Downloading Pentaho Content Using REST APIs

HITACHI
Inspire the Next

Change log (if you want to use it):

Date	Version	Author	Changes

Contents

This page intentionally left blank.

Overview

This document covers some best practices on downloading reports and other content from the Pentaho Repository using representational state transfer (REST) application programming interfaces (APIs). You will learn how to download reports to a folder outside the repository, and how to download the entire contents of your repository to an external location using client URL (cURL).

Our intended audience is Pentaho administrators or anyone with a background in cURL or REST APIs who is interested in downloading information to a location outside of the Pentaho Repository to avoid the need for using the Pentaho User Console (PUC) to access the information.

The intention of this document is to speak about topics generally; however, these are the specific versions covered here:

Software	Version(s)
Pentaho	7.x, 8.x

The Components Reference in Pentaho Documentation has a complete list of supported software and hardware.

Before You Begin

Before beginning, use the following information to prepare for the procedures described in the main section of the document.

This document assumes that you have knowledge of Pentaho and cURL and that you have already installed Pentaho and configured a report that you want to download.

Downloading from Pentaho Using APIs

Depending on your needs, you may find it useful to download Pentaho reports, or even your entire repository, using APIs. You can find details on these topics in the following sections:

Downloading Pentaho Interactive Reports to Outside Folder

You can retrieve and manipulate Pentaho Interactive Reports without accessing PUC if you download the reports to an existing folder on your server.

Vendor Sales Report

In this example, we use the Vendor Sales Report, an interactive report available in the Pentaho installation under `public/SteelWheels`.
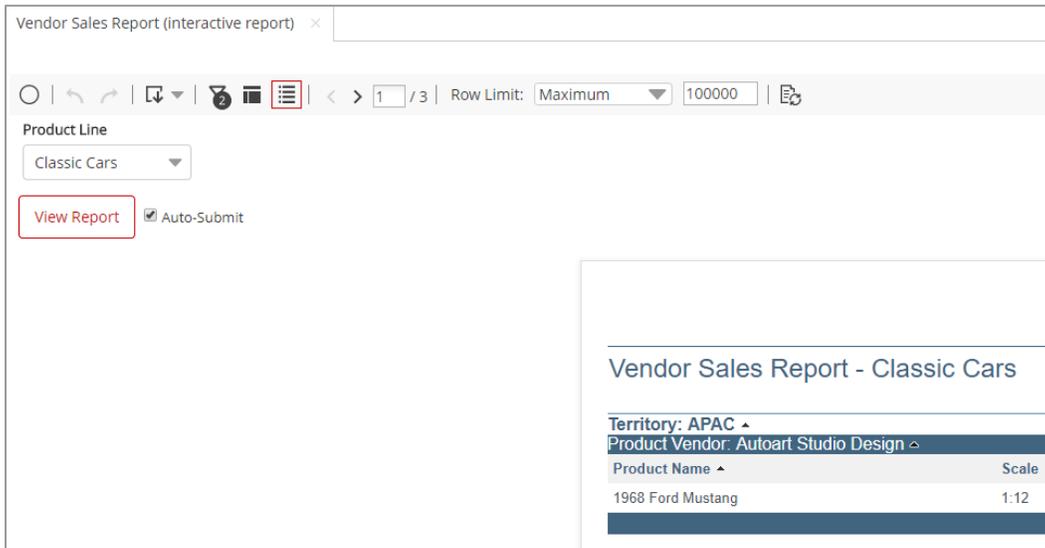
Figure 1: Vendor Sales Report (interactive report)

To download the report, use Pentaho's API and a URL appropriate for your situation. In this example, we use a URL with the parameters **Region** and **Product Line**:

```
http://localhost:8080/pentaho/api/repos/%3Apublic%3ASteel%20Wheels%3AVendor
%20Sales%20Report%20(interactive%20report).prpti/prpti.report?region=EMEA&P
roduct%20Line=Motorcycles&output-target=pageable/pdf
```

The request (see the cURL block below) contains all the parameters the report needs, and also contains the additional rendering parameter **output target**, which controls the rendering format (`HTML`, `PDF`, `XLS`, and so forth). For the output target, you can use the following:

