

KNIME Integrated Deployment Guide

KNIME AG, Zurich, Switzerland
Version 5.5 (last updated on)



Table of Contents

Introduction.....	1
Installation.....	2
Create a workflow	3
Capture segments of a workflow	3
Configuration of the Capture Workflow End node.....	5
Combine captured workflow segments	8
Save and execute a captured workflow.....	10
Write a workflow object as production workflow	10
Deploy a production workflow to KNIME Server	13
Read a production workflow	16
Execute a production workflow deployed to KNIME Server.....	18
Execute a production workflow locally.....	19
Extract the workflow summary	21

Introduction

The traditional data science process starts with raw data and ends up to the creation of a model. That model is then usually moved into daily production. Integrated Deployment allows not only the model but all the associated preparation and post-process steps to be identified and reused in production in an automatic way, and without having to shift between different tools.

The Integrated Deployment nodes allow you to capture the segments of the workflow needed for running in a production environment, the model or library itself as well as the data preparation. These captured subsets are saved automatically as workflows with all the relevant settings and transformations and can be run at any time, both on KNIME Analytics Platform for model validation and on KNIME Server for model deployment.

In this guide we will explain how to use the nodes that are part of the KNIME Integrated Deployment Extension, as well as other useful nodes that allow the use of this feature in different environments.



You can access more material about Integrated Deployment on this [website section](#).

Installation



If you have installed KNIME Analytics Platform 4.5 or greater you can skip this part.

You can install KNIME Integrated Deployment Extension from:

- KNIME Hub: go to the [KNIME Integrated Deployment Extension page](#) on KNIME Hub. Here, drag and drop the squared yellow icon to the workbench of KNIME Analytics Platform.
- KNIME Analytics Platform: go to *File* → *Install KNIME Extensions...* in the toolbar and find KNIME Integrated Deployment under KNIME Labs Extensions or type Integrated Deployment in the search bar.

For more detailed instructions refer to the [Installing Extensions and Integrations section](#) on the KNIME Analytics Platform Installation Guide.

Now the nodes are available in the node repository under *KNIME Labs* → *Integrated Deployment*, as shown in [Figure 1](#).

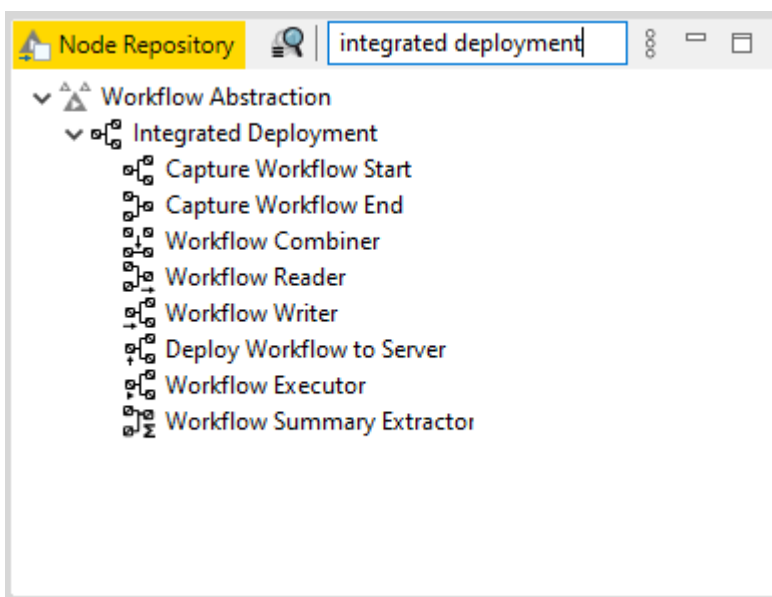


Figure 1. The Integrated Deployment nodes in the node repository

Create a workflow

In this section we will explain how to build a workflow, capture some segments, and finally combine them into a new workflow object. In this new workflow object you can also store the input and output data.



All the workflows shown in this section are available on [KNIME Hub](#).

Capture segments of a workflow

With the Integrated Deployment nodes you can capture the segments of the workflow that you need. To do this you can use [Capture Workflow Start](#) and [Capture Workflow End](#) nodes, as shown in [Figure 2](#).

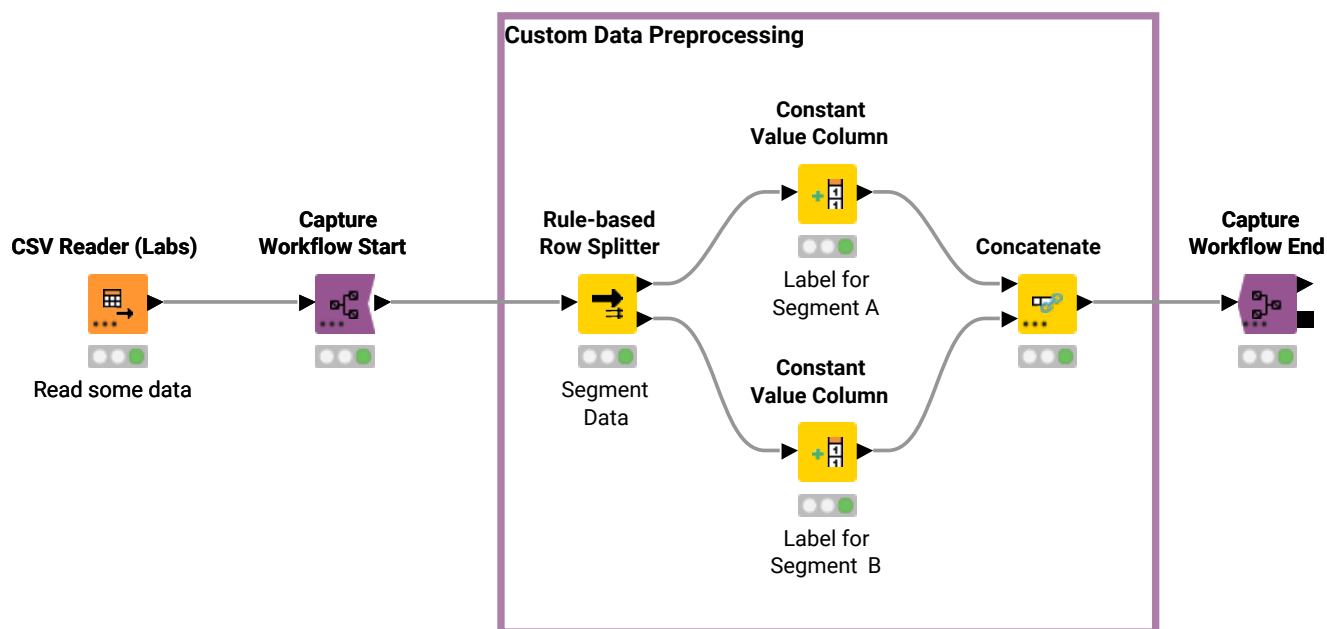


Figure 2. A segment of a workflow is captured

Capture Workflow Start node marks the start of a workflow segment to be captured. Capture Workflow End node marks the end of a workflow segment to be captured. The entire workflow segment within the scope of these two nodes is then available at the workflow output port of the Capture Workflow End node. Nodes that have out-going connections to a node that is part of the scope but are not part of the scope themselves are represented as static inputs but not captured.

In [Figure 3](#) an example of the corresponding captured workflow is shown. For an explanation on how to generate the captured workflow please refer to the [Store a captured workflow](#) section.

