

Text Data Bank

The Text Data Bank tool extracts values from any text content, including plain text, HTML, XML, etc., and makes the values available as parameters in other tools and configurations.

Sections include:

- [Understanding Text Data Bank](#)
- [Adding a Text Data Bank Tool](#)
- [Configuring the Text Data Bank](#)
- [Options](#)

Understanding Text Data Bank

The Text Data Bank can extract values from any text content (including plain text, HTML, XML, etc.). You can specify left- and right-hand boundaries that mark the beginning and end of the extracted text. Characters between the values specified in the left- and right-hand fields will be extracted. It is useful when you want to extract plain text (e.g., for use in another tool or configuration panel), and cannot use other extraction tools such as the Browser Data Bank, XML Data Bank, JSON Data Bank, and so forth. The Text Data Bank tool is typically configured as an output of another tool (most commonly, a Browser Playback or Message Responder tool) that delivers text output. In this case, the Text Data Bank's Text Content area will be populated with the text output of the chained tool after tool execution completes. Alternatively, you can create a "standalone" tool using the Text Data Bank tool, then specify a file that contains the desired text content (for instance, a log file or other file that is dynamically updated or created during every tool run).

You can extract any value that appears between specified text boundaries. This allows you to extract values that vary from tool run to tool run. For example, assume you want to extract a session ID, which changes from session to session, and this session ID always appears between "leftboundary" and "rightboundary". You can configure a Text Data Bank tool to extract whatever value appears between "leftboundary" and "rightboundary". Every time this tool is run, the value that appears in the specified location between the given boundaries will be extracted. This way, if the session ID between those boundaries changes from session to session, the value extracted will also change from session to session.

Extracted values are added to the column you specify, and can be used in any tool configuration fields that allow parameterized values.

If the specified boundaries change (as a result of application changes), the configured extractions need to be updated to specify the new boundary values.



What's Considered Text Content?

The Text Data Bank operates only on content considered to have a text MIME Type. A MIME Type is based on the specified ContentType or by the file extension. You can control which MIME Types are considered to be "text" by opening **Parasoft > Preferences > MIME Types** and selecting the "Text" option for the appropriate MIME Types.

Adding a Text Data Bank Tool

1. Right-click on a tool and choose **Add Output...** You can also right-click on the scenario folder and choose **Add New > Test...** to add a standalone instance of the tool. A standalone instance enables you to extract data from a file that is being dynamically updated or created for every tool run.
2. Choose a traffic source type in the left-side panel. For example, if you want to extract a part of the response header, choose **Response > Transport Header** in SOAtest or **Outgoing Response > Payload** in Virtualize.
3. Choose **Text Data Bank** and click **Finish**.

The data bank will be attached to the test or added to the suite. See [Configuring the Text Data Bank Tool](#) for next steps.

Configuring the Text Data Bank

If you added the tool as an output of an existing tool, configure the fields under [Tool Settings](#) to extract text data. Standalone instances of the tool will have an **Input** tab for specifying the source from which you to extract content, as well as a **Tool Settings** tab for configuring extractions.

Tool Settings

1. Enter the source content from which you want to extract into the Text Content field. If you added the Text Data Bank as an output, you can run the parent tool to populate the field. If you added the tool as a standalone tool, click on the **Input** tab to specify the file from which you want to extract content.
2. Select the text inside the Text Content field that you want to extract and click **Create Extraction from Selection**. You can also click **Add** to manually add and configure an extraction. See [Configuring Extractions](#) for details about manually configuring extraction settings.
3. You can change the default description (optional) and click **OK** to continue.

The data bank can be used as is, but more complex scenarios may require additional configuration. The value may change during subsequent tool runs, for example, so the specific value may not always be extracted. See [Configuring Extractions](#) for additional configuration information.

Configuring Extractions

1. Select the extraction and click **Modify** to open the tool configuration overlay.
2. Enter a description for the extraction in the **Description** field. The description is for identification purposes only and does not need to be unique.
3. Specify a name for the column holding the data being extracted in the **Column Name** field. The column refers to the fields in the Data Source Column Name column of the extractions table. Column names should be unique. Extractions with the same column name will be overwritten in order from first to last row in the extractions table.
4. Specify the start of the extraction in the **Left-hand text** field and the end of the extraction in the **Right-hand text** field. Characters that appear between the patterns specified in these fields will be extracted from the Text Content field. You can enable the **Regular expression** options for each field to programmatically specify the patterns. Using regular expressions enables the tool to correctly extract characters if the Text Content field changed, set end-of-line characters, etc. In the following example, the end of the extraction is marked with a regular expression that matches a carriage return or semicolon.

Modify Selected Extraction

Description
Extract csrfToken from headers

Column Name
Test 4: csrfToken

Extraction Boundaries

Left-hand text: csrfToken= Regular expression

Right-hand text: \\n\\n; Regular expression

Extract Option
URL Decoded

OK Cancel

5. If the extracted text is to be used for a URL, choose **URL Encoded** from the Extract Option menu.
6. Click **OK** to save your changes.

Options

Enable the **Remove tabs and newline characters before processing** option to normalize whitespace in extracted text. This option is enabled by default.