Table 1: Output Target Parameter Options

Parameter Option	Purpose
`table/html;page-mode=stream`	Outputs HTML as a single page; all report pagebreaks are ignored.
`table/html;page-mode=page`	Outputs HTML as a sequence of physical pages; manual and automatic pagebreaks are active.
`application/vnd.openxmlformats-officedocument.spreadsheetml.sheet;page-mode=flow`	Excel 2007 Workbook (`XLSX`)
`table/excel;page-mode=flow`	Excel 97 Workbook (`XLS`)
`table/csv;page-mode=stream`	`CSV` output
`pageable/pdf`	`PDF` output

The sample URL we are using does the following:

- Opens the report using the viewer endpoint.
- Runs a report called Vendor Sales Report (Interactive Report).
- Sets the `Product Line = Motorcycles`.
- Sets the `Region = EMEA`.

```
http://localhost:8080/pentaho/api/repos/%3Apublic%3ASteel%20Wheels%3AVendor
%20Sales%20Report%20(interactive%20report).prpti/prpti.view?region=EMEA&Pro
duct%20Line=Motorcycles
```

To download the content, use the following cURL command:

```
curl -u 'admin:password'
'http://localhost:8080/pentaho/api/repos/%3Apublic%3ASteel%20Wheels%3AVendo
r%20Sales%20Report%20(interactive%20report).prpti/prpti.report?region=EMEA&
Product%20Line=Motorcycles&output-target=pageable/pdf' --output
'C:\temp\Vendor_Sales_EMEA_Motorcycles.pdf'
```

This command has the following features:

- The endpoint used is `prpti.report`.
- The username and password are passed along to authenticate; this can vary based on your server setup.
- The cURL command is expecting an output folder.
- The filename can be dynamically created with a script; this can vary based on your needs.

If you have spaces between words, you must escape these prior to building your URL, or use `%20` within the URL.

Output is then stored in the `C:\TEMP` folder:
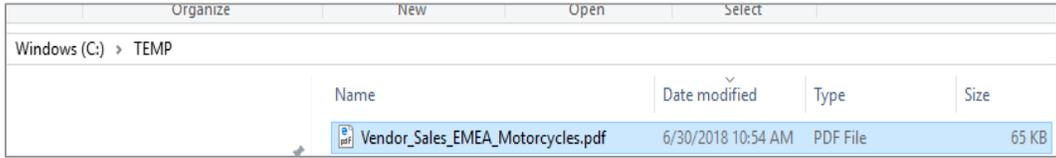
Figure 2: Stored Output

If you open the report, you can see the following content:
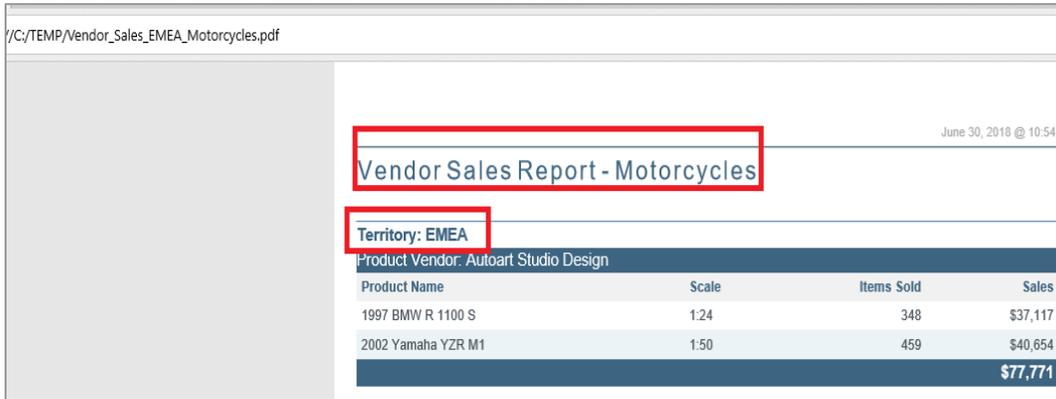
Figure 3: Report Content

