

KNIME Pro Guide

KNIME AG, Zurich, Switzerland
Version 1.15 (last updated on)



Table of Contents

KNIME Pro	1
Manage your KNIME Pro subscription	1
Connect to KNIME Community Hub and upload workflows	2
Upload items to KNIME Hub	2
Version items	4
Move items	5
Delete items from KNIME Hub	5
Application passwords	6
Execution of workflows	7
Parameterization of workflows with Configuration nodes	8
Schedule a workflow execution	9
Deploy a workflow as a data app	12
Jobs	15
Extensions available for execution on KNIME Community Hub	17
Data Apps Portal	21
Access data from different sources	22
Access data from Google Cloud	22
Access data from Google BigQuery	25
Access data from Azure	27
Access data from Snowflake	31
Access data from Amazon Athena	32
Advanced settings	36
IP whitelisting	36

KNIME Pro



KNIME Pro is currently in beta. If you are interested in trying it out, visit our website at [this page](#).

KNIME Pro is a paid service offered on KNIME Community Hub.

With KNIME Pro you can take advantage of the full capabilities of KNIME Community Hub, including:

- Running workflows on KNIME Community Hub
- Creating and sharing data apps
- Automating the execution of workflows on KNIME Community Hub

Manage your KNIME Pro subscription

You can manage your subscription on the *Subscription* page. On your Pro profile page, select *Subscription* on the side menu to access it.

Here you can:

- Show the cost breakdown: Click *Show cost breakdown* to see the amount to be paid for your KNIME Pro and the consumed credits for execution.
- End subscription: In the *Manage subscription* side panel, click *End your subscription now*. This action will take effect after the end of your current billing cycle. Your data will be kept for 30 days and you will be able to re-activate your account in that time period. After 30 days we will delete all your data so it is not recoverable anymore. Once you confirm with *Cancel subscription*, your Pro profile and its spaces will be deleted.



To re-activate your KNIME Pro please get in [contact with us](#).

- Past invoices: Here you can see an overview of all the invoices and download them by clicking *Download invoice*.
- Inspect your billing information and edit your payment details.

Connect to KNIME Community Hub and upload workflows

As a user of KNIME Community Hub you will be able to upload items, for example workflows or components, to your spaces.

This will allow you to have access to the items, and version so that you can keep track of your work.

To do this the process is the same as for simple KNIME Community Hub users, that are not Pro, as explained in the [KNIME Community Hub User Guide](#).

In the following sections you will find the fundamental steps that you need to do in order to connect, upload and version your workflows on the KNIME Community Hub.



The following steps assume that you have already [created an account](#) on KNIME Community Hub.

Upload items to KNIME Hub

You can upload items to KNIME Hub:

- From the [web UI](#) of KNIME Community Hub, if the items are saved on your local computer.
- From a local [KNIME Analytics Platform](#), if the items are on your KNIME Analytics Platform local space.


Upload from KNIME Community Hub web UI

Go to [KNIME Community Hub](#), sign in with your username and password, then navigate to the space where you want to upload your workflow (.knwf file) or file to. Click the *Upload* button and select the file you want to upload to the space from your local filesystem.

Upload from KNIME Analytics Platform

- Connect to the KNIME Community Hub from KNIME Analytics Platform: The first step is to connect to your KNIME Community Hub account on KNIME Analytics Platform. To do so, go to the Home page of KNIME Analytics Platform and sign in to KNIME Community Hub. You will be redirected to the sign-in page if you are not connected in the current

session already.

- Once you are signed in, you will see all the spaces you have access to. Select the Pro account on the left and then select the space you want to upload your items to.
- You can also create a new space by clicking the  button.
- Click on a space to access the items within the space. You can perform different types of operations on them:
 - Upload items to your KNIME Hub spaces, download, duplicate, move, delete, or rename your items. More information about this functionality is provided in the next section.
 - Open workflows as local copies or on KNIME Hub. You can find the respective buttons in the toolbar on top.
 - Create folders to organize your items in the space. Click the *Create folder* button in the toolbar on top.
- Now, you can upload the desired items to your KNIME Hub spaces. You can upload workflows or components to any of your spaces by right-clicking the item from the space explorer in KNIME Analytics Platform and selecting *Upload* from the context menu. A window will open where you will be able to select the location where you want to upload your workflow or component.

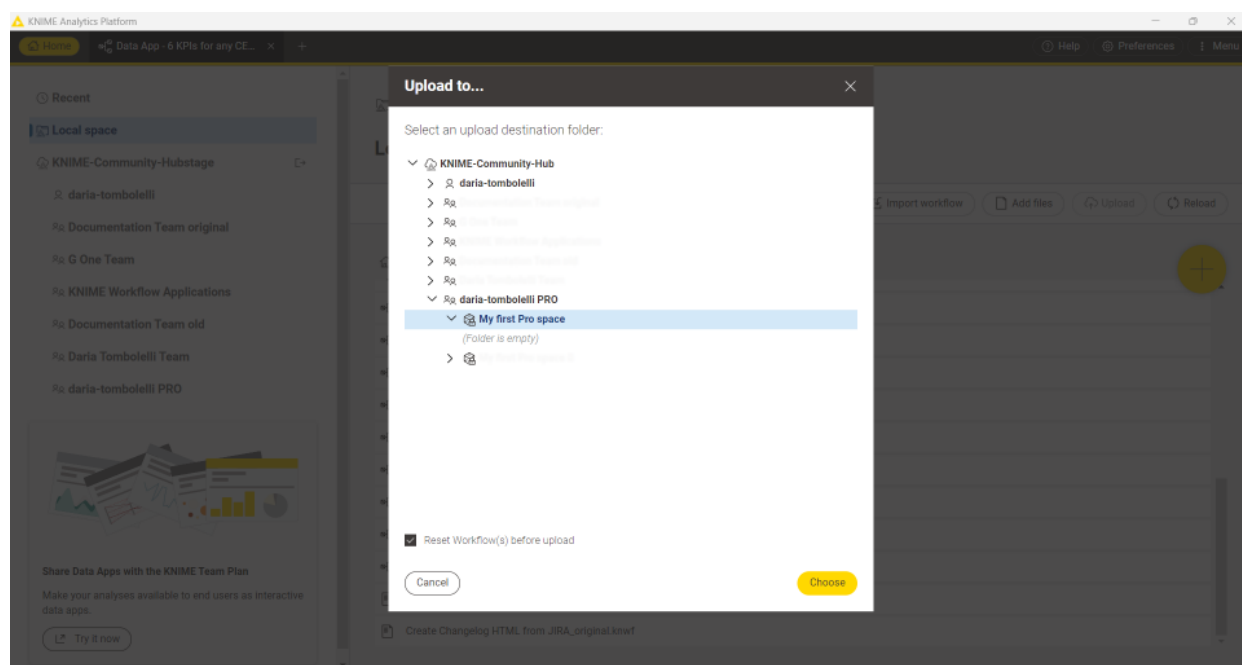


Figure 1. Upload a local item to KNIME Community Hub



Items that are uploaded to a public space will be available to everyone. Hence, be very careful with the data and information (e.g., credentials) you share.

Version items

When **uploading items to a space on KNIME Hub**, you will be able to keep track of their changes. Your workflow or component will be saved as a draft until a version is created.

When you create versions of the items, you can then go back to a specific saved version at any point in time in the future to download the item in that specific version.

Once a version of the item is created, new changes to the item will show up as draft.

Create a version of an item

1. Go to the item you want to create a version of
2. Click *Versions*
3. Click *Create version* in the *Versions history* panel on the right side of the page
4. A side panel opens where you can give the version a name and add a description
5. Click *Create* to create the version

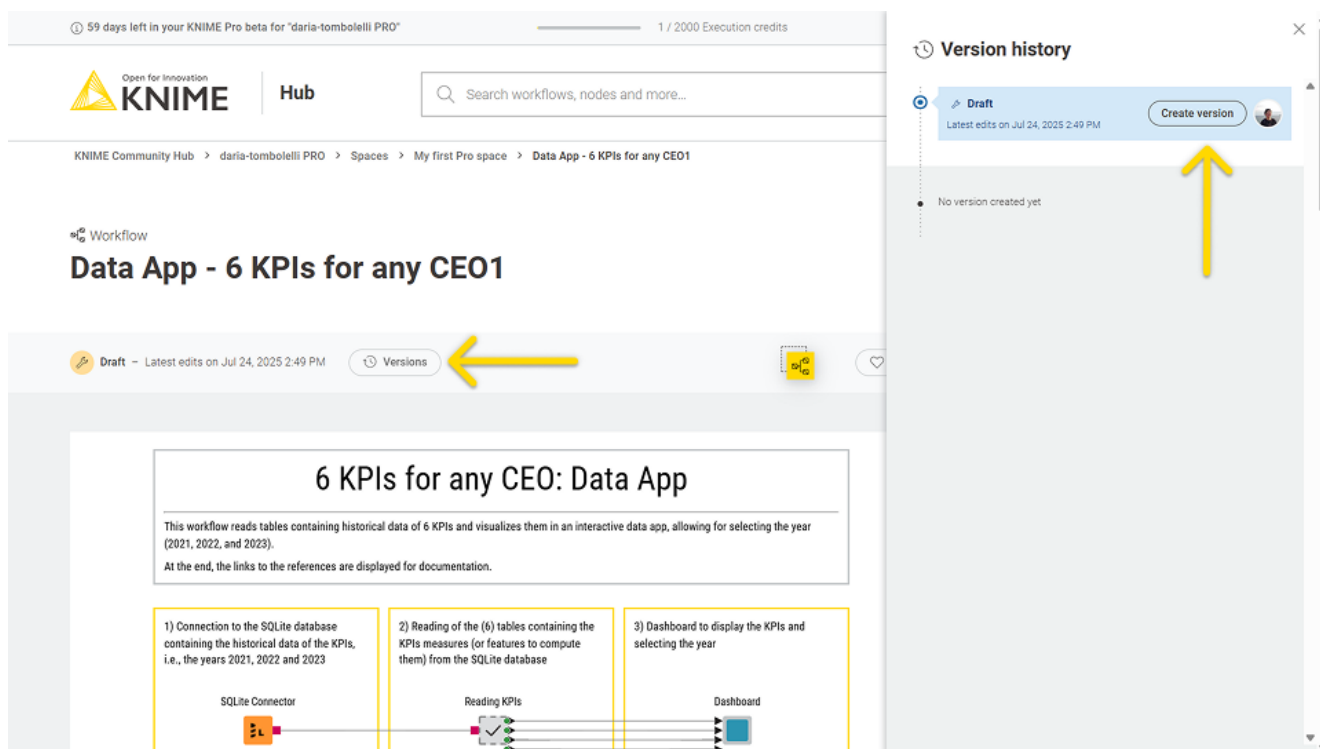




Figure 2. Versioning of a workflow

Show a version

In the *Versions history* panel, click the version you want to see, or click the  icon and select *Show this version*. You will be redirected to the item in that specific version.

To go back to the latest state of the item, click the selected version to unselect it.

Restore a version

To restore a version that you created, click the  icon in the version tile from the item history and select *Restore this version*.

The version will be restored as a draft.

Delete a version

In the *History* panel, click the  icon for the version you want to delete and click *Delete*.

Move items

You can move items that you have uploaded to KNIME Hub to a new location within the space that contains the item, or to a different space that you have access to. To do this, you need to be connected to the KNIME Hub on KNIME Analytics Platform. If you want to move items within the same space, drag the item in the space explorer, for example, to a subfolder. To move items from one space to another, right-click the item and select *Move to*. In the *Destination* window that opens, select the space to which you want to move the item to.




These changes will automatically apply to the space on KNIME Hub.


