Getting Started with SAP Data Intelligence 3.0 Trial edition
# Table of Contents

1. Solution Information .................................................................................................................. 3
   1.1. Product Overview .............................................................................................................. 3
   1.2. Architecture Overview ....................................................................................................... 4
   1.3. Installed Products .............................................................................................................. 4

2. Provisioning the Solution ........................................................................................................... 5
   2.1. Prerequisite: Cloud Provider Account .............................................................................. 5
       2.1.1. Amazon Web Services ............................................................................................... 5
       2.1.2. Microsoft Azure ........................................................................................................ 6
       2.1.3. Google Cloud Platform ............................................................................................. 6
   2.2. Accessing the SAP Cloud Appliance Library ..................................................................... 8
   2.3. Creating a Solution Instance ............................................................................................... 8

3. Accessing the Solution .............................................................................................................. 12
   3.1. Overview ............................................................................................................................ 12
   3.2. Determining the external IP addresses of your solution instance ..................................... 13
   3.3. Accessing SAP Data Intelligence on Amazon Web Services ............................................. 14
   3.4. Accessing SAP Data Intelligence on Microsoft Azure ...................................................... 15
   3.5. Accessing SAP Data Intelligence on Google Cloud Platform .......................................... 16
       3.5.1. Access via Windows Remote Desktop ......................................................................... 16
       3.5.2. Access via Web browser of local computer ................................................................. 18
   3.6. Accessing the Jump Box for troubleshooting .................................................................... 19

4. Licenses ..................................................................................................................................... 20
   4.1. Running Solution with a Product License Key .................................................................... 20
   4.2. Running Solution as a Trial for a Free Period .................................................................... 20

5. Security Aspects ....................................................................................................................... 21

6. Troubleshooting ....................................................................................................................... 22
   6.1.1. Pipeline graphs ............................................................................................................... 22
   6.1.2. Obtaining kubeconfig file ............................................................................................. 22
   6.1.3. Accessing SAP Data Intelligence namespace ............................................................... 23
   6.1.4. Accessing application logs ............................................................................................ 23
   6.1.5. Enable access to Kubernetes Dashboard ........................................................................ 24

7. Appendix ................................................................................................................................... 25
1 Solution Information

1.1 Product Overview

SAP Data Intelligence, trial edition is a pre-configured appliance for evaluating and testing SAP Data Intelligence, the successor to SAP Data Hub.

In addition to all the data sharing, pipelining and orchestration capabilities that were originally featured in SAP Data Hub, SAP Data Intelligence introduces a way to manage machine learning models with built-in tools for data governance, management, and transparency:

- **Data Pipelines** are data-driven applications consisting of reusable and configurable operators.
- **Data Workflows** orchestrate processes across a data landscape.
- **Data Governance** allows you to extract metadata from connected data stores. It supports discovery, profiling and search.
- **Automated Machine Learning** (AutoML) lets you automate the creation of complex machine learning pipelines with intuitive tools for configuring machine learning artifacts.

To learn more about SAP Data Intelligence, refer to:

1.2 Architecture Overview

The SAP Data Intelligence, trial edition currently is available on Amazon Web Services as well as Google Cloud Platform. While creating a solution instance of SAP Data Intelligence, trial edition, the SAP Cloud Appliance library creates:

<table>
<thead>
<tr>
<th></th>
<th>Amazon Web Services</th>
<th>Microsoft Azure</th>
<th>Google Cloud Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump Box to orchestrate the instances creation</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kubernetes cluster to run SAP Data Intelligence</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Windows Remote Desktop to easily connect to SAP Data Intelligence</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>

The Jump Box is used during the creation of the solution instance to run the installer for SAP Data Intelligence. As a user of the SAP Data Intelligence, trial edition you typically do not have to access the Jump Box except for troubleshooting (in rare cases). The Jump Box is automatically terminated 4 hours after the installation has finished successfully (afterwards it cannot be accessed anymore).

