



The All-in-One Platform for an efficient AI Evaluation

Building trust between businesses and AI
developers



About us

With **15 years** of proven expertise, we stand as a beacon of experience in the industry.

We are a team of passionate engineers, guided by a culture of craftsmanship and quality. We know how to work on existing systems, respecting their history, modernizing them seamlessly, and strengthening them with sound technical choices.

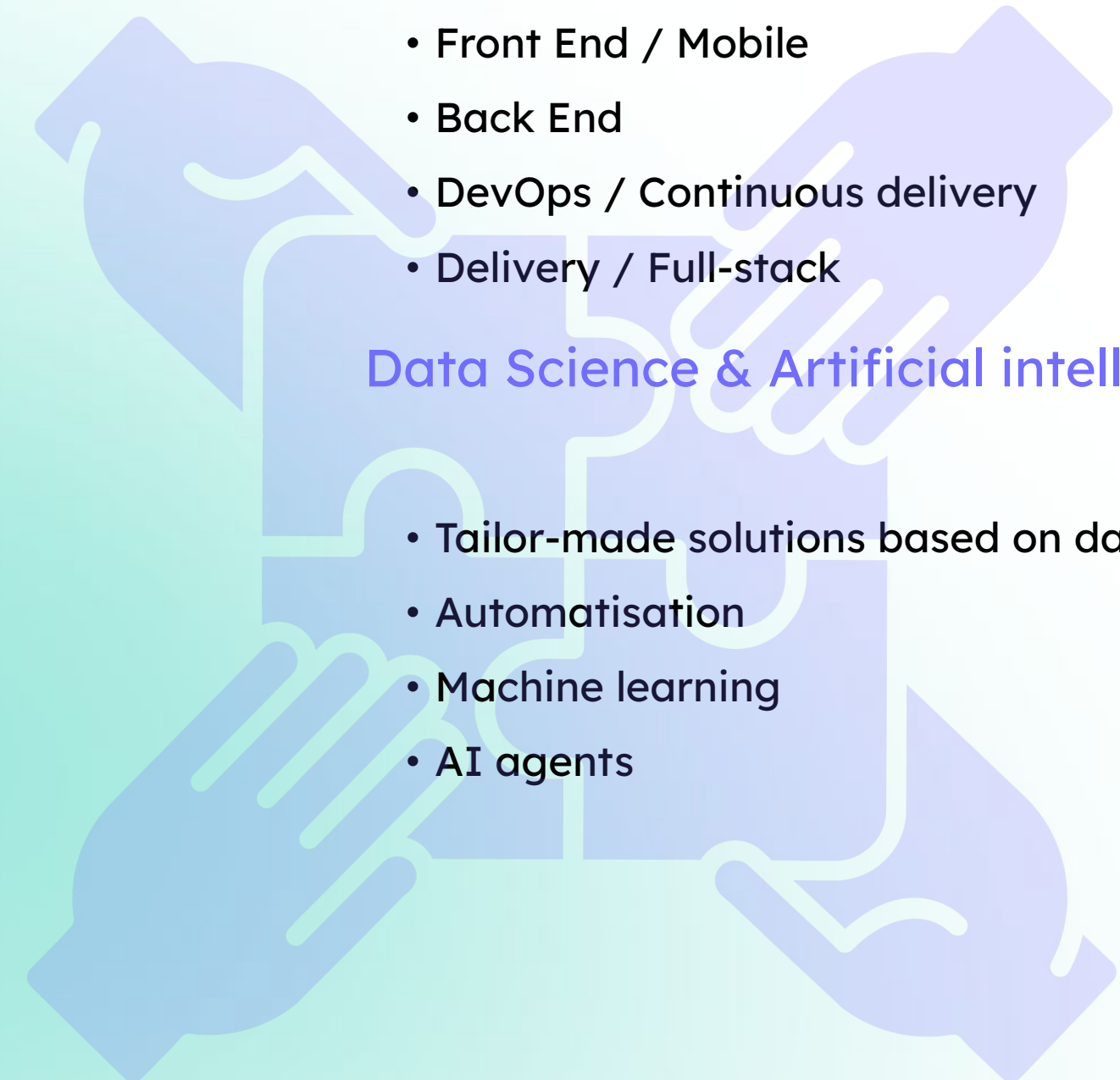
Our Tech Leads have over twenty years of experience in demanding environments (banks, scaleups, SaaS), and our team has already delivered more than thirty products and platforms.

