ESS Assistant – Privacy, Security & Technical overview Document

Version	Date	Author(s)	Changes
1.0	12 sept 2025	Ioana Bordea	
1.1	23 oct 2025	Jur van der Beek	Update to Ch. 6 logging
1.2	05 May 2026	Jur van der Beek	Update to LLM Model
1.3	03 Aug 2026	Jur van der Beek	Update new content

Overview

ESS Assistant – Privacy, Security & Technical overview Document 1

General Information 2

High level components 3

Happy flow 3

Model Access and Permissions 4

Token-Based Authentication: 4

Endpoints constraints: 4

No Direct Database Access: 4

Logging 4

General Information

This document provides a comprehensive overview of the security considerations, mechanisms, and practices implemented for our AI assistant integrated into the Employee Self Scheduling (ESS) application. The Assistant enhances the ESS platform by providing intelligent guidance and support for employees.

Our AI integration leverages the Semantic Kernel development kit, enabling secure and efficient communication between our API and a Large Language Model (LLM). The current LLM powering is **OpenAI's gpt-4.1,** a model that has been evaluated and positively selected for its performance and capability in our environment.

We practice continuous improvement: our team actively monitors advancements in AI modeling and regularly evaluates emerging options for performance, security, and compliance. Any future model replacements or upgrades will follow a structured evaluation process to ensure that security, accuracy, and reliability standards are consistently met or exceeded.

The following sections will detail how the ESS Assistant interacts with user data, how the underlying AI technology is integrated and secured, as well as the protective measures, both technical and procedural, that safeguard sensitive information within our scheduling ecosystem.

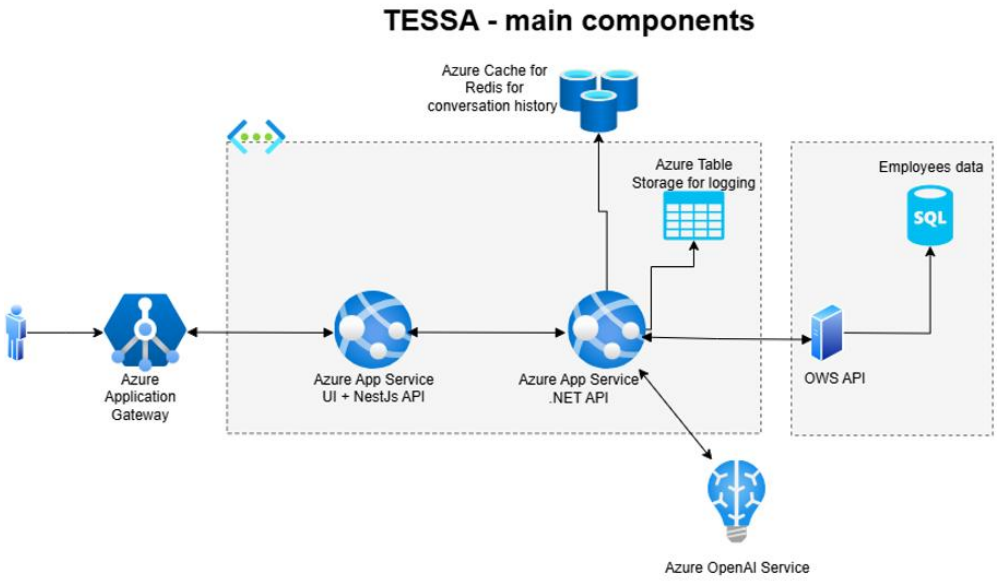
Intended purpose statement

The ESS Assistant is intended to make ESS tasks easier for employees by allowing them to ask scheduling and self-service questions in natural language.