SAP Data Intelligence leverages container technology, in particular Docker and Kubernetes. It allows to query large amounts of data (in distributed storages like Amazon S3, Azure Data Lake (ADL), Azure Storage Blobs (WASB), HDFS and Google Cloud Storage (GCS)) and to run highly scalable data flows and data-driven applications. As part of this pre-configured appliance, SAP Data Intelligence uses a Kubernetes cluster with three worker nodes.

The Windows Remote Desktop allows you to easily connect to SAP Data Intelligence without having to maintain your local hosts file (which is otherwise necessary on Google Cloud Platform; details are described in chapter 3.5.1).

1.3 Installed Products

The SAP Data Intelligence, trial edition leverages release 3.0 (Build 3.0.12), Full platform stack
2 Provisioning the Solution

2.1 Prerequisite: Cloud Provider Account

To use the SAP Cloud Appliance Library in order to create a solution instance of SAP Data Intelligence, trial edition, you need to have access to an Amazon Web Services account, Microsoft Azure account, or a Google Cloud Platform project. You need to link your account in the SAP Cloud Appliance Library to those.

2.1.1 Amazon Web Services

If you like to use Amazon Web Services, you need to create a user which you use to access your Amazon Web Services account through the SAP Cloud Appliance Library. It needs to have the following policies:

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AmazonEC2ContainerRegistryFullAccess</td>
<td>Provides administrative access to Amazon ECR resources</td>
</tr>
<tr>
<td>AmazonEC2FullAccess</td>
<td>Provides full access to Amazon EC2 via the AWS Management Console.</td>
</tr>
<tr>
<td>AmazonS3FullAccess</td>
<td>Provides full access to all buckets via the AWS Management Console.</td>
</tr>
<tr>
<td>AutoScalingFullAccess</td>
<td>Provides full access to Auto Scaling.</td>
</tr>
<tr>
<td>ElasticLoadBalancingFullAccess</td>
<td>Provides full access to Amazon ElasticLoadBalancing, and limited access to other services necessary to provide ElasticLoadBalancing features.</td>
</tr>
<tr>
<td>IAMFullAccess</td>
<td>Provides full access to IAM via the AWS Management Console.</td>
</tr>
</tbody>
</table>

In addition, you need to create two policies for AWS Elastic Container Service for Kubernetes and AWS CloudFormation:

- AWS Elastic Container Service for Kubernetes

```json
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "eks:*",
      "Resource": "*"
    }
  ]
}
```
• AWS CloudFormation

```json
{
"Version": "2012-10-17",
"Statement": [
{
"Effect": "Allow",
"Action": "cloudformation:*",
"Resource": "*"
}
]
}
```

For more information about how to link the SAP Cloud Appliance Library with an Amazon Web Services account, refer to the FAQ page.

2.1.2 Microsoft Azure

If you like to use Microsoft Azure, you need a subscription to a Microsoft Azure account. The subscription needs to fulfill the one of the following requirements:

- You are an administrator on the subscription, i.e. your user has the role Owner and has access to scope /subscriptions/<your subscription ID>.
- Your Microsoft Azure user has been assigned the role User Access Administrator and has access to scope /subscriptions/<your subscription ID>.

When you connect from SAP Cloud Appliance Library to Microsoft Azure, you need to use authorization type **Extended Authorization for Kubernetes Cluster**.

For more information about how to link the SAP Cloud Appliance Library with a Microsoft Azure account, refer to the FAQ page.

2.1.3 Google Cloud Platform

If you like to use Google Cloud Platform, you need to enable certain APIs and to create a service account with certain roles.

