Sitecore CMS 6.6-7.2

A Conceptual Overview of External Links Checker module

# Introduction

External Links Checker (ELC) allows you to verify all external links from the defined fields within a predetermined range of items.

How ELC works:

1. ELC collects external links.
2. Sends requests to the servers.
3. Gets responses from the servers.
4. Creates and displays a report with complete information about links (response code, link URL, container field, and container item).

# Compatibility information

ELC was tested with Sitecore CMS 6.6-7.2. ELC does not require any additional module. ELC requires Internet access to work correctly.

# Quick Start

To start using ELC, you should install the ExternalLinksChecker.zip Sitecore package.

To do this:

1. Navigate to the Development Tools menu of the Sitecore Desktop and run the Installation Wizard.
2. Open the package.
3. Install the package.

The External Links Checker package contains the following items:

Static files:

* + App\_Config\Include\ExternalLinksChecker.config
  + sitecore\_modules\Shell\ExternalLinksChecker folder with files
  + bin\ExternalLinksChecker.dll
* \sitecore\shell\Applications\Content Manager\Galleries\LinksChecker folder with sub items

Items:

1. Core database

* \sitecore\content\Applications\ExternalLinksChecker
* \sitecore\content\Documents and settings\All users\Start menu\Programs\ExternalLinkChecker

2. Master database

* \sitecore\system\Modules\ExternalLinkChecker (module settings)
* \sitecore\templates\ExternalLinkChecker (module templates)

Important

For the Sitecore installations higher than Sitecore CMS 7.0 rev. 130424 (7.0 Initial Release) add the following entry to the <configuration>/<runtime>/ section:

<assemblyBinding xmlns="urn:schemas-microsoft-com:asm.v1">

**<dependentAssembly>**

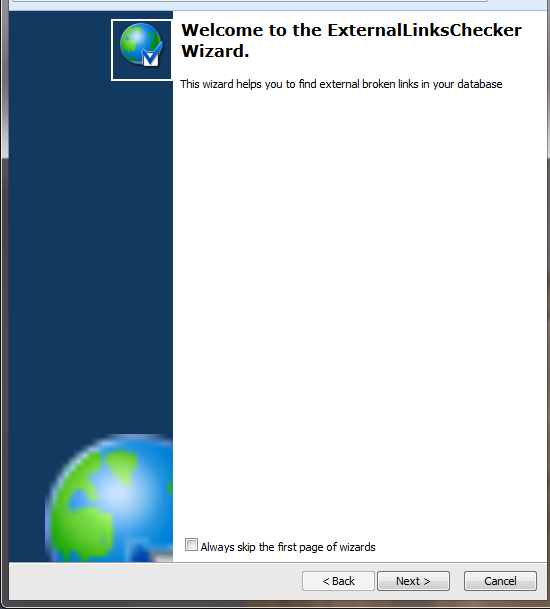
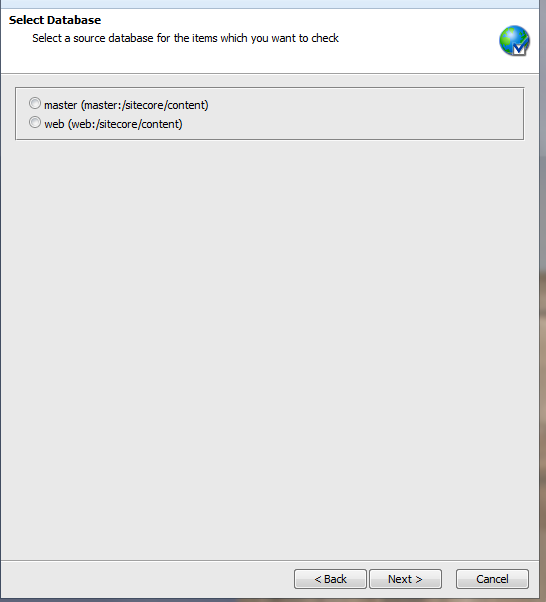
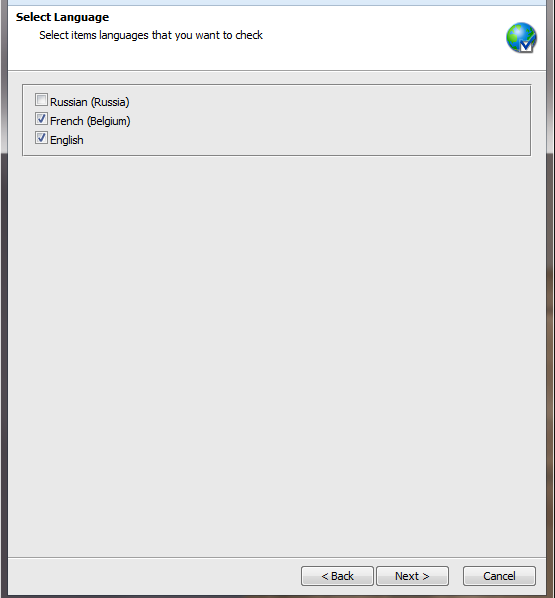
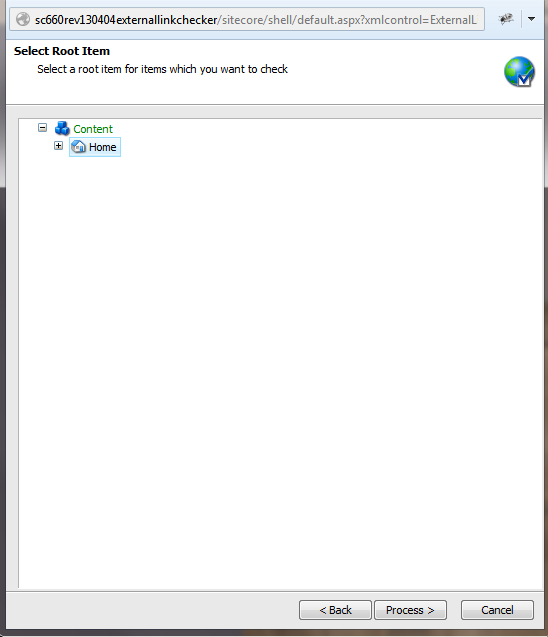
**<assemblyIdentity name="HtmlAgilityPack" publicKeyToken="bd319b19eaf3b43a" xmlns="urn:schemas-microsoft-com:asm.v1" />**

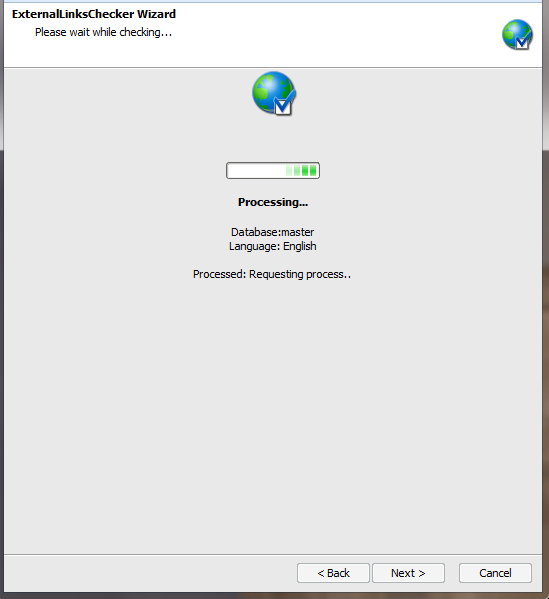
**<bindingRedirect oldVersion="1.4.0.0" newVersion="1.4.6.0" xmlns="urn:schemas-microsoft-com:asm.v1" />**

