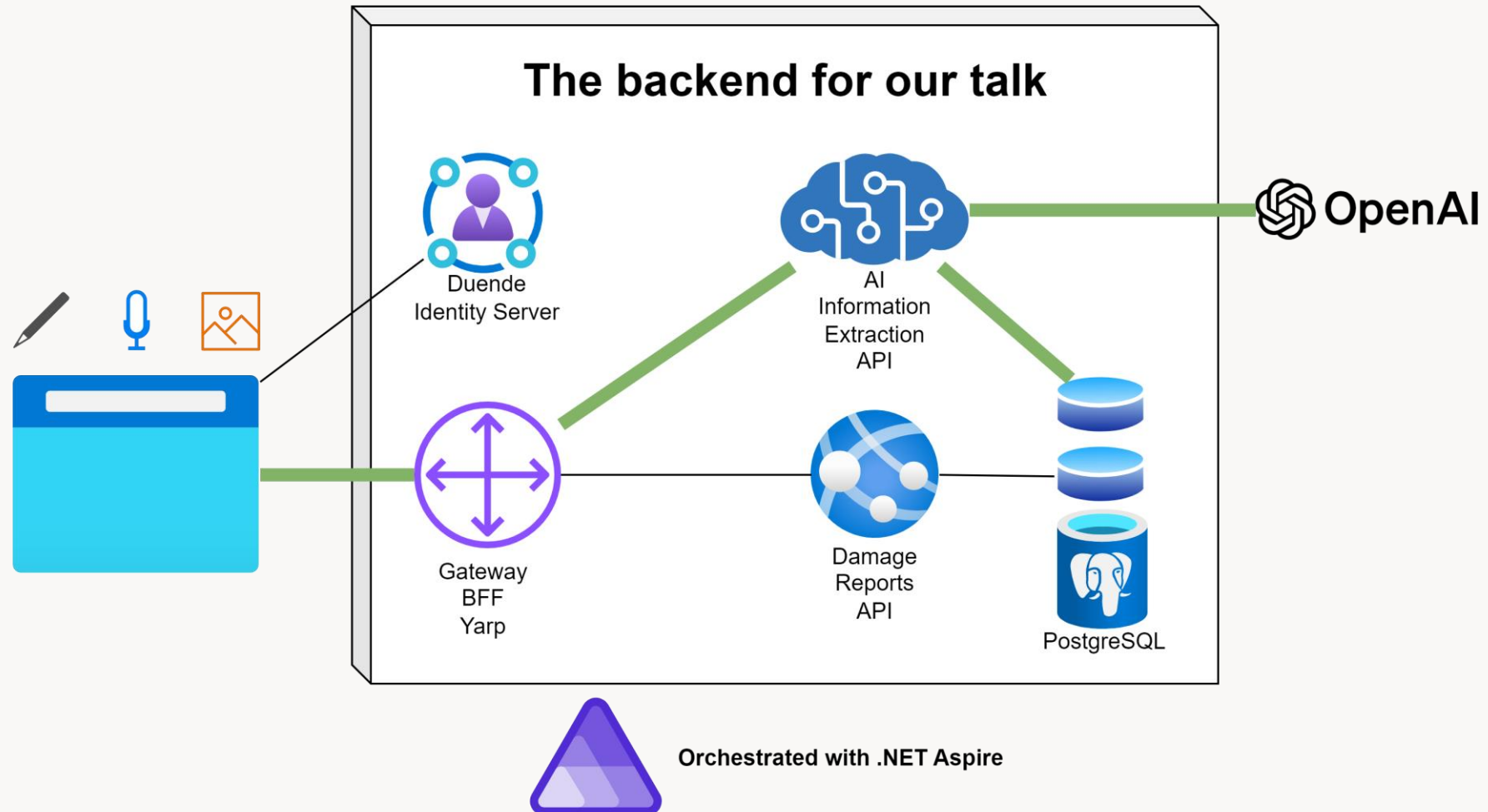
Können Sie die Mehrwertsteuer beim

ja ☐ nein ☐

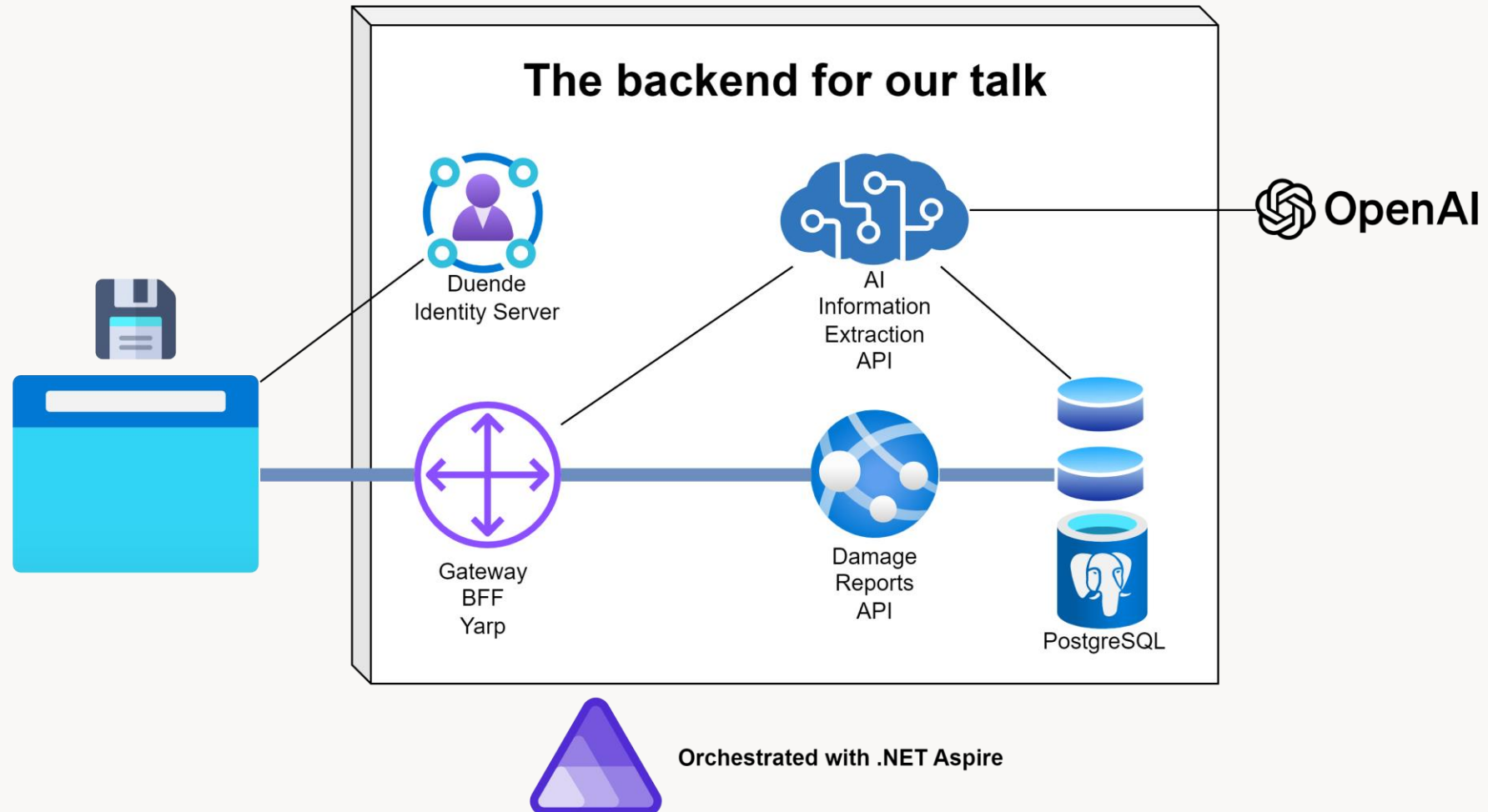
Backend Overview



While filling out the form



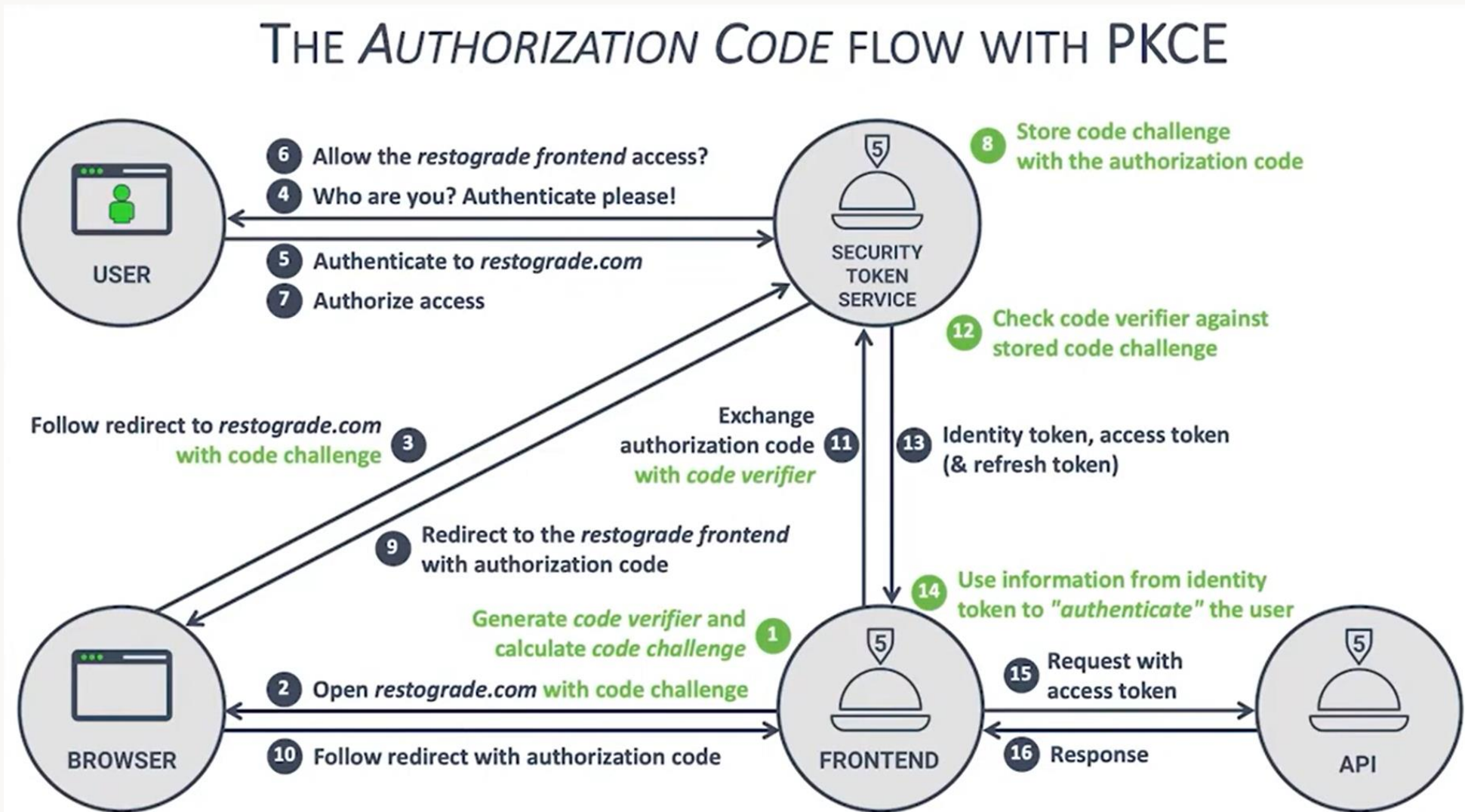
When reviewing and saving



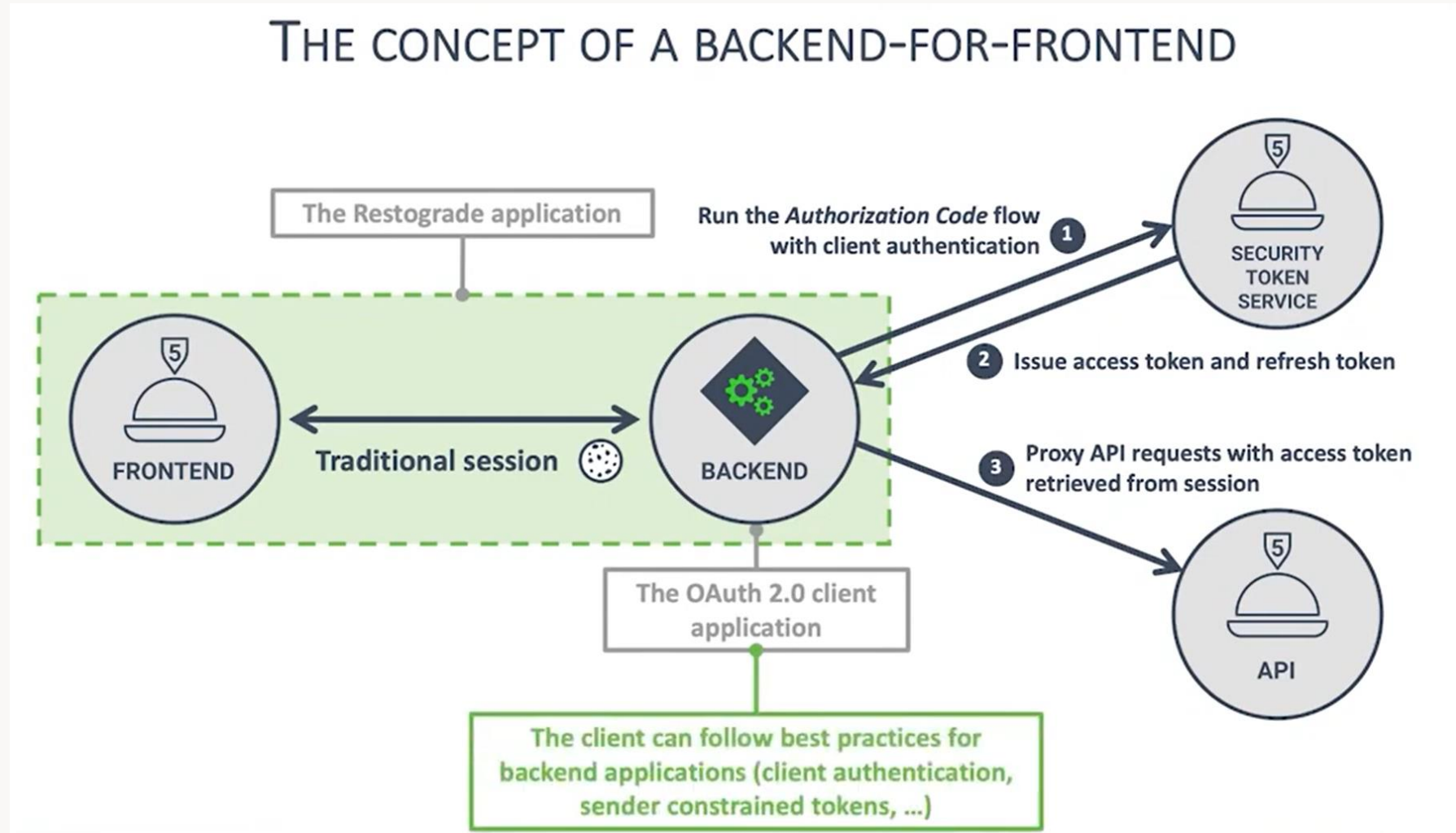
A close-up, low-angle shot of a computer keyboard. The keys are dark with light-colored characters. A semi-transparent dark gray rectangular box is centered over the keyboard, containing the text "Demo