Delete items from KNIME Hub


You can also delete items that you uploaded to KNIME Hub.

To do so you can:

- From KNIME Hub, sign in with your account and go to the item you want to delete. Click the  icon on the top right of the page and select *Move to bin*. This will move the item to the bin, where it will stay for 30 days before being permanently deleted.

You can:

- **Permanently delete** the item from the bin by going to the profile page, selecting *Recycle bin* from the left-hand menu, and clicking  > *Delete permanently* next to the item you want to delete.

- **Restore** the item from the bin within the 30 days time frame by going to the profile page, selecting *Recycle bin* from the left-hand menu, and clicking  > *Restore* next to the item you want to restore.


Application passwords

Application passwords can be used to provide authentication to KNIME Hub API, for example when using a KNIME Hub Authenticator node. To create a new application password:

1. Go to your profile page and select *Settings* → *Application passwords* from the menu on the left
2. Click *Create application password* and a side panel will show
3. In the side panel, you can give a name to the application password and click *Apply*
4. A new application password will be created and shown in the list of application passwords
5. You can copy the ID and the password of the application password you just created
6. You can use the ID and the password to authenticate to KNIME Hub API, for example, in the KNIME Hub Authenticator node



The application password is a one-time password, meaning that you will not be able to see the password again after you close the side panel.

From this page you can also click the  icon and click *Delete* from the menu that opens, in order to delete a specific application password.

Execution of workflows

Execution allows you to run and deploy your workflows directly on KNIME Hub.

Navigate to a workflow in one of your spaces, you will see a *Run* button and a *Deploy* button.

Click *Run* to simply execute the current workflow directly in the browser.

Here select:

- Which version of the workflow you want to run. Default is the latest state of your workflow.
- Which of the provided execution contexts you want to use.
- Enable email notifications if you want to be notified when the execution is finished or if it fails.

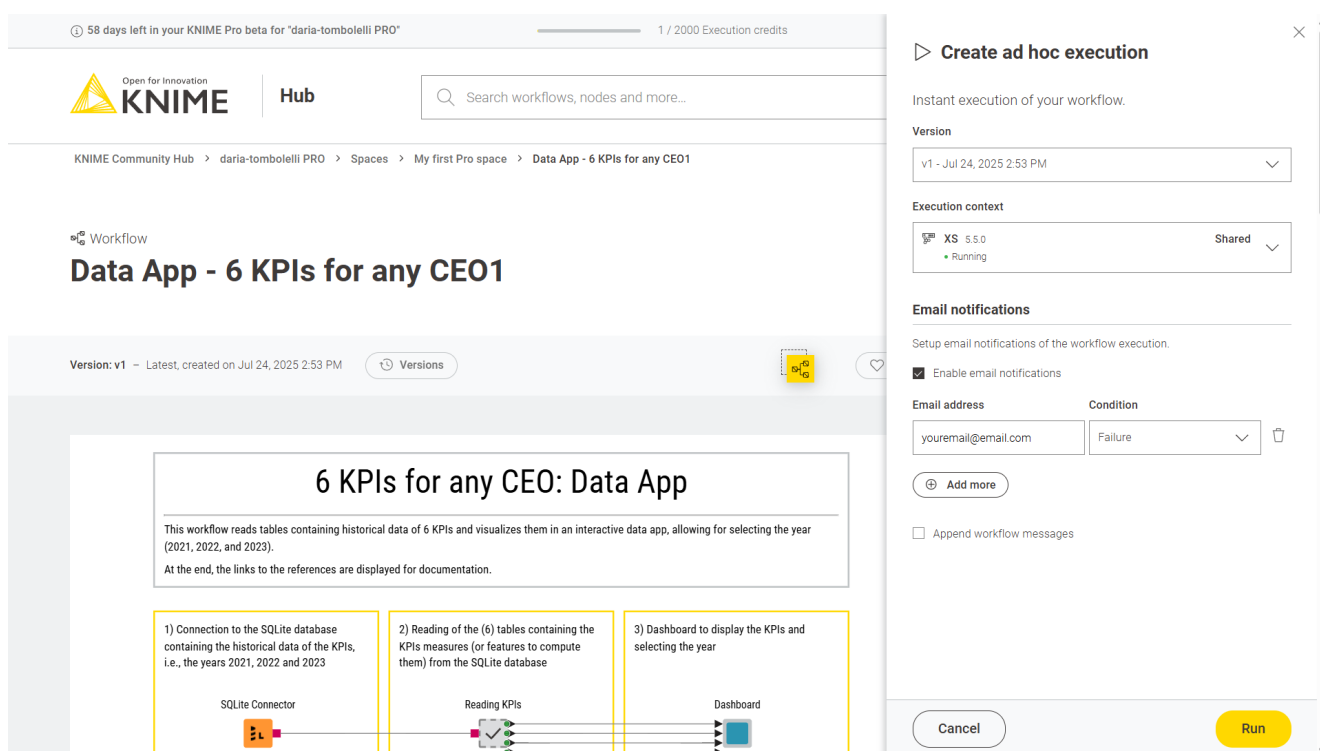


Figure 3. The ad hoc execution side panel

A new tab will open where the result of the workflow execution will be shown. If your workflow is a data app you will be able to interact with its user interface.

To know more about how to build a data app refer to the [KNIME Data Apps Beginners Guide](#).

To see all the ad hoc jobs created for a workflow navigate to the workflow page and select *Ad hoc jobs* from the right side menu.

Parameterization of workflows with Configuration nodes

When you execute a workflow on KNIME Hub you can also provide parameters to the workflow using **Configuration nodes**.

When building the workflow, add a Configuration node in the root of the workflow—meaning the Configuration node should not be nested in a component.

This allows you to perform a number of operations that can be:

- Pre-configured with certain parameters. This can be done by choosing the default value in the configuration dialog of this type of nodes.

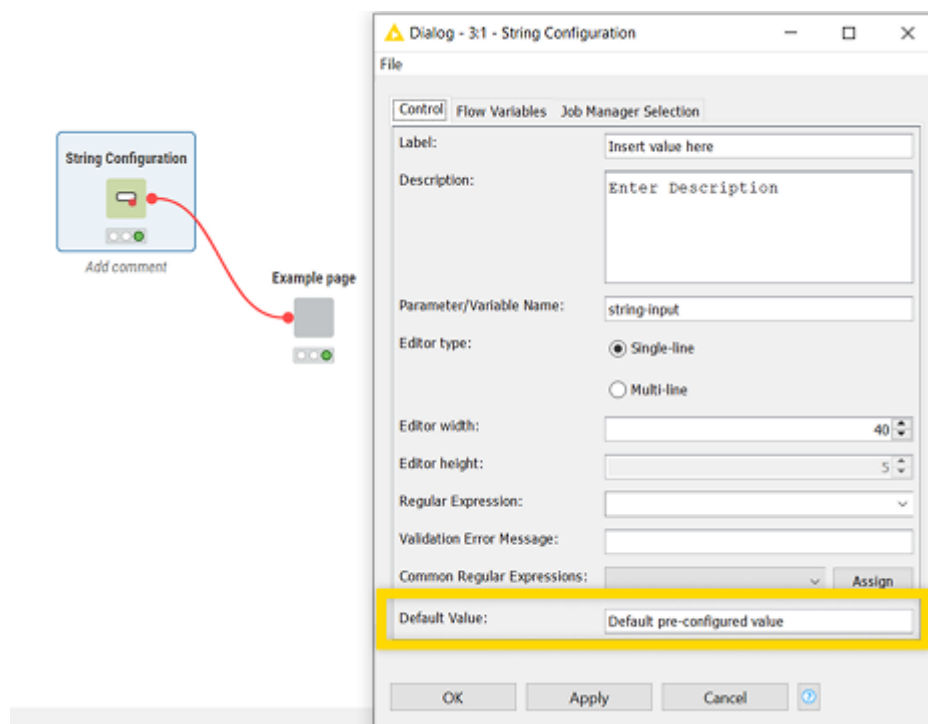


Figure 4. Configuration node

- Configured by the user at the time of execution or when creating a deployment. This can be done from the panel that opens when you click *Run* on the workflow page or when you create a **deployment**.

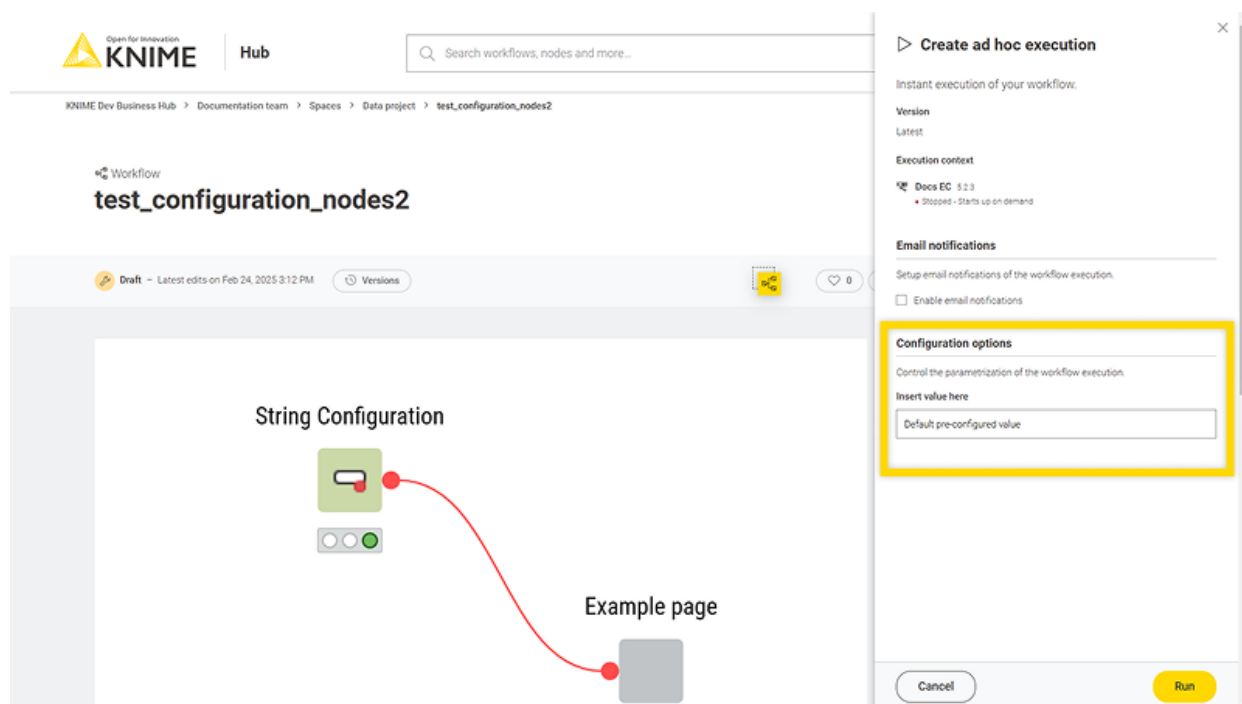


Figure 5. Workflow parameters set in the execution side panel

Schedule a workflow execution

A schedule allows you to set up a workflow to run automatically at selected times.

To create a schedule deployment of your workflow:

1. Navigate to the workflow in one of your spaces
2. In order to be able to create a schedule deployment you need to first have created at least one version of your workflow.



To know more about versioning read the [Versioning](#) section of this guide.

3. Once you have your stable version of the workflow, create a version (click *Versions* → *Create version*).
4. Now click the button *Deploy* and select *Create schedule*. This will open a side panel where you can set up the schedule deployment.

