

XML Validation Guide

Questions or comments about this document should be directed to:

E-mail CEPI@michigan.gov

Phone 517-335-0505



Table of Contents

Introduction	3
XML File Validation.....	5
Software Installation	5
Validating an XML File against the Schema File	10
XML Schema Validation	12
Software Installation	12
Plug-in Installation	18
Validating a Schema File	23
How to Read the Error Message	24

Introduction

If you are struggling with validating your XML files against a schema, please note that there are several freeware, shareware, and commercial products currently available for doing this. Following are two examples of this type of product:

1. XML Notepad: a free download from the Microsoft website
<http://www.microsoft.com/downloads/details.aspx?familyid=72d6aa49-787d-4118-ba5f4f30fe913628&displaylang=en>
2. Altova XMLSpy: this is an example of a for-purchase product
http://www.altova.com/products/xmlspy/xml_editor.html

This guide will provide instructions for using XML Notepad.

Below are the World Wide Web Consortium (W3C) definitions for the base data types. Following is a crosswalk of the data type descriptions to the data-type definitions used in the schema, and corresponding links to the W3C.

Data Type Description	W3C Data Type Definition	Link to W3C Description
Text (numbers and letters)	Definition: The string data-type represents character strings in XML. Pattern: A-Z (upper case) a-z (lower case) and numbers (0 -9)	http://www.w3.org/TR/xmlschema-2/#string
Numeric Text (can have leading zeros)	Definition: The string data-type represents character strings in XML. Pattern: Numbers only (0-9) Numbers may have leading zeroes Example : 000123.	http://www.w3.org/TR/xmlschema-2/#string
Alphabetic (letters only)	Definition: The string data-type represents character strings in XML. Pattern: A-Z (upper case) a-z (lower case)	http://www.w3.org/TR/xmlschema-2/#string
Long Text (more than 255 characters)	Definition: The string data-type represents character strings in XML. Pattern: A-Z (upper case) a-z (lower case) and numbers (0-9)	http://www.w3.org/TR/xmlschema-2/#string
Whole Number (integer)	Definition: Integer is derived from decimal by fixing the value of "fractionDigits" to be 0 and disallowing the trailing decimal point. This results in the standard mathematical concept of the integer numbers. The value space of integer is the infinite set {...,-2,-1,0,1,2,...}. The base type of integer is decimal Pattern: Numbers 0 - 9 without decimals	http://www.w3.org/TR/xmlschema-2/#integer
Decimal Number	Definition: Decimal represents a subset of the real numbers, which can be represented by decimal numerals. Pattern: Numbers 0-9 with decimal digits separated by a period.	http://www.w3.org/TR/xmlschema-2/#decimal
List Of Values (choose from list)	Pattern: Select one of the values provided in the list.	
Yes/No (or true false, or boolean)	Definition: Boolean has the value space required to support the mathematical concept of binary-valued logic: {true, false} Pattern: yes [or] no, true [or] false, 1 [or] 0	http://www.w3.org/TR/xmlschema-2/#boolean
Date (only)	Definition: The value space of date consists of top-open intervals of exactly one day in length on the timelines of date time, beginning on the beginning moment of each day (in each timezone), i.e. '00:00:00', up to but not including '24:00:00' (which is identical with '00:00:00' of the next day). Pattern: cccc-mm-dd	http://www.w3.org/TR/xmlschema-2/#date
Time (Only)	Definition: Time represents an instant of time that recurs every day. Pattern: hh:mm:ss	http://www.w3.org/TR/xmlschema-2/#time
Text (free form)	Definition: The string data-type represents character strings in XML. Pattern: A-Z (upper case) a-z (lower case) and numbers (0 -9)	http://www.w3.org/TR/xmlschema-2/#string

Data Type Description	W3C Data Type Definition	Link to W3C Description
School Year (9999-9999)	Definition: The string data-type represents character strings in XML. Pattern: Two sets of four-digit numbers from 0-9 separated by a hyphen. Example: 2001-2202	http://www.w3.org/TR/xmlschema-2/#string
Names (letters, apostrophes, commas, hyphens, periods)	Pattern: A-Z (uppercase) a-z (lower case) Comma, hyphen, period	http://www.w3.org/TR/xmlschema-2/#string
Numeric Text (0-6) (can have leading zeros)	Definition: The string data-type represents character strings in XML. Pattern: Numbers only (0 - 6). Numbers may have leading zeroes. Example: 000123.	http://www.w3.org/TR/xmlschema-2/#string
Addresses (numbers, letters, #)	Definition: The string data-type represents character strings in XML. Pattern: A-Z (uppercase), a-z (lower case), numbers (0 - 9), # sign.	http://www.w3.org/TR/xmlschema-2/#string

