Technical Fact Sheet
Gaia RCM V2 (IIA France)

Gaia RCM V2 is an AI agent based on large language models (LLMs) designed for IIA FRANCE members and partners. It is accessible via the Gaia platform.

Gaia RCM V2 primarily assists users in developing risk matrices, control frameworks, and audit programs. It relies on a curated repository of risks, controls, and audit tests developed by IIA FRANCE.

In accordance with Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13, 2024, laying down harmonized rules on artificial intelligence ("AI Act"), Gaia falls within the category of low-risk AI systems. Gaia does not fall within the categories of high-risk systems listed in Annex III, nor does it engage in any prohibited activities within the meaning of Article 5.

Under Article 50 of the Regulation, these systems must comply with transparency obligations towards users when they interact directly with AI, by clearly informing them that they are interacting with AI.

As the deployer, IIA FRANCE keeps this documentation up to date and available to the authorities upon request.

Providing this documentation to users ensures the system is accompanied by clear and appropriate information.

This technical sheet applies exclusively to the Gaia RCM V2 agent.

The following agents are covered in specific technical sheets:
- *Gaia*
- *Gaia Lex*
- *Gaia Observation*
- *Gaia RCM*
- *Gaia Writer*

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

1

Table of Contents

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

2

1 Definitions

Artificial intelligence system (AI system): within the meaning of Article 3.1 of the AI Act, an AI system is "a system based on AI techniques capable, for a given set of human-defined objectives, of generating results such as predictions, recommendations or decisions that influence the environments with which they interact".

Deployer: according to Article 3.4, any natural or legal person who uses an AI system under their authority, unless the system is used in the context of a strictly personal and non-professional activity.

Supplier: the entity that develops or places the AI model or system on the market.

General-purpose AI model (GPAI): as defined in Article 3.63, an AI model that can be used in a variety of applications for general purposes (e.g. GPT-4o). The specific obligations for GPAIs are detailed in Article 53 of the Regulation and concern the supplier.

RAG (Retrieval-Augmented Generation): an approach that combines a language model with a structured document database. Before generating a response, the system searches an internal corpus for the most relevant documents and then provides these elements to the model so that it can produce a contextualised response. This improves the reliability and traceability of responses (access to sources).

2 General Identification of the Gaia RCM V2 AI System

System name	Gaia RCM V2
Model provider	AzureOpenAI : OpenAI (gpt-4o, gpt-5.1, embeddings-3-large) on Microsoft Azure infrastructure (Europe)
Deployer	IIA FRANCE
Designated manager	Jean-Loup Grosse – IIA FRANCE Systems and Organisation Manager
Commissioning date	December 2025
Current version	Gaia 2.3.0 (December 2025) – Beta Version

3   Description of the Gaia RCM V2 Agent

Gaia RCM V2 is an AI agent accessible via the IIA FRANCE's Gaia platform. It assists users in preparing the work program and the risk/control/audit test matrix (RCM), based on their audit context.

- A conversational phase guides the user in defining their audit context.
    - At any point during the conversation, the user can validate the proposed audit context ("I validate the audit context").
    - The user is then prompted to initiate the generation of the RCM (by clicking the "Generate RCM" button).
- A reasoning model analyzes this audit context and queries a pre-established database of risks, controls, and tests using various strategies. It can also suggest additional elements based on its own capabilities.
- At the end of this phase (which takes 5 to 10 minutes), the following are presented for user review:
    - A list of audit objectives.
    - For each audit objective, a list of risks/controls/tests:
        - Either derived or inspired from the IIA FRANCE reference database and adapted to the audit context.
        - Or generated by the language model without reference to the IIA FRANCE database.
    - An executive summary.
- Users are encouraged to review all matrix elements, especially those generated by the model, before exporting to Excel.

4   Impact, Risks, and Mitigation Measures

Users concerned: IIA FRANCE members and partners with access to the Gaia RCM V2 agent.

4.1   Expected positive impacts

The deployment of the Gaia RCM V2 agent aims to generate several positive impacts, including:

- Upskilling users: Through structured, sourced, and contextualized responses, Gaia RCM V2 supports the understanding and adoption of professional best practices.
- Enhanced productivity: Gaia RCM V2 frees up time for analysis and human validation.