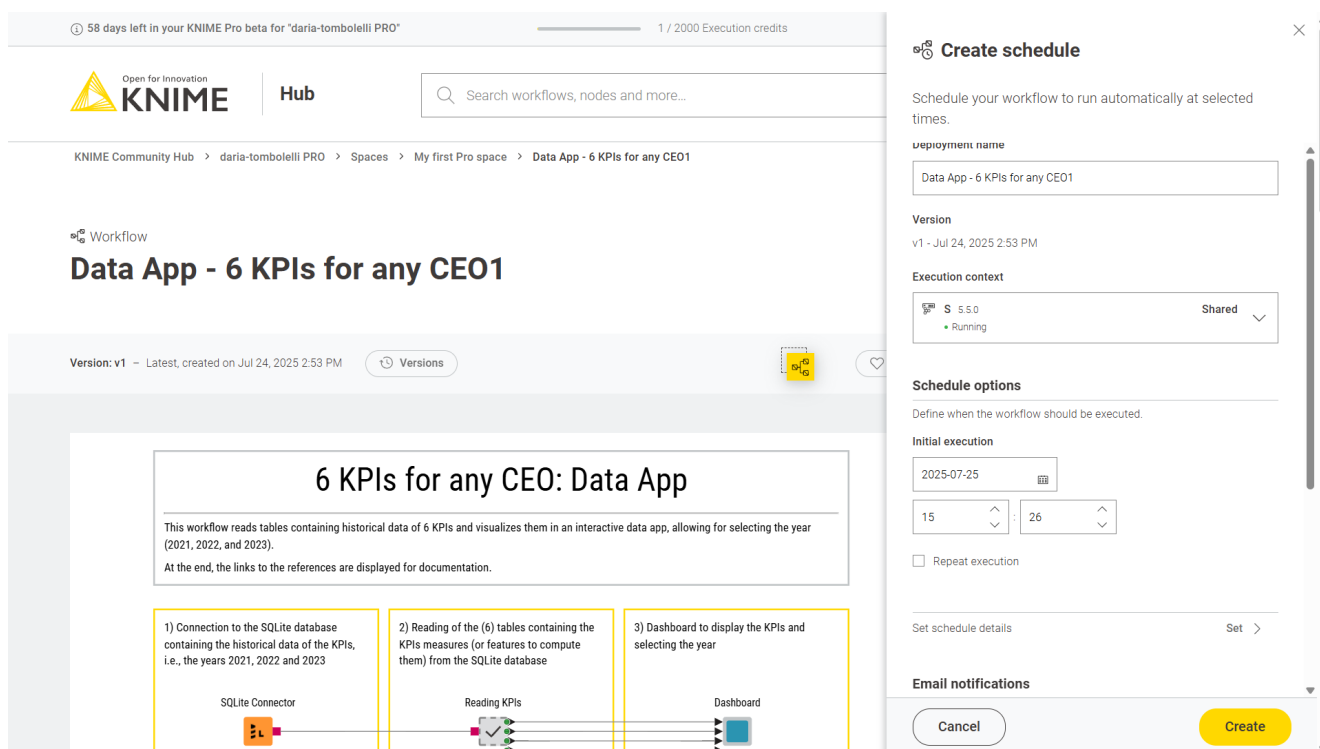


Figure 6. The create schedule side panel

In the panel you can select:

- A name for your schedule deployment
- Which version of the workflow you want to schedule for execution
- Select the size of the executor you want to use for the scheduled execution. You can choose between four executor sizes: *Extra Small*, *Small*, *Medium*, and *Large*. The executor size determines the amount of resources that will be allocated to the execution of the workflow.

i

To see all the deployments created for a workflow navigate to the workflow page and select *Deployments* from the right side menu. To see all the deployments you created navigate to your Pro profile page and select *Deployments* from the left side menu.

In the section **Schedule options** you can define when the workflow should be executed:

- Set up an initial execution, date and time
- Select *Repeat execution* to select if you want the workflow to run every specific amount of time (e.g. every 2 hours), or select start times when your workflow will be executed every day at the same time.
- Set up also a schedule ends date, either to never end or to end in a specific date and time.

Set additional schedule details

You can also set more advanced schedule details.

1. First, check the option *Repeat execution*, under *Schedule options*
2. Then click *Set* to set up recurring executions, retries, and advanced schedule details.



If you did not check the option *Repeat execution*, you will only find the set up options for retries and the advanced schedule details.

The workflow will run when **all the selected conditions are met**.

The screenshot displays the KNIME Hub interface. The main area shows a workflow titled "Data App - 6 KPIs for any CEO1" with a version of v1. The workflow description states: "This workflow reads tables containing historical data of 6 KPIs and visualizes them in an interactive data app, allowing for selecting the year (2021, 2022, and 2023). At the end, the links to the references are displayed for documentation." The workflow steps are: 1) Connection to the SQLite database, 2) Reading of the (6) tables containing the KPIs measures, and 3) Dashboard to display the KPIs and selecting the year. The 'Create schedule' side panel is open, showing options for recurring execution. The 'Days of the week' section is set to Monday through Friday (M, T, W, T, F). The 'Days' section shows a calendar grid with days 1 through 31, and the 'Months' section shows a calendar grid with months from January to December. The 'Timeframes' section is set to 'Between 00:00 and 23:59'. The 'Add timeframe' button is visible at the bottom of the side panel.

Figure 7. The create schedule side panel: set schedule details section.

In this example the workflow will run from Monday to Friday, every day of the month, except for the month of December.

Finally, in the section **Execution retries** you can set up the number of execution retries, and check the following options for advanced schedule details:

- *Reset before execution*: the workflow will be reset before each scheduled execution retries occur.
- *Skip execution if previous job is still running*: the scheduled execution will not take place if the previous scheduled execution is still running.
- *Disable schedule*: Check this option to disable the schedule. The scheduled execution will start run accordingly to the set ups when it is re-enabled again.

Set configuration options

This section will show dynamically when adding Configuration nodes in the root of the workflow.

See the [Parameterization of workflows with Configuration nodes](#) section for more information on how to add configuration nodes to your workflow.

Set advanced settings

Finally, you can select advanced settings for each schedule deployment.

Click *Set* under *Advanced settings* section to do so.

In the panel that appears you can configure additional settings:

- *Job lifecycle*: such as deciding in which case to discard a job, the maximum time a job will stay in memory, the job life time, or the options for timeout. Particularly important in this section is the parameter *Max job execution time*: the default setting can be changed according to your needs to keep control over the execution time.
- *Additional settings*: such as execution scope or the option to update the linked components when executing the scheduled job.

Deploy a workflow as a data app

Deploying a workflow as a data app allows you to share the data application with other users, who can interact with the workflow through a user interface.

To create a data app deployment of your workflow:

1. Navigate to the workflow in one of your spaces
2. In order to be able to create a data app deployment you need to first have created at least one version of your workflow.



To know more about versioning read the [Versioning](#) section of this guide.

3. Once you have your stable version of the workflow, create a version (click *Versions* → *Create version*).
4. Now click the button *Deploy* and select *Create data app*. This will open a side panel where you can set up the data app deployment.

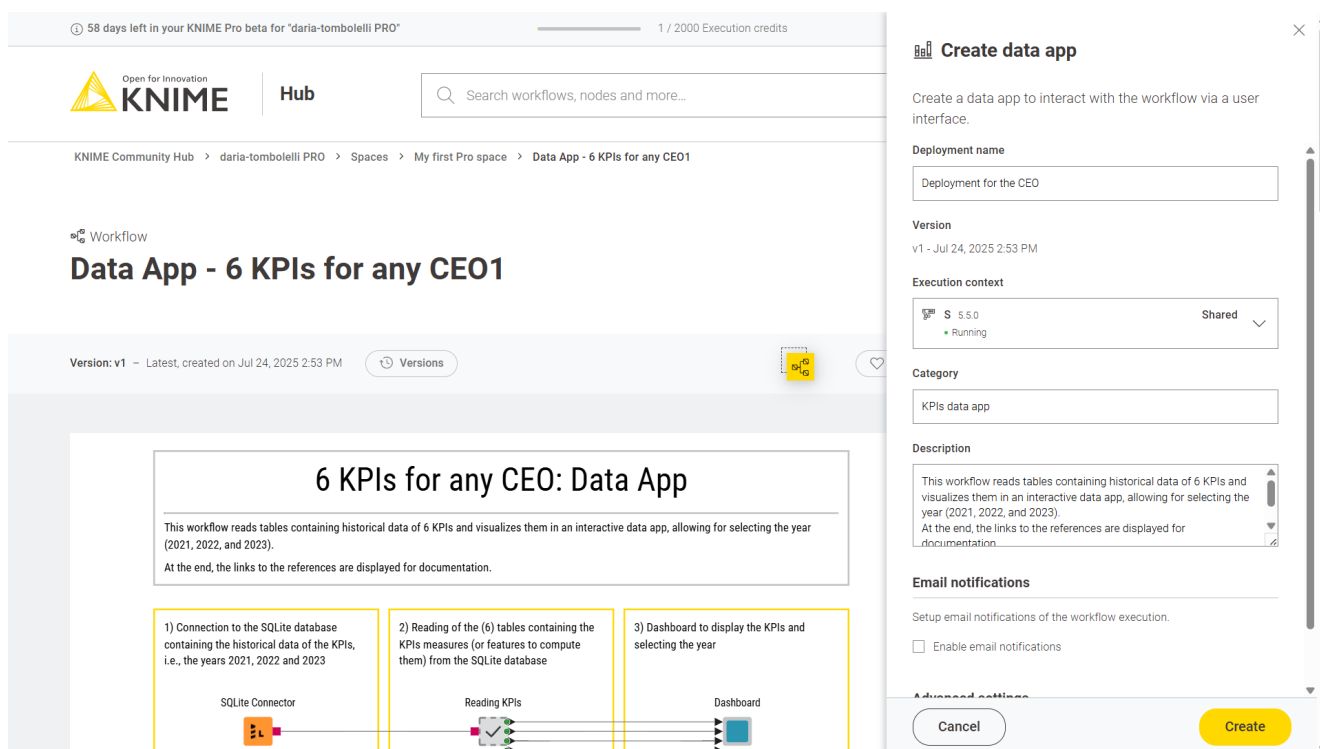


Figure 8. The create data app side panel

In the panel you can select:

- A name for your schedule deployment
- Which version of the workflow you want to schedule for execution
- Select the size of the executor you want to use for the scheduled execution. You can choose between four executor sizes: *Extra Small*, *Small*, *Medium*, and *Large*. The executor size determines the amount of resources that will be allocated to the execution of the workflow.

1. You can also select a *Category* and a *Description*.

Categories added here will be used to group the data apps in the [Data Apps Portal](#). Also descriptions added here will be visible in the corresponding deployment's tile in the [Data Apps Portal](#).

i

To see all the deployments created for a workflow navigate to the workflow page and select *Deployments* from the right side menu. To see all the deployments you created navigate to your Pro profile page and select *Deployments* from the left side menu.

Set configuration options

This section will show dynamically when adding Configuration nodes in the root of the workflow.

See the [Parameterization of workflows with Configuration nodes](#) section for more information on how to add configuration nodes to your workflow.

Set advanced settings


Finally, you can select advanced settings for each schedule deployment. Click **Set** under *Advanced settings* section to do so.

In the panel that appears you can configure additional set ups:

- *Job lifecycle*: such as deciding in which case to discard a job, the maximum time a job will stay in memory, the job life time, or the options for timeout. Particularly important in this section is the parameter *Max job execution time*: the default setting can be changed according to your needs to keep control over the execution time.
- *Additional settings*: such as report timeouts, CPU and RAM requirements, check the option to update the linked components when executing the scheduled job and so on.


Run a data app deployment

Once you have created a data app deployment to see the list of *Deployments* you can go on the specific workflow page.