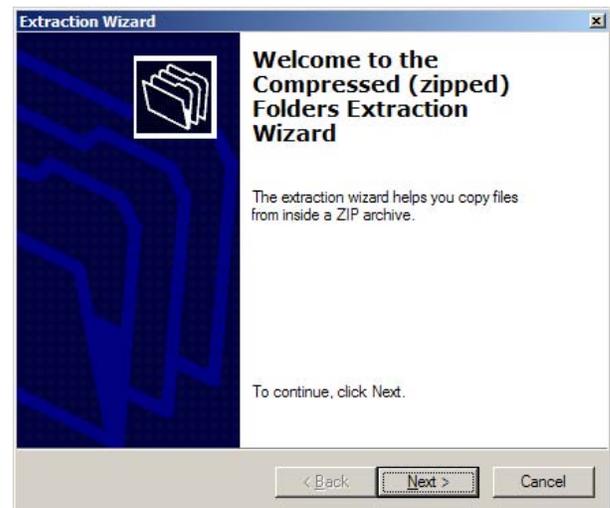
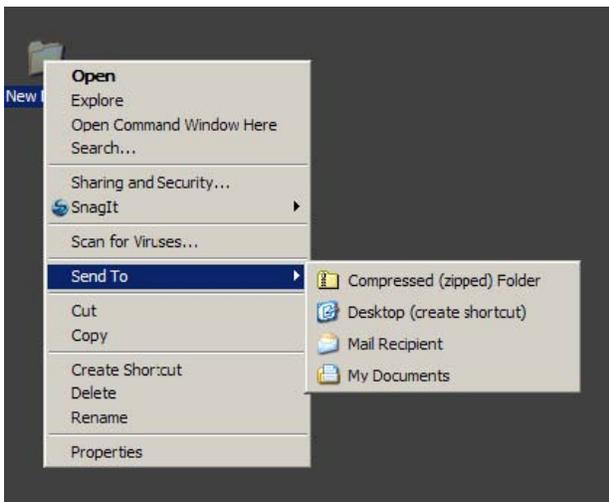
Using ZIP Files

ZIP functionality is native functionality for Windows XP and beyond. Below are instructions for using that feature.

To Zip:

1. Right click the file(s)/folder(s) you want to zip.
2. Hover over "Send To" in the context menu.
3. Click "Compressed (zipped) Folder."

The zipped file(s) will be located in the same folder from where the file(s) was selected.



To Unzip:

1. Right click on any ZIP file
2. Click "Extract All..."
3. Follow the instructions on the Extraction Wizard.

If you require additional archiving functionality, there are commercial products available for purchase. In addition, several freeware options exist. A few of the more popular applications include 7-Zip, IZarc, and Filzip.