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

4

- Support for quality and compliance: The generation, validation, and source access features help strengthen methodological rigor and reduce oversights or inaccuracies.

4.2 Potential negative risks/impacts

Despite these benefits, several risks must be considered:

- Risk of factual or interpretative errors: some responses may be incomplete, approximate or out of context, particularly in the case of ambiguous wording or limitations inherent in AI models.
- Risk of uncritical acceptance: users may be tempted to consider responses as accurate without validation, which could lead to erroneous decisions or the dissemination of incorrect information.
- Risk of poor wording or omission of essential elements: particularly when generating audit documents (observations, matrices, supporting documents), certain key components may be missing or poorly structured.
- Risk of excessive dependence: intensive use of agents could reduce critical thinking or direct consultation of reference materials.
- Residual risk of bias or hallucination: as with any conversational model, the production of erroneous information remains possible, even with a high-performance RAG.
- Risk of inappropriate use: agents could be used for purposes that do not comply with their intended purpose (e.g. outside the intended professional or regulatory framework), or to produce inappropriate or unauthorised content.

4.3 Mitigation measures

To limit these risks, several technical, organizational, and methodological measures are implemented:

- Use of RAG: all responses are based on a validated body of documentation (IPPF, standards, IIA FRANCE documents), limiting hallucination and enhancing reliability.
- Disclaimers: A disclaimer is present at the beginning of each conversation, in the RCM generation interface, and included in every Excel export.
- Instructions: System prompts are designed to limit risks from the outset. They guide the behavior of Gaia RCM V2 by specializing it in audit and internal control topics, defining expected response types, and integrating professional best practices into their core instructions.
- Production Guidance: The generation of the audit context follows a structured process and requires final validation before generating the RCM.

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

5

- <u>User Training and Awareness:</u> Users have access to integrated assistance for prompt formulation within Gaia. Training sessions on Gaia are offered by IIA FRANCE.
- <u>Human Review and Supervision:</u> Users are encouraged to assess the various elements of the matrix, especially those generated by the model, before exporting to Excel. A mechanism for detecting inappropriate content (in both prompts and generated responses) is implemented.

5   Privacy and data protection

5.1   Non-training of models

The models used are hosted in Europe on Azure infrastructure (AzureOpenAI). Prompts (questions) and completions (answers) are not used to improve or train OpenAI or Microsoft models[1].

5.2   Hosting

Gaia's infrastructure is hosted on a dedicated Azure environment for IIA FRANCE in Europe (Azure West Europe): data, servers, and LLM inference services.

5.3   Data Processing and storage

Gaia RCM V2 (beta) does not store either the questions or the complete responses[2].

The full matrix is not stored, nor is the executive summary. The conversation (generation of the audit context) and the RCM are stored locally in the user's web browser.

The user can delete the conversation from the local browser storage at any time using the "New Conversation" and "Delete All Locally Saved Conversations" buttons.

The following anonymized data is retained for platform improvement:
- A brief intent deduced from a few words related to the audit objective (for example, "procurement and inventory audit").
- The macro process and the industry involved (for example, "Deliver Physical Products" and "Manufacturing").
- A condensed and sanitized version of risks/controls/tests not sourced from the IIA FRANCE database (purple cells in the RCM).
- Any feedback and evaluations (see below).

---

[1] https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/openai/data-privacy
[2] Except for inappropriate content, see "Inappropriate Content Detection and Prevention."

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

6

5.4 Personal and confidential data

No personal data (name, email, etc.) is stored in clear text in Gaia RCM V2.

Only a pseudonymized identifier, generated from the user's email address using a hashing function combined with encryption, is associated with each request. This identifier is used solely for aggregated usage statistics (typically the number of unique users on the platform).

In accordance with Article 4 of the GDPR, this identifier constitutes pseudonymized personal data. It does not, in its current form and for third parties, allow direct identification of an individual without additional information.

In line with the previous section, "Data Processing and Storage," any personal or confidential data transmitted by the user during usage (via prompts) is not stored in Gaia[3].

5.5 Inappropriate content detection and prevention

A request and response analysis system detects inappropriate content and prevents its generation using classification models. These models cover 4 categories defined by Azure: hate, sexual content, violence, and self-harm.