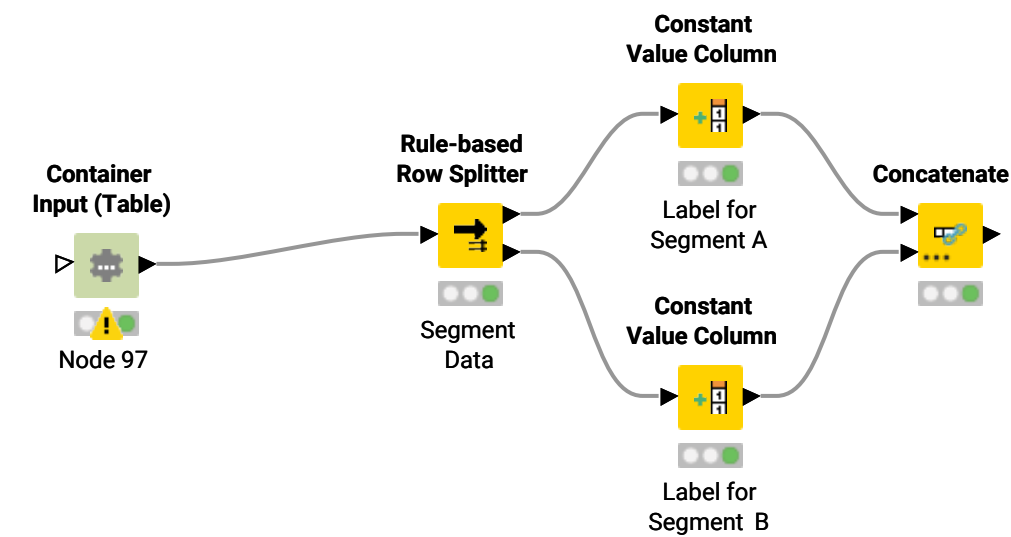


Figure 3. The captured workflow

Configuration of the Capture Workflow End node

Input data

You can store input data with your captured workflow segment. These data will be used by default when running the captured workflow segment unless you provide a different input. Storing input data will add data to the Container Input (Table) node in the captured workflow segment. In this way data are saved with the workflow when **writing it out** or **deploying it to KNIME Server** and are available when executing the workflow **locally** or **on KNIME Server**.

To store input tables, right-click the Capture Workflow End node and choose *Configure...* from the context menu. In the node configuration dialog that opens, shown in **Figure 4**, check *Store input tables* option. Here you can also choose the maximum number of rows to store with the captured workflow.

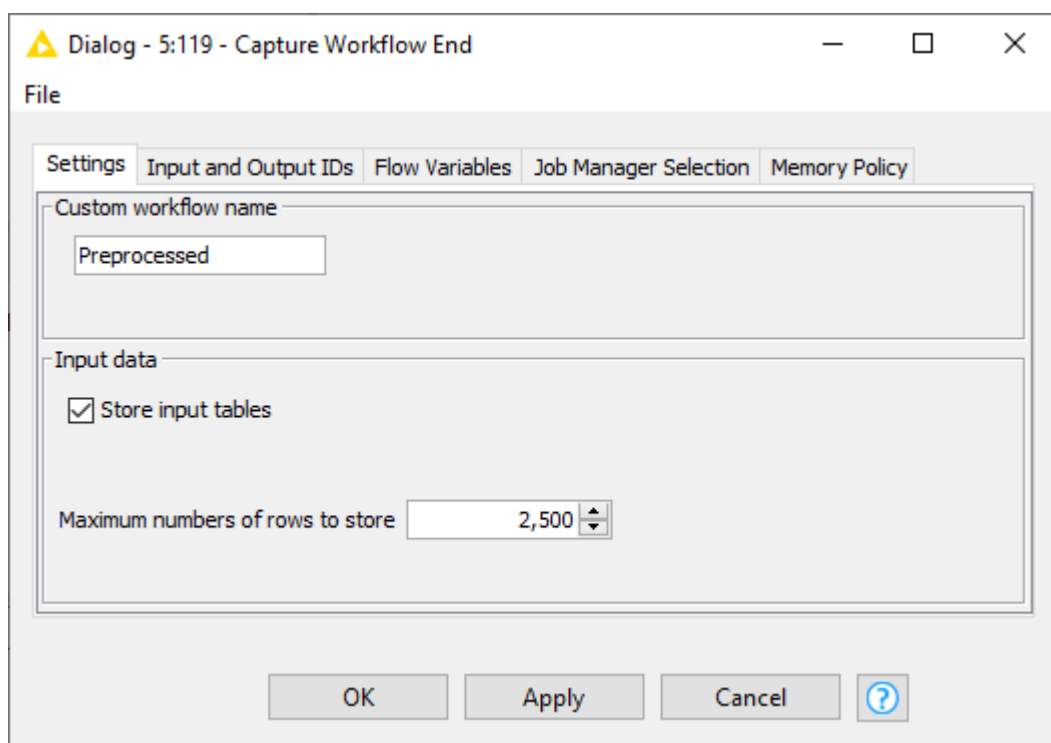


Figure 4. The Capture Workflow End node configuration dialog

Add or remove input or output ports in Capture Workflow nodes

In Capture Workflow Start and Capture Workflow End nodes you can choose to have how many input and output ports of different types as necessary, as shown in [Figure 5](#).

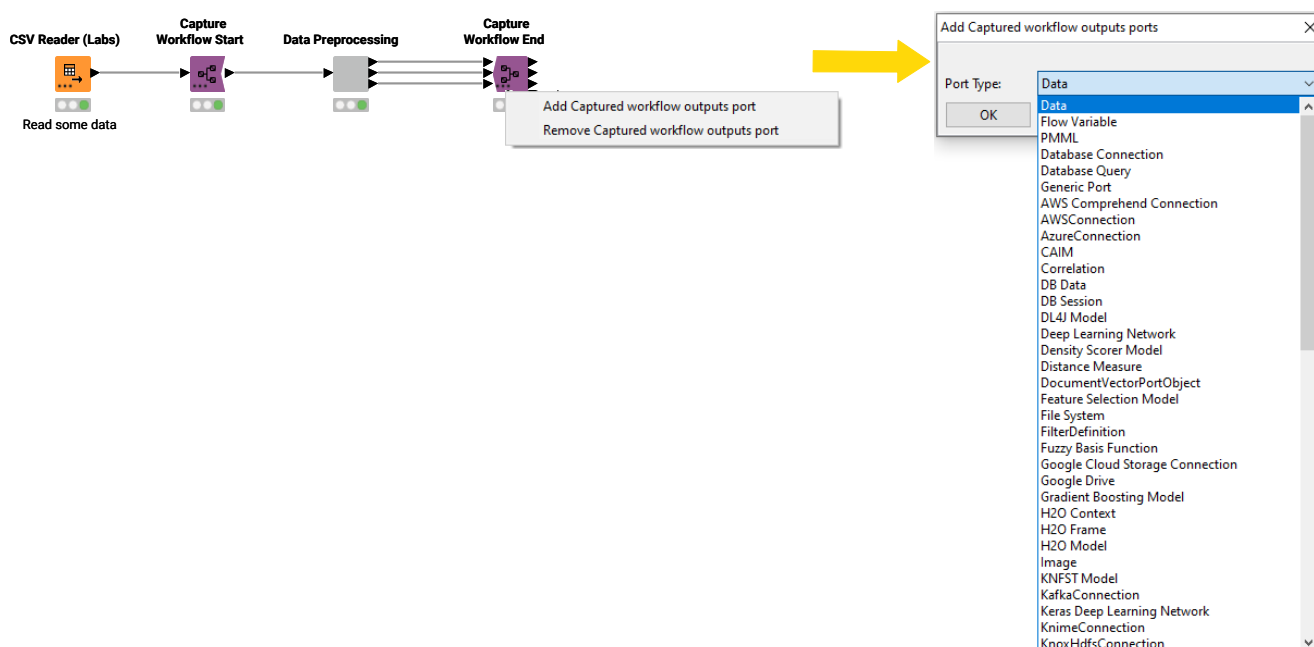


Figure 5. Add/remove input/output ports to and from Capture Workflow Start/End nodes

- To add input ports:
 1. Click the three dots in the left-bottom corner of the Capture Workflow Start/End node
 2. Choose *Add Captured workflow inputs port* from the context menu that opens
 3. Choose the port type from the drop-down menu in the window that opens
- To remove input ports:
 1. Click the three dots in the left-bottom corner of the Capture Workflow Start/End node
 2. Click *Remove Captured workflow inputs port*.

Rename captured workflow and input/output ports

In the Capture Workflow End node you can change the name of the captured segment of the workflow, as well as inputs and output ports. Right-click the Capture Workflow End node and choose *Configure...* from the context menu. In the node configuration dialog that opens, you can give the captured workflow segment a custom name in the *Settings* tab, as shown in [Figure 4](#). To rename input and output ports instead go to *Input and Output IDs* tab, as shown in [Figure 6](#).

