Recommender System

FAQ

Issue 01

Date 2020-07-23

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: https://www.huawei.com

Email: support@huawei.com

Contents

1 Basic Questions 1
1.1 What Is RES? 1
1.2 Related Services 1
1.3 How Do I Get Started with RES? 2
1.4 Obtaining a Pair of AK/SK 2
1.5 How Can I Create a Recommendation Job? 2
1.6 Will the Created Scenario Be Published Immediately? 3
1.7 The Minimum Number of Concurrent Online Instances Supports Auto Scaling. Is It OK to Set the Minimum Number of Concurrent Online Instances? 3
1.8 Is There Any Sample Data to Help Me Further Understand RES? 3
1.9 What Are Region and AZ? 3
2 Billing 4
2.1 How Do I View the Jobs that Are Being Billed in RES? 4
2.2 How Do I View RES Billing Details? 4
2.3 How Is the Data Source Billed? 5
2.4 How Are the Intelligent and Custom Scenarios Billed? 5
3 Data Source 6
3.1 What Is the Offline Data? 6
3.2 How Do I Upload Data to OBS? 6
3.3 How Do I Upload Real-Time Data? 6
3.4 How Do I Use Offline Data and Nearline Real-Time Data Together? 7
3.5 What Is Data Exploration? How Is Nearline Real-time Data Displayed in Data Exploration Reports? 7
4 Intelligent Scenarios 8
4.1 What Are the Main Application Scenarios of the "You May Also Like" Template? 8
4.2 What Are the Main Application Scenarios of the "Frequently Used Together" Template? 8
4.3 What Are the Main Application Scenarios of the "Most Searched" Template? 8
5 Custom Scenarios 9
5.1 What Are the Differences Between the Recommendation Engine and Ranking Engine? 9
5.2 What Custom Strategies Does RES Support? 9
5.3 Do I Need to Redeploy an Online Service After Re-executing the Retrieval Strategy Referenced by the Online Service? 10

A Description..... 11

1 Basic Questions

1.1 What Is RES?

Recommender System (RES) is powered by big data and AI technologies. It provides enterprises with recommendation platforms and algorithm-related services to build personalized recommendation applications with ease and improve CTR rates, retention rates, and user experience.

1.2 Related Services

Services related to RES:

- Data Lake Insight (DLI)

  DLI performs offline and nearline computing tasks in RES. For more information on DLI, see the **Data Lake Insight User Guide**.

- Object Storage Service (OBS)

  OBS stores the data sources required by RES. OBS is highly secure, reliable, and cost-efficient. For more information on OBS, see the **Object Storage Service User Guide**.

- Data Ingestion Service (DIS)

  DIS provides real-time logs of the data sources required by RES. For more information on DIS, see the **Data Ingestion Service User Guide**.

- Identity and Access Management (IAM)

  IAM authenticates unified access to HUAWEI CLOUD for RES and provides entrusted authorization of OBS and DIS. For more information on IAM, see the **Identity and Access Management User Guide**.

- ModelArts

  ModelArts is a one-stop development platform for AI developers. RES's ranking strategy uses the deep learning and computing capabilities of ModelArts to generate and train RES's ranking models. For more information on ModelArts, see the **User Guide (Senior AI Engineers)**.

1.3 How Do I Get Started with RES?

For details, see **Operation Procedure** in *RES User Guide*.

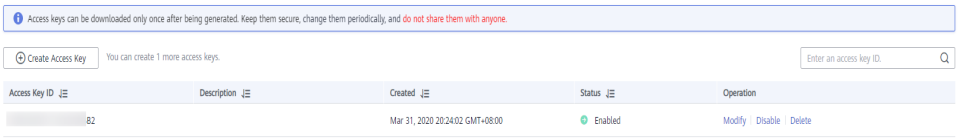
1.4 Obtaining a Pair of AK/SK

RES is dependent on other services. Data uploading and offline calculation tasks in RES cannot be completed without other services. An AK/SK pair is required to use other services.