**</dependentAssembly>**

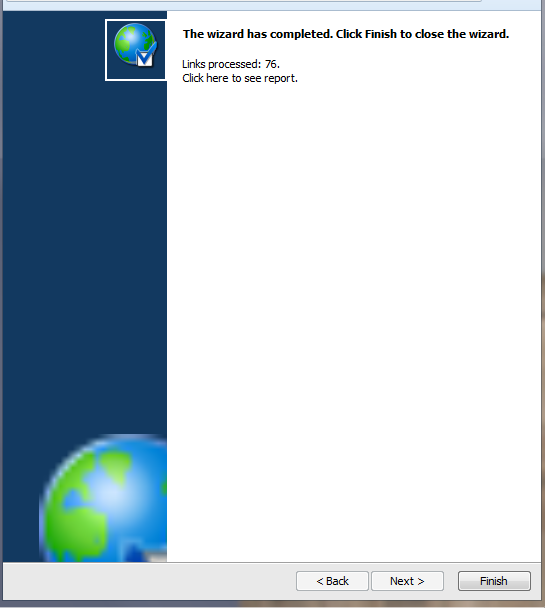
</assemblyBinding>

You can start working with the module after you install the package.

1. In the Sitecore desktop click All Applications – ExternalLinksChecker. This opens the ExternalLinksChecker wizard.  
   
2. Select the database.  
   
3. Select language:  
   
4. Select the Root item to define item range where ELC will look for the link fields.  
   
5. Click Process.
6. ELC begins collecting links and requesting pages.

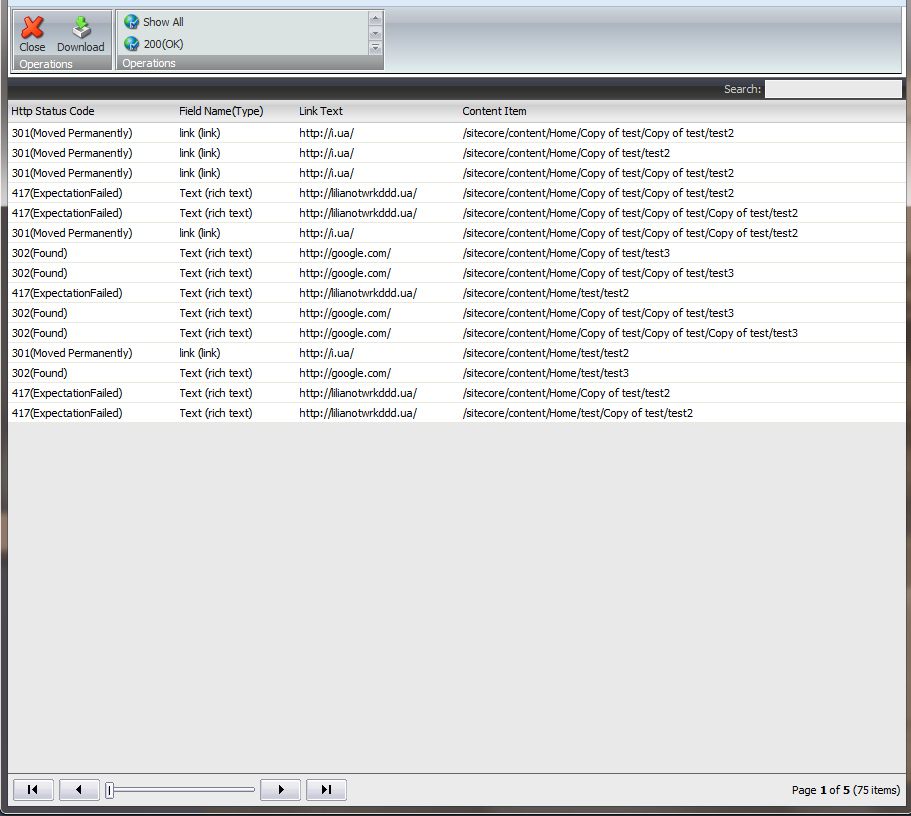


1. When ELC finishes collecting links, you can see the report. To do that, click Click here to see a report link:



On the Report page you can search using the fields Name, Code or Content Item, sort the results by code, field or item. You can use quick sort by code options in Operations chunk

You can download the report in the .csv format.



