

IDOL Education

Software Version 12.9.0

Release Notes



Document Release Date: June 2021
Software Release Date: June 2021

Legal notices

Copyright notice

© Copyright 2021 Micro Focus or one of its affiliates.

The only warranties for products and services of Micro Focus and its affiliates and licensors (“Micro Focus”) are as may be set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Micro Focus shall not be liable for technical or editorial errors or omissions contained herein. The information contained herein is subject to change without notice.

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 updated documentation, visit <https://www.microfocus.com/support-and-services/documentation/>.

Support

Visit the [MySupport portal](#) to access contact information and details about the products, services, and support that Micro Focus offers.

This portal also provides customer self-solve capabilities. It gives you 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 MySupport portal to:

- Search for knowledge documents of interest
- Access product documentation
- View software vulnerability alerts
- Enter into discussions with other software customers
- Download software patches
- Manage software licenses, downloads, and support contracts
- Submit and track service requests
- Contact customer support
- View information about all services that Support offers

Many areas of the portal require you to sign in. If you need an account, you can create one when prompted to sign in. To learn about the different access levels the portal uses, see the [Access Levels descriptions](#).

Contents

New in this Release	4
Resolved Issues	6
Notes	7
Requirements	9
Minimum System Requirements	9
Software Dependencies	9
Supported Operating System Platforms	10
Documentation	11

New in this Release

The following new features were released in Education version 12.9.0.

Enhancements to Education Server and Education SDK

- The overhead of Lua post-processing has been reduced.
- The `EnableUniqueMatches` configuration parameter now provides more consistent behavior. When `EnableUniqueMatches` is set to **True**, Education now generates matches for a given offset, and then imposes the uniqueness constraint.

For example, if a document contains repeated instances of the name *John James Smith*, and Education is configured to match *John James Smith*, *John James*, and *James Smith*, and with `AllowOverlaps` set to **False** and `EnableUniqueMatches` set to **True**, Education now returns a single match for *John James Smith*. Previously, it returned all three matches because of the way `EnableUniqueMatches` interacted with the match generation process.

- The `Exclusion` parameter has been added to Education pre-filter configuration. This parameter specifies that you want to exclude match windows that were found by a previous pre-filter task. Education compares any candidate windows that have been found by an earlier task to the regex or dictionary for the exclusion task. It discards any windows that match the exclusion filter.
- The `InvalidRegexBeforeMatch` and `InvalidRegexAfterMatch` parameters have been added to Education pre-filter configuration. These parameters specify a regular expression pattern that cannot be part of a valid match for the pre-filter. If Education finds the specified pattern in a pre-filter window, it truncates the window so that the match window does not include the non-valid pattern.
- Packaged post-processing Lua scripts have been updated so that you are not required to modify the `LUA_PATH` environment variable to run post-processing.
- You can now configure the pre-filter to stop processing a document after a threshold amount of data has been output from the pre-filter task without generating an actual Education match. This option allows you to quickly reject documents that spuriously match the pre-filter, but which are unlikely to contain real matches.

To activate the pre-filter, you can use the new `PrefilterMaxReturnedBytes` configuration parameter, or the Education SDK functions `EdkSessionSetMaxPrefilterBytesPerDoc (C)`, `ITextExtractionSession.SetMaxPrefilterBytesPerDoc (.NET)`, or `TextExtractionSession.setMaxPrefilterBytesPerDoc (Java)`.

- The `number iban.ecr` grammar now includes entities for many new countries, and the definition for Costa Rica has been updated to match the current standard.

Enhancements to Education Server

- ACI servers can now communicate with a License Server instance that requires a GSSAPI authenticated connection (that is, where the License Server configuration has `RequireGSSAuth` set to **True**). To allow this communication, you must set `GSSServiceName` in the `[License]` section of the configuration file to the GSSAPI service name of the License Server. For example:

```
[License]  
GSSServiceName=LICENSESERVER
```

- The third-party OpenSSL library has been updated to 1.1.1.j.
- The Admin, Help, LoadDoc, and LoadImage actions now set the HTTP response header Content-Security-Policy, to improve security against cross-site scripting. If the Content-Type header is set, these actions also set the header X-Content-Type-Options: nosniff.

Enhancements to Eduction SDK

- The Eduction SDK is now available for Apple M1 and Linux ARM64 (centos AArch64) platforms.

IMPORTANT: To use Eduction SDK on the Apple M1 platform, you must update your license key.

- You can now adjust the values to use for the MaxMatchesPerDoc, MinScore, and PostProcessingThreshold configuration parameters in an SDK session.

The SDK values override the values Eduction obtains from the engine configuration that you use to create the session.

The following table describes the functions and methods for each available language.