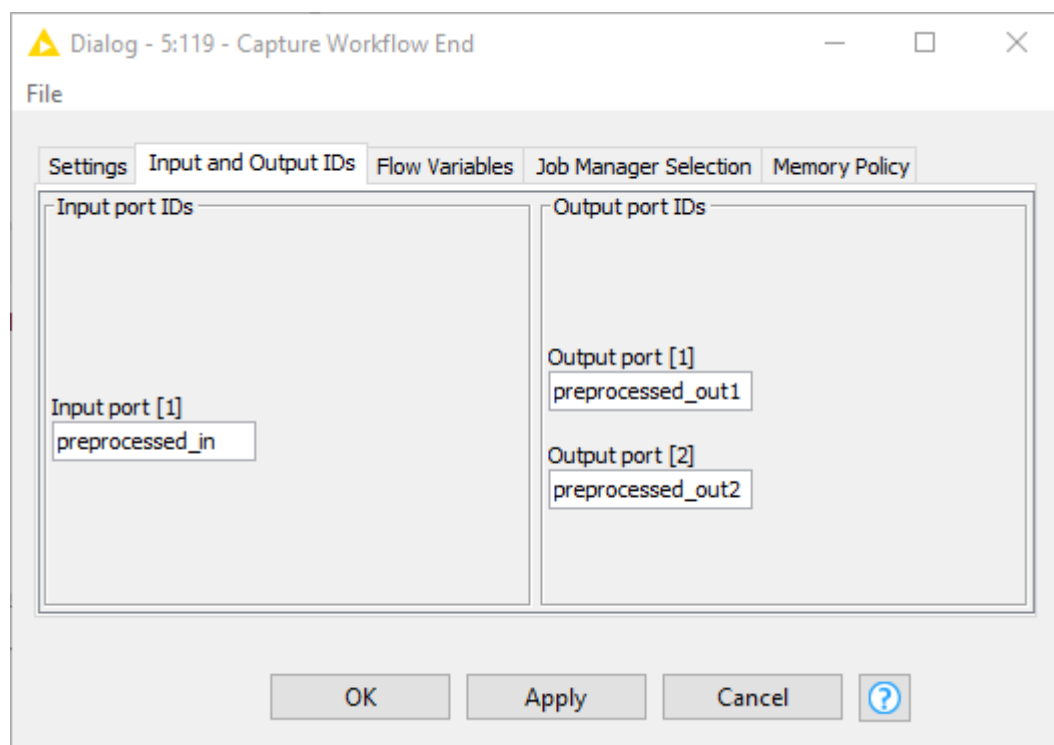


Figure 6. The Capture Workflow End node configuration dialog, Input and Output IDs tab.

Combine captured workflow segments

You can capture more than one segment of the workflow with multiple couples of Capture Workflow Start/End nodes and combine all the different captured segments into a single new workflow object. To combine the different captured segments use the **Workflow Combiner node**. The Workflow Combiner node can take as many as necessary workflow objects as input and generates a combined captured workflow. To add or remove input ports click the three dots in the left-bottom corner of the Workflow Combiner and choose *Add(Remove) workflow model port* in the context menu. An example with the relative captured workflow is shown in **Figure 7**.

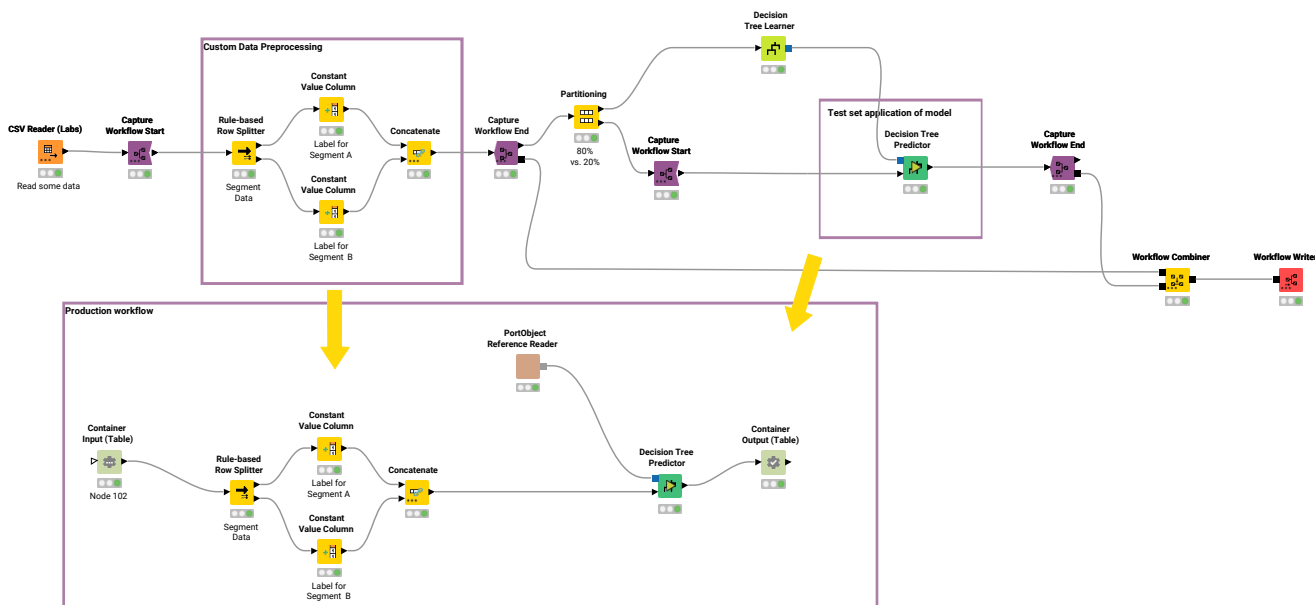


Figure 7. Different segments of a workflow are captured and combined in a new workflow

Input/output port mapping

Free output ports from one workflow segment are connected to the free input ports of the consecutive workflow segment. However the pairing of the output and input ports can be manually configured.

Workflow segments connected to consecutive ports to the Workflow Combiner node form a pair. For each pair you can select if and how to connect the outputs of the first workflow segment to the inputs of the consequent workflow segment. By default the inputs of each workflow segment are automatically connected with the outputs of the predecessor. In case the default pairing cannot be applied, e.g. because of a non-matching number of inputs and outputs or incompatible port types, the node requires manual configuration in order to be executed.

If the default configuration is not applicable or you want to manually configure it, right-click the Workflow Combiner node and choose *Configure...* from the context menu to open the

node configuration dialog, shown in [Figure 8](#)