Click the  icon on the row corresponding to the deployment you want to execute, and click *Run*.

Manage access to data apps

You can change who has access to data apps.

1. In the *Deployments* section on the workflow page, click the  icon.
2. Select *Manage access* from the menu.
3. A side panel opens. Here, you can type the name of the user you want to share the data app with and select the access level you want to give them.

You have the following options:


- *Manage*: The user can change access settings, delete the deployment, and run it.
- *Run & Inspect*: The user can run the deployment and [inspect the executed workflow](#), but cannot change access settings or delete it.


- *Run only*: The user can run the deployment but cannot change access settings or delete it. This also shows the data app in the user's profile in the *Data Apps Portal* section.


Jobs

Every time a workflow is executed a job is created on KNIME Hub.

To see the list of all the jobs that are saved in memory for the ad hoc execution of a specific workflow go to the workflow page and on the right side menu click *Ad hoc jobs*.


To see the list of all the jobs that are saved in memory for each of the deployments created for a specific workflow, go to the workflow page and on the right side menu click *Deployments*. You can expand each deployment with the  icon on the left of the deployment.

Also you can go to your Pro profile page and find a list of all deployments you created. Also here you can click the  icon corresponding to a specific deployment to see all its jobs.

On each job you can click the  icon on the right of the corresponding job line in the list and perform the following operations:

- Open: For example you can open the job from a data app and look at the results in a new tab
- Save as workflow: You can save the job as a workflow in a space
- Inspect: A job viewer opens in a new tab. Here, you can investigate the status of jobs.
- Delete: You can delete the job
- Download logs: You can download the log files of a job - this feature allows for debugging in case the execution of a workflow did not work as expected.

Inspect an executed workflow

You can inspect the status of an executed workflow, for example if the workflow execution did not succeed, or the execution is taking a long time. Click the  icon for the desired job and select *Inspect*.



Be aware that this action will start up your executor, which, once the inspection is started will start the billing.

Here you can hover over info and error icons at the node's traffic light, to visualize the message, check where the execution stopped or the status of specific nodes. You can also

inspect the data at any node's output port.

The screenshot displays the KNIME workspace with a workflow containing a **CSV Reader** node and a **Filter for job** node. A red box highlights the **Filter for job** node with the message: "Contains one node with execution failure (Breakpoint #3) Breakpoint #3: Empty table".

The **CSV Reader** node's left sidebar shows its configuration, including a note about file handling and a list of file system connection options (e.g., "They are located on the machine that is running this node").

The output table, titled "1: File Table", shows 32,561 rows and 15 columns. The columns are: RowID, age, workclass, fnlwgt, education, education-num, marital-status, occupation, relations, race, sex, capital-gain, capital-loss, hours-per-week, native-country, and income. The table displays data for rows 1 through 11.

#	RowID	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relations	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
1	Row0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	
2	Row1	50	Self-emp-not-inc	83311	Bachelors	13	Married-div-s...	Exec-manage...	Husband	White	Male	0	0	13	United-States	
3	Row2	38	Private	215646	HS-grad	9	Divorced	Handlers-clea...	Not-in-family	White	Male	0	0	40	United-States	
4	Row3	53	Private	234721	11th	7	Married-div-s...	Handlers-clea...	Husband	Black	Male	0	0	40	United-States	
5	Row4	28	Private	338409	Bachelors	13	Married-div-s...	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	
6	Row5	37	Private	284582	Masters	14	Married-div-s...	Exec-manage...	Wife	White	Female	0	0	40	United-States	
7	Row6	49	Private	160187	9th	5	Married-spou...	Other-service	Not-in-family	Black	Female	0	0	16	Jamaica	
8	Row7	52	Self-emp-not-inc	209642	HS-grad	9	Married-div-s...	Exec-manage...	Husband	White	Male	0	0	45	United-States	
9	Row8	31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family	White	Female	14084	0	50	United-States	
10	Row9	42	Private	159449	Bachelors	13	Married-div-s...	Exec-manage...	Husband	White	Male	5178	0	40	United-States	
11	Row10	37	Private	280454	Some-college	10	Married-div-s...	Exec-manage...	Husband	Black	Male	0	0	80	United-States	

Figure 9. Inspect node messages and data at the output port of a node

You can also navigate inside components and metanodes and inspect the nodes.

The screenshot displays the KNIME workspace with a workflow containing a **Refresh Button Widget**, a **String Widget**, a **Component Input**, a **Row Filter**, a **Breakpoint** node, and a **Table View**. A red box highlights the **Breakpoint** node with the message: "Execute failed: Empty table".

The **Breakpoint** node's left sidebar shows its configuration, including a note about halting execution and a list of input ports.

The output table, titled "1: Output table", shows no data and 15 columns. The columns are: age, workclass, fnlwgt, education, education-num, marital-status, occupation, relationship, race, sex, capital-gain, capital-loss, hours-per-week, native-country, and income. The table displays data for rows 1 through 11.

age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
Number (Integer)	String	Number (Integer)	String	Number (Integer)	String	String	String	String	String	Number (Integer)	Number (Integer)	Number (Integer)	String	String

Figure 10. Navigate inside components and metanodes to inspect the encapsulated nodes

Extensions available for execution on KNIME Community Hub

All the extensions in the following list are available to run workflows on KNIME Community Hub.

- KNIME Expressions
- KNIME Executor connector
- KNIME Remote Workflow Editor for Executor
- KNIME Remote Workflow Editor
- KNIME Hub Additional Connectivity (Labs)
- KNIME Hub Integration
- KNIME Active Learning
- KNIME AI Assistant (Labs)
- KNIME Autoregressive integrated moving average (ARIMA)
- KNIME Basic File System Connectors
- KNIME Views
- KNIME Nodes for Wide Data (many columns)
- KNIME Big Data Connectors
- KNIME Databricks Integration
- KNIME Extension for Big Data File Formats
- KNIME Extension for Apache Spark
- KNIME Extension for Local Big Data Environments
- KNIME Chromium Embedded Framework (CEF) Browser
- KNIME Integrated Deployment
- KNIME Base Chemistry Types & Nodes
- KNIME Amazon Athena Connector
- KNIME Amazon DynamoDB Nodes
- KNIME Amazon Cloud Connectors
- KNIME Amazon Machine Learning Integration
- KNIME Amazon Redshift Connector And Tools
- KNIME Conda Integration
- KNIME Columnar Table Backend

- KNIME Streaming Execution (Beta)
- KNIME BigQuery
- KNIME JDBC Driver For Oracle Database
- KNIME Microsoft JDBC Driver For SQL Server
- KNIME Vertica Driver
- KNIME Database
- KNIME Data Generation
- KNIME Connectors for Common Databases
- KNIME Distance Matrix
- KNIME Deep Learning - Keras Integration
- KNIME Deep Learning - ONNX Integration
- KNIME Deep Learning - TensorFlow Integration
- KNIME Deep Learning - TensorFlow 2 Integration
- KNIME Email Processing
- KNIME Ensemble Learning Wrappers
- KNIME Expressions
- KNIME Azure Cloud Connectors
- KNIME Box File Handling Extension
- KNIME Chemistry Add-Ons
- KNIME External Tool Support
- KNIME H2O Snowflake Integration
- KNIME H2O Machine Learning Integration
- KNIME H2O Machine Learning Integration - MOJO Extension
- KNIME Extension for MOJO nodes on Spark
- KNIME H2O Sparkling Water Integration
- KNIME Itemset Mining
- KNIME Math Expression (JEP)
- KNIME Indexing and Searching
- KNIME MDF Integration
- KNIME Office 365 Connectors

- KNIME SharePoint List
- KNIME Open Street Map Integration
- KNIME SAS7BDAT Reader
- KNIME Excel Support
- KNIME Power BI Integration
- KNIME Tableau Integration
- KNIME Textprocessing
- KNIME Twitter Connectors
- KNIME External Tool Support (Labs)
- KNIME Google Connectors
- KNIME Google Cloud Storage Connection
- KNIME Javasnipet
- KNIME Plotly
- KNIME Quick Forms
- KNIME JavaScript Views
- KNIME JavaScript Views (Labs)
- KNIME JSON-Processing
- KNIME Extension for Apache Kafka (Preview)
- KNIME Machine Learning Interpretability Extension
- KNIME MongoDB Integration
- KNIME Neighborgram & ParUni
- KNIME Network Mining distance matrix support
- KNIME Network Mining
- KNIME Optimization extension
- KNIME Python Integration
- KNIME Interactive R Statistics Integration
- KNIME Report Designer (BIRT)
- KNIME Reporting
- KNIME REST Client Extension
- KNIME Salesforce Integration

- KNIME SAP Integration based on Theobald Xtract Universal
- KNIME Git Nodes
- KNIME Semantic Web
- KNIME Snowflake Integration
- KNIME Statistics Nodes
- KNIME Statistics Nodes (Labs)
- KNIME Timeseries nodes
- KNIME Timeseries (Labs)
- KNIME Modern UI
- KNIME Parallel Chunk Loop Nodes
- KNIME Weak Supervision
- KNIME Webanalytics
- KNIME XGBoost Integration
- KNIME XML-Processing
- SmartSheet extension
- KNIME AI Extension (Labs)
- KNIME Web Interaction (Labs)
- KNIME Nodes for Scikit-Learn (sklearn) Algorithms
- Geospatial Analytics Extension for KNIME
- KNIME Giskard Extension
- KNIME Presidio Extension
- RDKit Nodes Feature
- Vernalis KNIME Nodes
- Slack integration
- Continental Nodes for KNIME
- Lhasa public plugin



To know more about the different extensions see the [Extensions and Integrations Guide](#).

Data Apps Portal

To access the *Data Apps Portal* you can either go to <https://apps.hub.knime.com> and sign in with your KNIME Hub account or click on the *Data Apps Portal* link in the menu on your profile page in the KNIME Hub.

Every signed in user can access to this page to see all the data apps that have been shared with them, execute them at any time, interact with the workflow via a user interface, without the need to build a workflow or even know what happens under the hood.

To execute a data app click the corresponding tile. A new page will open in the browser showing the user interface of the data app.