Application development

- UX Design
- Front End / Mobile
- Back End
- DevOps / Continuous delivery
- Delivery / Full-stack

Data Science & Artificial intelligence

- Tailor-made solutions based on data
- Automatisation
- Machine learning
- AI agents



Our expertise

We operate in multiple sectors ranging from finance to video games to create and deploy digital solutions for every needs.

Application development

- UX Design
- Front End / Back End
- DevOps / Continuous delivery
- Delivery / Full-stack

Product Management

- Pre-sales consulting
- Marketing collateral
- User training
- Support and operations

Data Science & Artificial intelligence

- Tailor-made solutions based on data
- Automatisation
- Machine learning
- AI agents



AI Agents, The Big Gap!

Ministerial role [\[edit\]](#)

On 11 September 2025, Diella was formally appointed "Minister of State for Artificial Intelligence".^[11] The appointment followed a presidential decree authorizing the Prime Minister to oversee the creation and operation of a virtual AI minister. Procurement responsibilities are planned to be transferred gradually to the system to reduce political influence in tender procedures. The appointment is part of broader anti-corruption reforms and measures intended to align Albania with [European Union accession](#) requirements.^{[12][13]}

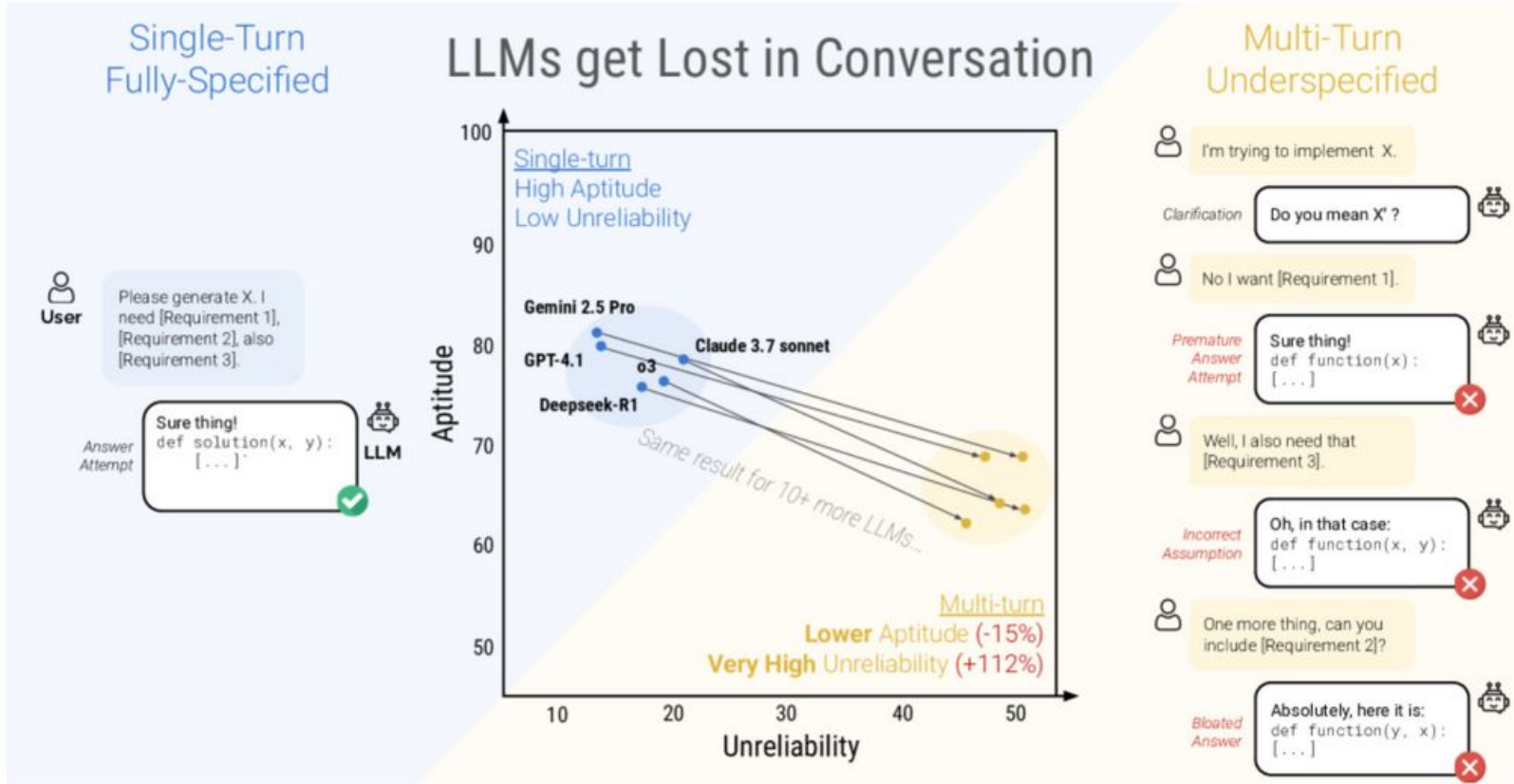
Prime Minister [Edi Rama](#) stated that Diella would help ensure that "public tenders will be 100% free of corruption".^[14]