**Required APIs**

To create a solution instance of SAP Data Intelligence, you need to enable the following APIs in your Google Cloud Platform project.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Resource Manager API</td>
<td>The Google Cloud Resource Manager API provides methods for creating, reading, and updating project metadata.</td>
</tr>
<tr>
<td>Compute Engine API</td>
<td>Creates and runs virtual machines on Google Cloud Platform.</td>
</tr>
<tr>
<td>Kubernetes Engine API</td>
<td>The Google Kubernetes Engine API is used for building and managing container-based applications, powered by the open source Kubernetes technology.</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Container Registry API</td>
<td>Google Container Registry provides secure, private Docker image storage on Google Cloud Platform. Our API follows the Docker Registry API specification, so we are fully compatible with the Docker CLI client, as well as standard tooling using the Docker Registry API.</td>
</tr>
<tr>
<td>Google Cloud Storage</td>
<td>Google Cloud Storage is a RESTful service for storing and accessing your data on Google’s infrastructure.</td>
</tr>
</tbody>
</table>

**Service account and roles**

The service account which you use to access your Google Cloud Platform project through the SAP Cloud Appliance Library needs to have the following roles.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Instance Admin (v1)</td>
<td>Full control of Compute Engine instances, instance groups, disks, snapshots, and images. Read access to all Compute Engine networking resources.</td>
</tr>
<tr>
<td>Compute Network Admin</td>
<td>Full control of Compute Engine networking resources.</td>
</tr>
<tr>
<td>Compute Security Admin</td>
<td>Full control of Compute Engine security resources.</td>
</tr>
<tr>
<td>Kubernetes Engine Admin</td>
<td>Full management of Kubernetes Clusters and their Kubernetes API objects.</td>
</tr>
<tr>
<td>Service Account User</td>
<td>Create VMs and other GCP tasks with a service account. Users cannot impersonate the account directly as they can with Service Account Actor role.</td>
</tr>
<tr>
<td>Storage Admin</td>
<td>Full control of GCS resources.</td>
</tr>
</tbody>
</table>

For more information about how to link the SAP Cloud Appliance Library with a Google Cloud Platform project, refer to the FAQ page.
2.2 Accessing the SAP Cloud Appliance Library

Open the SAP Cloud Appliance Library in your web browser using the following link: https://cal.sap.com.
If you are a first-time user of SAP Cloud Appliance Library, familiarize yourself with its basic concepts and how to work with the user interface by reading the documentation.

2.3 Creating a Solution Instance

The following steps guide you through the creation of a solution instance of SAP Data Intelligence, trial edition 3.0. The screenshots have been taken for Google Cloud Platform (and might look slightly different when using Amazon Web Services or Microsoft Azure as cloud provider). Wherever you need more information, refer to the documentation.

1. Navigate to solution **SAP Data Intelligence, trial edition 3.0**

2. Press the **Create Instance** button and then switch to **Advanced Mode** (pressing the button in the lower right of the screen).

3. Select the cloud provider account. Move to step 2 by pressing the **Step 2** button.

4. Enter **Name**, **Number of Instances** (1), **Region**, **Zone**\(^1\), **Network** and **Subnet**. Ensure that the checkbox **Public Static IP Address** is marked. Move to step 3 by pressing the **Step 3** button.

5. Next check the parameters of the **Jump Box** and the **Windows Remote Desktop**\(^2\). Normally you do not have to change anything. Move to step 4 by pressing the **Step 4** button.

---

\(^1\) only on Google Cloud Platform  
\(^2\) only on Google Cloud Platform
### Access Points

<table>
<thead>
<tr>
<th>Virtual Machine</th>
<th>Service</th>
<th>Port Range</th>
<th>IP Range</th>
<th>Type</th>
<th>Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump Box</td>
<td>SSH</td>
<td>22</td>
<td>0.0.0.0/0</td>
<td>Default</td>
<td>✓</td>
</tr>
<tr>
<td>Windows Remote Desktop</td>
<td>RDP</td>
<td>3389</td>
<td>0.0.0.0/0</td>
<td>Default</td>
<td>✓</td>
</tr>
</tbody>
</table>

