



nocware® Playground – Service Description

ambiFOX stellaneo

Fleehook 1
48683 Ahaus

Fon +49 (0) 2561 8693 0

post@ambifox.com
www.ambifox.com



ambiFOX



Document of - Provider -
ambiFOX stellaneo
Fleehook 1
48683 Ahaus

1 **PREAMBLE**.....3

2 **PLAYGROUND OVERVIEW**.....3

2.1 **CUSTOM AI AGENTS (NOCKIS)**4

3 **NOCKI CATEGORIES**5

3.1 **CUSTOM CHATBOT**.....5

3.2 **WEBSEARCHER**6

3.3 **DATA PROCESSOR**7

4 **SUPPORTED LLMS**.....9

5 **API KEYS AND TOKEN USAGE**.....10

6 **PACK MODE (MULTI-AGENT-WORKFLOWS)**.....10

7 **A/B SPLIT TESTING**.....12

Version	Date	Content	
0.9	09.12.2024	First draft	Simon Klein-Menting



1 Preamble

The nocware Playground is designed to make the power and benefits of AI agents accessible to a wide range of users in a simple and intuitive manner. It provides an environment for exploring and testing various scenarios and use cases with AI chatbots, empowering both technically skilled users and those with minimal technical expertise. As a personalized sandbox platform, the Playground enables users to create, customize, and optimize AI agents—referred to as “nockIs”—for diverse applications. Users can deploy nockIs, adjust their settings, and save optimized workflows for future use (coming soon).

The platform supports multiple state-of-the-art AI models, including OpenAI, Anthropic, Mistral, and Gemini. Accessible via modern web browsers with user login, the Playground adopts an API-first approach, ensuring all features are available programmatically. Additionally, it is built as a scalable and tenant-ready standalone environment, ready to meet the needs of organizations of all sizes. Future enhancements include an open workflow engine, enabling seamless integration and extension into external systems.

2 Playground overview

The Playground's Overview page serves as the central hub for managing all available nockIs, presented in a visually grid-based tile system. Each tile provides key information about the respective nockI, including its **name**, **avatar**, and **category**, allowing users to quickly identify and access their agents.

The Overview page also includes essential management features:

- **Create New nockI:** A streamlined process for adding new nockIs directly from the interface.
- **Delete nockI:** Easily remove agents no longer in use.
- **View nockI ID:** Display the unique identifier for each nockI, enabling integration with external systems or workflows.

Designed as a **White Label product**, the Playground offers extensive customization options to adapt its appearance to the user's branding needs. This includes support for both **dark mode** and **light mode**, ensuring optimal usability across different environments. Additionally, the Playground supports **language localization**, allowing users to switch seamlessly between English and German interfaces.

With its intuitive layout and versatile features, the Playground Overview ensures efficient management of nockIs while providing a flexible and customizable experience for users.



2.1 Custom AI agents (nockIs)

Users can deploy personalized AI agents, referred to as “nockIs.” Each nockI is defined by the following attributes:

- **Display Name:** A unique identifier for the agent.
- **Avatar Image:** A visual representation of the nockI.
- **Short Description:** A brief overview of the nockI’s purpose.

nockIs are organized into categories to reflect their intended use.

- **Custom Chatbot:** Tailored conversational agents.
- **Websearcher:** AI-driven search tools.
- **Data Processor:** Tools for analyzing and processing data.

Created nockIs are displayed on an overview page and can be further customized on a detail page that includes:

- A **chat window** for interacting with the nockI, supporting text prompts, image uploads, and file uploads.
- A **settings section** for no-code configuration of category-specific options.

Additional features:

- Rename or delete a nockI.
- Retrieve the nockI ID directly from the interface.

Note: User inputs, including chat logs, images, and files, are stored within the Playground environment and processed by the provider for analytical purposes.

Coming soon: Enhanced guidance, such as pre-configured nockI templates for specific use cases and example configurations, to streamline the setup process.