Obtaining an Access Key

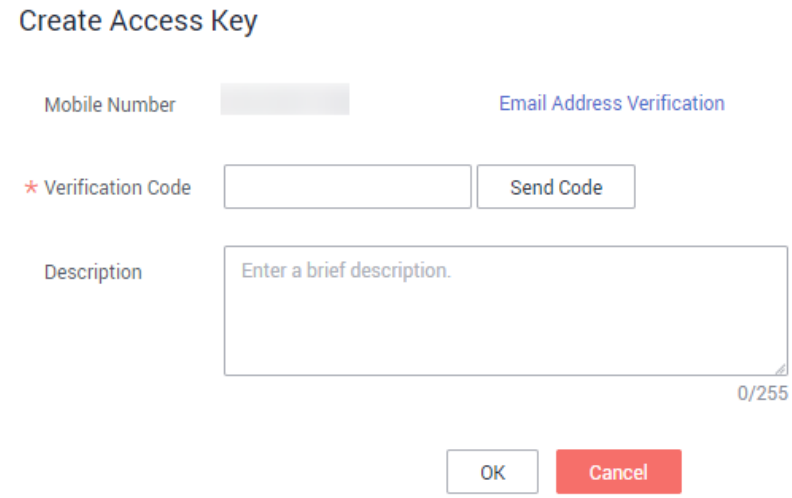
1. Log in to **HUAWEI CLOUD**, enter the **My Credentials** page, and choose **Access Keys > Create Access Key**. See **Figure 1-1**.

   Figure 1-1 Clicking **Create Access Key**

2. In the **Create Access Key** dialog box that is displayed, enter the verification code received by SMS or email. See **Figure 1-2**.

   Figure 1-2 Creating an access key

3. Click **OK** and save the access key file as prompted. The access key file is saved in the default download folder of the browser. Open the **credentials.csv** file to view the access key (**Access Key Id** and **Secret Access Key**).

1.5 How Can I Create a Recommendation Job?

RES supports the following job creation modes:

- Create jobs to view recommendations and effect evaluation results on the RES management console. For details, see the **RES User Guide**.
- Submit a task through the API and obtain results. For details, see the **RES API Reference**.

1.6 Will the Created Scenario Be Published Immediately?

Newly created intelligent scenarios and custom scenarios are not published immediately.

In a new scenario, you need to complete related recommendation configurations to complete creation of the scenario. After configuration, the scenario is ready to be published. You need to click **Publish** in the **Operation** column of the target scenario before running jobs involved in the scenario.

1.7 The Minimum Number of Concurrent Online Instances Supports Auto Scaling. Is It OK to Set the Minimum Number of Concurrent Online Instances?

Currently, the recommended maximum number of concurrent online instances is 5. The scaling resource has a certain delay, which slows down the prediction APIs. You are therefore advised to predict and enter possible concurrency specifications. The system reserves resources in advance based on concurrency specifications to avoid peaks.

If you need more specifications, **submit a service ticket** to Huawei technical support.

1.8 Is There Any Sample Data to Help Me Further Understand RES?

RES provides full data for testing. You can perform sample tests by referring to **Getting Started**.

1.9 What Are Region and AZ?

A region and availability zone (AZ) identify the location of a data center. You can create resources in a specific region and AZ.

For details about the region and AZ, see **Region and AZ**.

2 Billing

2.1 How Do I View the Jobs that Are Being Billed in RES?

On the RES management console, click **Overview** in the left menu bar. In the **Overview** area, you can view the running jobs in intelligent scenarios and custom scenarios. Access the corresponding management page based on site requirements and stop or delete corresponding scenarios to stop billing.

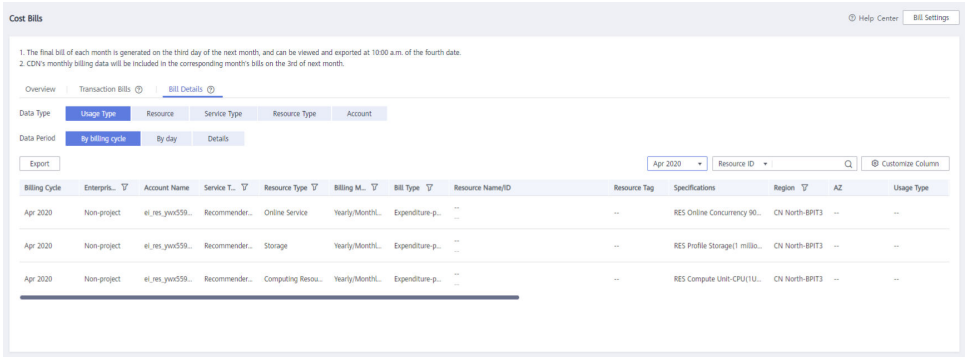
2.2 How Do I View RES Billing Details?

On the **Billing Center** page, you can query RES expenditure details by billing cycle and product type. This section describes how to view billing details. For more information, see **Viewing Cost Bills**.