Select size and access points of the virtual machines:

### Sizes

<table>
<thead>
<tr>
<th>Virtual Machine</th>
<th>Size</th>
<th>Activate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump Box</td>
<td>n1-standard-2 (2 cores, 7.5GB memory, SSD)</td>
<td>✓</td>
</tr>
<tr>
<td>Windows Remote Desktop</td>
<td>n1-standard-4 (4 cores, 15GB memory, HDD)</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Expandable Storage

<table>
<thead>
<tr>
<th>Volume</th>
<th>Default Size</th>
<th>Additional Size</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump Box</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS Volume</td>
<td>60 GB</td>
<td>0 GB</td>
<td>60 GB</td>
</tr>
<tr>
<td>Windows Remote Desktop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS Volume</td>
<td>50 GB</td>
<td>0 GB</td>
<td>50 GB</td>
</tr>
<tr>
<td>frontend_config</td>
<td>2 GB</td>
<td>0 GB</td>
<td>2 GB</td>
</tr>
</tbody>
</table>
6. Next check the parameters of SAP Data Intelligence (Kubernetes Cluster, Cluster Subnets, Access Points). Normally you do not have to change anything. Move to step 5 by pressing the Step 5 button.

4. Kubernetes Cluster

<table>
<thead>
<tr>
<th>Number of Nodes</th>
<th>Size</th>
<th>Cluster Subnets</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4-core (8 GB memory, 100 GB)</td>
<td>[hrm-console-100.277.184.203:30000], [hrm-console-100.277.184.203:30000]</td>
</tr>
</tbody>
</table>

Access Points

<table>
<thead>
<tr>
<th>Service</th>
<th>Port Range</th>
<th>IP Range</th>
<th>Type</th>
<th>Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTTPS</td>
<td>443</td>
<td>0.0.0.0/0</td>
<td>Default</td>
<td>☑</td>
</tr>
</tbody>
</table>

7. Next define a solution password. We will refer to this as `<Master Password>` for the rest of this document. Move to step 6 by pressing the Step 6 button.

5. Solution Password

Set the master password for the solution instance. You can check the Getting Started Guide to learn where to use this password.

- Password: ********
- Retype Password: ********

The valid characters are: A-Z, a-z, 0-9, $, _, Â
The first 3 characters cannot be one of the following: A-Z, ã, ë, 
The password must contain at least 1 lowercase letter(s)
The password must contain at least 1 uppercase letter(s)
The password must contain at least 1 digit(s)
The password must be between 8 and 9 characters long

---

3 On Amazon Web Services only
8. Finally set up the schedule of the solution instance. You can define a schedule to suspend it. And you can set a termination date. Finally press the **Review** button.

9. Then (assuming you are happy with all parameters) press the **Create** button. **It will take between 3 to 5 hours to create the solution instance.** The startup time strongly depends on cloud provider and region.

   You can store and/or download the private key. This is needed in case you want to access the Jump Box on operating system level which is typically only needed for troubleshooting activation issues.
3 Accessing the Solution

3.1 Overview

As already described in chapter 0
**Architecture Overview**, a solution instance of SAP Data Intelligence, trial edition consists of:

- a virtual machine to orchestrate the solution instance creation (Jump Box)
- a Kubernetes cluster to run SAP Data Intelligence
- a Windows Remote Desktop to connect to SAP Data Intelligence (**only on Google Cloud Platform**)

Subsequently you learn how to access your solution instance.

- Chapter 3.2 tells you how to determine the external IP addresses of your solution instance.
- Chapter 3.3 explains how you can access SAP Data Intelligence on Amazon Web Services.
- Chapter 3.4 explains how you can access SAP Data Intelligence on Microsoft Azure.
- Chapter 3.5 shows how you can access SAP Data Intelligence on Google Cloud Platform.
- Chapter 3.6 outlines how to access the Jump Box for troubleshooting.