3 nocKI Categories

The nocware Playground currently supports three categories of nocKIs:

- **Custom Chatbot**
- **Websearcher**
- **Data Processor**

3.1 Custom Chatbot

The “Custom Chatbot” category offers a range of configuration options to tailor the agent’s behavior and functionality. This category is particularly suited for use cases such as:

- **Brainstorming sessions:** Generating creative ideas and exploring concepts.
- **Personal assistants:** Supporting work tasks, such as drafting emails, scheduling, or providing information.

Through a defined role, the chatbot can be optimized for specific purposes, ensuring it adapts perfectly to user needs.

Configuration Options

- **LLM Selection:** Users can choose from various large language models (LLMs)
- **System Prompt:** Users can view and modify the system prompt that defines the chatbot’s core behavior.
- **User Prompt Field**
 - An additional free-text field allows users to customize the chatbot by adding specific instructions. The user prompt is treated additively with the system prompt, enabling a high degree of personalization. Examples include:
 - “Respond as concisely as possible.”
 - “Reply in a specific language.”
 - “Act according to a particular role or persona.”
 - “Include follow-up questions in your responses.”



- **Temperature Setting:** A slider allows users to adjust the chatbot’s creativity by setting the temperature between 0 and 1.
- **Low Temperature (e.g., 0.2):** Produces more deterministic and focused responses, ideal for factual or technical tasks.
- **High Temperature (e.g., 0.8):** Generates more creative and varied responses, suitable for brainstorming or storytelling.

These settings provide users with individual tools to adapt the chatbot to their unique requirements, ensuring flexibility and support for the individual use cases.

3.2 Websearcher

noCKIs in the “Websearch” category are optimized to perform internet searches, acting as an innovative type of search engine. These noCKIs allow users to configure both the **System Prompt** and **User Prompt** for personalized behavior. Additionally, users can select from the previously described LLMs to power the Websearch functionality.

The settings section for this category is divided into two subcategories:

- **Web Researcher 1**
- **Web Researcher 2**

Web Researcher 1 settings

- **Temperature:** Adjustable to control the creativity of responses.
- **Result Count:** Number of search results, configurable between 1 and 5.
- **Search Tools:** Users can select the preferred search tool from Taviliy, DuckDuckGo, or Google.
- **Grading Score:** A slider (0 to 1) determines how closely the results align with the user query.

If the results fail to meet the Grading Score threshold, **Web Researcher 2** is activated.



Web Researcher 2 Functionality

Web Researcher 2 can be configured to use a different LLM, search tool, and result count. It processes the same query to provide an alternative output, either correcting or challenging the results of Web Researcher 1. This layered approach ensures improved search accuracy and quality, offering users a robust and dynamic web search experience.

Use Cases

The “Websearch” category offers a modern way to gather results from the web while simultaneously validating their quality. By leveraging the dynamic interaction between Web Researcher 1 and Web Researcher 2, users can efficiently search for information and have it reviewed in one streamlined process.

3.3 Data processor

The “Data Processor” category enables users to expand the knowledge base of their nockIs by uploading custom data files, such as text documents or PDFs. These files are organized into a dedicated **File Collection**, tied to the specific nockI being configured. By training the nockI on this data, users empower it to provide responses enriched with the uploaded information, making it highly specialized for their unique use cases.

The data processor nockI offers a range of advanced configuration options to fine-tune both its chatbot functionality and its data-driven capabilities. Upon creating a new nockI in this category, a default system prompt is assigned, which users can edit to better match their requirements. Additionally, a **User Prompt** field allows for further customization of the chatbot’s behavior, enabling instructions such as tone preferences, specific roles, or context-sensitive guidelines. The temperature setting provides further control, allowing users to balance between deterministic and creative responses.