Downloading Inventory Report from Report Designer

You can retrieve and manipulate [Pentaho Report Designer](#) reports without accessing PUC if you download the reports to an existing folder on your server. In this example, we use the Inventory List, available in the Pentaho installation under `public/SteelWheels`.
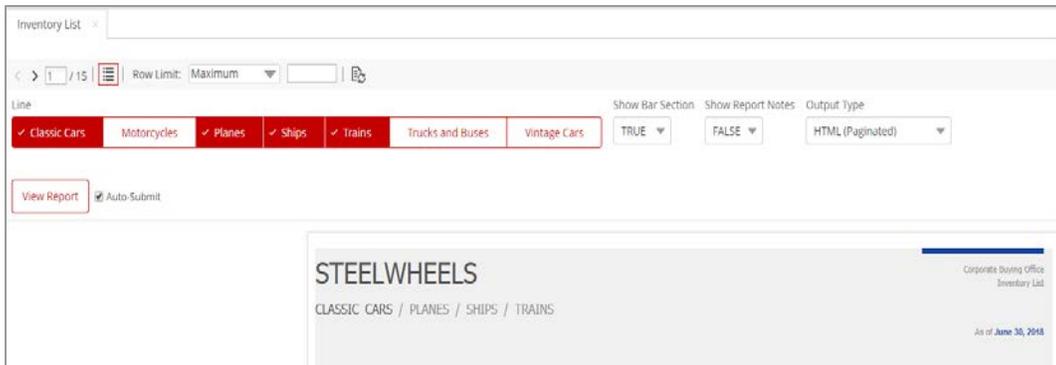
Figure 4: Inventory Report

To download the report, use Pentaho's API and a URL appropriate for your situation. In this example, we use a URL with the parameters **Line**, **Show Bar Section**, **Show Report Notes**, and **Output Type**:

```
http://localhost:8080/pentaho/api/repos/:public:Steel%20Wheels:Inventory%20
List%20(report).prpt/generatedContent?line=Motorcycles&hideBarSection=TRUE&
showNotes=TRUE&output-target=pageable/pdf
```

Note that the parameters shown on PUC are the parameters' **labels**, which can be different from the parameters' **names**. To be sure you refer to the parameters by their correct name, open the **Top Customers** report in Pentaho Report Designer and review the parameters' names as shown:
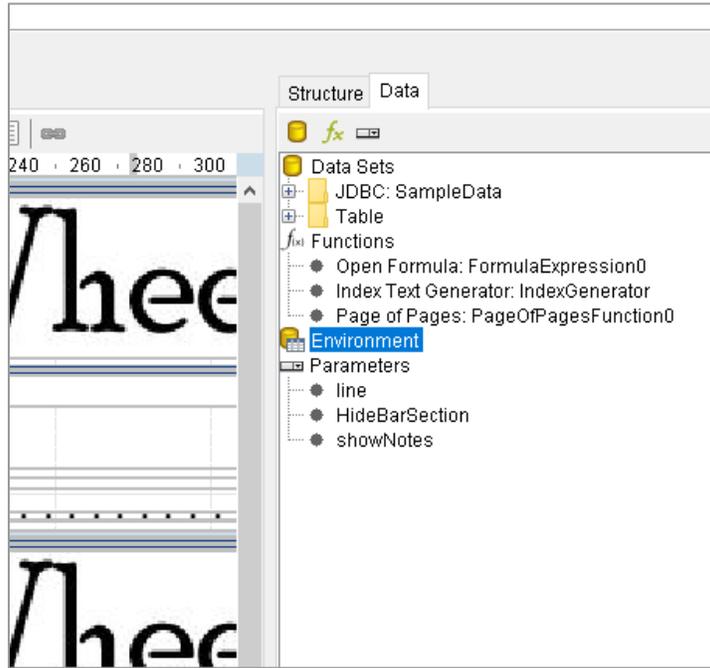
Figure 5: Parameter Names

- **Line** = `line`
- **Show Bar Section** = `HideBarSection`
- **Show Report Notes** = `showNotes`

The request contains all parameters the report expects, plus the additional rendering parameter output-target, which controls the rendering format (`HTML`, `PDF`, `XLS`, and so on). For the output target, you can use the following options:

Note that there are more options in this table than there were for the [Interactive Report](). This is because the Interactive Report is limited to the parameters listed in that table.

Table 2: Output Target Parameter Options

Parameter Option	Purpose
`table/html;page-mode=stream`	Outputs HTML as a single page; all report pagebreaks are ignored.
`table/html;page-mode=page`	Outputs HTML as a sequence of physical pages; manual and automatic pagebreaks are active.
`application/vnd.openxmlformats-officedocument.spreadsheetml.sheet;page-mode=flow`	Excel 2007 Workbook (`XLSX`)
`table/excel;page-mode=flow`	Excel 97 Workbook (`XLS`)

Parameter Option	Purpose
`table/csv;page-mode=stream`	CSV output
`pageable/pdf`	PDF output
`table/rtf;page-mode=flow`	Rich text format
`pageable/text`	Plain text
`pageable/xml`	Pageable layouted XML
`table/xml`	Table-XML output
`pageable/X-AWT-Graphics;image-type=png`	A single report page as PNG.