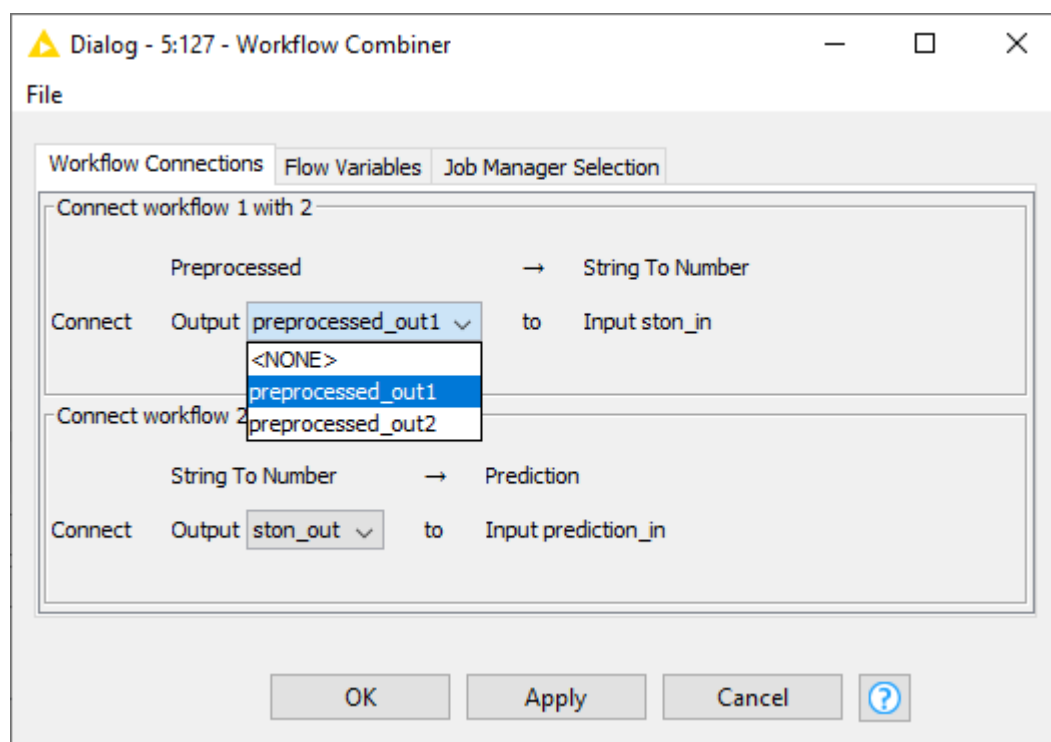


Figure 8. The Workflow Combiner node configuration dialog

Here, you can choose the pairing manually. For example, the first pane in [Figure 8](#) shows the pairing between a workflow segment (Preprocessed) that has two data output ports `preprocessed_out1` and `preprocessed_out2` with a workflow segment (String To Number) connected to the consequent port of the Workflow Combiner node. From the drop-down menu you can choose to connect any or `<NONE>` of the output ports of the Preprocessed workflow segment to the input port of the String To Number workflow segment.

Please, note that only compatible port types are shown and can be connected.

Save and execute a captured workflow

In this section we will explain how to save, deploy and execute the captured workflow, in a local or connected file system or on KNIME Server.



All the workflows shown in this section are available on [KNIME Hub](#).

The captured workflow can be:

1. Written locally or to a file system using the Workflow Writer node
2. Deployed to KNIME Server with the Deploy Workflow to Server node
3. Directly executed with the Workflow Executor node

Write a workflow object as production workflow

Write locally

You can write the newly stored workflow object to the [local file system](#) as production workflow using the [Workflow Writer node](#). Connect the Workflow Writer node editor to the output workflow object port (squared solid black port) of the workflow you want to write, i.e. the output port of Workflow Combiner node or Capture Workflow End node, as shown in [Figure 9](#).

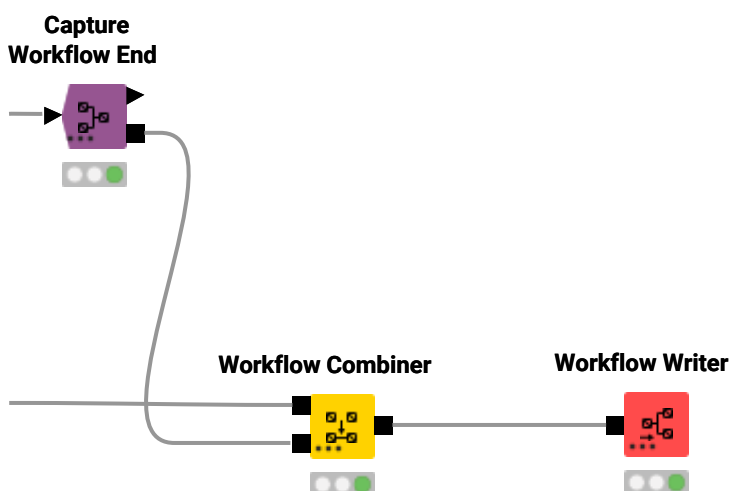


Figure 9. An example workflow to write a production workflow to the local file system

Now, you can open the Workflow Writer node configuration dialog, shown in [Figure 10](#).

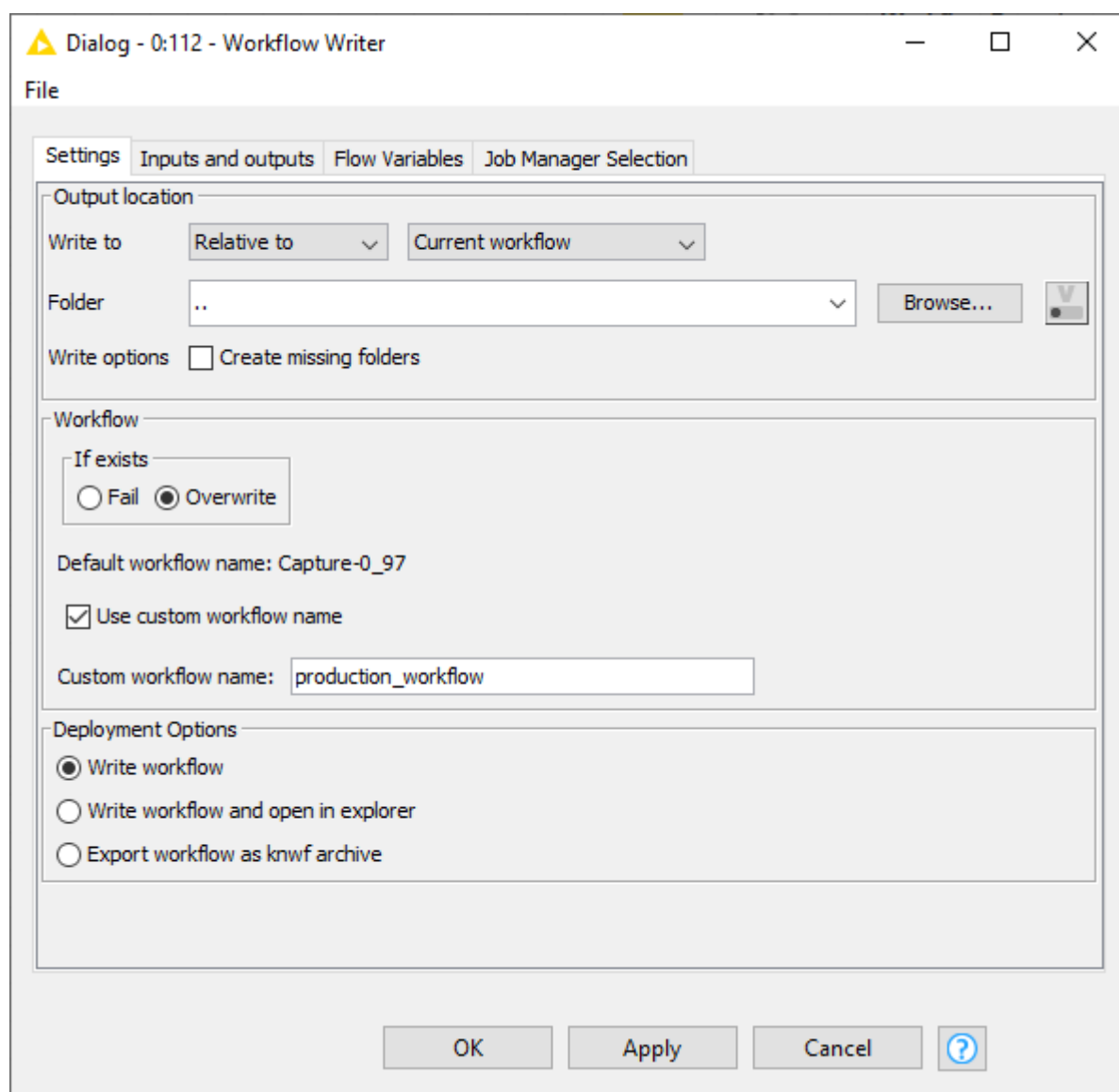


Figure 10. The Workflow Writer node configuration dialog

Here, you can setup:

- **Output location:** You can choose to write a captured workflow with a specific path relative to the local file system, a specific mountpoint or relative to current mount point, current workflow or current workflow data area.
- **Workflow:** You can choose to overwrite the workflow if one with the same name already exists in the chosen location or to fail. You can also give the captured workflow a custom name.
- **Deployment options:** You can choose between writing the workflow, writing the workflow and opening it in explorer, or export the workflow as knwf archive.