Query method:

1. On the **Billing Center** page, choose **Bills > Cost Bill**. Click the **Bill Details** tab.
2. The **Bill Details** tab page lists the product types under the account and the bill details of each task. Select **Data Type** and **Data Period**. Drag the progress bar at the bottom of the page to view **Total Usage**, **Amount Due**, and **Account Overdue**.

Figure 2-1 Cost bill



2.3 How Is the Data Source Billed?

Data sources are billed for three items.

- **OBS storage space**: Upload offline data sources to OBS for storage. The price is billed based on the storage space. For details, see **OBS Pricing Details**.
- **Calculation fees**: After data is successfully imported to the data source, you are billed per million profiles per hour. After the data source is deleted, billing stops.
- **Profile storage fees**: You need to select a training flavor when creating an offline job or importing real-time data. You are billed based on the training flavor.

  NOTE

  If you want to stop the data source billing, stop importing the nearline data source and delete the corresponding data source and data storage in OBS.

2.4 How Are the Intelligent and Custom Scenarios Billed?

Pay-per-use billing is used in each scenario. The price depends on the selected flavor. A training job is billed based on the resources consumed during each execution. If a job in this scenario is in the **Stopped** state, for example, the job fails to be executed, billing will be stopped.

Each time a training job is executed, you will be billed. **Running** scenarios are billed.

3 Data Source

3.1 What Is the Offline Data?

Offline data is the data that stored in the following types of tables:

- User Feature List
- Item Feature List
- User Behavior List

For details on the fields in each table, see **Offline Data Sources** in *RES User Guide*.

3.2 How Do I Upload Data to OBS?

To use RES, you need to upload data to an OBS bucket. You can log in to **OBS Console** to create an OBS bucket, create a folder in the OBS bucket, and upload data. For details on how to create an OBS bucket and upload an object, see **Creating a Bucket** and **Uploading an Object**.

NOTE

When creating an OBS bucket, ensure that your OBS bucket and RES are in the same region.

3.3 How Do I Upload Real-Time Data?

RES allows you to upload real-time data through SDK. For details, see **Uploading Real-Time Data** in *RES User Guide*.

3.4 How Do I Use Offline Data and Nearline Real-Time Data Together?

During RES initialization, you need to provide offline data sources in batches and upload them to OBS in the format required by RES to complete data detection and import.

You are advised to use RES SDKs to upload nearline real-time data sources. All data updated from this operation will take effect immediately.

3.5 What Is Data Exploration? How Is Nearline Real-time Data Displayed in Data Exploration Reports?

Data exploration is used to mine and analyze data from data sources. It focuses on feature distribution range, statistics, and feature completeness, helping you better understand data and guide the configuration of feature engineering and related algorithms.

The data exploration task is an offline analysis task with a start time. Incremental data is saved to the database in real time. Therefore, you can execute the data exploration task periodically to overwrite the nearline real-time data.

4 Intelligent Scenarios

4.1 What Are the Main Application Scenarios of the "You May Also Like" Template?

The **You May Also Like** template is applicable to homepage recommendation scenarios and suitable for end users who are not clear about what they are looking for. RES can learn about users' long- and short-term behaviors and train jobs based on the behaviors to make personalized recommendations.

4.2 What Are the Main Application Scenarios of the "Frequently Used Together" Template?

The **Frequently Used Together** template is used to mine related content and behavior based on existing items and match associated items in multiple dimensions to make recommendations.

4.3 What Are the Main Application Scenarios of the "Most Searched" Template?

The **Most Searched** template is tailored to homepage recommendations and the-best-hit scenarios. With this template, you can collect statistics on popular content or items and attract new customers.

5 Custom Scenarios

5.1 What Are the Differences Between the Recommendation Engine and Ranking Engine?

Recommendation Engine

A recommendation engine uses recommendations as its service logic. Input the candidate sets generated based on your configurations in the engine. It will then generate the recommendation result sets.

Ranking Engine

A ranking engine uses ranking as its service logic. Input the candidate sets generated based on your configurations in the engine. It will then generate the recommendation result sets.

5.2 What Custom Strategies Does RES Support?

Currently, the following strategies are supported by RES:

- Retrieval strategies
- Filter rules
- Feature engineering
- Ranking strategies
- Nearline strategies
- Effect evaluation

For details on recommendation strategies, see **Algorithms and Parameters** in the *RES User Guide*.

5.3 Do I Need to Redeploy an Online Service After Re-executing the Retrieval Strategy Referenced by the Online Service?

No. If a retrieval strategy is executed again, a new candidate set is directly referenced by the online service. Therefore, there is no need to redeploy the online service.

A Description

Release Date	Description
2020-05-08	This is the first official release.