Enhanced Data Interaction

The category introduces a **Document Grader**, a powerful feature that assesses whether user queries can be answered directly using the uploaded data. Users can enable or disable this feature via a simple checkbox. When activated, the Document Grader works alongside a **Question Rewriter**, which reformulates user questions to enhance the chances of retrieving relevant information from the data. The



Question Rewriter includes customizable options for selecting the LLM, defining a system prompt, setting the number of rewrites, and adjusting the temperature for varied outputs.

Responses generated by the nocKI are handled by an **Answer Generator**, which draws on the enriched data to produce precise and relevant answers. Both the LLM and system prompt for the Answer Generator can be tailored by the user to achieve optimal results. Additionally, a **Hallucination Check** can be activated to ensure the reliability of the responses. This feature uses a separate configurable model to verify whether the generated output aligns with the source data, reducing the risk of incorrect or fabricated information.

File Collection and Processing

As users upload files, the system creates a **Vector Database** that organizes the data for efficient querying and retrieval. During the upload process, users can define settings for **File Split Size** (the maximum number of characters per segment) and **Overlapping Size** (the amount of overlap between segments) to optimize data segmentation and ensure comprehensive indexing. This enables the nocKI to accurately reference the provided data when processing user queries.

Coming Soon

Looking ahead, the Data Processor category will introduce several new features to enhance usability and scalability:

- **User-Created Collections:** Users will be able to create and manage collections on a dedicated page. These collections can be named, stored, and assigned to specific nocKIs as needed, offering greater organizational flexibility.
- **Shared Collections:** Collections will no longer be limited to individual users but will become accessible across multiple associated accounts, allowing for collaborative use cases.
- **Integration with External Data Sources:** Using tools like Apache Airflow, the system will support logic-based integration with external data pipelines, enabling users to pull data from diverse sources into their nocKIs effortlessly.

The Data Processor category empowers users with unparalleled control and customization, transforming nocKIs into powerful tools capable of leveraging custom data for advanced applications.



4 Supported LLMs

The nocware Playground provides access to a wide range of cutting-edge large language models (LLMs) from leading providers, ensuring users have the flexibility to select the most suitable model for their specific needs. The following models are supported:

- **OpenAI**
 - GPT-4o
 - GPT-4o-mini
 - o1-pro

- **Anthropic:**
 - Claude 3.5 Sonnet
 - Claude 3.5 Haiku

- **Mistral:**
 - Mistral 7B
 - Mistral 7B Instruct
 - Mixtral 13B
 - Mistral Lite

- **Gemini:**
 - Gemini 1.5 Flash
 - Gemini 1.5 Pro
 - Gemini Lite

Each model is optimized for different use cases, ranging from conversational AI and brainstorming to data processing and technical support. By offering this diverse selection, the Playground ensures users can harness the strengths of state-of-the-art models to achieve their goals effectively.



5 API Keys and Token usage

The nocware Playground integrates seamlessly with large language models (LLMs) by leveraging API keys provided by the respective model providers. For each customer, token usage is systemically limited to ensure optimal performance and cost management.

Token usage for individual nockI messages and conversation sessions is displayed directly in the Playground's frontend, separated into:

- **Input Tokens:** Tokens used to process the user's input, per message and session
- **Output Tokens:** Tokens generated as part of the nockI's response, per message and session

Coming soon: Future enhancements will include a monthly token usage overview and a quota management system and automated capping to prevent customers from exceeding their allocated limits. Additionally, the platform plans to provide direct cost transparency by converting token usage into monetary values, as well as user notifications for high consumption patterns, ensuring customers maintain full control over their resource usage.

6 Pack mode (Multi-Agent-Workflows)

The "Pack Mode" feature introduces a unique way to leverage the strengths of multiple nockIs by organizing them into collaborative workflows within a single interface. This mode allows users to position several pre-configured nockIs side by side, creating a structured flow where the output from one nockI is seamlessly transferred to the next for further processing. Since AI agents excel when defined for specific roles, Pack Mode empowers users to build specialized multi-agent workflows tailored to their needs.

