

IBM Technology



G-Cloud 15 IBM - Watsonx Orchestrate Service Definition Document

Email: ukcat@uk.ibm.com

Contact: Anne-Marie Wheeler

Service Description

IBM watsonx Orchestrate

This Service Description describes the Cloud Service. The applicable order documents provide pricing and additional details about Client's order. Upon acceptance of Client's order, this Service Description and the Cloud Services Agreement found at <http://www.ibm.com/terms/csa> (or an equivalent agreement for Cloud Services between the parties) apply to Client's use of the Cloud Services.

1. Cloud Service

IBM watsonx Orchestrate is a software-as-a-service Agentic AI platform that has multi-agent orchestration and provides Clients with access to a catalog of prebuilt agents and capabilities to discover and build agents, conversational AI assistants and automations.

1.1 Offerings

The Client may select from the following available offerings:

1.1.1 IBM watsonx Orchestrate Essentials (Classic Orchestrate, Non-Agentic)

IBM watsonx Orchestrate Essentials is a subscription Cloud Service offering that includes the following:

- Up to 1000 Resource Units per month;
- Access to Skill studio to build Skills and Skill Flows
- Assistant Builder access to design, build and use sophisticated conversational AI assistants
- IBM Robotic Process Automation as a Service Platform

The Assistant Builder provides a low-code, self-service experience to build conversational AI assistants. End users interact with AI assistants to initiate conversations using natural language and can optionally execute skills and automations. The number of unique active users per month (MAUs) interacting with the AI assistants consume Resource Units. Additional Resource Units are consumed if the users execute published Skills and automations from the watsonx Orchestrate Skills catalog. The conversion ratios are as follows:

- 10 unique Monthly Active User interactions: 1 Resource Unit (RU)
- 11 unique Voice Monthly Active User interactions: 1 Resource Unit
- 100 Skill Runs: 1 Resource Unit (RU)

The IBM Robotic Process Automation (RPA) as a Service Platform

- 1 x IBM Robotic Process Automation as a Service Environment
- Unlimited use of IBM Robotic Process Automation as a Service Studio
- Unlimited number of IBM Robotic Process Automation as a Service Unattended Bot Agents

Use of IBM RPA with the Cloud Service is restricted to building bots as watsonx Orchestrate Skills.

1.1.2 IBM watsonx Orchestrate Standard (Classic Orchestrate, Non-Agentic)

IBM watsonx Orchestrate Standard is a subscription Cloud Service offering that includes everything in watsonx Orchestrate Essentials (excluding the included Resource Units per month) and, in addition, includes the following:

- Up to 8500 Resource Units per month;
- Access to Skill studio to build and use automations for workflows, decision services and document processing workloads.

The Skill studio provides users with a low code experience to build automations using workflows, AI decisions and document processing. These automations are stored as Projects in the Skill studio and can be enhanced and published as Skills in the Skills catalog for end users to consume. Resource Units are consumed when users execute the Skills from the Skills catalog. When an end user invokes a Skill from the Skills catalog, in addition to the Skill Run, the number of workflows, decision executions and pages processed are also counted towards Resource Unit consumption. The conversion ratios are as follows:

- 100 workflow activity executions: 1 Resource Unit (RU)

- 100 pages processed: 1 Resource Unit (RU)
- 100 decision executions: 1 Resource Unit (RU)

1.1.3 IBM watsonx Orchestrate Agentic Essentials

IBM watsonx Orchestrate Agentic Essentials is a subscription Cloud Service offering that includes the following:

- Up to 600 Resource Units per month;
- Access to Agent Builder provides a low-code, self-service experience to build AI assistants and custom agents.
- 10 GB Documents uploaded with agents

Once AI agents or assistants are published, the number of unique active users per month (MAUs) interacting with the AI assistants and pages processed by Intelligent Document Processing (IDP), consume Resource Units.