In the *Inputs and outputs* tab of the Workflow Writer node configuration dialog, shown in [Figure 11](#), you can also choose to add input and output nodes.

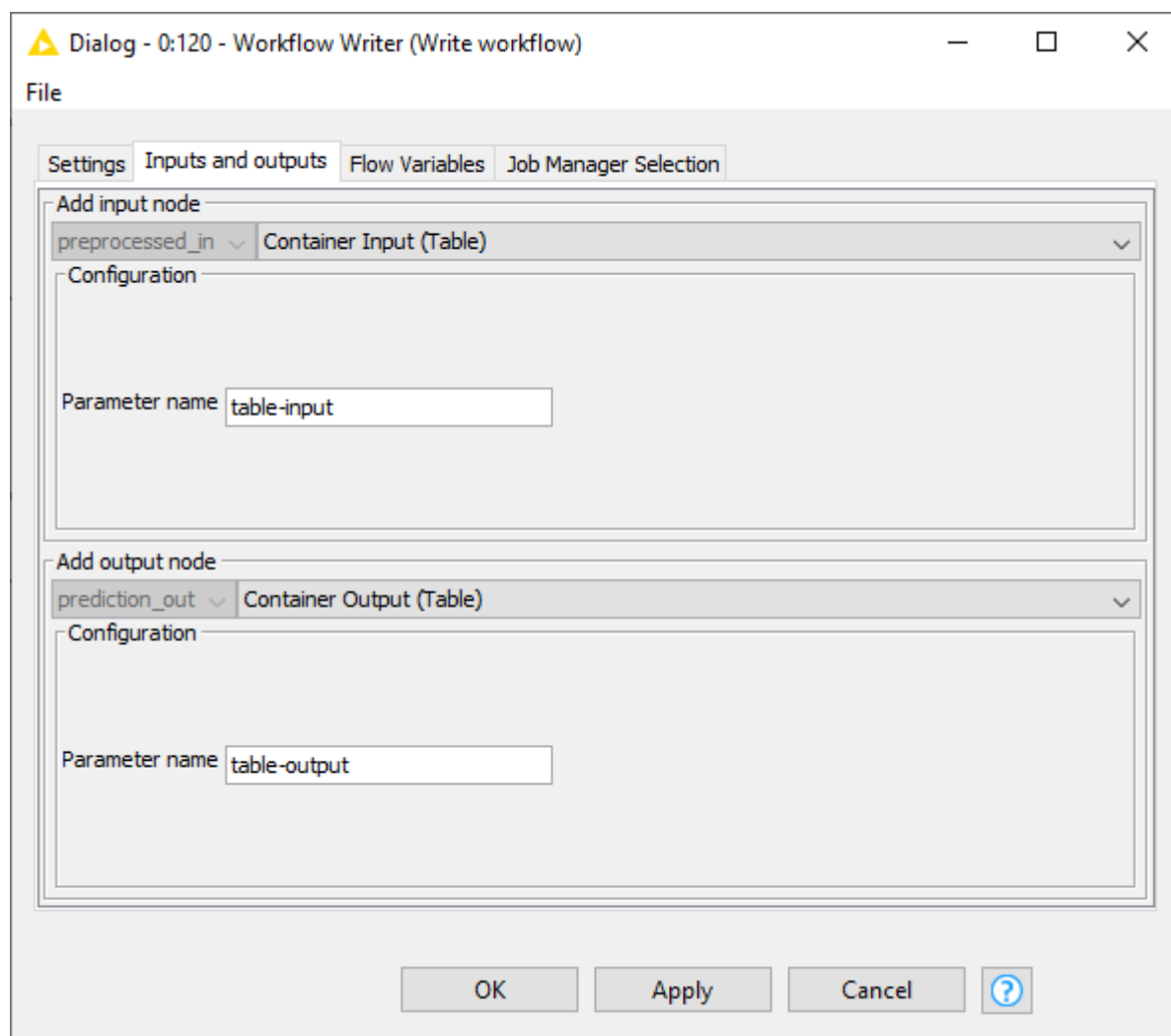


Figure 11. The Workflow Writer node configuration dialog, Input and outputs tab

Write to a connected file system

You can also connect the Workflow Writer node to a **connected file system**, e.g. SharePoint, Amazon S3, Google Cloud Storage. First you need to authenticate to the chosen connected file system and then connect to it. Then, right-click the three dots in the left-bottom corner of the Workflow Writer node and from the context menu choose *Add File System Connection port*. A squared solid cyan connection port will appear and it can be connected to the relative output file system connection port of the chosen connected file system connector node. An example workflow with SharePoint connected file system is shown in **Figure 12**.

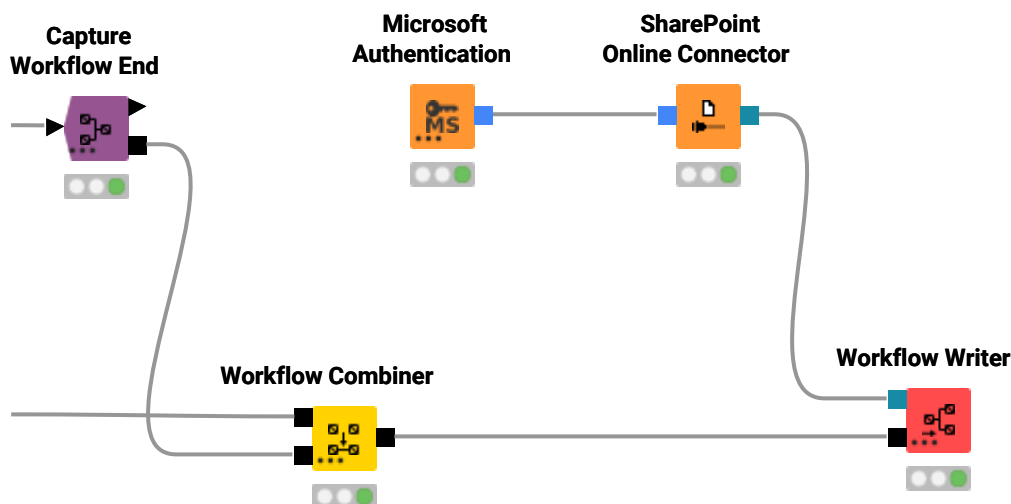


Figure 12. An example workflow to write a production workflow to a connected file system

Now the Workflow Writer node configuration dialog shows the remote file system in the *Output location* section of the *Settings* tab and you can choose to save the captured workflow into a location on that file system.



Please note that if you are writing to a connected file system the workflow can only be written or exported as `knwf` archive, and it is not possible to prompt its opening as a new explorer tab on KNIME Analytics Platform.

Deploy a production workflow to KNIME Server

You can deploy a production workflow to KNIME Server with the **Deploy Workflow to Server node**, following these steps:

1. Set up a connection to the server with the KNIME Server Connection node
2. Connect the KNIME Server Connection node to the Deploy Workflow to Server node via the squared solid light blue KNIME Server connection port
3. Connect the Deploy Workflow to Server node to the output workflow object port of the workflow you want to write, i.e. the squared solid black output port of Workflow Combiner node or Capture Workflow End node, as shown in **Figure 13**.

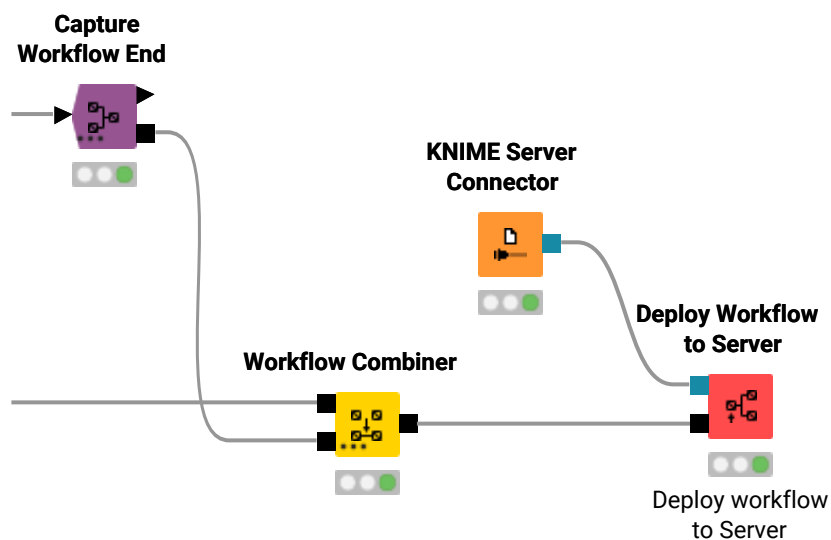


Figure 13. An example workflow to deploy a production workflow to KNIME Server

4. Open the Deploy Workflow to Server configuration dialog, shown in [Figure 14](#), by right-clicking the node and choosing *Configure...* from the context menu.