Since there are a few differences between a solution instance of SAP Data Intelligence, trial edition running on Amazon Web Services, Microsoft Azure and Google Cloud Platform, please be careful to read the chapter relevant for you.
3.2 Determining the external IP addresses of your solution instance

If you like to access your solution instance, you might need to know the external IP addresses or domain names. You can look them up by opening your solution instance in the SAP Cloud Appliance Library:

The following IP addresses / domain names are available in detail:

<table>
<thead>
<tr>
<th>Amazon Web Services</th>
<th>Microsoft Azure</th>
<th>Google Cloud Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP Data Intelligence External IP Address</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kubernetes Master IP Address</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jump Box Internal IP Address</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jump Box External IP Address</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Windows Remote Desktop Internal IP Address</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Windows Remote Desktop External IP Address</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transaction Coordinator Internal IP Address</td>
<td>X</td>
<td>-</td>
</tr>
</tbody>
</table>
3.3 Accessing SAP Data Intelligence on Amazon Web Services

On Amazon Web Services you can access SAP Data Intelligence, trial edition through the web browser of your local computer:

1. Press the Connect button in the SAP Cloud Appliance Library.

2. In the popup window press Open.

3. Your (default) web browser opens and you can logon to SAP Data Intelligence. The application is signed using a self-signed certificate and for that reason you will see an error message that the web browser is not able to verify the security certificate of *.elb.amazonaws.com. Confirm this error message and choose to trust this certificate.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
<td>*.elb.amazonaws.com</td>
<td>SAP Data Intelligence logon screen</td>
</tr>
<tr>
<td>Tenant name</td>
<td>default</td>
<td>All standard applications are in the default tenant. The system tenant is only used for administration tasks.</td>
</tr>
<tr>
<td>Username</td>
<td>system</td>
<td>SAP Data Intelligence user for default tenant SAP Data Intelligence user for the system tenant</td>
</tr>
<tr>
<td>Password</td>
<td>&lt;Master Password&gt;</td>
<td>The initial master password of the system you provided in the SAP Cloud Appliance Library when creating the instance.</td>
</tr>
</tbody>
</table>
3.4 Accessing SAP Data Intelligence on Microsoft Azure

**Note:** SAP CAL does not configure network security groups on Microsoft Azure. In Microsoft Azure network security groups are used to filter network traffic to and from Azure resources in an Azure virtual network. If you are using an enterprise account, then it is possible your network administration team has set in place firewall rules that may prevent access to SAP Data Intelligence Trial Edition. In such cases please contact your network administration to allow inbound **HTTPS access to the Kubernetes Load Balancer on port 443** and (Optional) **SSH access to the jumpbox on port 22**. For more information refer to the **Security Aspects** section and the **Microsoft Azure documentation**.

On Microsoft Azure you can access SAP Data Intelligence, trial edition through the web browser of your local computer:

1. Press the **Connect** button in the SAP Cloud Appliance Library

2. In the popup window press **Open**

3. Your (default) web browser opens and you can logon to SAP Data Intelligence. If you see an error message that the web browser is not able to verify the security certificate of ***.cloudapp.azure.com**, you need to confirm this error message (reason: the appliance uses self-signed certificates).

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
<td>*.cloudapp.azure.com</td>
<td>SAP Data Intelligence logon screen</td>
</tr>
<tr>
<td>Tenant name</td>
<td>default</td>
<td>All standard applications are in the default tenant. The system tenant is only used for administration tasks.</td>
</tr>
</tbody>
</table>
| Username      | system      | SAP Data Intelligence user for default tenant
SAP Data Intelligence user for the system tenant |
| Password      | <Master Password> | The initial master password of the system you provided in the SAP Cloud Appliance Library when creating the instance. |
3.5 Accessing SAP Data Intelligence on Google Cloud Platform

On Google Cloud Platform you can access SAP Data Intelligence, trial edition through the Windows Remote Desktop included in the appliance or through the web browser of your local computer.