By default ELC checks the General link, link and Rich text fields. You can change the list in the ExternalLinksChecker.config file, <FieldCheckTypes> node. For more information on how to extend the list read more in Add Custom Fields section.

# Configuration

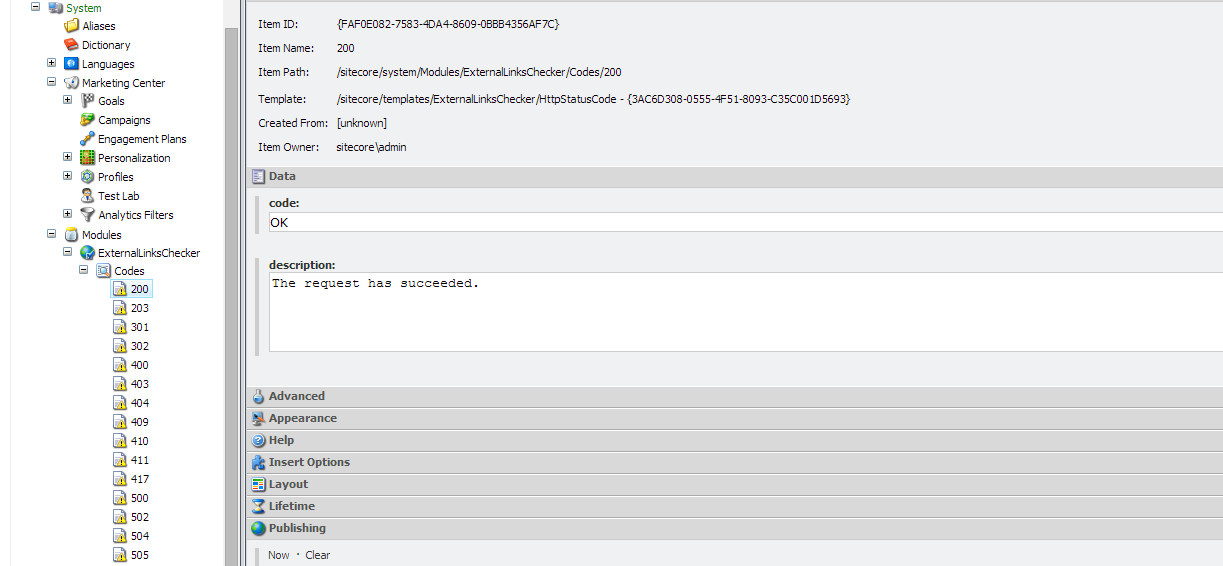
### Response Codes

You can configure, which httpResponse codes you want to display in the Report and what you want to see in the tooltip. By default ELC displays all responses. You can disable it using the <setting name="ShowAllResults" value="true"/> setting:

If the ShallAllResults setting is set to false, ELC shows only codes defined under the HttpStatusCodeFolder. By default it is the /sitecore/system/Modules/ExternalLinksChecker/Codes folder in master database, but you can change it using the <setting name="HttpStatusCodeFolder" value="/sitecore/system/Modules/ExternalLinksChecker/Codes"/> setting.

You must name the code item by the code number. The item contains two fields:

* + Code (string code equivalent)
* Description (tool tip code description)

Target Databases

You can configure which databases are available for selecting in the Wizard. All target databases are defined as items in the SourceDatabasesRoot folder, which is the /sitecore/system/Modules/ExternalLinksChecker/SourceDatabases folder in the Master database by default. You can change the folder path in the configuration file using the <setting name="SourceDatabasesRoot" value="/sitecore/system/Modules/ExternalLinksChecker/SourceDatabases"/> setting:

The target database item contains two fields:

* DatabaseName
* Root (defines a default root item for the current target database)

### Configuration Settings

You can configure most ECL settings in the ExternalLinksChecker.config file. Here some of main settings:

DefaultSettingDatabase

The DefaultSettingDatabase setting defines the database name where response codes and target databases are stored. By default it is the Master database.