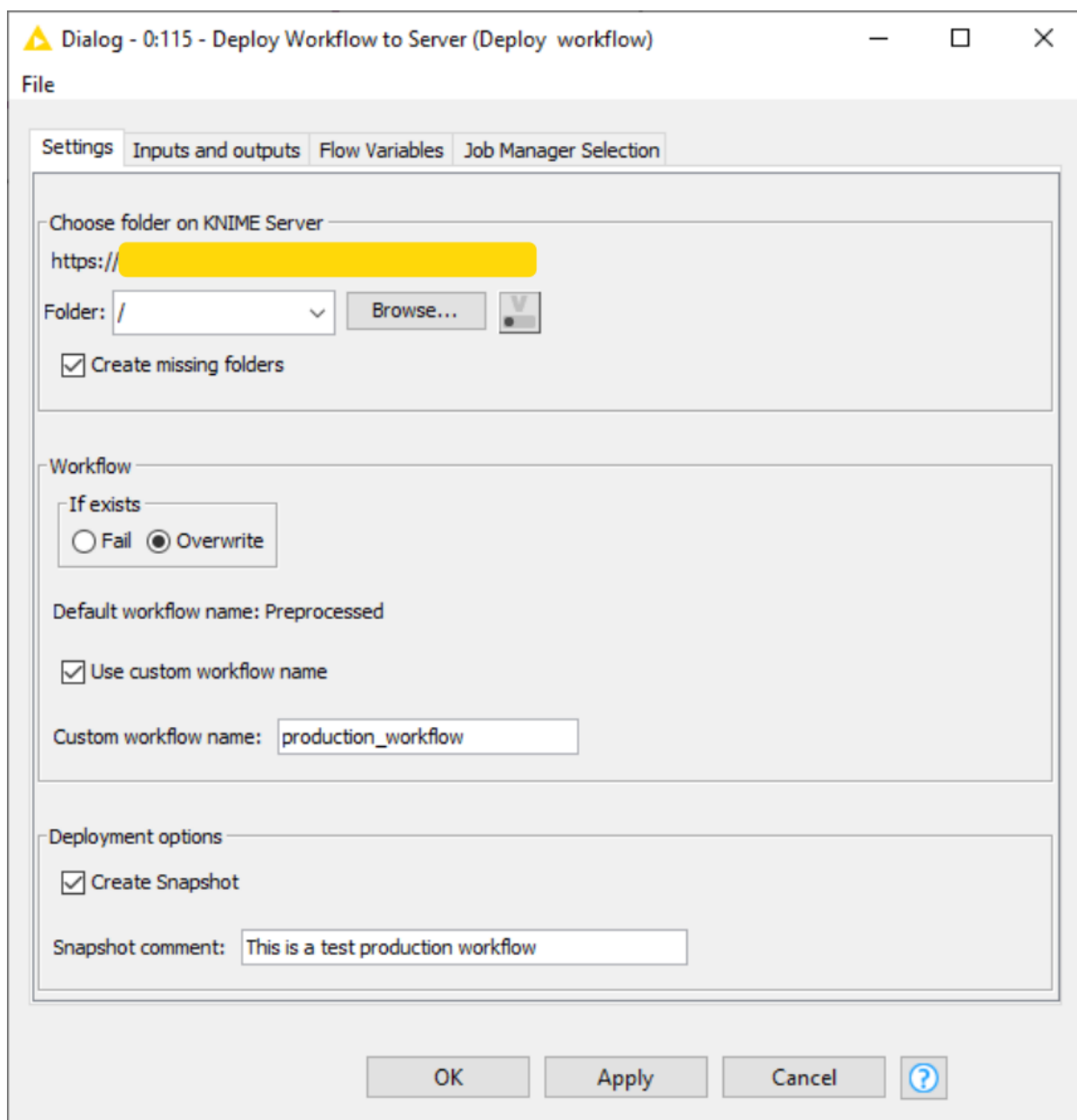


Figure 14. The Deploy Workflow to Server node configuration dialog

In the Deploy Workflow to Server node configuration dialog you can setup the following options:

- Choose folder on KNIME Server: You can choose the folder on the server file system to which deploy the newly captured workflow.
- Workflow: You can choose to overwrite a workflow if one with the same name already exists in the chosen location or to fail. You can also give the written workflow a custom name.
- Deployment options: You can choose to create a snapshot and add an optional snapshot comment in order to keep track of the changes made to the workflow on the server.

In the *Inputs and outputs* tab of the Deploy Workflow to Server node configuration dialog, shown in [Figure 15](#), you can also choose to add input and output nodes to the workflow. This will add Container Input/Output (Table) nodes to the production workflow deployed to KNIME Server.