Feature	C function	Java method ¹	.NET method ²
MaxMatchesPerDoc	EdkSessionSetMaxMatchesPerDoc	setMaxMatchesPerDoc	SetMaxMatchesPerDoc
MinScore	EdkSessionSetMinScore	setMinScore	SetMinScore
PostProcessingThreshold	EdkSessionSetPostProcessingThreshold	setPostProcessingThreshold	SetPostProcessingThreshold

- You can now explicitly reset the input text. Eduction also implicitly updates the input text when new text is added. To reset the input text, use one of the following options:
 - C: EdkResetInputText function.
 - Java: resetInputText method in the TextExtractionSession interface.
 - .NET: ResetInputText method in the ITextExtractionSession interface.
- Eduction now uses less memory when using streaming input or adding input text chunks whilst processing matches.

¹These methods are part of the TextExtractionSession interface.

²These methods are part of the ITextExtractionSession interface.

Resolved Issues

The following issues were resolved in Education version 12.9.0.

- Education stopped unexpectedly when the configured `ResourceFile` for a pre-filter task was not set to an existing file.
- When `MaxMatchesPerDoc` was turned on, Education sometimes returned fewer results than expected because counting was done before post-processing, which could then eliminate matches.
- When `MaxMatchesPerDoc` was turned on for table extraction, counting was not done for the whole table, and did not limit matches.
- Underscores did not behave as configured when set as a Tangible character.

Resolved Issues for Education SDK

- On non-Windows platforms, the `edktool` benchmark mode measured elapsed CPU time rather than wall clock time, so with multiple threads running it logged a larger time for each thread than it actually took.
- The `edkjni` DLL did not work correctly with later versions of JDK 8 on Windows platforms.

Notes

These notes provide extra information about installing and using Education.

Deprecated Features

The following features have been deprecated.

Category	Deprecated Feature	Deprecated Since
C Education SDK	EdkSessionSetMatchTimeout EdkSessionSetRequestTimeout EdkGetNextMatchTimed	12.8
.NET Education SDK	ITextExtractionSession::SetMatchTimeout ITextExtractionSession::SetRequestTimeout ITextExtractionSession::SetStartTime	12.8
Java Education SDK	TextExtractionSession<T>::setMatchTimeout TextExtractionSession<T>::setRequestTimeout	12.8
Java Education SDK	EDKMatch::getTimedOut function	12.6
edktool configuration	SearchFields parameter. This parameter applies only to IDX document input, which was deprecated in 12.5.0.	12.6
Configuration	[PostProcessingTasks] configuration section. Use the PostProcessingTaskN and PostProcessThreshold parameters in the [Education] section.	12.5
edktool	The OutputScores and OutputSimpleMatchInfo configuration parameters. NOTE: The default values of the parameters OutputSimpleMatchInfo and OutputScores have been updated, to False and True respectively.	12.5
edktool	XML configuration files are now deprecated. Use CFG format configuration files instead.	12.5
edktool	Education on IDX documents. The following configuration parameters that apply only to IDX are	12.5

	also deprecated as part of this change: <ul style="list-style-type: none">• AllowDuplicates• CantHaveFieldCSVs• Databases• DocumentDelimiterCSVs	
--	---	--

The following notes apply to Eduction Server only:

- The autpassword command-line tool has been updated to provide hardened AES encryption for passwords. Micro Focus strongly recommends that you reencrypt all passwords in configuration files by using the updated tool.

The older AES encryption format and basic encryption methods are now deprecated. Passwords that you have encrypted with older versions continue to work, but the server logs a warning. Support for these older encryption methods will be removed in future.

Requirements

This section describes the system requirements, supported platforms, and software dependencies for Education 12.9.0.

Minimum System Requirements

The following are minimum system requirements for Education 12.9.0.

- 512 MB RAM
- 1 GB disk space
- 800 MHz or higher processor
- 400 MB of space designated for /tmp or /var/tmp (UNIX), or user-defined temporary directory (Windows)

Software Dependencies

Java

Development with the Education Java SDK requires the Java 8.0 JDK and JRE.

Microsoft Visual C++ Redistributable Package

To run Education version 12.9.0 on the Microsoft Windows operating system, you might need to install Microsoft Visual C++ Redistributable packages. The Education SDK and Education Server stand-alone zip packages include the required redistributable files for Microsoft Visual C++ 2017. You can also update your packages by using the latest version at: <http://support.microsoft.com/kb/2019667>.

UNIX Libraries

To run Education version 12.9.0 on UNIX platforms, the server must have the following minimum versions of libraries:

- GLIBC_2.3.2
- GLIBCXX_3.4.21
- GCC_4.8.0

NOTE: The Education SDK and Education Server stand-alone zip packages provide these libraries in the `libgcc_s` and `libstdc++` shared libraries.

You might need to set the `LD_LIBRARY_PATH` to include the `InstalLDir/bin` directory, to ensure that Education can access the installed shared libraries.

You can also copy the shared libraries to the component working directory.

Supported Operating System Platforms

The following operating system platforms are supported by Education 12.9.0.

- Windows x86 64
- Linux x86 64
- Linux ARM 64 (centos AArch 64)
- Solaris x86 64
- Solaris SPARC 64
- Mac OS X (Education SDK only)
- Apple M1

Documentation

The following documentation was updated for Education version 12.9.0.

- *Education Server Reference*
- *Education User and Programming Guide*