The conversion ratios are as follows:

- 100 unique Monthly Active User interactions: 15 Resource Units (RU)
- 10 unique Voice Monthly Active User interactions: 1 Resource Unit (RU)
- 100 pages processed: 1 Resource Unit (RU)

1.1.4 IBM watsonx Orchestrate Agentic Standard

IBM watsonx Orchestrate Agentic Standard is a subscription Cloud Service offering that includes everything in watsonx Orchestrate Agentic Essentials (excluding the included Resource Units per month) and, in addition, includes the following:

- Up to 6000 Resource Units per month;
- Includes IBM pre-built domain agents

1.2 Additional Offerings

1.2.1 IBM watsonx Orchestrate Add-on Resource Unit

IBM watsonx Orchestrate Add-on Resource Unit is an optional subscription for Essentials and Standard editions, providing additional Resource Units per month in packs of 1000. Unused Resource Units do not roll over to the next month.

1.2.2 IBM watsonx Orchestrate Add On Documents

IBM watsonx Orchestrate Add Documents is an optional subscription for Agentic Essentials and Agentic Standard editions, providing additional Document entitlements in packs of 1,000.

1.2.3 IBM watsonx Orchestrate Add On Domain Agents

IBM watsonx Orchestrate Add Domain Agents is an optional subscription for Agentic Essentials, providing Clients with pre-built, ready to use domain agents.

1.2.4 Non-IBM Products Add On Domain Agents

Non-IBM Products Add on Domain Agents are optional subscriptions which provide Client with pre-built, ready to use domain agents from various third parties. The Non-IBM agents and tools included in this Cloud Service are Non-IBM Products and are identified as such in the catalog. By using agents and tools that are Non-IBM Products, Client accepts the applicable license terms, including any associated obligations and use restrictions specified in the catalog.

1.2.5 IBM watsonx Orchestrate Class Inference

IBM watsonx Orchestrate Class Inference allows Clients to make calls to a foundation model to generate a response ("Inference"), powered by IBM's watsonx platform and IBM's foundation models. This provides Clients the conversational search features used to create generative AI flows for tasks like Retrieval Augmented Generation (RAG) using Large Language Models (LLM). A collection of open-source and IBM foundational models are categorized into three classes: Class 1, Class 2, and Class 3. The LLM models in each class may be updated at IBM's discretion. Supported foundation models made available in this Cloud Service are selected by IBM from those listed at this URL:

<https://www.ibm.com/docs/en/watsonx/saas?topic=solutions-supported-foundation-models>.

Client will be charged in Resource Units for the number of Code Tokens consumed during Inference generation, at a rate of one Resource Unit for each 1,000 Code Tokens.

2. Data Processing and Protection Data Sheets

IBM's Data Processing Addendum at <http://ibm.com/dpa> (DPA) and the Data Processing and Protection Data Sheet(s) (referred to as data sheet(s) or DPA Exhibit(s)) in the links below applies to IBM's processing of personal data on behalf of Client.

IBM watsonx Orchestrate

<https://www.ibm.com/terms/?id=4A4E5A6099FC11ECAB2681C6B89D4420>

IBM Robotic Process Automation as a Service (applicable only to Classic Orchestrate, Non-Agentic)

<https://www.ibm.com/terms/?id=7D6B3530A02A11EA88F5BE8EBF8F323B>

3. Service Levels and Technical Support

3.1 Service Level Agreement

IBM provides the Client with the following availability service level agreement (SLA). IBM will apply the highest applicable compensation based on the cumulative availability of the Cloud Service as shown in the table below. The availability percentage is calculated as the total number of minutes in a contracted month, minus the total number of minutes of Service Down in the contracted month, divided by the total number of minutes in the contracted month. The Service Down definition, the claim process and how to contact IBM regarding service availability issues are in IBM's Cloud Service support handbook at https://www.ibm.com/software/support/saas_support_overview.html.

Availability	Credit (% of monthly subscription fee*)
Less than 99.9%	2%
Less than 99.0%	5%
Less than 95.0%	10%

* The subscription fee is the contracted price for the month which is subject to the claim.

3.2 Technical Support

Technical support for the Cloud Service, including support contact details, severity levels, support hours of availability, response times, and other support information and processes, is found by selecting the Cloud Service in the IBM support guide available at <https://www.ibm.com/support/home/pages/support-guide/>.

3.2.1 Optional Support Offerings

Additional support options are available for this Cloud Service. If Client selects to order an optional support offering, the terms defined in the Optional Support Offerings Service Description at <https://www.ibm.com/terms/?id=i126-9789> will apply to the support offering specified in Client's TD.