3.5.1 Access via Windows Remote Desktop

To access SAP Data Intelligence through the Windows Remote Desktop included in the appliance, proceed as follows:

1. Press the Connect button in the SAP Cloud Appliance Library (alternatively you can look up the external IP address of the Windows Remote Desktop and initiate the remote desktop connection “manually” from your local computer).
2. In the popup window press Connect.

3. Your web browser downloads the remote desktop connection. Open this and then logon to the Windows Remote Desktop.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>administrator</td>
<td>Windows Remote Desktop user</td>
</tr>
<tr>
<td>Password</td>
<td>&lt;Master Password&gt;</td>
<td>The initial master password of the system you provided in the SAP Cloud Appliance Library when creating the instance.</td>
</tr>
</tbody>
</table>

Attention: We recommend using Google Chrome while working with SAP Data Intelligence, trial edition. To download it, use this link.

5. Now you can open SAP Data Intelligence with the web browser installed on the Windows Remote Desktop (Microsoft Internet Explorer, Google Chrome, Mozilla Firefox) by opening https://sapdataintelligence/.

If you see an error message that the web browser is not able to verify the security certificate of https://sapdataintelligence, you need to confirm this error message (reason: the appliance uses self-signed certificates).

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
<td><a href="https://sapdataintelligence/">https://sapdataintelligence/</a></td>
<td>SAP Data Intelligence logon screen</td>
</tr>
<tr>
<td>Tenant name</td>
<td>default system</td>
<td>All standard applications are in the default tenant. The system tenant is only used for administration tasks.</td>
</tr>
<tr>
<td>Username</td>
<td>system</td>
<td>SAP Data Intelligence user for default tenant</td>
</tr>
<tr>
<td>Password</td>
<td>&lt;Master Password&gt;</td>
<td>The initial master password of the system you provided in the SAP Cloud Appliance Library when creating the instance.</td>
</tr>
</tbody>
</table>
3.5.2 Access via Web browser of local computer

If you have deactivated the Windows Terminal Server during solution instance creation or you prefer to access SAP Data Intelligence, trial edition with the web browser of your local computer instead of using the Windows Remote Desktop included in the appliance, you need to maintain your local host file. The reason is that SAP Data Intelligence, trial edition uses the same hostname sapdataintelligence for all solution instances. This hostname is not fully qualified and you hence need to map the actual (external) IP address against it. The necessary steps are described subsequently.

1. Look up the external IP addresses for your solution instance in the SAP Cloud Appliance Library as described in chapter 3.2 Determining the external IP addresses of your solution instance. Identify the external IP address of SAP Data Intelligence (SAP Data Intelligence External IP Address). Example: Let's assume that the SAP Data Intelligence External IP Address is 123.456.78.91.
2. Open a text editor (e.g. Notepad in case of Microsoft Windows) as administrator (e.g. for Microsoft Windows search for Notepad on your computer, open the context menu of Notepad and click Run as administrator).
3. Click File → Open and enter the following path:
   - For Microsoft Windows operating system: c:\windows\system32\drivers\etc\hosts
   - For Linux operating system: /etc/hosts
4. Ensure to select All Files (*.*)
5. Open the hosts file and add the following lines to it:

   # SAP Data Intelligence, trial edition
   <External IP Address> sapdataintelligence

6. Replace <SAP Data Intelligence External IP Address> with the external IP addresses which you have retrieved from SAP Cloud Appliance Library. Example:

   # SAP Data Intelligence, trial edition
   123.456.78.91 sapdataintelligence

7. Save the hosts file and exit the text editor.
8. Now you can open SAP Data Intelligence with the web browser of your local computer (Microsoft Internet Explorer, Google Chrome, Mozilla Firefox) by opening https://sapdataintelligence/. If you see an error message that the web browser is not able to verify the security certificate of https://sapdataintelligence/, you need to confirm this error message (reason: the appliance uses self-signed certificates).