XML File Validation

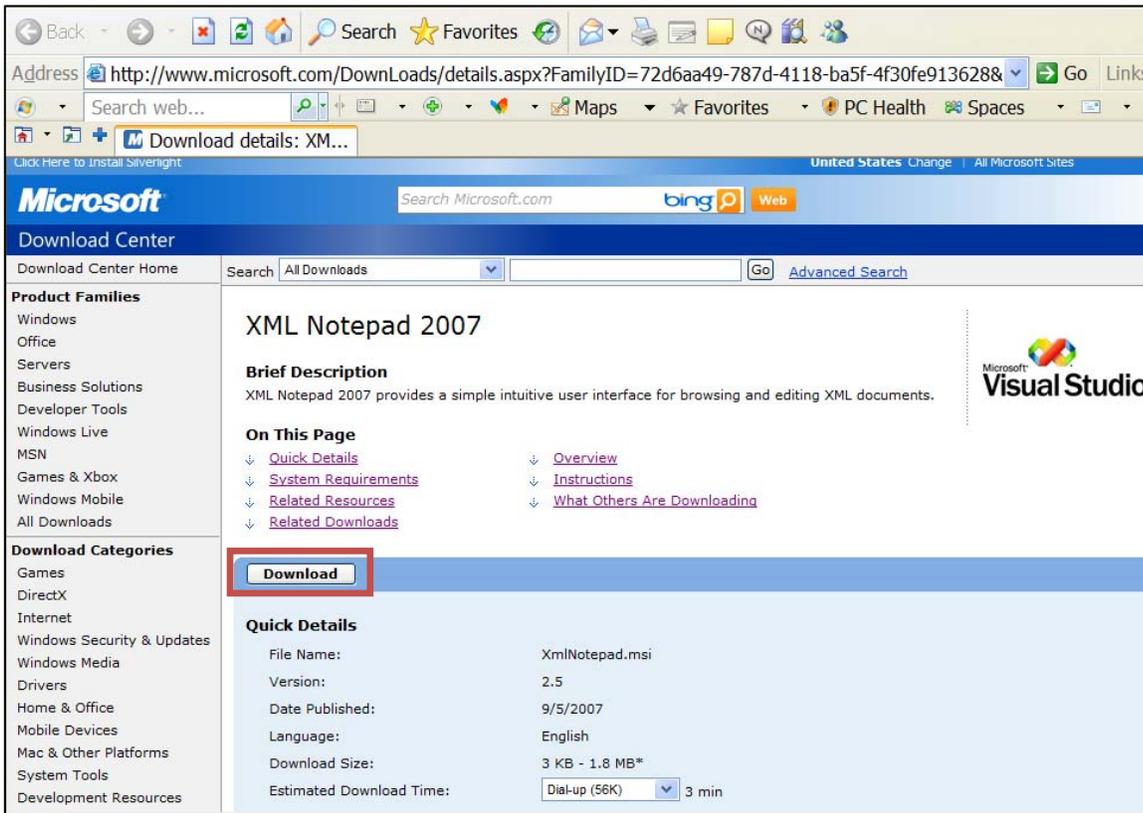
Software Installation

In order to validate XML files, you have to download and install the tool XML Notepad on to your desktop. This is a one-time installation. Once installed, XML Notepad can be used to validate all XML files.

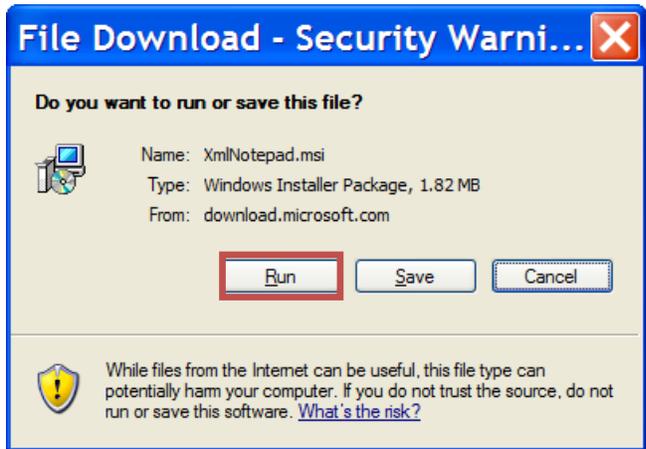
The system requirements for this download are Windows XP, Vista, or 7 and .Net Framework v2.0.

Go to the Web page: <http://www.microsoft.com/Downloads/details.aspx?FamilyID=72d6aa49-787d-4118-ba5f-f30fe913628&displaylang=en>

Click on the "Download" button.



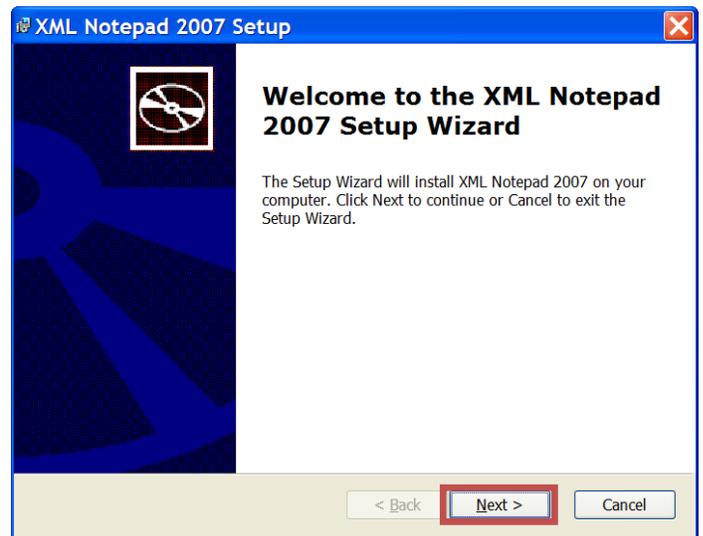
To install the application upon the completion of the download click the "Run" button. If you wish to save the file for later use click "Save."



You may be prompted with a security warning when you begin the installation process. Click "Run" to continue.