Time" in a large, bold, white sans-serif font, and "Let's see some code" in a smaller, white sans-serif font below it. The lighting is dramatic, with strong highlights on the edges of the keys and the text box.

Demo Time
Let's see some code

Open ID Connect - Authorization Code Flow



Open ID Connect – Backend for Frontend manages Tokens



Final thoughts on AI Information Extraction

- Provide scope: focus on parts of the problem so that the MLM can provide better responses
- Provide context: present the current values so that the MLM can derive information from it
- Structured output: only accept JSON so that your app can easily parse the results
- Expect errors: the LLM might provide invalid responses – validate it
- Try local models: Llama 3.2 and OpenAI Whisper might be everything that you need
- This use case is rather easy: no RAG embeddings required, no context window management

Yesterday, .NET Aspire 9 RC1 was released

Welcome to the Upgrade Assistant!

This experience will guide you through the process of upgrading your project towards newer technologies. [Learn more](#)

Ready for upgrade? Select how you want to upgrade MyAspireApp.AppHost

▼ Aspire upgrades



Upgrade to latest .NET Aspire version

Upgrade your .NET Aspire AppHost and projects to latest packages and APIs

▼ Upgrade project to a newer .NET version



In-place project upgrade

Upgrades project and its components in place using transformations applicable for the project.

