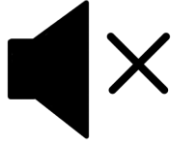


Taskflow for Orchestrating your Data Integration tasks

Prakash Jain
Informatica Global Customer Support

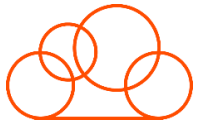


Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available to view on our **INFASupport YouTube channel** and **Success Portal**. The link will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

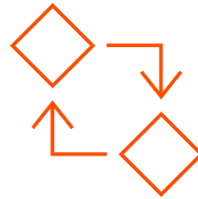
Feature Rich Success Portal



Bootstrap trial and
POC Customers



Enriched Customer
Onboarding
experience



Product Learning
Paths and Weekly
Expert Sessions



Informatica
Concierge with
Chatbot integrations



Tailored training and
content
recommendations

More Information



Success Portal

<https://success.informatica.com>



Communities & Support

<https://network.informatica.com>



Documentation

<https://docs.informatica.com>



University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informativa product or functionality described today remain at the sole discretion of Informativa and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Agenda

- Introduction
- Taskflow steps
- Taskflow templates
- Taskflow properties
- Parameters in taskflows
- Expression Editor
- Running a taskflow
- Demo
- Q&A

Introduction to Taskflows

Taskflow are used to control the execution sequence of a Data Integration task. You can run Data Integration tasks in parallel, use advance decision-making criteria, time tasks, and perform other advanced orchestrations.

Example 1: You want to run two mapping tasks in sequence. However, you do not want to run the second mapping task as soon as the first task ends if you get a warning on the first mapping task. Instead, you want to run the second mapping task to run after two hours. You can create a task flow to orchestrate this scenario.

Example 2: You may want to re-try your Data Integration task few times if they fail due to some network glitch or other issue. In this case, you can use taskflow for performing this retry attempts before notifying the end user if they still fail.

Prerequisites

- Advanced taskflow feature is enabled in the Org license.
- You cannot create a mapping task or a synchronization task during the taskflow creation. You must have these already available before creating a taskflow.

Taskflow Steps

Assignment

Assignment step is used to set a value for a field. A field is a data holder that carries data around a taskflow. You can use input fields and temporary fields to set a value for a field.

Input fields provide input when you run the taskflow. The taskflow uses temporary fields internally to handle data.

Data Task

Data Task step is used to add a mapping task, synchronization task, or PowerCenter task to a taskflow. You can configure how the taskflow handles errors and warnings, perform actions based on a schedule, and override runtime parameters.

Notification Task

Notification Task step is used to send an email notification to specified recipients.

You can configure the Notification Task step to send an email notification. For example, you can send an email notification to inform recipients about the number of success rows and error rows that were encountered in a Data Task step of a taskflow.

Taskflow Steps

Command Task

Use a Command Task step to run shell scripts or batch commands from a file on the Secure Agent machine. For example, you can use a command task to move a file, copy a file, zip or unzip a file, or run clean scripts or SQL scripts as part of a taskflow.

You can use the Command Task outputs like status, Exit code etc to orchestrate subsequent tasks in the taskflow.

File Watch Task

Use a File Watch Task step to listen to files in a defined location and monitor file events. In the File Watch Task step, you can select an existing file listener with the connector source type. You can use file events to orchestrate taskflow execution.

For example, you can wait for a file to arrive at a particular location and then consume the file in a subsequent step.

Ingestion Task

Use an Ingestion Task step to leverage a file ingestion task for taskflow orchestration. Select an existing file ingestion task.

You might want to perform data integration operations after moving files to an intermediate location and before transferring the files to the target. In this case, you can use the Ingestion Task step in conjunction with the Data Task step.

Taskflow Steps

Subtaskflow

Use a Subtaskflow step to embed and reuse an existing taskflow. You can configure input fields to provide input when you run the taskflow. You can also enable fault handling to determine the reason for a taskflow failure.

Decision

Use a Decision step when you want a taskflow to take different paths based on the value of a specific field.

Parallel Paths

Use a Parallel Paths step when you want a taskflow to run multiple items at the same time. For example, you can run three mapping tasks simultaneously. The taskflow runs all items in the Parallel Paths step and then moves to the next step.

Jump

Use a Jump step when you want to jump from one part of the taskflow to another.

End

Use an End step to define the HTTP status code that must be used when a taskflow completes.

Taskflow Steps

Wait

Use a Wait step when you want to pause taskflow execution for a specific duration.

Throw

Use a Throw step to catch a fault, return the fault details, and prevent the execution of the subsequent steps in a taskflow. The Throw step is an interrupting step, which means that if a fault occurs, the Throw step stops the execution of the taskflow and sets the taskflow status to failed.

Different steps in taskflow in action

Taskflow templates

Taskflow Basic

Use this template when you need a basic canvas with a Start step and an End step. You can add other steps at any point on the canvas on using any of the task templates.