In a typical workflow, users can begin with a nockI designed to generate an initial result, such as summarizing text or extracting key data points. This output can then be refined or analyzed further by a second nockI, which has been specifically optimized to handle such data. Additional nockIs can continue the chain, each performing a distinct role, culminating in a final result presented to the user or forwarded for further use in external systems.



Enhanced Use Cases

The Pack Mode is ideal for advanced applications such as:

- **Data Processing Pipelines:** Transforming raw inputs into structured outputs through a series of specialized nockIs.
- **Text Enrichment and Analysis:** Generating detailed insights by combining multiple steps, such as summarization, sentiment analysis, and data visualization.
- **Decision Support Systems:** Refining and validating results through collaborative agent workflows before finalizing recommendations.

Integration with External Systems

While the current implementation requires users to manually transfer data between nockIs via copy-paste, the feature is designed with future scalability in mind. Planned updates will include automated workflows that eliminate the need for manual input. Additionally, Pack Mode will support data exchange with external systems such as APIs, databases, or third-party applications, enabling seamless integration into larger operational contexts.

User-Friendly Design

To ensure accessibility, Pack Mode features a clean and intuitive interface. Future updates will introduce visual tools such as drag-and-drop mechanisms and flow diagrams, making it easy for users to configure, visualize, and manage complex workflows without requiring advanced technical knowledge.

Coming Soon

The roadmap for Pack Mode includes the following enhancements:

1. **Automated Workflows:** Enabling nockIs to communicate directly, transferring data without manual intervention.
2. **Shared and Global Workflows:** Allowing workflows to be shared across user groups and integrated into wider organizational processes.
3. **Advanced Integration Options:** Incorporating tools like APIs and external data pipelines to expand the functionality of Pack Mode beyond the nocware Playground.



By empowering users to orchestrate collaborative workflows among noCKIs and external systems, Pack Mode represents a significant step forward in achieving seamless, multi-agent collaboration.

7 A/B Split Testing

The “A/B Split Tests” feature is designed to facilitate side-by-side comparisons of two noCKIs, enabling users to evaluate and refine their performance under identical conditions. By synchronizing the user input across both noCKIs, this feature provides a controlled environment to observe and analyze how different configurations or models respond to the same prompts.

This capability is particularly useful for comparing variations in chatbot behavior, such as:

- Testing the impact of different **LLMs** (e.g., GPT vs. Claude).
- Evaluating how changes to the **System Prompt** influence responses.
- Measuring the effect of adjustments to parameters like **Temperature** or **User Prompt**.

How It Works

In the A/B Split Test interface, two noCKIs are displayed side by side. When the user enters an input, the system automatically synchronizes the query, sending it to both agents simultaneously. Their responses are then displayed in parallel, allowing users to compare the outputs in real-time.

This functionality ensures that all variables other than the specific differences between the two noCKIs remain constant, creating a reliable framework for identifying which agent performs better for a given use case. Whether testing creative writing capabilities, factual accuracy, or role-based interactions, A/B Split Tests offer valuable insights into optimizing chatbot configurations.

Key Benefits

- **Data-Driven Decision-Making:** By observing responses side by side, users can make informed choices about the optimal settings, prompts, or models for their needs.
- **Performance Tuning:** The feature allows users to fine-tune noCKIs iteratively by testing and refining one parameter at a time.
- **Enhanced Workflow Optimization:** Results from A/B tests can inform broader workflows, ensuring the most effective agents are integrated into larger processes.



Future Enhancements

Planned updates for A/B Split Tests include:

1. **Performance Metrics:** Automatic evaluation of responses based on predefined criteria, such as response length, relevance, or sentiment.
2. **Multi-Agent Comparisons:** Expanding beyond two agents to enable multi-agent split tests for comprehensive evaluations.
3. **Integration with External Systems:** Facilitating export of test results to analytics platforms or collaboration tools for deeper analysis.

A/B Split Tests provide a structured and efficient way to experiment with different nockI configurations, making them an essential tool for users aiming to maximize the effectiveness of their AI agents.