Informatica Intelligent Data Management Cloud CLAIRE®

Security & Privacy Overview

Informatica

Table of Contents

Metadata and Data Used for CLAIRE Training

Metadata and Data Stored in Intelligent Data Management Cloud™

Informatica Intelligent Data Management Cloud (IDMC) stores different types of information:

Metadata

- Operational and Usage Metadata: Includes information extracted from service and activity logs, such as connection and schedule data. It also includes information about how the cloud services are used, such as transactions conducted by the customer in its data marketplace.
- Technical Metadata: Includes data schemas, rules and data profile statistics, and design metadata that define integration tasks and processes, such as data sync, data replication, mappings and templates, taskflows, process definitions, and data lineage.
- Customer Business Metadata: Information related to customer data, including data classifications and glossaries designated by the customer.

The collection of this metadata by IDMC is necessary to provide IDMC cloud services and can't be disabled.

Data

Informatica stores business data from your applications and endpoints accessed through IDMC.

Encryption for Data and Metadata at Rest

Any data or metadata persisted in the customer's IDMC multi-tenant data repository is encrypted using the AES encryption algorithm which uses a 256-bit key. The key is unique to the customer tenant. By default, the key is rotated once a year, but customers can configure it to be rotated every 90, 120 or 180 days.

Customer Managed Encryption Keys

IDMC has the ability to support Customer Managed Keys. While Informatica uses strong encryption practices, customers may want to utilize their own encryption keys to safeguard their data. With this feature, a customer can hold their encryption keys and maintain the authority to encrypt and decrypt their data within our cloud environment. This gives our customers confidence that customer data remains confidential and protected, even from Informatica as the service provider.

Organization-Specific vs. Customer-Agnostic CLAIRE Models

Informatica IDMC implements different AI models with different scopes to deliver CLAIRE functionality:

- Organization-specific models: CLAIRE models tailored to meet the unique needs of individual customers. These models are trained using data and metadata specific to

each customer, allowing Informatica to provide services customized to that particular organization.

- Customer-agnostic models: CLAIRE models that serve all customers, including CLAIRE GPT. They are trained on Informatica product documentation, public data sets, and aggregated anonymized metadata.

Assets Used to Train CLAIRE models

CLAIRE customer-agnostic models are trained on metadata, which includes technical metadata, operational and usage metadata, and customer business metadata.

This metadata is fully anonymized and aggregated before being used for training, ensuring that no sensitive information is exposed during the process.

Informatica never accesses or uses customer data stored in IDMC for training CLAIRE customer-agnostic models. This approach ensures that the customer's data remains private and secure.

Informatica does not access customer data endpoints (such as Salesforce, SAP, or Snowflake) or other cloud applications, databases, and data warehouses, for CLAIRE or large language model (LLM) training.

The only exception is fine-tuning CLAIRE MDM match and merge tenant specific models, which is explicitly performed by the customer. This data is not used for cross-training CLAIRE models and is only utilized by organization-specific CLAIRE models.

Informatica does not reuse data from prompt conversations that occur within CLAIRE GPT or any other Informatica IDMC service that uses CLAIRE GPT. Informatica does not use specific customer prompt commands for cross-training CLAIRE models, only for organization-specific models.

Finally, other sources of public data are used to train Informatica customer-agnostic large language models, such as:

- General world knowledge: wiki data, books, public articles
- Verticalized domain knowledge: industry terms, industry metrics, industry systems
- Informatica documentation and Knowledge Base articles

The following table shows which assets are used to train CLAIRE:

Asset Type	Opt-out	Cross-Training	Org-Training
Technical metadata	Yes	Yes	Yes
Operational and usage metadata	Yes	Yes	Yes
Customer business metadata	Yes	Yes	Yes
Customer data stored on IDMC	N/A	No	Yes (User explicit)
Customer data on applications	N/A	No	No

Prompt commands on Claire AI	Yes	No	Yes (Feedback)
General world knowledge data	N/A	Yes	N/A
Verticalized domain knowledge data	N/A	Yes	N/A
Informatica knowledge base data	N/A	Yes	N/A

CLAIRE Model Fine Tuning

Selected CLAIRE models can be refined by users using their own data. Neither the data nor the refined models are shared with other tenants. They are stored within the context of the encrypted tenant available only to the specific customer.