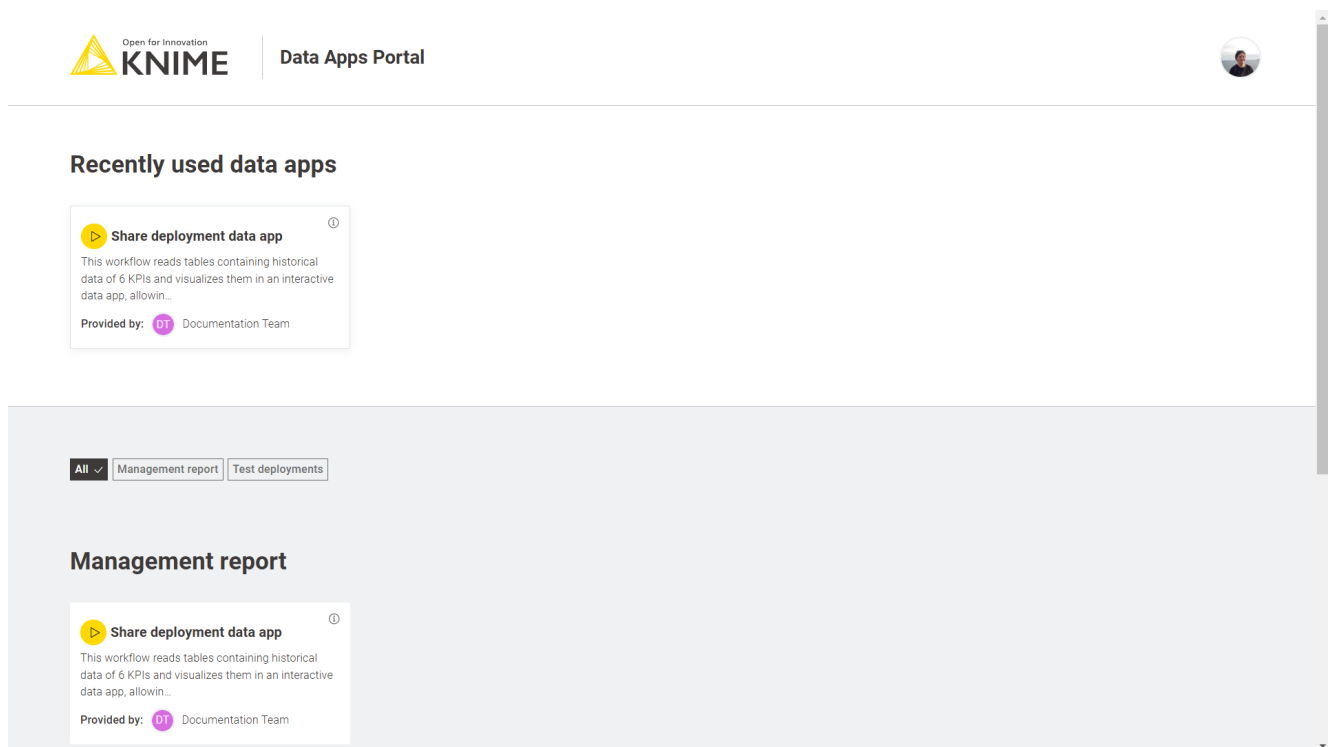


Figure 11. The Data Apps Portal



To share a data app deployment with someone follow the instructions in the [Manage access to data apps](#) section.

Access data from different sources

This section provides information on how to build your workflows to connect to different data sources and which kind of authentication method is recommended for a particular data source, to then run the workflows ad hoc or schedule them on KNIME Community Hub.

Access data from Google Cloud

Refer to the [KNIME Google Cloud Integration User Guide](#) to know how to create an API Key and to learn more about the general usage of the Google Authenticator node.

To access data from Google Cloud the first step when building your workflow is to use a Google Authenticator node. One of the possibilities is then to use the API Key authentication method. When using the API Key authentication method in the Google Authenticator node, the API key file will need to be stored somewhere where it is accessible to the node in order for the execution to be successful.

For example, if your folder structure on the KNIME Community Hub (KCH) looks something like the one below, meaning that your API Key file is stored in the following path:

```
<space> -> <folder> -> <api-key.fileformat>
```

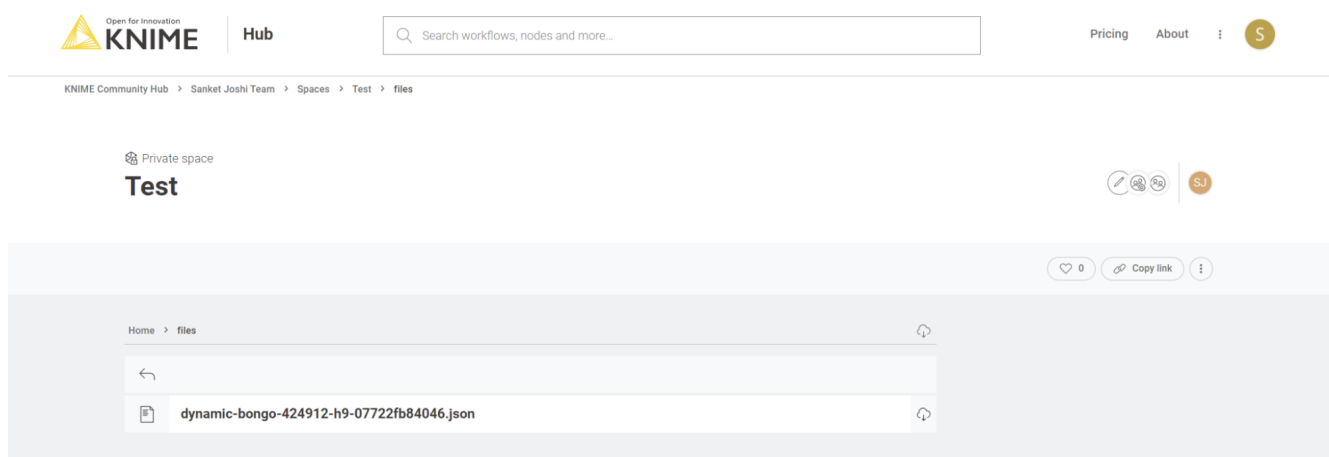
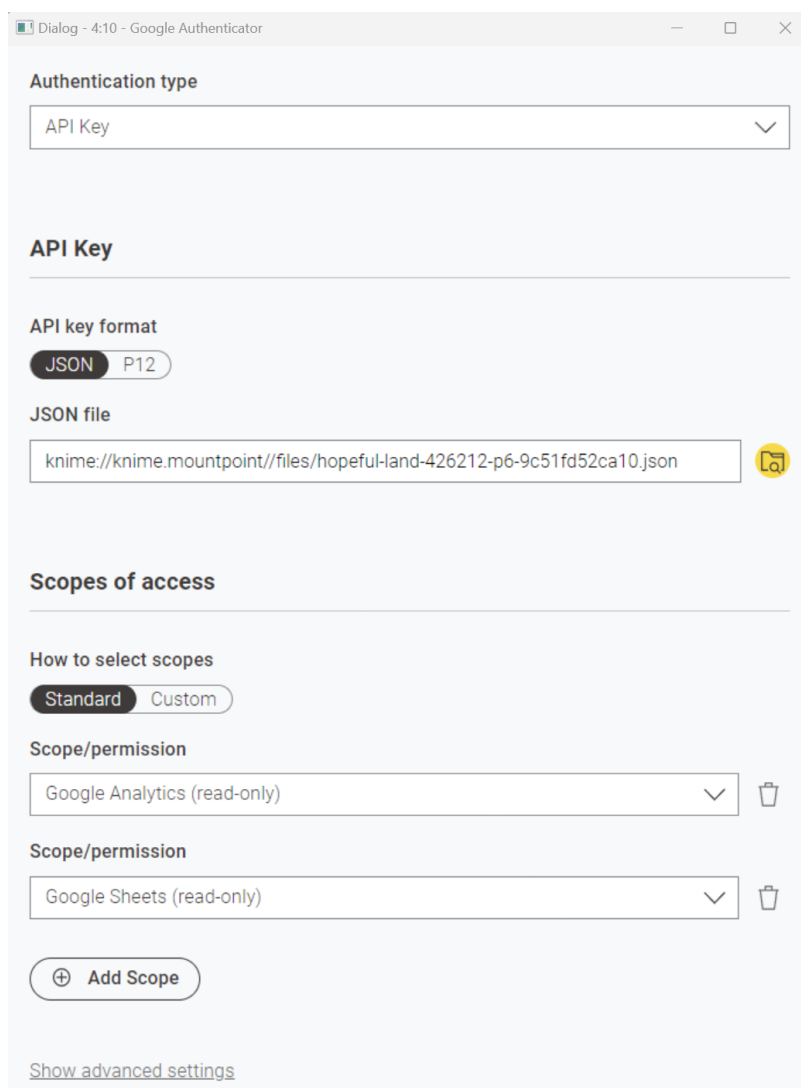


Figure 12. Example folder structure for API Key file storage

then, you can access the API key using either one of the following paths in the Google Authenticator node configuration dialog.

- **Mountpoint relative:**

```
knime://knime.mountpoint//<folder>/<api-key.fileformat>
```

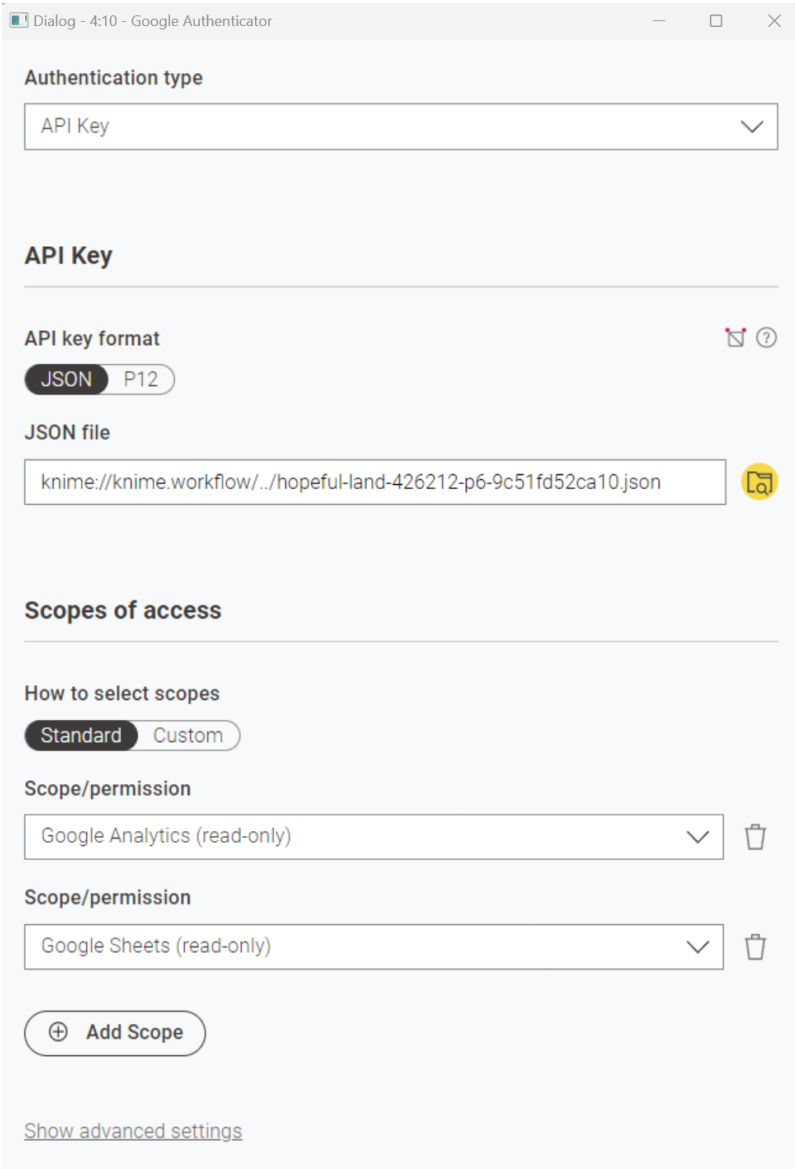



The image shows a 'Dialog - 4:10 - Google Authenticator' window. It has a title bar with standard window controls. The main content is divided into sections: 'Authentication type' with a dropdown menu set to 'API Key'; 'API Key' section with 'API key format' buttons for 'JSON' (selected) and 'P12'; 'JSON file' with a text input field containing 'knime://knime.mountpoint/files/hopeful-land-426212-p6-9c51fd52ca10.json' and a file selection icon; 'Scopes of access' section with 'How to select scopes' buttons for 'Standard' (selected) and 'Custom'; two 'Scope/permission' dropdown menus, both set to 'Google Analytics (read-only)' and 'Google Sheets (read-only)' respectively, each with a trash icon; an 'Add Scope' button with a plus icon; and a 'Show advanced settings' link at the bottom.