When triggered, the response generation is blocked, and Gaia RCM V2 displays an error. In cases of repeated abuse detection, IIA FRANCE may suspend or revoke access to Gaia RCM V2 for the concerned identifier.

Additionally, if inappropriate content is detected by the classification model, Microsoft may conduct further review:
- Initially, through automated processes using a more advanced model (LLM).
- If necessary, by a human operator if the automated review is deemed insufficient by the model.

The automated review adheres to the principles outlined in sections 5.1 (no model training), 5.3 (no storage of prompt/response data), and 5.4 (personal data). In cases of escalation for human review, the prompt and response are stored by Microsoft (for a maximum of 30 days) in the European region. Microsoft personnel authorized to analyze inappropriate content are located within the European Economic Area (EEA).

5.6 Reporting responses

Gaia RCM V2 includes a feature for users to evaluate the quality of RCM elements using *"Like"/"Dislike"/"Needs Review"* buttons placed under each cell in the matrix.

---

[3] Except for inappropriate content, see "Inappropriate Content Detection and Prevention."

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

Additionally, at the time of Excel export, the user is invited to rate their experience: a score between 1 and 5 and an optional comment.

This information is stored on the platform during the Excel export for statistical analysis and to support future developments of Gaia RCM V2. It is not retained if the RCM matrix is not exported to Excel.

The guided conversational phase for generating the audit context also offers a *"Like"/"Dislike"* reporting feature.

When a user clicks on one of these buttons, the ongoing conversation is automatically sent in clear text via email to the designated IIA FRANCE contact for human review. No conversation data is stored in Gaia's infrastructure when using this feature; only the email sent contains the conversation.

5.7 Usage metrics and logs

The following anonymized metrics are collected and, where applicable, regularly analyzed by the designated IIA FRANCE representative:
- Daily usage count of Gaia RCMV2
- Number of unique users
- Number of accesses to source documents
- Number of accesses to IIA FRANCE training materials
- Number of response reports (or flagged responses)
- Number and typology of encountered errors
- Detection of inappropriate content
- Average rating
- Anonymized comments
- Performance of the RAG system and the reasoning LLM model

To enable these analyses, the following events are recorded, potentially associated with a non-reversible pseudonymized identifier:
- Use of the Gaia RCM V2 agent, associated intent (summarized in a maximum of 5 words, e.g., "procurement and inventory audit"), and associated process and industry (closed lists, e.g., "Deliver Physical Products" and "Manufacturing").
- Consultation of a source document.
- Redirection to an IIA FRANCE training session.
- Use of the Like/Dislike/Needs Review features.
- Errors during response generation.
- A condensed and sanitized version of risks/controls/tests not sourced from the IIA FRANCE database (purple cells).
- List of risk/control/test elements from the IIA FRANCE database used to build the response.

Prompts, responses, audit context, and the RCM are never stored in these usage traces, and no personal data is retained in clear text.

6   Obligations Related to General-Purpose AI (Chapter V of Regulation (EU) 2024/1689)

The language/embedding models used by Gaia RCM V2 (gpt-4o, gpt-5.1, and embeddings-3-large) are general-purpose AI models (General Purpose AI models – GPAI) as defined in Article 3, point 63 of Regulation (EU) 2024/1689.

IIA FRANCE is not a GPAI provider under the regulation; it acts solely as a deployer, using models supplied by OpenAI and distributed through Microsoft Azure.

In accordance with Articles 52 to 55, the production and publication of the "model summary" (GPAI fact sheet) are the responsibility of the provider. These fact sheets will be appended to this document once officially published.

In the meantime, the publicly available documentation for each model is indicated below:

Model	Technical documentation	System Card
gpt-4o	https://platform.openai.com/docs/models/gpt-4o	https://openai.com/index/gpt-4o-system-card
gpt-5.1	https://platform.openai.com/docs/models/gpt-5.1	https://cdn.openai.com/pdf/4173ec8d-1229-47db-96de-06d87147e07e/5_1_system_card.pdf
embeddings-3-large	https://openai.com/index/new-embedding-models-and-api-updates	Not applicable (embedding model)

Version : v01	26/01/2026	Rédacteur : Jean Loup Grosse – Responsable Systèmes et Organisation IIA FRANCE

9