### Name | Value | Description
--- | --- | ---
URL | https://sapdataintelligence/ | SAP Data Intelligence logon screen
Tenant name | default system | All standard applications are in the default tenant. The system tenant is only used for administration tasks.
Username | system | SAP Data Intelligence user for the default tenant<br>SAP Data Intelligence user for the system tenant
Password | <Master Password> | The initial master password of the system you provided in the SAP Cloud Appliance Library when creating the instance.
3.6 Accessing the Jump Box for troubleshooting

You can access the Jump Box on operating system level via the secure shell protocol. This is only necessary for troubleshooting. For a detailed description, go to this FAQ wiki page and check this question: **How to connect to a running instance via the secure shell protocol (SSH)?**

On operating system level the following users and user groups are available:

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>root</td>
<td>root / super user with access to all commands and files</td>
</tr>
</tbody>
</table>

**Attention:** The Jump Box is automatically terminated 4 hours after the installation has finished successfully (afterwards it cannot be accessed anymore).

**Using kubectl on the Jump Box**

The Kubernetes command line interface (**kubectl**) is preinstalled on the Jump Box. This allows you to easily connect to the containers (pods) of the Kubernetes cluster used by SAP Data Intelligence.

If you, for example, run a **kubectl get nodes** command, you get a list of the nodes (virtual machines) of the Kubernetes cluster.

```
NAME              STATUS    ROLES           AGE     VERSION
node-lkgint-166719- -clus-sapci-d7ac1561-2im1 Ready <none> 3h v1.8.2-gke.1
node-lkgint-166719- -clus-sapci-d7ac1561-zj4 Ready <none> 3h v1.8.2-gke.1
node-lkgint-166719- -clus-sapci-d7ac1561-vgj6 Ready <none> 3h v1.8.2-gke.1
```

For more information refer to the [Kubernetes documentation](#).
4  Licenses

4.1 Running Solution with a Product License Key

To use the instance created from this solution under your own SAP Product License Agreements, you have to unlock the solution in SAP Cloud Appliance Library. As a prerequisite, you must own the required SAP Product licenses and need to purchase the SAP Cloud Appliance Library subscription package. For more information, see Unlocking Solutions.

4.2 Running Solution as a Trial for a Free Period

This solution can be used with a Free Trial License Agreement for the configured free period. You will be asked to agree to the Free Trial License Agreement during the instantiation steps.

Once you create the solution instance, the SAP system will generate a 30-day temporary license key that is sufficient for exploration purpose. After the expiration of the pre-installed temporary license a valid license is required to keep using the solution instance.

Note that you are not allowed to install license keys under the Free Trial License Agreement. You can continue the solution with the Product License Agreement any time prior to the end of the free period by unlocking it.
5 Security Aspects

Be aware that creating your instances in the public zone of your cloud computing platform is convenient but less secure. Ensure that only port 22 (SSH) is opened when working with Linux-based solutions and port 3389 (RDP) when working with Windows based solutions. In addition, we also recommend that you limit the access to your instances by defining a specific IP range in the Access Points settings, using CIDR notation. The more complex but secure alternative is to set up a virtual private cloud (VPC) with VPN access, which is described in this tutorial on SCN.

The list below describes the ports opened for the security group formed by the server components of your solution instance:

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Port</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSH</td>
<td>22</td>
<td>Used for SSH connection to Jump Box</td>
</tr>
<tr>
<td>RDP</td>
<td>3389</td>
<td>Used for remote desktop connection to Windows Remote Desktop</td>
</tr>
<tr>
<td>TCP</td>
<td>443</td>
<td>Used for HTTP(S) connection to SAP Data Intelligence</td>
</tr>
</tbody>
</table>

For more information about security vulnerabilities, see this community page.
6 Troubleshooting

6.1.1 Pipeline graphs

In the Pipeline Modeler application you may view the error message of all dead graphs by clicking on the title of the graph which open an overview tab. Alternatively you may download a Diagnostic Information of the dead graph as a .tar.gz archive file by clicking on the download button to the far right of the graph title.

