



Hewlett Packard
Enterprise

HPE Web Connector

Software Version: 11.4.0

Release Notes

Document Release Date: June 2017

Software Release Date: June 2017

Legal notices

Warranty

The only warranties for Hewlett Packard Enterprise Development LP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HPE shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted rights legend

Confidential computer software. Valid license from HPE required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright notice

© Copyright 2017 Hewlett Packard Enterprise Development LP

Trademark notices

Adobe™ is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

This product includes an interface of the 'zlib' general purpose compression library, which is Copyright © 1995-2002 Jean-loup Gailly and Mark Adler.

Documentation updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent software updates, go to <https://downloads.autonomy.com/productDownloads.jsp>.

To verify that you are using the most recent edition of a document, go to <https://softwaresupport.hpe.com/group/softwaresupport/search-result?doctype=online help>.

This site requires that you register for an HPE Passport and sign in. To register for an HPE Passport ID, go to <https://hpp12.passport.hpe.com/hppcf/login.do>.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HPE sales representative for details.

Support

Visit the HPE Software Support Online web site at <https://softwaresupport.hpe.com>.

This web site provides contact information and details about the products, services, and support that HPE Software offers.

HPE Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Access product documentation
- Manage support contracts
- Look up HPE support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HPE Passport user and sign in. Many also require a support contract.

To register for an HPE Passport ID, go to <https://hpp12.passport.hpe.com/hppcf/login.do>.

To find more information about access levels, go to <https://softwaresupport.hpe.com/web/softwaresupport/access-levels>.

To check for recent software updates, go to <https://downloads.autonomy.com/productDownloads.jsp>.

Contents

New in this Release	5
Resolved Issues	7
Supported Operating System Platforms	8
Notes	9
Documentation	10

New in this Release

This section lists the enhancements to HPE Web Connector version 11.4.0.

- The connector supports XHTML pages.
- The connector can extract text from web pages, add the text to the DRECONTENT document field, and ingest metadata-only documents. Using the connector to extract text, rather than sending HTML files to CFS (KeyView), can result in better text extraction because the connector processes HTML pages using the HTML Document Object Model (DOM). To configure the connector to extract text, set the new configuration parameter `IngestAsPlainText` to `TRUE`.
- The connector can process metadata fields, extracted from a HTML page using the parameters `MetadataSelector` and `MetadataFieldName`, so that they contain only plain text. To do this, set the new configuration parameter `MetadataSelectorExtractPlainText` to `TRUE`.
- The connector now applies XSL transformations to XML documents where a relevant XSL stylesheet (`<?xml-stylesheet ... type="text/xsl" ... ?>`) is provided.
- The new configuration parameter `FailOnScriptTimeout` has been added, so that you can configure whether the connector considers having to terminate a script as an error.
- The new configuration parameters `MaxLinkChanges` and `MaxLinkChangePercentage` have been added. These specify the maximum number or proportion of the links on a page that can no longer appear, compared to the last time the page was synchronized, for the connector to delete unreachable pages that it has seen before. You can use these parameters to prevent the connector sending ingest-deletes when pages fail to load correctly (for example because a script fails to complete).
- JavaScript processing is faster and detection of script inactivity has been improved.
- The `identifiers` fetch action supports a new parameter, `FilterTypes`, which accepts a comma-separated list of the types of items to return identifiers for.
- The connector supports progress reporting for the `collect` fetch action.
- The connector supports the following Lua functions:
 - `get_log_service`, and the new class `LuaLogService`. You can use these when you want to write log messages to a custom log file (instead of the standard ACI server log files).
 - `parse_document_csv`, `parse_document_idx`, and `parse_document_xml`. These functions parse CSV, IDX, or XML files into documents and call a function on each document. `parse_document_idx` and `parse_document_xml` can also parse a string or file that contains a single document and return a `LuaDocument` object.
 - New functions and classes for parsing and manipulating JSON. The new functions are `parse_json`, `parse_json_array`, and `parse_json_object`. The new classes are `LuaJsonArray`, `LuaJsonObject`, and `LuaJsonValue`.
- You can now call the Lua method `insertJson` on `LuaField` objects as well as `LuaDocument` objects. You can also pass it one of the new `LuaJsonArray` or `LuaJsonObject` objects instead of a string.
- You can now configure action authorization more flexibly. The `[AuthorizationRoles]` configuration section has been added. You can add subsections to create roles, which can use a combination of

existing roles (equivalent to the existing `AdminClients`, `QueryClients`, and so on), or a specific set of actions. For each role, you can specify the client IPs and hosts, SSL identities, and GSS principals to use to identify users that have particular permissions to run actions.

If you want to use only SSL and GSS authorization, you can disable the client settings by setting the appropriate client configuration parameters to `""`. For example, `AdminClients=""` disables client authorization for administrative actions, and ensures that users must meet the SSL or GSS requirements.

- You can now set `SSLCertificate` to be a chain certificate in PEM format (consisting of the end-entity certificate, any intermediate certificates, and ending with the root CA certificate). This option allows a complete certificate to be returned to the connected peer.
- You can now set `SSLCheckCertificate` to `False` even when `SSLCACertificate` or `SSLCACertificatePath` are set. This allows the component to fill in any chain required for the `SSLCertificate` by using the certificates that you specify in `SSLCACertificate` and `SSLCACertificatePath`, without requiring a certificate from the connected peer.
- The `GSSAPILibrary` configuration parameter has been added to the `[Paths]` section. You can set this parameter to the path to the GSSAPI shared library or DLL that the application uses. Depending on your system configuration, HPE Web Connector attempts to detect the appropriate library to use. However, if you use Kerberos or GSSAPI security in your setup, HPE recommends that you set an explicit value for this parameter.

Resolved Issues

This section lists the resolved issues in HPE Web Connector version 11.4.0.

- The `GetVersion` action could incorrectly report the operating system on Microsoft Windows 10 and Microsoft Windows Server 2016.

Supported Operating System Platforms

The following operating system platforms are supported by HPE Web Connector 11.4.0.

- Windows x86 64
- Linux x86 64

The most fully tested versions of these platforms are:

Windows

- Windows Server 2012 x86 64
- Windows 7 SP1 x86 64
- Windows Server 2008 R2 x86 64
- Windows Server 2008 SP2 x86 64

Linux

For Linux, the minimum recommended versions of particular distributions are:

- Red Hat Enterprise Linux (RHEL) 6
- CentOS 6
- SuSE Linux Enterprise Server (SLES) 10
- Ubuntu 14.04
- Debian 7

Notes

- The Lua function `get_log(config, logstream)` has been deprecated. HPE recommends that you use the new function `get_log(log_type)` instead.
- The Lua function `string_uint_less` has been removed.

Documentation

The following documentation was updated for this release.

- *HPE Web Connector Administration Guide*
- *HPE Web Connector Reference*