Figure 13. Configuration dialog of Google Authenticator for mountpoint relative path

- **Workflow relative:**

```
knime://knime.workflow/../../<api-key.fileformat>
```



The screenshot shows a 'Dialog - 4:10 - Google Authenticator' window. It has a title bar with standard window controls. The main content is divided into sections: 'Authentication type' with a dropdown menu set to 'API Key'; 'API Key' section with 'API key format' buttons for 'JSON' (selected) and 'P12'; 'JSON file' with a text input field containing the path 'knime://knime.workflow/ ../hopeful-land-426212-p6-9c51fd52ca10.json' and a file icon; 'Scopes of access' section with 'How to select scopes' buttons for 'Standard' (selected) and 'Custom'; and a list of 'Scope/permission' items, each with a dropdown menu (showing 'Google Analytics (read-only)' and 'Google Sheets (read-only)') and a trash icon. At the bottom of the list is an 'Add Scope' button with a plus icon. A 'Show advanced settings' link is at the very bottom.

Figure 14. Configuration dialog of Google Authenticator for workflow relative path

- **Workflow Data Area:**

```
knime://knime.workflow/data/<api-key.fileformat>
```

To use this path, you need to make sure you have stored the JSON or p12 file in the workflow's **Data** folder.

If the folder does not exist, create one, store the file inside, and then upload the workflow to the KNIME Community Hub.

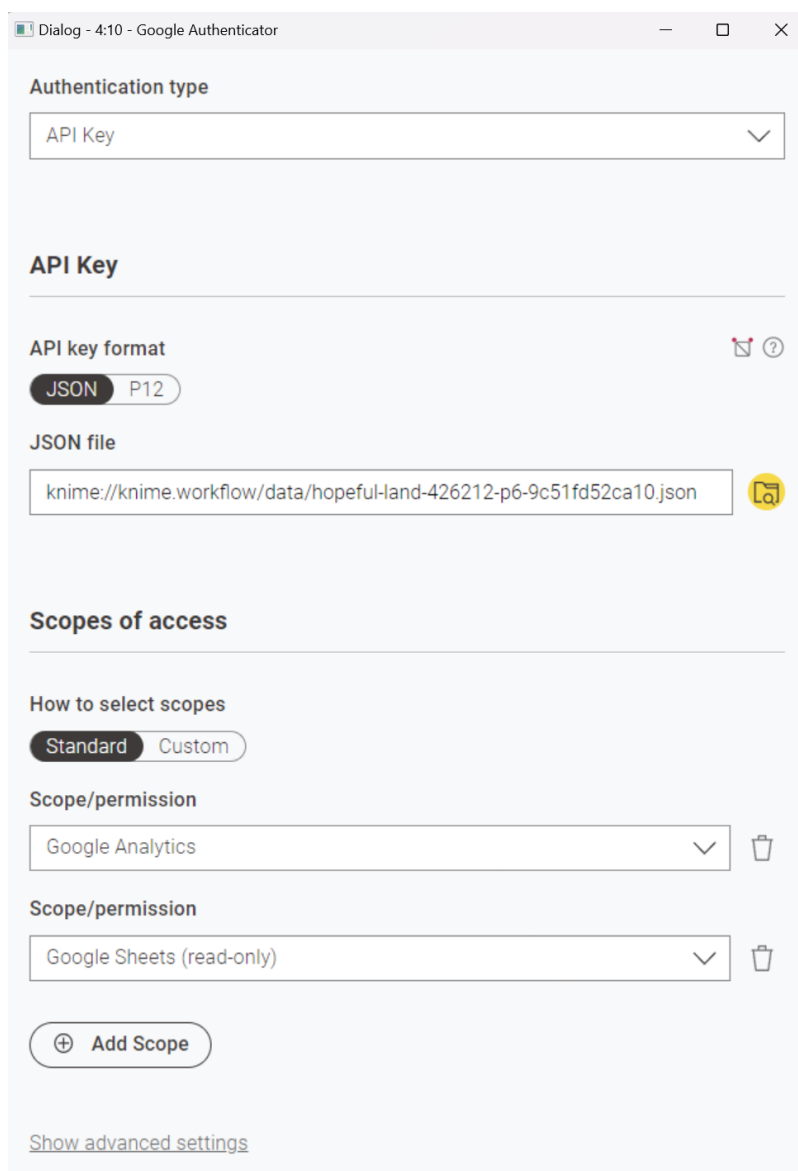


Figure 15. Configuration dialog of Google Authenticator for workflow data area path

Access data from Google BigQuery

In order to have access to data on Google BigQuery, the first step is to authenticate with Google, using the Google Authenticator node as described above. Then you can connect to Google BigQuery using the [Google BigQuery Connector](#) node.

However, due to license restrictions, you first need to install version 1.5.4 of the Google Big Query JDBC driver, which you can download [here](#). Further details about how to configure the driver, its limitations and support can be found in the [Google documentation](#).

Once you have downloaded the driver, extract it into a directory and register it in your KNIME Analytics Platform as described [here](#).

Your driver registration dialog should look like this, where <path-to-

folder\SimbaJDBCDriverforGoogleBigQuery42_1.5.4.1008 is the directory that contains all the JAR files.

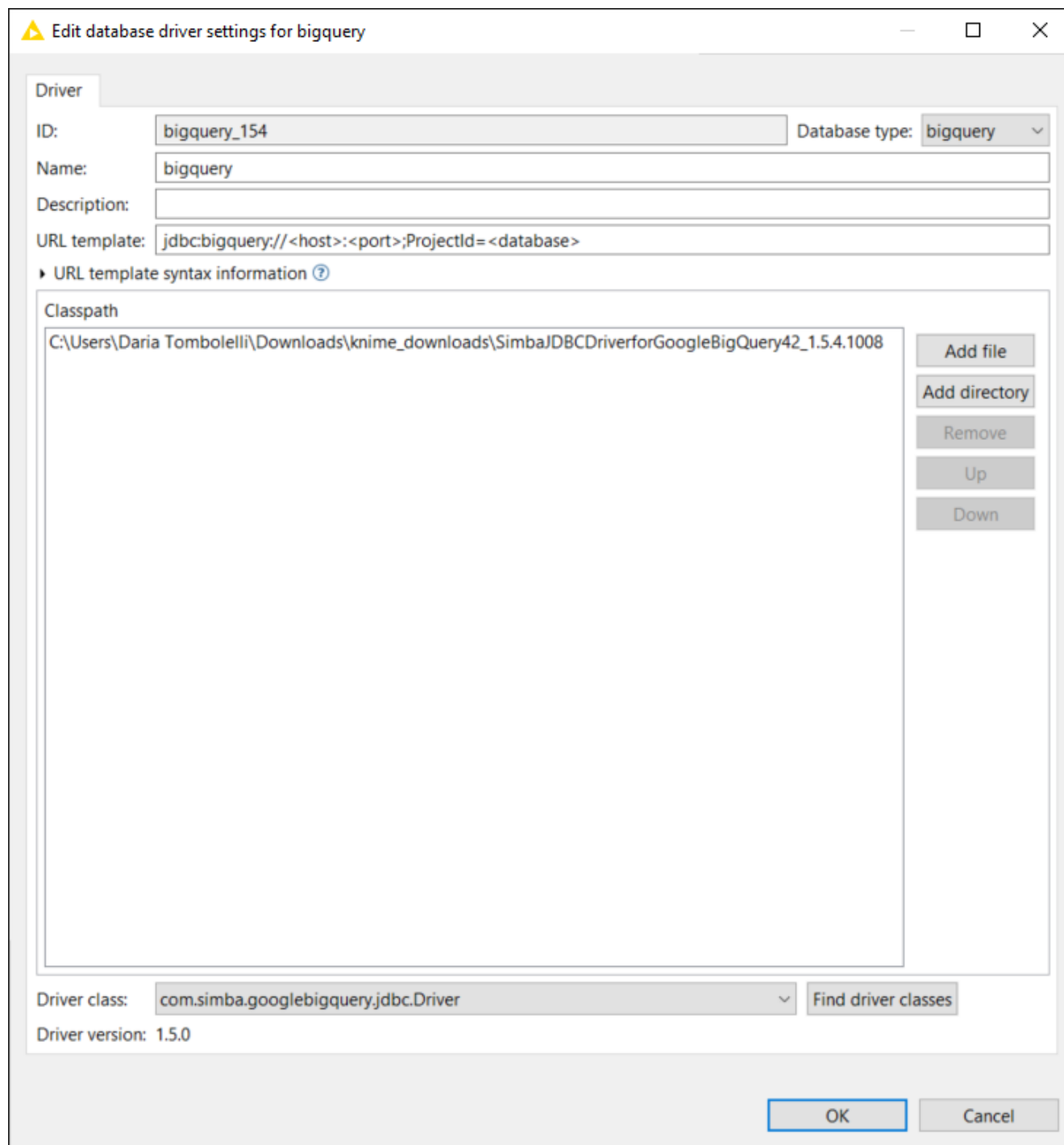


Figure 16. Configuration dialog to register a new database driver

Then you can configure the Google BigQuery Connector node. Make sure to select the option *Use latest driver version available as Database Driver* in the Google BigQuery Connector node. In this way the node will use the latest driver version available locally or on the KNIME Community Hub disregard of where the node is executed.