<setting name="DefaultSettingDatabase" value="master"/>

HttpWebRequest.Timeout

The HttpWebRequest.Timeout setting defines timeout for the link request in milliseconds. By default it is 8000.

<setting name="HttpWebRequest.Timeout" value="8000"/>

CheckerThreadPriority

The thread priority of checker jobs.

Valid values: Lowest, BelowNormal, Normal, AboveNormal, Highest.

Default value: BelowNormal

SourceDatabasesRoot

For more about this setting, see the [Target Database](#_Target_Databases)s section.

SourceDatabaseTemplateID

This is the guid of the default template for source database items (stored under /sitecore/system/Modules/ExternalLinksChecker/SourceDatabases/).

SourceRootItem

This is the path to the default Root Item. It can be changed in the appropriate /sitecore/system/Modules/ExternalLinksChecker/SourceDatabases/ database item

Default value: /sitecore/content

ShowAllResults and HttpStatusCodeFolder

For more information about Read more about ShowAllResults and HttpStatusCodeFolder, see the Response Codes section.

CheckerMaxThreadCount

This setting lets you set Thread Count for request processing ThreadPool.

Default value: 20

AllowedProtocols

This setting specifies request/response protocols (used in external links) allowed for checking. The protocols should be split by | character.

Default value: http|https

HttpWebRequest.Timeout

The setting specifies the number of milliseconds before Sitecore considers a page request as expired. Default value: 8000

### Add Custom Fields

You can add or remove some field types from the types for checking list. Field types are defined in the FieldCheckTypes node. By default they are link, general link and rich text.

<FieldCheckTypes>

<fieldType fieldTypeName="link" type="ExternalLinksChecker.TypesForChecking.CheckLinkField, ExternalLinksChecker"/>

<fieldType fieldTypeName="General Link" type="ExternalLinksChecker.TypesForChecking.CheckLinkField, ExternalLinksChecker"/>

<fieldType fieldTypeName="rich text" type="ExternalLinksChecker.TypesForChecking.CheckRichTextField, ExternalLinksChecker"/>

</FieldCheckTypes>

To remove some type, you need to comment out the field type definition:

<!-- <fieldType fieldTypeName="General Link" type="ExternalLinksChecker.TypesForChecking.CheckLinkField, ExternalLinksChecker"/> -->

To add a new type you, need to create a custom class that implements the ExternalLinksChecker.TypesForChecking.ITypeForChecking interface.

It expects the CheckFieldValue method to be realized.

public List<ResponseResults> CheckFieldValue(Field field)

{}

The method must parse the field value, get all external links and request a result from the server. You can use RequestUtil for this:

string code = RequestUtil.GetResponseCode(url);

After that collect and return results.

# How To Remove the External Links Checker

To remove (uninstall) the External Links Checker module, you need to remove all items and files installed by the package:

Static files:

* + App\_Config\Include\ExternalLinksChecker.config
  + sitecore\_modules\Shell\ExternalLinksChecker folder with files
  + bin\ExternalLinksChecker.dll
  + \sitecore\shell\Applications\Content Manager\Galleries\LinksChecker folder with sub items

Items:

1. Core database

* /sitecore/content/Applications/ExternalLinksChecker
* /sitecore/content/Documents and settings/All users\Start menu/Programs/ExternalLinkChecker

1. Master database

* /sitecore/system/Modules/ExternalLinkChecker (module settings)
* /sitecore/templates/ExternalLinkChecker (module templates)