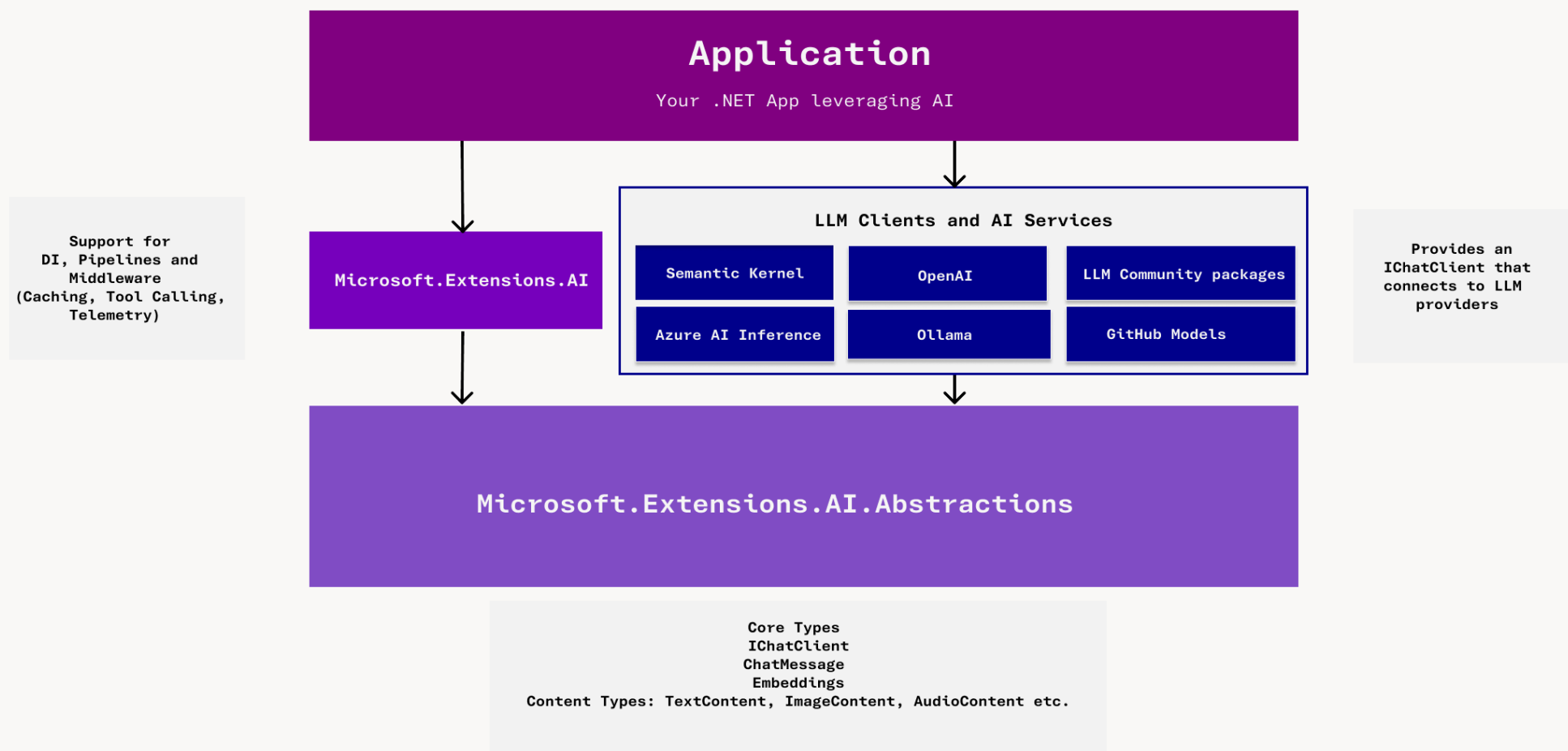
▼ NuGet upgrades



NuGet central package management (CPM)

Convert selected projects to use NuGet central package management (CPM) feature to simplify managing common dependencies across all projects

What about Microsoft.Extensions.AI?



Sources

- [OpenAI API for .NET](#) – GitHub
- [.NET Aspire Overview](#) – Microsoft Learn
- [Smart Form Filler](#) – Thinktecture Labs, GitHub
- [What's new in .NET Aspire 9 RC1](#) – Microsoft Learn
- [Introducing Microsoft.Extensions.AI Preview – Unified AI Building Blocks for .NET](#) – Luis Quintanilla, Microsoft .NET Blog
- [The insecurity of OAuth 2.0 in frontends](#) – Philippe de Ryck, NDC Security 2023
- [Securing SPAs and Blazor Applications using the BFF \(Backend for Frontend\) Pattern](#) – Dominick Baier, NDC Security 2023
- [Reading JSON and binary Data from Multipart/Form data](#) – Andrew Lock's Blog

Thanks!