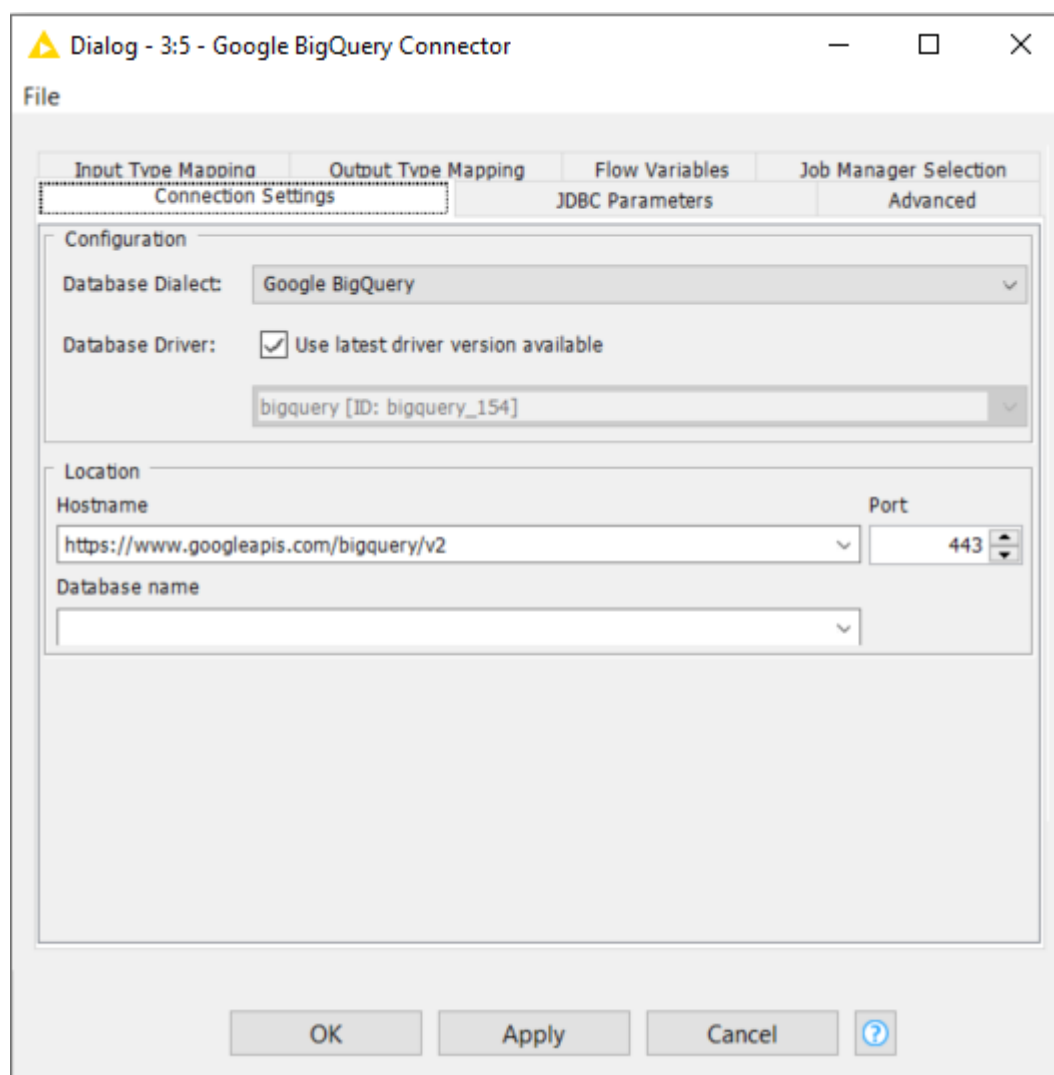


Figure 17. Configuration dialog of the Google BigQuery Connector node

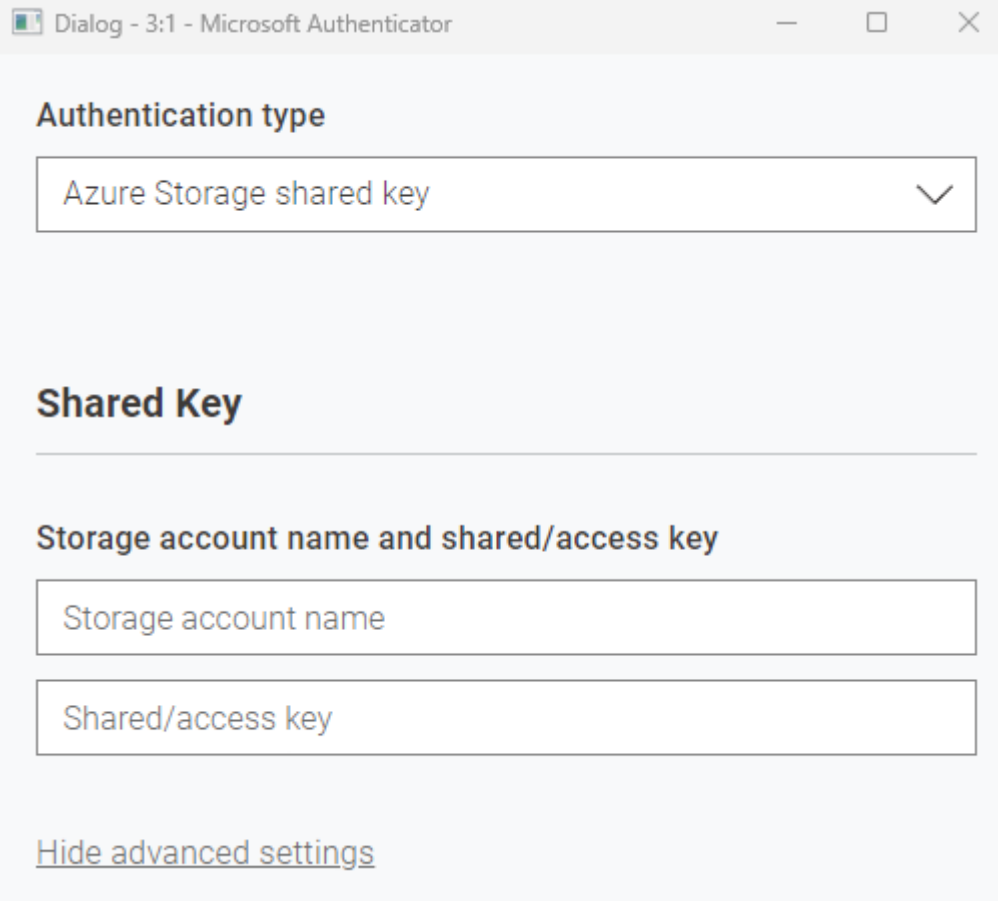
Access data from Azure

Refer to the [KNIME Azure Integration User Guide](#) for setting up the Azure integration with KNIME.

To access data from Azure, the first step when building your workflow is to use a [Microsoft Authenticator](#) node.

When configuring the node, select **Azure Storage Shared Key** as authentication type.

Enter the **Storage Account Name** and **Shared/Access key** and connect to Azure.

The image shows a Windows-style dialog box titled "Dialog - 3:1 - Microsoft Authenticator". It has standard minimize, maximize, and close buttons in the top right corner. The dialog is divided into sections. The first section is titled "Authentication type" and contains a dropdown menu with "Azure Storage shared key" selected and a downward arrow. The second section is titled "Shared Key" and is separated by a horizontal line. Below this, there is a section titled "Storage account name and shared/access key" which contains two text input fields: "Storage account name" and "Shared/access key". At the bottom of the dialog, there is a link that says "Hide advanced settings".

Dialog - 3:1 - Microsoft Authenticator

Authentication type

Azure Storage shared key

Shared Key

Storage account name and shared/access key


Storage account name


Shared/access key

[Hide advanced settings](#)

Figure 18. Configuration dialog of Microsoft Authenticator node

The credentials could be stored inside the node, which is not recommended, or passed to the node using the **Credentials Configuration** node. This will allow you to pass the credentials while deploying the workflow, as shown in the below snapshot.



 **Create schedule**


Schedule your workflow to run automatically at selected times.

Schedule options



Define when the workflow should be executed.

Initial execution

2024-07-03





11

:

00

☐ Repeat execution

Set schedule details Set >

Email notifications

Setup email notifications of the workflow execution.

☐ Enable email notifications

Configuration options

Control the parametrization of the workflow execution.

Label

Username

Password

Advanced settings

Change job lifecycle behavior, timeouts, Set >

Cancel

Create

Figure 19. Create a schedule deployment using credentials

Once authenticated, use **Azure Data Lake Storage Gen2 Connector** or **Azure Blob Storage Connector** nodes and select the working directory from where you want to access the data.

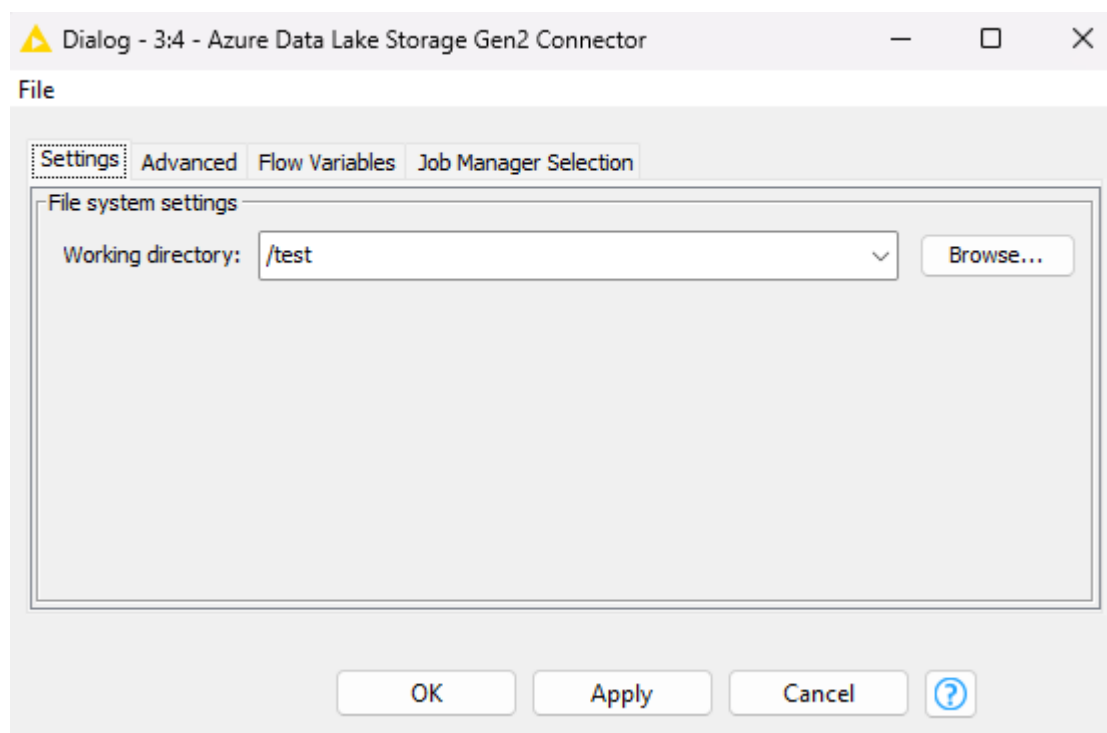


Figure 20. Configuration dialog of Azure Data Lake Storage Gen2 Connector node

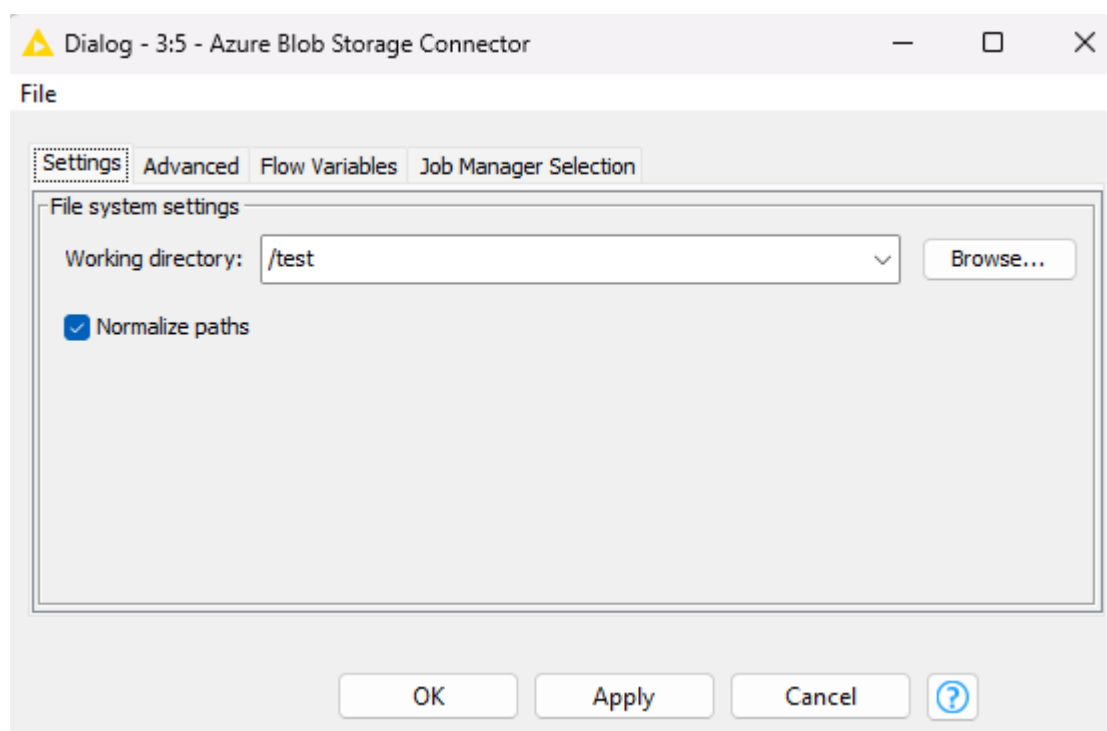


Figure 21. Configuration dialog of Azure Blob Storage Connector node

Enter the file name in the respective file reader node and read the file.