Parallel Tasks

Use Parallel Tasks template if your major requirement is to run two or more data integration tasks in parallel. You start with a taskflow that contains a Start step, a Parallel Paths step, and an End step.

Parallel Tasks with Decision

Use Parallel Tasks with Decision template if your major requirement is to run two or more data integration tasks in parallel and then make a decision based on the outcome of any task. You start with a taskflow that contains a Start step, a Parallel Paths step, a Decision step, and an End step.

Sequential Tasks

Use Sequential Tasks template if your major requirement is to run two data integration tasks, one after the other. You start with a taskflow that contains a Start step, two Data Task steps, and an End step.

Sequential Tasks with Decision

Use Sequential Tasks with Decision template if your major requirement is to run two Data Integration tasks sequentially and then make a decision based on the output of either task. You start with a taskflow that contains a Start step, two Data Task steps, a Decision step, and an End step.

Single Task

Use Single Task template if your major requirement is to run one data integration task on a daily or weekly schedule, for example. You start with a taskflow that contains a Start step, a Data Task step, and an End step.

Taskflow properties

Taskflow binding

- The Binding property defines how a taskflow is initiated

REST/SOAP

- If you select the REST/SOAP binding type, you can run the taskflow by using a REST or SOAP endpoint. You can use the Allowed Groups and the Allowed Users fields to define the user groups and users who can run a published taskflow as an API.

Event

- If you select the Event binding type, the taskflow is invoked when the specified event occurs. For example, a taskflow can be invoked upon an event such as arrival of a file in a file system. The Event Source Name field is available where you can select the file listener that you created for the event.

Taskflow access:

If the taskflow uses the REST/SOAP binding type, you can define the user groups and users who can run a published taskflow as an API in the Allowed Groups and the Allowed Users fields.

Taskflow properties

Input fields

- Use the Input Fields to add fields that a taskflow uses at the beginning of a step.
- When you select the binding type for a taskflow as Event and select a file listener, Data Integration creates an input field to store details of the files that arrived, were updated, or were deleted as part of the file listener event.

Temporary fields

- Create temporary fields for use in a taskflow step. A taskflow uses temporary fields internally. Temporary fields do not appear in the input or output of a taskflow.

Advanced properties

You can configure a taskflow to suspend on a fault that occurs at the taskflow level. You can also configure the taskflow to send an email notification to specified recipients when it is suspended on a fault.

Demo on Taskflow properties

Parameters in Taskflows

Parameters in taskflows

You can use a taskflow to pass input and in-out parameters to a task. The mapping must be created with input parameters or in-out parameters.

Input parameters

- Input parameter is a placeholder for a value or values in a mapping. You define the value of the parameter when you configure the mapping task.

- You can use a taskflow to override the following subset of mapping input parameters:

--Source object. Changes the object that the mapping task reads from.

-Source connection. Changes the connection that the mapping task uses to read from the source.

-Target connection. Changes the connection that the mapping task uses to write to the target.

-Target object. Changes the object that the mapping task writes to.

Parameters in Taskflows

In-Out parameters

- In-out parameter is a placeholder for a value that you can pass in to or out of a mapping.
- Unlike input parameters, an in-out parameter can change each time a task runs.
- You can use a taskflow to override any type of in-out parameters that a mapping task supports.