4. Charges

The charge metric(s) for the Cloud Service are specified in the Transaction Document.

The following charge metrics apply to this Cloud Service:

- Access is the right to access functionality of the Cloud Services. For this Cloud Service, Access is applicable only for the use of Domain Agents.
- Document is a finite volume of data that is enveloped within a document header and trailer record that marks its beginning and end or any electronic representation of a physical document processed by the Cloud Services. Documents are used for agent knowledge and are limited to 50 per agent with a maximum file size of 25MB. The maximum number of files uploaded in one batch is limited to 20 and the total file size of all uploaded files in one batch must not exceed 50MB. In addition, the size of the total number of documents that can be uploaded across all agents is limited to 10GB.
- Instance is each access to a specific configuration of the Cloud Services.

- Resource Unit is an independent measure of a resource managed by, processed by, or related to the use of the Cloud Service.

The Resources for the purpose of the Resource Unit calculation are the following cumulatively:

- Skill Run - A Skill Run is the execution of a Skill. A Skill is a capability that can be used to accomplish an element of work, using natural language processing and capabilities to discover and build conversational AI assistants and automations, and may be a single action (like adding a row to a table) or a multi-step action needed to complete a task (like finding contacts, creating a table of those contacts, and then sending an email to those contacts). Skills include those made available with the Cloud Service in a catalogue of prebuilt skills, and custom skills built by or for Client. If multiple Skills are dynamically sequenced to run in a Skill flow, then execution of each Skill within that Skill flow is counted as a separate Skill Run.
- Monthly Active User - Monthly Active Users are the number of Active Users in any calendar month. An Active User is a unique person who accesses the Cloud Service in any manner directly or indirectly (for example, through a multiplexing program, device or application server) through any means. Client's development personnel accessing the Cloud Service for the purposes of internal development and testing are not counted as Active Users. Monthly Active Users (MAU) consume messages when they interact with an AI assistant or agent. A message is a query sent to an AI assistant or an AI agent prompting a response or action processed by the Cloud Service. Each active user has a 50 message limit. Every time a user sends a prompt and receives a response, this counts as one complete message in the system. When a user exceeds 50 messages, the MAU count is incremented. MAUs in Agentic Plans include the Agent Tokens for LLM models embedded in watsonx Orchestrate, tools, tool flows and conversational search. When the number of Resource Units, using MAU to RU and pages to RU processed conversion ratios listed in Section 1.1.3, exceed the entitled limit, Client needs to purchase additional Resource Units.
- Voice Monthly Active User - Voice Monthly Active Users are the number of an Active Users in any calendar month that use a voice or SMS channel to interact with the Cloud Service user interface. For clarity, if a Monthly Active User is also a Voice Monthly Active User, that Active User is counted under both categories.
- Workflows, decision executions and pages processed applicable for Classic Orchestrate Non-Agentic Standard Plan as described in Section 1.1.2.
- Code Token - A Code Token is a collection of characters that has semantic meaning for a model. Tokenization is the process of converting the words in a task into code tokens.

5. Additional Terms

For Cloud Service Agreements (or equivalent base cloud agreements) executed prior to January 1, 2019, the terms available at <https://www.ibm.com/acs> apply.

5.1 Third Party Applications

IBM watsonx Orchestrate capabilities involve connecting to third party applications through Skills. The Client will therefore need to accept any applicable terms when signing up for those third-party applications and provide appropriate connection information to enable the Cloud Service to access the Client's content in such applications. Certain end point applications may be subject to usage limitations imposed by the application provider. This may result in a lower number of permitted calls to an application than the relevant Cloud Service monthly limit.

5.2 Limitations

In order to protect the Cloud Service from issues arising when a Client generates too much traffic, the amount of data transfers that a Client can make in a month is limited to 10 GB.