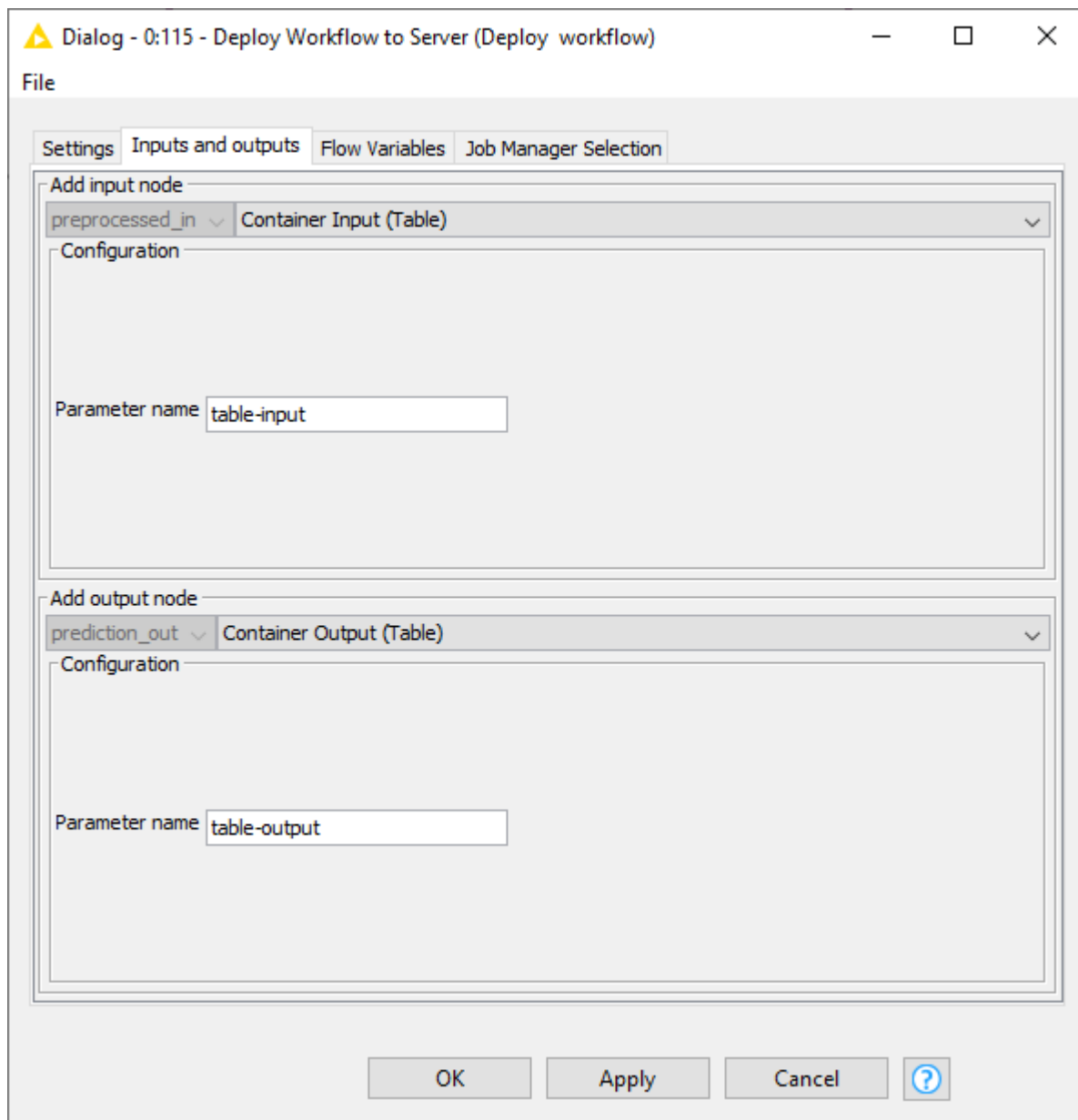


Figure 15. The Deploy Workflow to Server node configuration dialog, *Input and outputs* tab



If the captured workflow deployed to Server is executed on the Server (see [Execute a captured workflow deployed to KNIME Server](#) section) via Call Workflow node the Container Input/Output (Table) nodes are mandatory.

Read a production workflow

You can read a workflow that has been saved in any [standard](#) or [connected](#) file system or

deployed to KNIME Server by using the Workflow Reader node.

Select the production workflow to read

You can read workflows from standard file systems just by configuring the Workflow Reader node to read from either *Local File System* or *Mountpoint and Relative to* in the node configuration dialog as shown in [Figure 16](#).

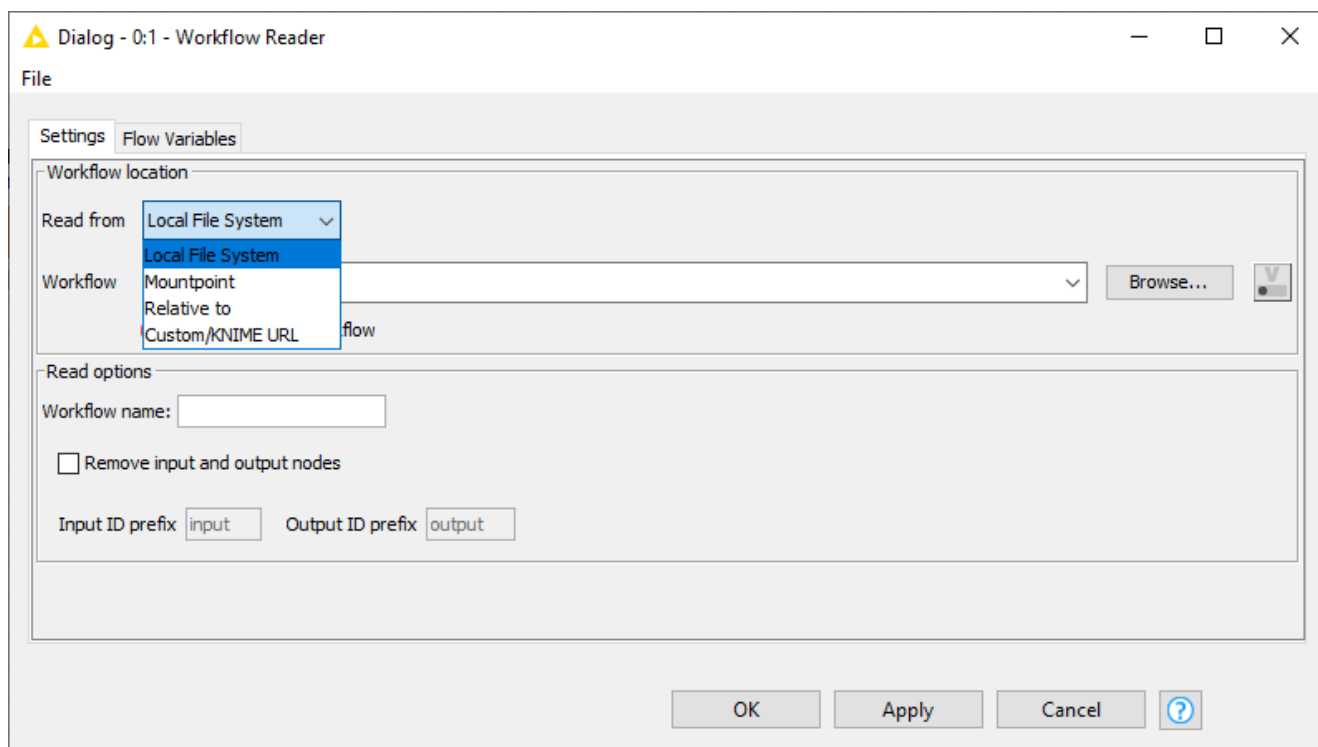


Figure 16. The configuration dialog of a Workflow Reader node

Alternatively you can read workflows stored on a remote file system or deployed to a KNIME Server. To do so you need to add a connection port to the Workflow Reader node by clicking the three dots on the left bottom corner of the node and use a Connector node to connect to the desired remote file system.

An example to read a workflow deployed to KNIME Server is shown in [Figure 17](#).

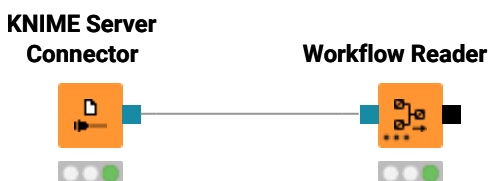


Figure 17. The Workflow Reader node to read a workflow deployed to a KNIME Server

Please be aware that if you are reading your workflow from a file system that is not a KNIME file system, i.e. a remote file system that is not KNIME Server or your local system standard file system, you will be able to read the workflows that are saved as `.knwf` files.



It is not possible to read **KNIME components**.

Configuration options of the Workflow Reader node

In the *Settings* tab of the configuration dialog of the Workflow Reader node, not only you can choose the workflow location from which read the workflow but also, under *Read options* pane you can:

- Under *Workflow name*: choose a name for the workflow you read in
- Check *Remove input and output nodes*: remove all Container Input and Container Output nodes from the workflow you read in. These nodes are then implicitly represented, allowing you to, e.g., combine the resulting workflow segment with other workflow segments via the new implicit input and output ports via the Workflow Combiner node, or to execute the workflow segment providing the desired input data via the Workflow Executor node. You can then assign an *Input ID prefix* and/or an *Output ID prefix* to the implicit input and output ports of the workflow segment. The input and output IDs will then be created as follows: "<Input ID prefix>n" and "<Output ID prefix>m" with *n* and *m* run from 1 to the maximum number of input/output ports.

Execute a production workflow deployed to KNIME Server

You can execute a production workflow:

- On **KNIME Server** or on **KNIME WebPortal**
- On KNIME Server from the KNIME Analytics Platform client. In this way you can also provide additional data in case you want to run the production workflow on new data. To do this you can use a Call Workflow node, connected to the KNIME Server Connection node via the KNIME Server connection port. If needed, i.e. if no **input data was stored** with the production workflow, input data have to be provided, as shown in **Figure 18**.