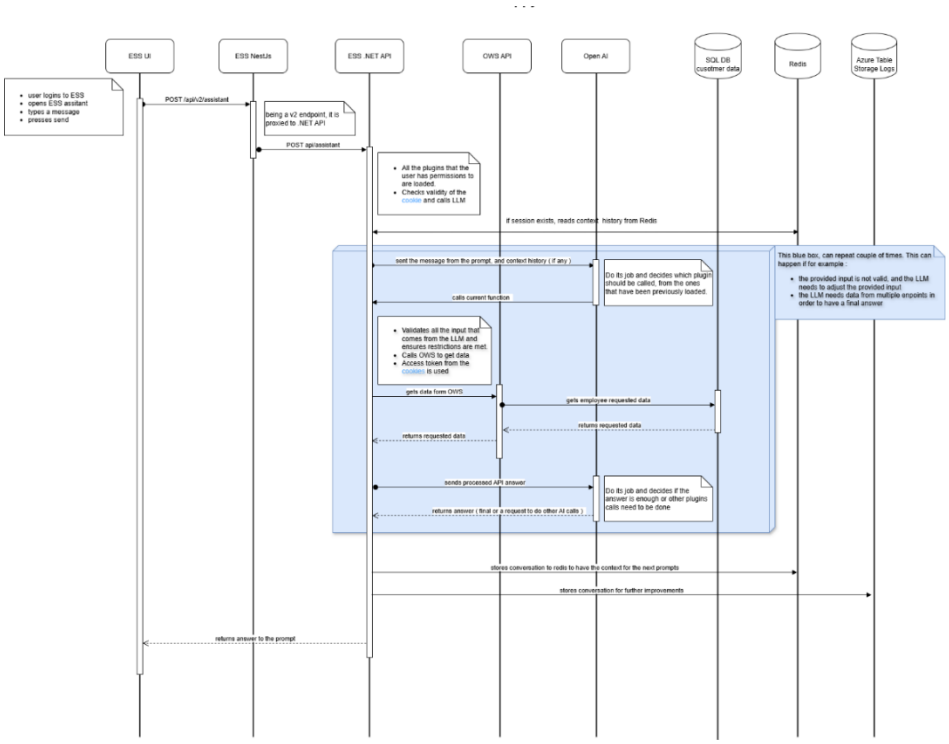
Details

- The assistant helps employees retrieve their own schedule information, leave requests, work preferences, public holidays, balances and supported ESS reports.
- Where configured, the assistant can help create, delete or adjust supported self-service records, such as leave requests, work preferences and realizations, within normal ESS permissions.
- The assistant responds in a conversational way and may support the employee in their preferred language.
- The assistant is limited by the available ESS permissions, supported plugins and customer configuration.
- The assistant should not be used for HR advice, employment decisions, performance assessments, disciplinary matters, health advice, legal advice, or manager/planner decision making.

High level components



Happy flow



Model Access and Permissions

This section outlines the model access and permissions framework that governs how the language model interacts with backend systems. It details the token-based authentication process, the constraints on accessible endpoints, and the strict data-handling rules that prevent direct database access.

Token-Based Authentication:

Access to backend endpoints requires a valid access token.

- The LLM does not store or know the access token.
- Token validation and permission checks are the same as for regular UI-driven calls.

Endpoints constraints:

There is a single assistant endpoint available in our API that returns data relevant to the userinteracting in the chat. Based on the incoming prompt, the LLM determines which of the allowed
OWS endpoints to call.
The LLM is not able to directly call the OWS API. This is only possible through so called plugins. Plugins available to the LLM:

- Are part of the supported features.
- Match the permissions granted to the authenticated employee. (Only the endpoints for which the user has permissions to are known for the LLM.)

No Direct Database Access:

Data remains within the customer database and is only retrieved via authenticated API calls.
The LLM cannot directly query the database.

Conversation Context and Caching

To maintain a smooth and coherent flow in user interactions, this system leverages Redis for caching both prompts and responses. By preserving the conversation context in Redis, each new interaction will include the relevant chat history for the current user, ensuring continuity.

Redis Cache Usage:

- Both prompts and results are stored in Redis to preserve the conversation context. On each interaction, the chat history relevant to the current user is supplied to the LLM.
- Context data is stored up to one hour after the last interaction on the chat. After expiration, the cached data is removed.

User-Specific Context:

- The LLM cannot access chat histories of other users.

Logging

Every user interaction is recorded into dedicated logs in a manner that prevents any alteration or deletion, providing a reliable trace for later analysis and enhancement. Importantly, the large language model (LLM) itself is restricted from accessing these logs in any form — it can write to them but cannot read from them — minimizing the risk of data leakage and preserving user confidentiality.

Write-Only Logging:

- All interactions are saved to logs in a write-only manner for future improvement and analysis.
- The LLM has **no read access** to logs.

The logs will be stored for 90 days. After this period, they will be automatically and permanently deleted and therefore will no longer be recoverable.

Microsoft agreement

Microsoft has contractually agreed not to use the data used as input in their OpenAI models. Information about how data provided by us to the Azure OpenAI service is processed, used, and stored can be found here: https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/openai/data-privacy?tabs=azure-portal

Role Allocation

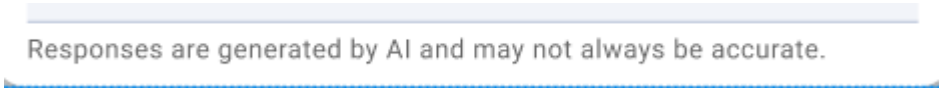
ORTEC designs and provides the ESS Assistant as part of ESS7, including the assistant endpoint, allowed plugins, product configuration, documentation and support route. The customer decides whether and how to make the assistant available to its employees, which employee groups receive access, and how local communication and support are organized. Microsoft / Azure OpenAI provides the underlying model and cloud service used by ORTEC. ORTEC remains responsible for how the assistant is integrated into ESS and how customer data is passed through the integration.

Risk Assessment

The assistant supports employees through a conversational interface for ESS tasks such as shift information, leave requests, work preferences, holidays, balances and realizations.
The assistant retrieves data through authenticated API calls and allowed plugins. It is not designed to directly query the database or execute actions outside the permitted ESS functionality.
The assistant should be described as supportive and informational. Its output is not a decision, approval, rejection, ranking, performance assessment, or staffing instruction.
Human users remain responsible for checking the result and using the regular ESS or organizational channels when needed.

Employee Notice

At the start of the conversation as well as always under the chat we show that the responses are AI Generated and may not be accurate. The responsibility of verifying the accuracy is the responsibility of the employees.



Human oversight, escalation and switch-off

Human oversight is provided through normal ESS workflows, the ability to ignore or correct assistant output, regular support channels, feedback controls, and technical configuration that can restrict or disable access.

Details

- Employees can choose not to use the assistant and can continue to use the standard ESS interface.
- Users can provide feedback through the assistant and through regular ORTEC/customer contacts.
- Operational issues should follow the normal ESS support process, with AI-specific incidents escalated to the ORTEC product owner and relevant technical owner.

• Access can be controlled through permissions such as the ESS Assistant permission and customer-specific rollout groups.
• If needed, the organisation can stop rollout or remove access while an issue is investigated.