5.3 AI Technology Terms

a. Foundation and Embedding Models

The Cloud Service may include generative AI foundation models and/or embedding models as described in the documentation for the Cloud Service. Each foundation model and embedding model included in the Cloud Service will contain a model information document that provides information about that model ("Model Information"), including details on the model provider and

applicable license terms. The available IBM-developed foundation models and embedding models are considered part of the IBM Cloud Service. Client may not use the IBM-developed models separately from the IBM watsonx.ai platform. Some models included in this Cloud Service are Non-IBM Products, as identified in the Model Information. By using a model that is a Non-IBM Product, Client accepts the applicable license terms, including any associated obligations and use restrictions specified in the Model Information.

Client is prohibited from extracting the Model Weights. "Model Weights" means the numerical values associated with the parameters of the training, optimization and performance of a model.

All Client inputs to any model, including by inferencing or prompting the model, and any content provided by Client for further training of a model, including fine-tuning or prompt-tuning, are Client Content. Except as set forth in (c) (Output Indemnity) below, Client is solely responsible for Outputs generated by any model. IBM does not claim to have ownership rights to the Output. Unless separately authorized by Client through a voluntary feedback submission, IBM will not use the Client Content or model Outputs to train any foundation models or embedding models.

b. **AI Use Restrictions**

Client agrees not to, and not to direct or allow third parties, use foundation models in connection with the use of this Cloud Service: (i) for mass surveillance, racial profiling, or any use that violates or encourages the violation of basic human rights or other applicable laws and regulations; (ii) to distribute false, misleading, disparaging or obscene information or content; (iii) to provide fully automated decision making in connection with use cases involving critical processes or the risk of loss of life, property or impact on an individual's legal rights; (iv) in a manner that impersonates another for deceptive purposes or conceals the fact a user is interacting with AI; or (v) to distribute or intentionally generate malware or other harmful code.

This Cloud Service is not intended for use in a high risk context as defined under the EU AI Act or other applicable regulations. If you chose to use it in such, please contact ChiefPrivacyOffice@ca.ibm.com.

c. **Output IP Indemnity**

IBM's obligation to protect Client against third party claims as set forth in the applicable base agreement includes third party claims that the result generated by the Covered Model in response to a Client's prompt or other input (collectively, "Output") infringes that third party's patent or copyright provided that (in addition to the exclusions and Client obligations provided in the relevant base agreement):

- (1) Client has a paid subscription to the Cloud Service or the Covered Model at the time Client generates the Output.
- (2) Client enables all available filters and logging features recommended by IBM in the product documentation or delivered with the Covered Model or in the Cloud Service, and Client supplies information collected by such filters and features as requested by IBM in connection with a third party claim;
- (3) Client was using a supported version of the Covered Model;
- (4) the Output at issue was generated in connection with an Indemnified Use Case;
- (5) Client takes mitigating actions reasonably requested by IBM, which may include ceasing use of the allegedly infringing Output; and
- (6) Client had no reasonable basis to know that the Content or Output was infringing or likely to cause an infringement.

"Covered Model" means: an IBM developed foundation model provided with the Cloud Service ("IBM Covered Model"); and a third-party foundation model provided with the Cloud Service and identified in the product documentation as being an indemnified third party model ("Third Party Covered Model")

"Indemnified Use Cases" means use within Client's enterprise (a) of an IBM Covered Model in a manner consistent with such model's intended use as described in its published Model Information or other Transaction Documents (such as license information or program specifications), and (b) of a Third Party Covered Model to analyze, summarize or classify Content or extract information or insights from Content provided that Client has accepted and complied with the applicable third party license.

All other terms of '(a) Foundation and Embedding Models' and of the Liability and Intellectual Property Protection terms in the base agreement remain in effect.

d. **Liability for Infringement Claims Against Third Party Covered Models**

IBM's entire liability associated with its obligation to defend Client against claims that a Third Party Covered Model or its Output infringes a third party patent or copyright and pay amounts related to such claim as set forth in the applicable base agreement is limited to the amounts paid by Client that are attributable to the applicable Third Party Covered Model in the 12 months preceding the claim. All other terms of the Liability and Intellectual Property Protection terms in the base agreement remain in effect.

5.4 Pre-Trained Content

Pre-Trained Content is a library of terms and phrases for specific domains provided with the Cloud Service used to train Watson Natural Language processing (NLP) models in this Cloud Service and is IBM Content.