Localization and Residency

Residency Observation

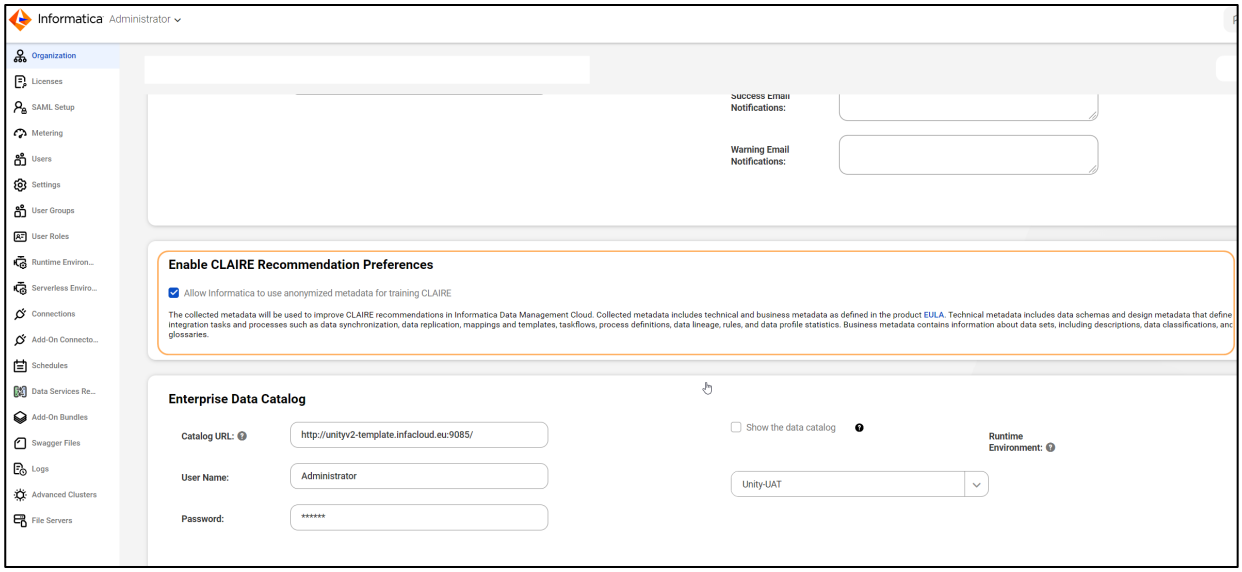
At Informatica, we recognize and prioritize the critical significance of complying with data residency regulations. In pursuit of this imperative, Informatica orchestrates multiple locales for LLM model training, ensuring that neither metadata nor data traverses beyond the delineated boundaries.

To adhere to regulatory frameworks, Informatica deploys distinct instances of CLAIRE models across designated geographies, encompassing the Americas, European Union, United Kingdom, APAC, and beyond. This approach underscores our commitment to safeguarding data integrity and complying with regional compliance standards.

Opt-in, Opt out, and Disposition of Data

Opting Out of Metadata Sharing

Organizations can choose to opt out of providing metadata for CLAIRE's training at any time through Administrator. This allows for complete control over the use of customers' metadata at any point in time.



Note that if you opt out, you might lose some of the functionality provided by CLAIRE.

Disposition of Data

Informatica's policy is to retain Processed Customer Data and Customer-Specific Metadata for at least thirty (30) days after termination or expiration of Customer's subscription to the Cloud Service, and to delete Processed Customer Data and de-identify or delete Customer-specific Metadata within sixty (60) days of termination or expiration of Customer's subscription to the Cloud Service, solely except as otherwise provided herein or to the extent such Metadata are included in backup and disaster recovery logs the integrity of which requires that they remain unmodified.