The sample URL we are using:

- Opens the report using the viewer endpoint.
- Runs a report called `Inventory List.prpt`.
- Sets the `Line = Motorcycles`.
- Sets `Show Bar Section = True`.
- Sets `Show Report Notes = True`.
- Sets the `Output Type = PDF`.

```
http://localhost:8080/pentaho/api/repos/:public:Steel%20Wheels:Inventory%20
List%20(report).prpt/viewer?line=Motorcycles&hideBarSection=TRUE&showNotes=
TRUE&output-target=pageable/pdf
```

To download the content, use the following cURL command:

```
curl -u 'admin:password'
'http://localhost:8080/pentaho/api/repos/:public:Steel%20Wheels:Inventory%2
0List%20(report).prpt/generatedContent?line=Motorcycles&hideBarSection=TRUE
&showNotes=TRUE&output-target=pageable/pdf' --output
'C:\temp\Motorcycles_InventoryList.pdf'
```

If you have spaces between words, you must escape these prior to building your URL, or use `%20` within the URL.

This command has the following features:

- The endpoint used is `generatedContent`.
- The username and password are passed along to authenticate; this can vary based on your server setup.
- The cURL command is expecting an output folder.
- The filename can be dynamically created with a script; this can vary based on your needs.

Output is then stored in the `C:\TEMP` folder:
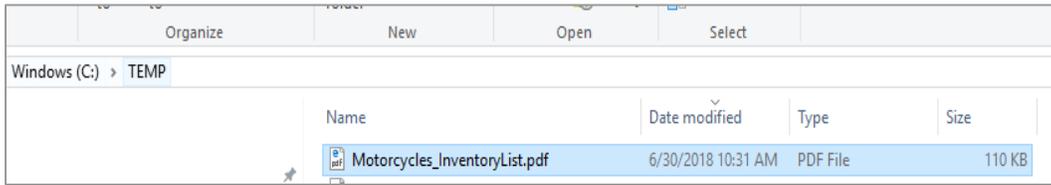
Figure 6: Stored Output

If you open the report, you can see the following content:
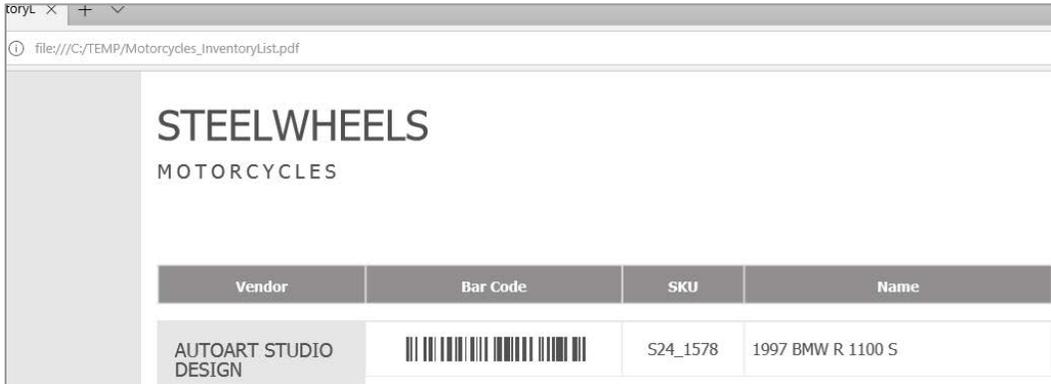
Figure 7: Report Content

Downloading Entire Repository Using cURL

You may need to download your entire repository to a folder for backup purposes. There are two ways to do it: using the import/export utility, or automate the download using the server's API with a cURL command. To automate with cURL:

1. Use the following URL: `http://localhost:8080/pentaho/api/repo/files/backup`.
2. Using cURL, change the URL to the following syntax:

```
curl -H 'user-agent: FireFox' -u 'admin:password' -X GET
'http://localhost:8080/pentaho/api/repo/files/backup' --output
'C:\TEMP\sampleBackup.zip'
```

3. This URL expects the following:
   a. **User Agent** = `firefox`
   b. **Output location** = `C:\TEMP`
   c. **Filename** = `sampleBackup.zip`

Related Information

Here are some links to information that you may find helpful while using this best practices document:

- [Pentaho Components Reference](#)
- [Pentaho Interactive Reports](#)
- [Pentaho Report Designer](#)