Diella 	
Other names	Diella AI
Developer(s)	AKSHI
Initial release	19 January 2025; 8 months ago
Stable release	MinR version (Diella 2.0) / 12 September 2025; 12 days ago
Preview release	VA version (Diella 1.0) / 19 January 2025; 8 months ago

VS

Our experiments confirm that all major LLMs, whether open- or closed-weight, exhibit significantly lower performance in multi-turn conversations compared to single-turn exchanges, with an average drop of 39% across six generation tasks.

Microsoft Research - May 2025

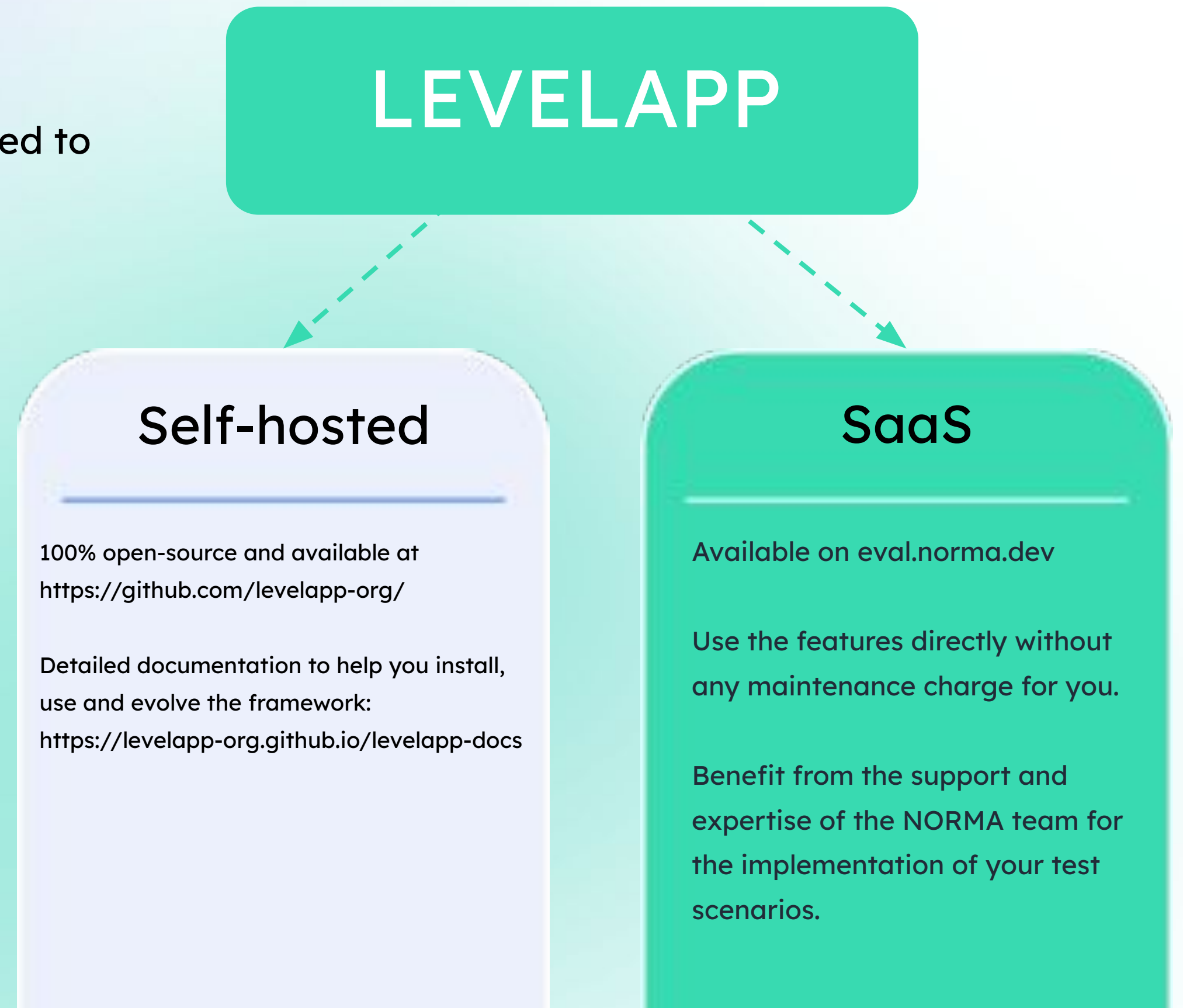


Our Solution

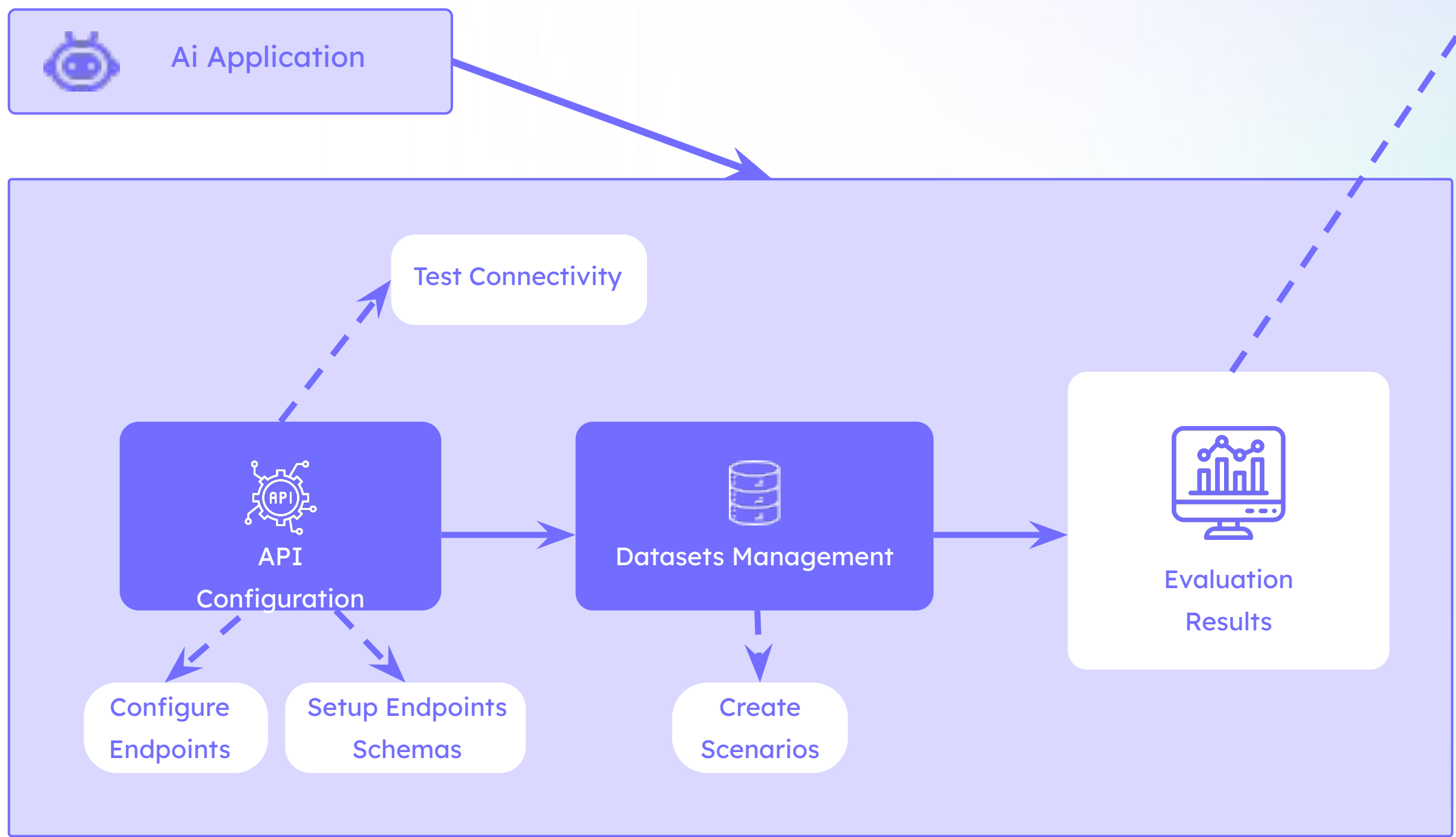
LevelApp is an open-source framework designed to test agentic AI applications across complex, multi-round scenarios. It easily integrates into technical teams' CI/CD workflows.