Informatica will promptly comply to the extent practicable with written requests to destroy Processed Customer Data within shorter time periods than those indicated above and provide written certification of destruction of Processed Customer Data upon Customer's written request.

For more details on Informatica data retention and data destruction policies, please visit Informatica Security Addendum.

Development and Design Principles

Ethical and Responsible AI

Many Informatica products and services feature technology that enables our users to process information with an increasing degree of autonomy. At Informatica, we understand the profound impact of artificial intelligence (AI) that makes this automation possible, and we guide our AI development with an ethical, responsible, and comprehensive set of principles.

These principles are designed to ensure that the AI technologies we create and deploy are developed and used in a way that respect human rights, contribute to societal benefits, uphold privacy and security, prioritize transparency and explainability, and strive for inclusivity and diversity.

We aim to democratize AI, providing tools that are accessible to all users regardless of technical expertise. Our commitment also extends to not designing AI for deployment in ways that can potentially cause harm or undermine the values that we stand for.

CLAIRE Principles

Informatica's main principles when designing and developing CLAIRE copilot and CLAIRE GPT applications and features are as follows:

- Focus on Enhancing Human Productivity in Data Management: We aim to develop AI technologies for data management - making it easy for data teams and business users to manage their data effectively. By narrowing our focus, we aspire to deliver impactful solutions while tailoring our technologies to the unique needs and challenges within this area.
- Ensure Data Security and Accountability: We pledge to create AI technologies that prioritize data privacy and security and balance them against functionality of our product features. AI development oversight will include third-party audits, robust feedback mechanisms, and a dedicated oversight team. We will maintain documentary evidence of how our AI was trained. This will ensure transparency in our processes and trust in our operations.
- Provide Transparency and Explainability: We aim to create AI models that are not just effective but also understandable. We will leverage advanced explainability frameworks and tools to provide insights into how our models make decisions, providing users where appropriate an understanding of how our AI application reached its conclusions.
- Design Delightful User Experiences: We aim to harness AI to augment human productivity by crafting thoughtfully designed user experiences that delight end-users.
- Democratize AI Responsibly: We commit to making AI accessible to a broad range of users while maintaining a strong focus on ethical and privacy considerations. We will balance openness with robust control mechanisms designed to prevent misuse of technology and protect data privacy.

Risk Mitigation

Informatica acknowledges the potential risks and ethical issues associated with certain AI applications. Accordingly, Informatica will not develop or deploy AI:

- In ways that are designed to cause or create an undue risk of harm to individuals or society.
- For purposes that are prohibited in the major jurisdictions in which we do business or otherwise constitute a clear threat to the safety, livelihoods and rights of people.

Informatica's commitment is to use AI to make the world a better place, not to harm or disadvantage any individual or group. We believe in using AI responsibly, and we are dedicated to following these principles as we develop and deploy our AI technologies. We understand that the field of AI is rapidly evolving, and thus, we will reassess and update these principles to keep pace with technological advancements and emerging ethical considerations. We firmly believe that by adhering to these principles, we can drive progress while ensuring the responsible use of AI.

Informatica IDMC Certifications and Compliance

Certifications and Compliance

The security of customer data is a critical objective of the IDMC platform. Informatica established a risk-based information security program protecting Informatica and its customers' data security and privacy.

Informatica has voluntarily undertaken, and/or is required by contractual obligation to perform in accordance with the below listed standards, which are measured through internal security teams and champions, third parties, and external assessments partners such as AICPA accredited external audit firms.

Among others, Informatica is SOC Type II, ESN, UK Cyber Essential Plus or FedRAMP certified.

For a complete list of certifications, assessments and standards for IDMC please visit Informatica Trust Center.

Worldwide Headquarters 2100 Seaport Blvd., Redwood City, CA 94063, USA Phone: 650.385.5000, Toll-free in the US: 1.800.653.3871 IN09_1217_3407
www.informatica.com linkedin.com/company/informatica twitter.com/Informatica