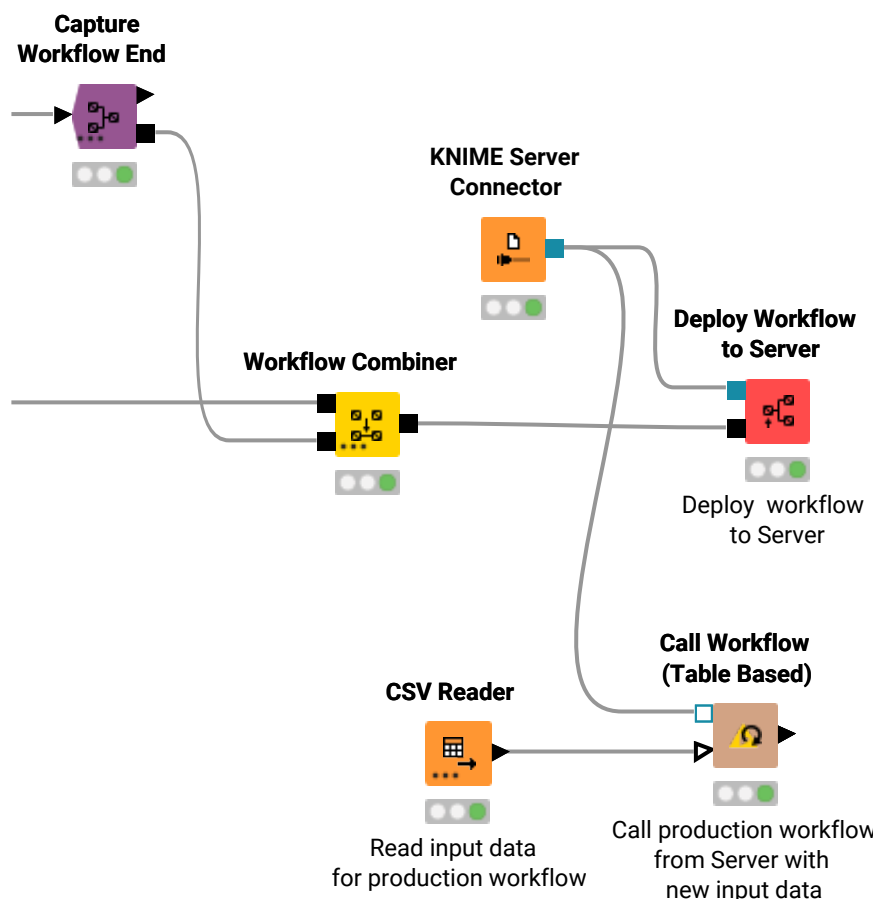


Figure 18. An example workflow to execute the production workflow from KNIME Server with some input data

Execute a production workflow locally

You can also execute the workflow locally by using the [Workflow Executor node](#).



The Workflow Executor node can only execute a workflow that is captured from the same workflow the Workflow Executor node is part of.

The node ports match the input and output ports of the workflow. Connect the Workflow Executor node to the output workflow object port of the workflow you want to write, i.e. the output workflow object port of Workflow Combiner node or Capture Workflow End node, and, if needed, the input data port to a data table. An example is shown in [Figure 19](#).

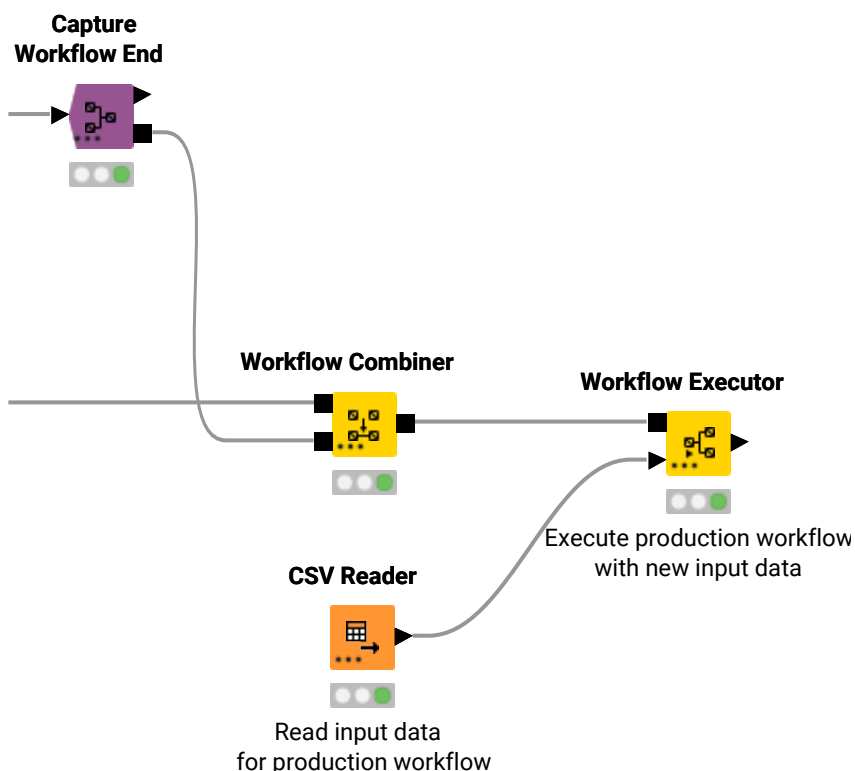


Figure 19. An example workflow to execute the production workflow locally

You can adjust the needed input and output ports to match those of the workflow to be executed by opening the node configuration dialog shown in [Figure 20](#), and clicking *Auto-adjust ports (carried out on apply)*.

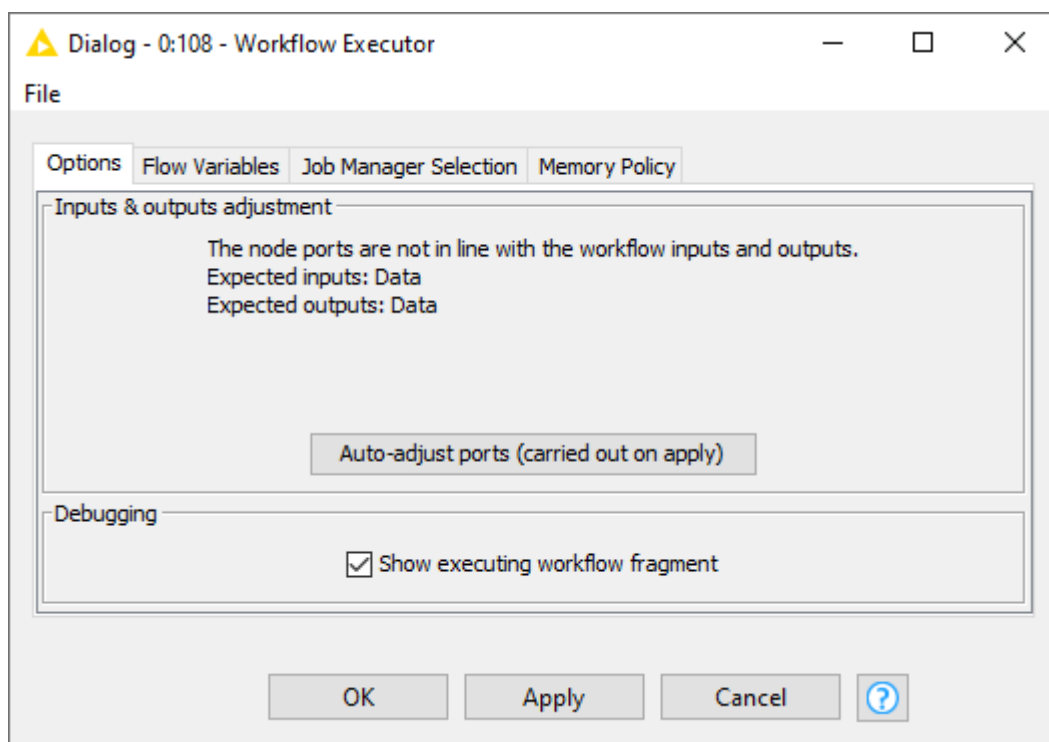


Figure 20. The Workflow Executor node configuration dialog

Extract the workflow summary

You can generate a workflow summary using the Workflow Summary Extractor node. It creates a document in either XML or JSON format which contains workflow meta data and makes it available at the node output as data table. The workflow summary is a detailed and structured description of a workflow including its structure, configuration of the nodes, port specifications, node and workflow annotations.

You can connect the input workflow port of the Workflow Summary Extractor node to any other output port of the same type to select the workflow to be processed.

In the configuration dialog of the Workflow Summary Extractor node you can choose the format of the output between JSON and XML and assign a name other than the default one to the column of the final table which will contain the workflow summary extracted.

KNIME AG
Talacker 50
8001 Zurich, Switzerland
www.knime.com
info@knime.com