LevelApp features:

- Evaluation of each step:
 - Extraction
 - Classification
 - Generation
- Batch testing & continuous integration
- Detection of weaknesses or regressions



How does it work?



Evaluations / Results / 1

← Projects / Customer Support Bot

lonos-starter-pack-agentic-chatbot Active

AI agent for handling customer inquiries

● API Configured 1 Datasets 5 Scenarios 3 Evaluations

⚙️ API Configuration 📄 Datasets ▶ Evaluations

Test Name

Staging Testing 09 Sep

📅 Created: 09/09/2024, 10:03:29

API Host

🔗 `http://127.0.0.1:8000/api/v1/chat`

📈 PROVIDER : OPENAI

84.7/100

Justifications:

- Strong technical accuracy and comprehensive responses
- Good boundary setting for off-topic questions
- Consistent professional tone across interactions

📈 PROVIDER : IC

82.3/100

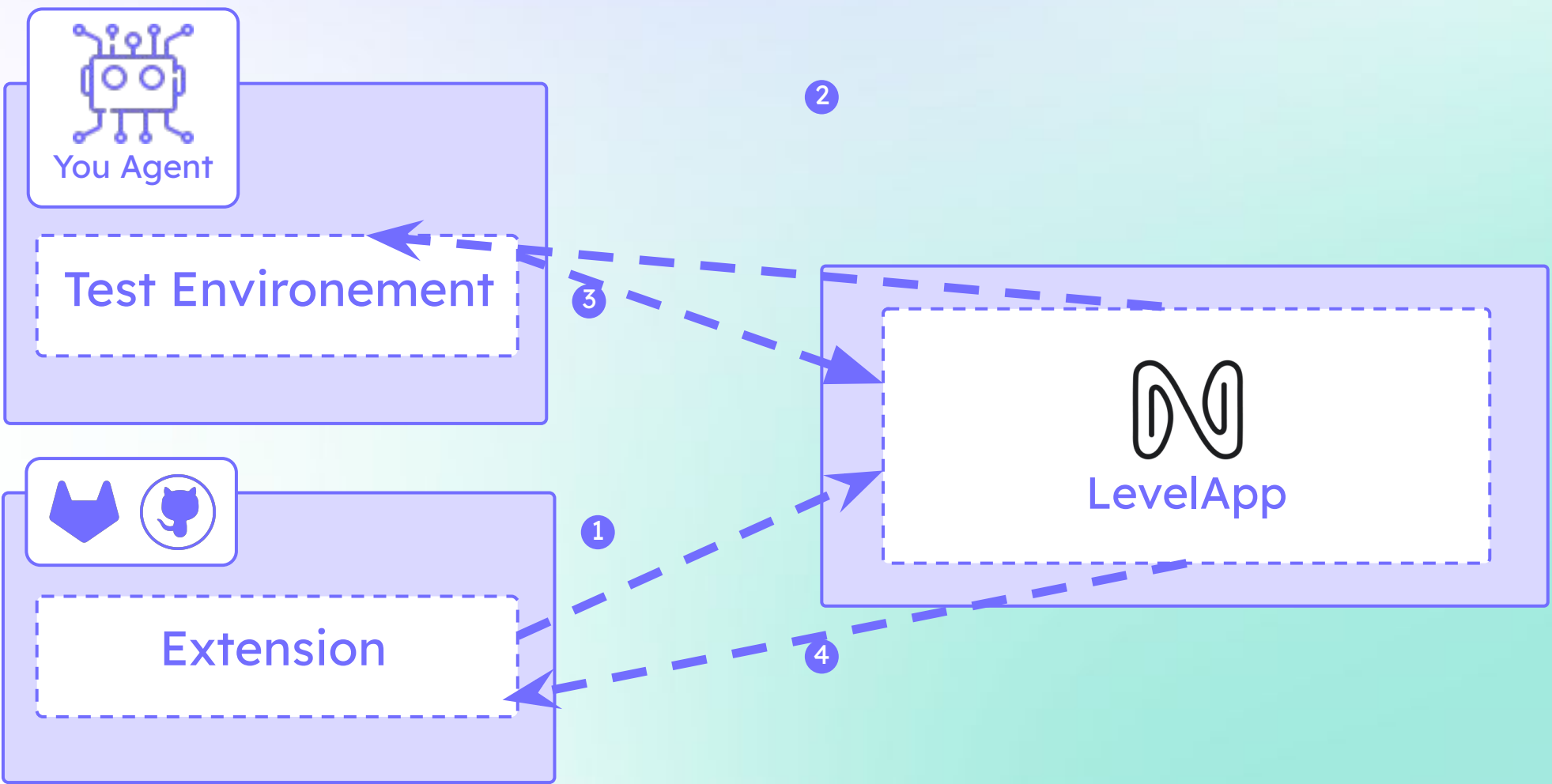
Justifications:

- Accurate service in
- Appropriate scope
- Some responses c

Batch Test Results

Scenario Name	Status	Avg Score	Execution Time
IONOS Web Hosting Inquiry	passed	89/100	2.3s
Weather Question (Off-topic)	passed	85/100	1.8s
Domain Setup Technical Support	failed	72/100	3.1s

CI/CD integration



github-actions bot commented on Jun 20 • edited

Evaluation ongoing

- **Test Name:** Run new table
- **API Endpoint:** https://eval-assistant-gateway-46wq2721.uc.gateway.dev
- **Type:** multiAgent

🔧 If you need to make changes, update your branch and rerun the workflow.

github-actions bot commented on Jun 20 • edited

Fetches evaluation results

- **User ID:** 01E3okNgmKaPUQXbCikgsazgTNi1
- **Project ID:** f5ee240b-da4d-4bb1-8599-f175939eaa3f
- **Batch ID:** batch-d9d6d121-74d5-457c-81fe-b1882be2300c
- **GPT global average score:** 12%
- **Ionos global average score:** 12%
- **Metadata global average score:** 33%

[View results in dashboard](#)

Result table

Scenario name	GPT scenario average score	Ionos scenario average score	Metadata scenario average score
Scenario 1	22%	22%	33%
Scenario 2	0%	0%	33%
Scenario 3	13%	13%	33%

🔧 If you need to make changes, update your branch and rerun the workflow.

Evaluation results

Norma Eval

Dashboard / Projects / Ionos-starter-pack-agentic-chatbot / Evaluations / Results / 1

PLATFORM

Overview

Projects

Evaluations

Settings

← Projects / Customer Support Bot

IONOS Service Questions

Active

AI agent for handling customer inquiries

API Configured

1 Datasets

3 Scenarios

1 Evaluations

API Configuration

Datasets

Evaluations

Test Name

IONOS Service Questions

Created: 09/09/2024, 10:03:29

API Host

http://127.0.0.1:8000

AVG OPENAI

84.7/100

Justifications:

- Strong technical accuracy and comprehensive responses
- Good boundary setting for off-topic questions
- Consistent professional tone across interactions

AVG IONOS

82.3/100

Justifications:

- Accurate service information and
- Appropriate scope management
- Some responses could be more

Batch Test Results

Scenario Name	Status	Avg Score	Execution Time
IONOS Web Hosting Inquiry	passed	89/100	2.3s
Weather Question (Off-topic)	passed	85/100	1.8s
Domain Setup Technical Support	failed	72/100	3.1s

IONOS Web Hosting Inquiry

Conversation turns and evaluation scores

Status

passed

Average Score

89/100

Execution Time

2.3s

Conversation Turns

1 Turn 1

User Message

What web hosting plans does IONOS offer?

Expected Response

IONOS offers several web hosting plans including Essential, Business, and Expert packages. Each plan comes with different storage, bandwidth, and features. Would you like me to explain the specific features of each plan?

AI Response

IONOS provides multiple web hosting solutions: Essential (€4/month), Business (€8/month), and Expert (€12/month) plans. Each includes different storage capacities, email accounts, and performance features. Can I help you choose the right plan?

OpenAI Judge

92/100

Excellent response covering all key points with specific pricing. Maintains helpful tone and offers follow-up assistance.

IONOS Judge

88/100

Good coverage of hosting plans with pricing details. Could be more detailed about specific features but maintains professional tone.

2 Turn 2

User Message

Tell me about the Business plan features.

Expected Response

The IONOS Business plan includes 100GB SSD storage, unlimited bandwidth, up to 100 email accounts, free SSL certificate, and 24/7 customer support. It's perfect for growing businesses and costs €8/month.

Thank you !

**Meet us at the stand F22 to
continue the discussion**