![Pipeline graph error message]

6.1.2 Obtaining kubeconfig file

To collect logs from the underlying Kubernetes pods/containers you will first need a workstation or jumpbox with the kubectl command line tool installed.

Each cloud vendor has different mechanisms to download the kubeconfig file for your Kubernetes cluster that is running the SAP Data Intelligence Trial Edition.

For Google Cloud: If you deployed the Trial Edition in Google Cloud then you have the option of using a preconfigured jumpbox. If you wish to use your own workstation then you will have to install Google Cloud SDK or access to Google Cloud Shell, then run the command gcloud container clusters-get credentials

For Azure: Install the Azure CLI tool or access to the Azure Cloud Shell, then run the command az aks get-credentials.

For Amazon Web Services (AWS): Install either AWS CLI or AWS shell, then run the command aws eks update-kubeconfig
6.1.3 Accessing SAP Data Intelligence namespace

Regardless of cloud vendor, the SAP Data Intelligence will be deployed in the Kubernetes namespace with the name `datahub`.

```
kubectl get namespaces
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>STATUS</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>datahub</td>
<td>Active</td>
<td>23d</td>
</tr>
<tr>
<td>datahub-system</td>
<td>Active</td>
<td>22d</td>
</tr>
<tr>
<td>default</td>
<td>Active</td>
<td>22d</td>
</tr>
<tr>
<td>ingress-nginx</td>
<td>Active</td>
<td>22d</td>
</tr>
<tr>
<td>kube-node-lease</td>
<td>Active</td>
<td>23d</td>
</tr>
<tr>
<td>kube-public</td>
<td>Active</td>
<td>23d</td>
</tr>
<tr>
<td>kube-system</td>
<td>Active</td>
<td>23d</td>
</tr>
</tbody>
</table>

6.1.4 Accessing application logs

The following table illustrate most common used pod names and how they relate to the individual applications in SAP Data Intelligence:

<table>
<thead>
<tr>
<th>Application name</th>
<th>Pod name (Excluding suffix identifier)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipeline Modeler</td>
<td>pipeline-modeler (Front end) vflow-graph (Pipeline execution instance)</td>
</tr>
<tr>
<td>Metadata Explorer</td>
<td>datahub-app-data</td>
</tr>
<tr>
<td>Vora Disk engine</td>
<td>vora-disk-0</td>
</tr>
<tr>
<td>Issues with building docker images</td>
<td>Kaniko</td>
</tr>
</tbody>
</table>

Note that some applications are user specific, while others are shared across the tenant. For that reason, you may see multiple pods that belong to the same application.

To view a complete list of pods deployed in your Kubernetes cluster use the command:

```
kubectl get pods --namespace datahub
```

To view the logs of a pod use the command:

```
kubectl logs <pod-name-with-suffix> --namespace datahub
```
6.1.5 Enable access to Kubernetes Dashboard

If you prefer interacting with the Kubernetes cluster using a web user interface instead of using the `kubectl` command line tool then you may consider deploying the Kubernetes Dashboard. A detailed step by step guide can be found on the official SAP Data Intelligence Administration guide. Note that the Kubernetes Dashboard is an open-source tool and not created or maintained by SAP.

**Caution:** Exposing the Kubernetes Dashboard to the public internet poses significant security risks to your Kubernetes cluster.
## Appendix

### Installed Software Components

<table>
<thead>
<tr>
<th>Name</th>
<th>Release</th>
<th>Support Package Stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP DATA INTELLIGENCE</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Platform full</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Detailed information (incl. licenses) about 3rd party software used can be found in the [Free & Open Source Notices](#).