If Client elects to use the Pre-Trained Content, then the following term will also apply:

IBM grants to Client a revocable (solely in the event of termination or expiration of this Agreement), nonexclusive, non-assignable, worldwide, paid-up license to use, execute, reproduce, perform and modify the terms and phrases provided with the Pre-Trained Content, for internal (to Client) use only, for the sole purpose of enhancing, extending or customizing the training of the IBM Cloud Service. Any Client enhancements or extensions of Pre-Trained Content will be treated as Client Content. Pre-Trained Content may not be further distributed or used separately from the IBM Cloud Service. IBM has no responsibility for the training data resulting from Client's use of and modification to the Pre-Trained Content and IBM warranties will not apply. IBM has no responsibility to maintain or provide support for the modified library of Pre-Trained Content containing Client's enhancements, extensions or customizations.

5.5 Non-IBM Products Agents and Tools in the Catalog

This Cloud Service includes access to a catalog which contains agents and tools from third parties that are Non-IBM Products, as identified in their relevant Catalog Information. By using agents and tools that are Non-IBM Products, Client accepts the applicable third party terms, including any associated obligations and use restrictions specified in the non-IBM Product catalog information.

5.6 Enabling Software

- a. IBM Robotic Process Automation is provided to Client under the following terms (applicable only to Classic Orchestrate, Non-Agentic plans):

Enabling Software	Applicable License Terms (if any)
IBM Robotic Process Automation as a Service Attended Bot Agent	https://ibm.biz/BdGvCW
IBM Robotic Process Automation as a Service Unattended Bot Agent	https://ibm.biz/BdGvQd
IBM Robotic Process Automation as a Service Studio	https://ibm.biz/BdGvQD
IBM Robotic Process Automation as a Service Vault	https://ibm.biz/BdGvQH
IBM Robotic Process Automation as a Service Launcher	https://ibm.biz/BdGvQr

- b. IBM watsonx Developer Edition is provided to Client under the following terms:

License to use watsonx Orchestrate Developer Edition is provided to users of watsonx Orchestrate (all plans). Watsonx Orchestrate Developer Edition is installable software used to create agents.

Enabling Software	Applicable License Terms
IBM watsonx Orchestrate Developer Edition	https://www.ibm.com/terms/?ref=L-YRMZ-PB6MHM

- c. IBM Voice Gateway is provided to Client under the following terms:

License to use the Voice Gateway package is provided to users of the Standard and Agentic Standard Plans of watsonx Orchestrate only for answering phone calls charged by Orchestrate Voice MAUs. Voice Gateway is installable software that orchestrates Speech to Text, watsonx Assistant, and Text to Speech to connect to Client's assistant in a telephony conversation.

Enabling Software	Applicable License Terms (if any)
IBM Voice Gateway	https://www.ibm.com/terms/?id=L-RPAL-CEY27U

5.7 Early Release Functionality

The Cloud Service may provide beta or experimental functionality within the service and will be identified as such in the IBM cloud user interface. Client may suggest that IBM enhance the beta or experimental functionality (Feedback). Client is under no obligation to provide Feedback and IBM is free to use all Feedback that Client provides. Beta or experimental functionality may not be at a level of performance or compatibility of generally available services, is not fully tested, and may not comply with the normal Cloud Service security practices. Such features or functions are provided to Client in accordance with the Preview Cloud Services terms as defined in the base Agreement and are not designed for any type of Content that contains personal data or for use in a production environment. Client agrees not to provide or input Content that contains personal data in a beta or experimental service. Client is responsible to remove content Client wishes to retain prior to expiration or termination of the beta or experimental functionality.

5.8 Additional Terms related to watsonx Assistant functionality in watsonx Orchestrate

5.8.1 Non-IBM Services

If Client uses the Third Party Cloud Integration (defined below) through this Cloud Service, the following terms apply and prevail over anything to the contrary in this Service Description, the base IBM services agreement under which this Cloud Service is provided (e.g. the IBM Cloud Services Agreement) or the IBM Cloud Service Description:

- a. The watsonx Assistant Enterprise edition of this Cloud Service enables the ability for Client to connect via weblink to, and transmit Client's Content to, a limited version of Segment, called Segment Limited Use for IBM (the "Non-IBM Service") through the Cloud Service's interface ("Third Party Cloud Integration").
- b. IBM does not provide and is not responsible for the Non-IBM Service or its availability or operation. The Third Party Cloud Integration is provided "AS IS", with no warranties or conditions of any kind, express or implied. Client's use of the Non-IBM Service is subject to the terms of a separate agreement with Segment, and Client is responsible for complying with the terms of such third party agreement and its usage of the Non-IBM Service in connection with this Cloud Service.
- c. The Segment Limited Use for IBM offering consists of Segment Connections with one or more Sources of watsonx Assistant Enterprise edition ("Watson Assistant Segment Source") and three categories of available Segment Destination, such as analytics, raw data and storage destinations, specifically for IBM watsonx Assistant Enterprise edition Clients. "Source" means a source, such as a Client's website or mobile app from which the Client instructs Segment to collect Client's Content and route it to a Destination. For purposes of this Non-IBM Service, the only Source in the Segment Limited Use for IBM offering is IBM watsonx Assistant Enterprise edition. "Destination" means a database or third party application to which Client instructs Segment to send Client's Content after being collected from a chosen Source.
- d. IBM will pay to Segment, on Client's behalf, for Client's consumption of Monthly Tracked Users, only for the watsonx Assistant Segment Source, and only for the Segment Limited Use for IBM offering. If Client wishes to add additional Sources, Client should contact Segment directly to subscribe to a Segment commercial offering. "Monthly Tracked Users" or "MTU"s mean the number of unique user IDs plus the number of unique anonymous IDs which accesses the Segment Limited Use for IBM offering in any manner directly or indirectly during the billing period.

5.8.2 Continuous Delivery

This Cloud Service operates under a continuous delivery model, which allows for updates without causing downtime to the Client.

5.8.3 High Availability

The Cloud Service offers in-region data redundancy enabling high availability protection. IBM provides automatic data replication for Client databases which contain training and/or custom model data at no additional cost. Replication is completed across in-region availability zones within IBM Cloud data centers.

5.8.4 Backup & Restore

Clients are responsible for backing up and restoring their own data, including training and/or custom model data as well as any Client generated custom models. For Client backup and restore instructions, please see the Cloud Service documentation.

5.8.5 Disaster Recovery

In-region Business Continuity is completed by leveraging the automatic replication across in-region availability zones within IBM Cloud data centers.

Clients are responsible for multi-region Disaster recovery. The responsibilities include backing up, restoring and synching of their own security policies, training and/or custom model data as well as any Client generated custom models. In addition, the Client is responsible for routing and/or load balancing across the regions. For Client backup and restore instructions, please see the Cloud Service documentation.

6. Overriding Terms

6.1 Mistral Commercial Models

The following takes precedence and prevails over any conflicting term in Section 5.3 (AI Technology Terms) of this Service Description and the Liability and Intellectual Property Protection terms of the governing base agreement.

The Cloud Service includes access to Mistral commercial models which are considered part of the Cloud Service and are a Third Party Covered Model.

Mistral retains ownership of the (i) the original, unmodified version of the Mistral model, and (ii) any customized and/or fine-tuned versions of such model, except for in the case of both (i) and (ii) the Retained Rights (as defined below).

“Retained Rights” means (i) Model Inputs, (ii) Generated Output, and (iii) any intellectual property rights developed by or on behalf of Client, including rights related to using or otherwise exploiting the Mistral models (as opposed to the Mistral models themselves).

“Model Inputs” are any data, content and software uploaded, exposed, made available, shared with, or input into a Mistral model (including for the avoidance of doubt a fine-tuned Mistral model) for any purpose whatsoever, including but not limited to, for fine-tuning, usage, training, improvement. Model Inputs are Client Content.

“Generated Output” means any text, images, and other outputs generated by a Mistral model (including a fine-tuned Mistral Model) as a result of Client’s use of the Mistral model (including a fine-tuned Mistral model).

Neither IBM nor Mistral claims to have a license to or ownership of the Model Inputs or Generated Output. Client has no obligation to provide IBM or Mistral with any fine-tuned Mistral model.

Unless separately authorized by Client, neither IBM nor Mistral will leverage, use (except to deliver the Cloud Service to the Client), or otherwise exploit the fine-tuned Mistral model.