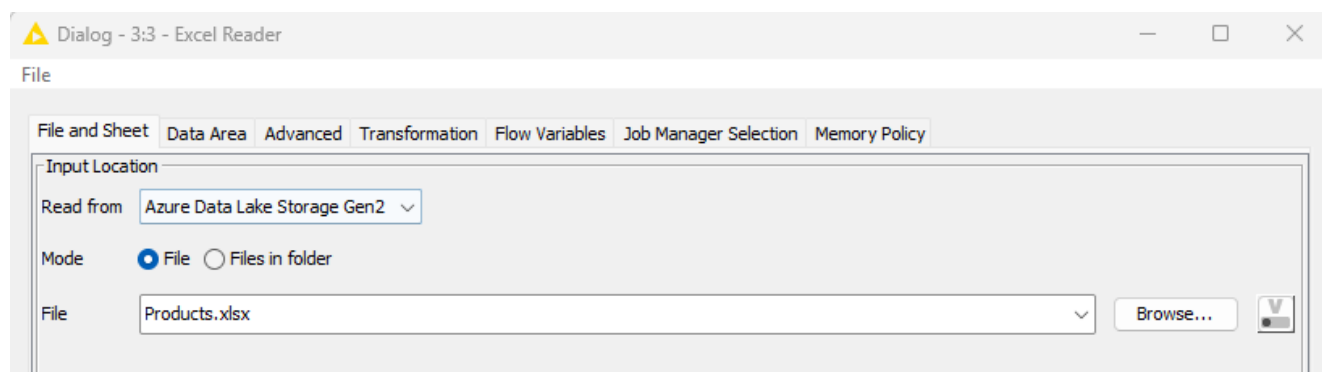



Figure 22. Configuration dialog of the File Reader node


Access data from Snowflake

Refer to the [KNIME Snowflake Extension Guide](#) for setting up Snowflake with KNIME.

To access data from Snowflake, the first step when building your workflow is to use a [Snowflake Connector](#) node.

The credentials could be stored inside the node, which is not recommended, or passed to the node using the [Credentials Configuration](#) node. This will allow you to pass the credentials while deploying the workflow.



 **Create schedule**


Schedule your workflow to run automatically at selected times.

Schedule options



Define when the workflow should be executed.

Initial execution

2024-07-03





11

:

00

☐ Repeat execution

Set schedule details Set >

Email notifications

Setup email notifications of the workflow execution.

☐ Enable email notifications

Configuration options

Control the parametrization of the workflow execution.

Label

Username

Password

Advanced settings

Change job lifecycle behavior, timeouts, Set >

Cancel

Create

Figure 23. Create a schedule deployment using credentials

Access data from Amazon Athena

In order to have access to data on Amazon Athena, the first step is to authenticate with Amazon, using the [Amazon Authenticator](#). Then you can connect to Amazon Athena using

the [Amazon Athena Connector](#) node.

However, due to license restrictions, you first need to install version 2.1.1 of the Amazon Athena JDBC driver, which you can download [here](#). Further details about how to configure the driver, its limitations and support can be found in the [Amazon Athena documentation](#).

Once you have downloaded the driver, extract it into a directory and register it in your KNIME Analytics Platform as described [here](#).

Your driver registration dialog should look like this, where `<path-to-file>\AthenaJDBC42-2.1.1.1001.jar` is the JAR file you downloaded.

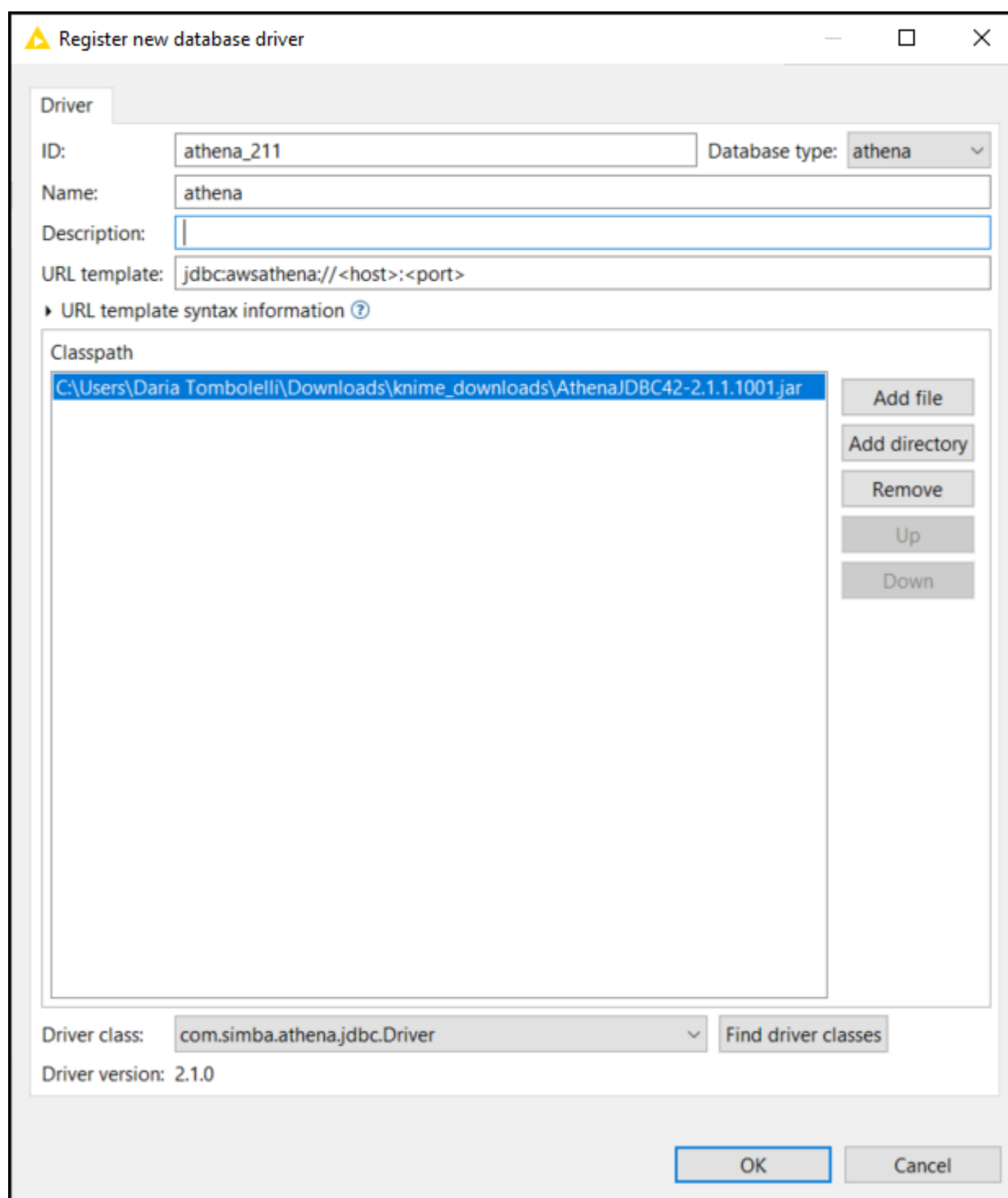


Figure 24. Configuration dialog to register a new database driver

Then you can configure the Amazon Athena Connector node. Make sure to select the option *Use latest driver version available as Database Driver* in the Amazon Athena Connector node. In this way the node will use the latest driver version available locally or on the KNIME Community Hub disregard of where the node is executed.

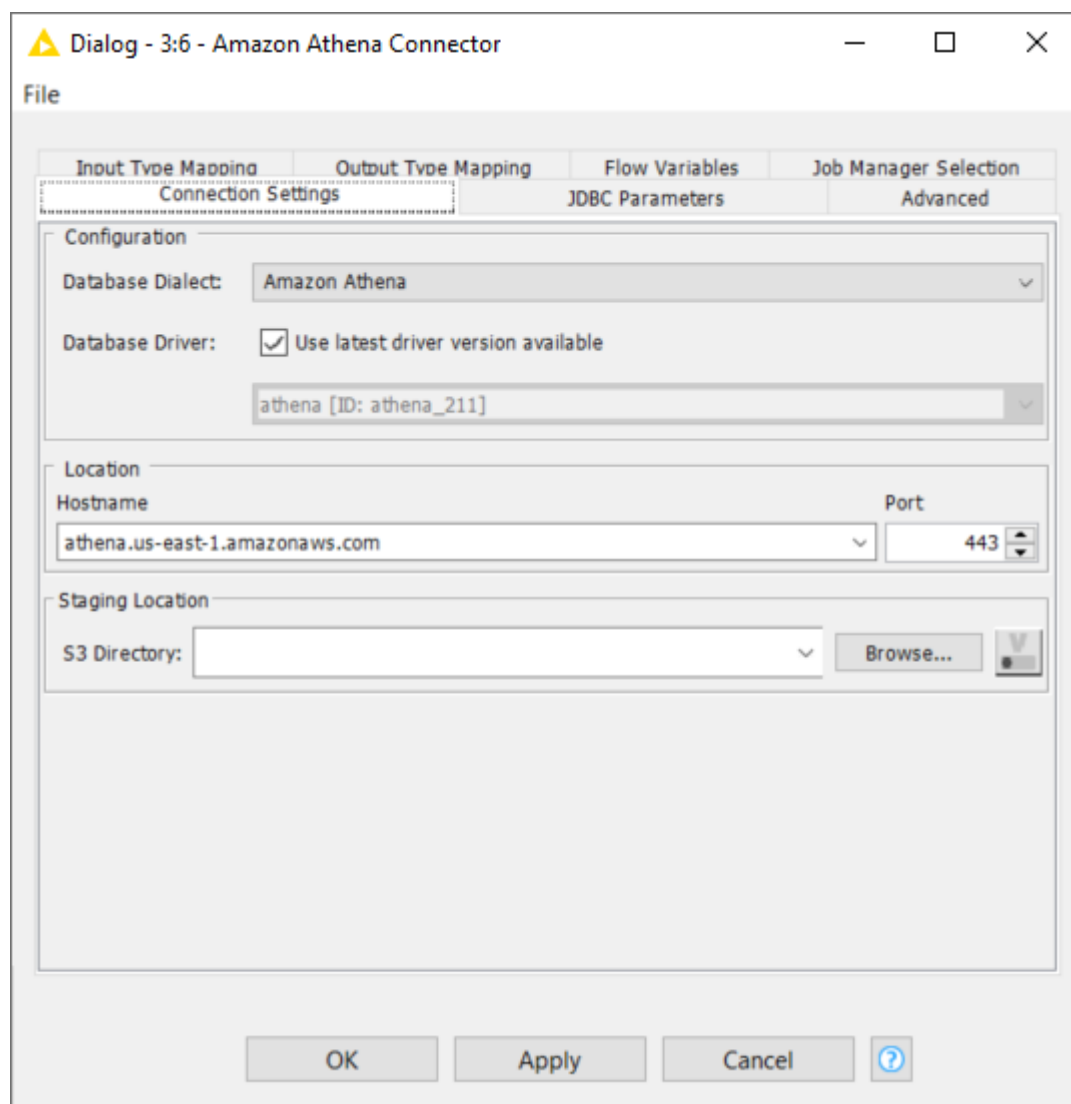


Figure 25. Configuration dialog of the Amazon Athena Connector node

Advanced settings

IP whitelisting



Relying solely on IP whitelisting for security is insufficient and potentially risky.

KNIME Community Hub has a dedicated static IP address that executors use: 3.66.133.75. Whitelisting this IP address ensures that your workflows can reach the remote data location.

As this IP address is **shared** between **all executors** on Community Hub, IP whitelisting in itself as a secure connection is **not recommend** (as others on the Hub might also access your data).

It is **strongly advised** to use it in **combination** with other **security mechanisms**. For example: credential based authentication. In this case, rotating/changing the credentials periodically is recommended as a best practice.

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