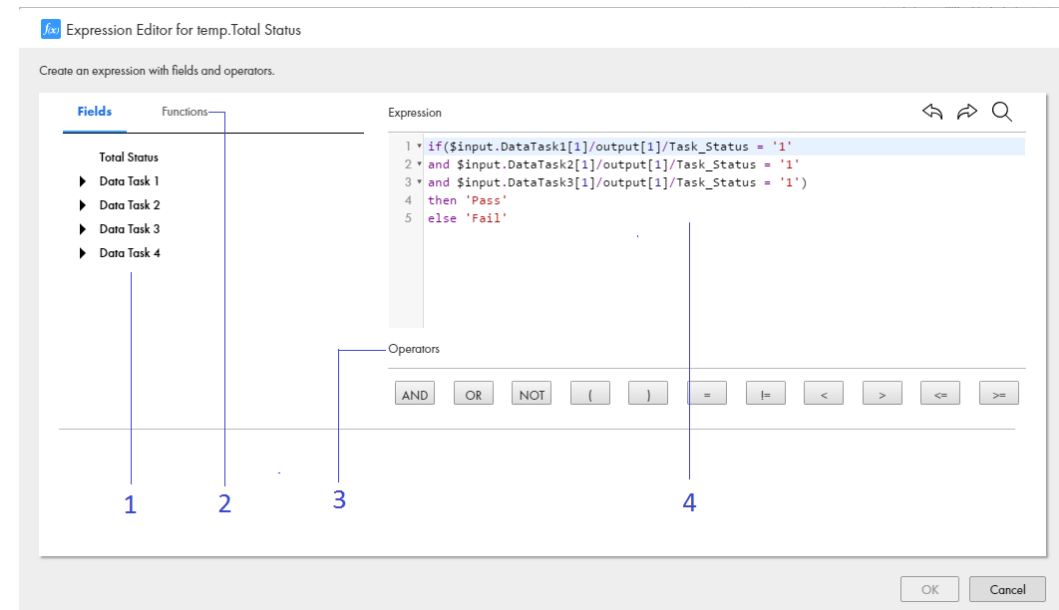
Expression Editor

The expression editor help you to create Xpath / Xquery expression to be performed on the field value. XQuery is for XML what SQL is for databases. i.e., to query element or attribute etc., construct XML and more. You may use fields, functions and operators to create expressions.

- To add a field, click the Fields tab, drill down to the field that you want to use, and click Add.
- To add an operator, click an operator in the **Operators** section. You can also manually enter an operator. For example, manually enter the If operator.
- To add a function, click the **Functions** tab, drill down to the function you want to use, and click **Add**.
- To add a comment, enter the comment in the Expression section with the following syntax:(:<comment>:).

Example:

(:This is a sample comment:).



Expression Editor

- XQuery version 3.0 to create expressions in the Expression Editor details are here <https://www.w3.org/TR/xquery-30/>
- You may use tools like BaseX or online XQuery runners(<http://videlibri.sourceforge.net/cgi-bin/xidelcgi>) to validate your XQuery code.
- Use validation panel to understand if there are any errors with your expression.
- You may use functions available in the expression editor panel to build your XQuery expression.

Running few sample XQuery expression using BaseX

Running a Taskflow

Taskflow Designer:

To run a taskflow from the taskflow designer, open the taskflow and click Run in the upper-right part of the page.

API

To run a taskflow as an API, you must first publish the taskflow as a service, and then run it. When you publish a taskflow, Data Integration generates the service URL and the SOAP service URL. You can use these endpoint URLs to run the taskflow as an API.

File listener event

You can invoke a taskflow through a connector file listener. Within the taskflow, define the binding type as Event and select the connector file listener as the event source. When you publish the taskflow, the taskflow subscribes to the connector file listener that is defined in it.

Schedule

- To run a taskflow on a schedule, create a schedule in Administrator and associate the taskflow with the schedule.

Running a Taskflow as API

- To run a taskflow as an API, you must first publish the taskflow as a service.
- When you publish a taskflow, Data Integration generates the service URL and the SOAP service URL.
- You can use these endpoint URLs to run the taskflow as an API.
- You can dynamically provide input parameters for the tasks that the taskflow contains and perform orchestration.
- You receive the taskflow run ID as the response.
- The service URL uses the following format:

<Informatica Intelligent Cloud Services URL>/active-bpel/rt/<API_name>

The SOAP service URL uses the following format:

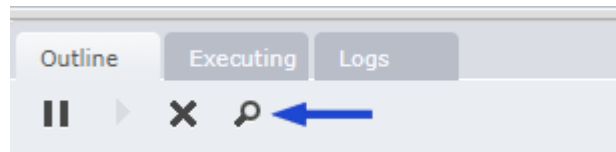
<Informatica Intelligent Cloud Services URL>/active-bpel/soap/<API_name>

- Use the run ID to monitor the taskflow run in the My Jobs page.

Designing Taskflow with re-try logic for a failed data task

Debugging Taskflow

- You may click on My Jobs to see all the running taskflow by a logged-in user run.
- Clicking on View Subtask display list of all Data Integration task invoked from Taskflow. You may view rows processed, StartTime, EndTime, Status etc. for each of the data integration task.
- You may go to Monitor app to see all the run of a taskflow in the organization. Upon opening a taskflow run, it would display a graphical representation taskflow with path it as taken and status of each data task step.
- One can click on “Advanced view” to view bpel view of the task with variable information. The taskflow process logs can be downloaded by clicking on glass icon in the outline tab.



Debugging Taskflow

- The variable section holds all the variable value used by the taskflow. Few variable are created for internal use.
- The logs tab in the advanced view shows status of the each bpel step with time stamp.
- This screen also shows any error while executing the taskflow.

References

Task flow online document:

- <https://network.informatica.com/onlinehelp/IICS/prod/CDI/en/index.htm#page/mm-cloud-taskflows/Taskflows.html>

Few knowledge-based articles on Taskflow:

- <https://knowledge.informatica.com/s/article/000097644>
- <https://knowledge.informatica.com/s/article/630573>
- <https://knowledge.informatica.com/s/article/000138268>
- https://knowledge.informatica.com/s/article/FAQ-Is-it-possible-to-fail-taskflow-when-any-of-the-Data-task-fails?language=en_US&type=external

Q & A



Thank You