Click "Next" to proceed with the installation process.



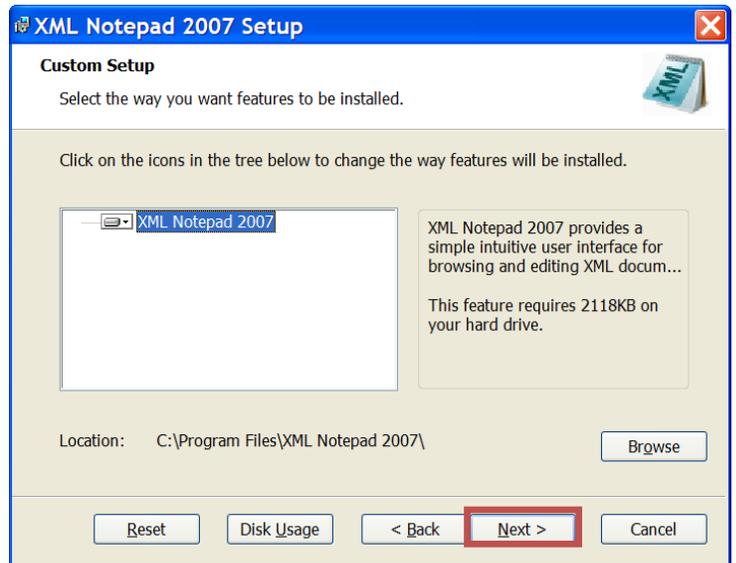
Check the box "I accept the terms..." and then click the "Next" button.

The "Next" button will activate once you accept the terms.

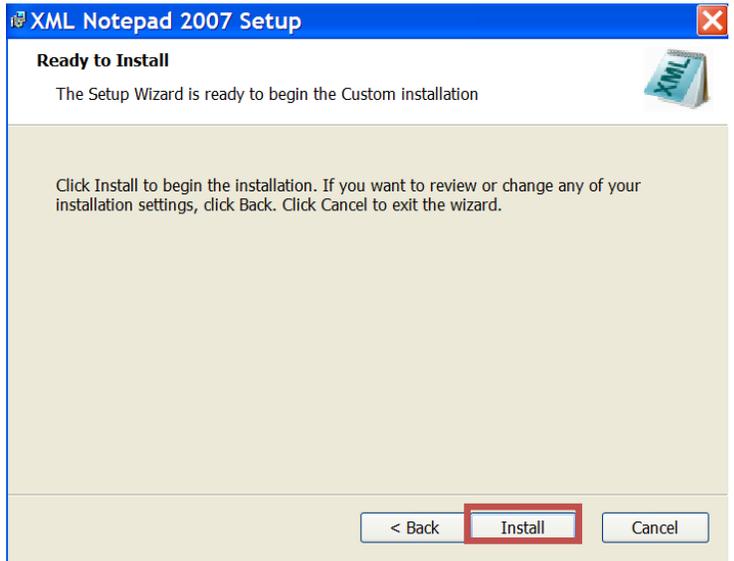
The default destination folder should be:

C:\Program Files\XML Notepad 2007\

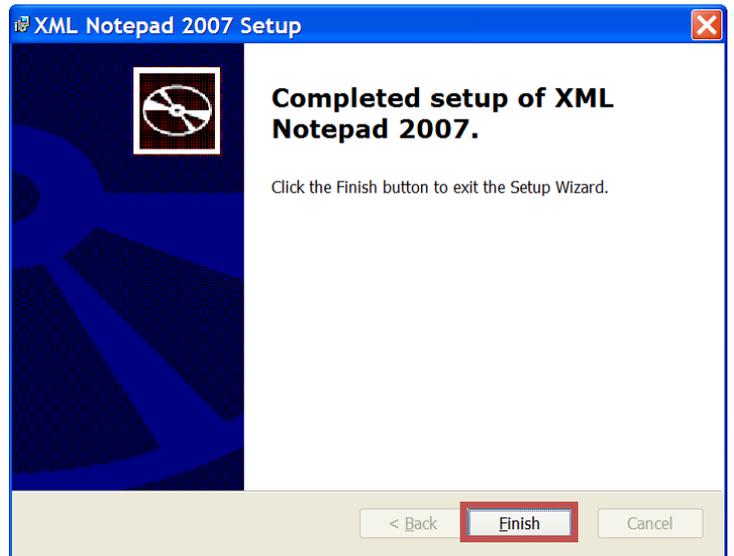
Advanced users may wish to customize their installation from this menu. To continue with the default installation click "Next."



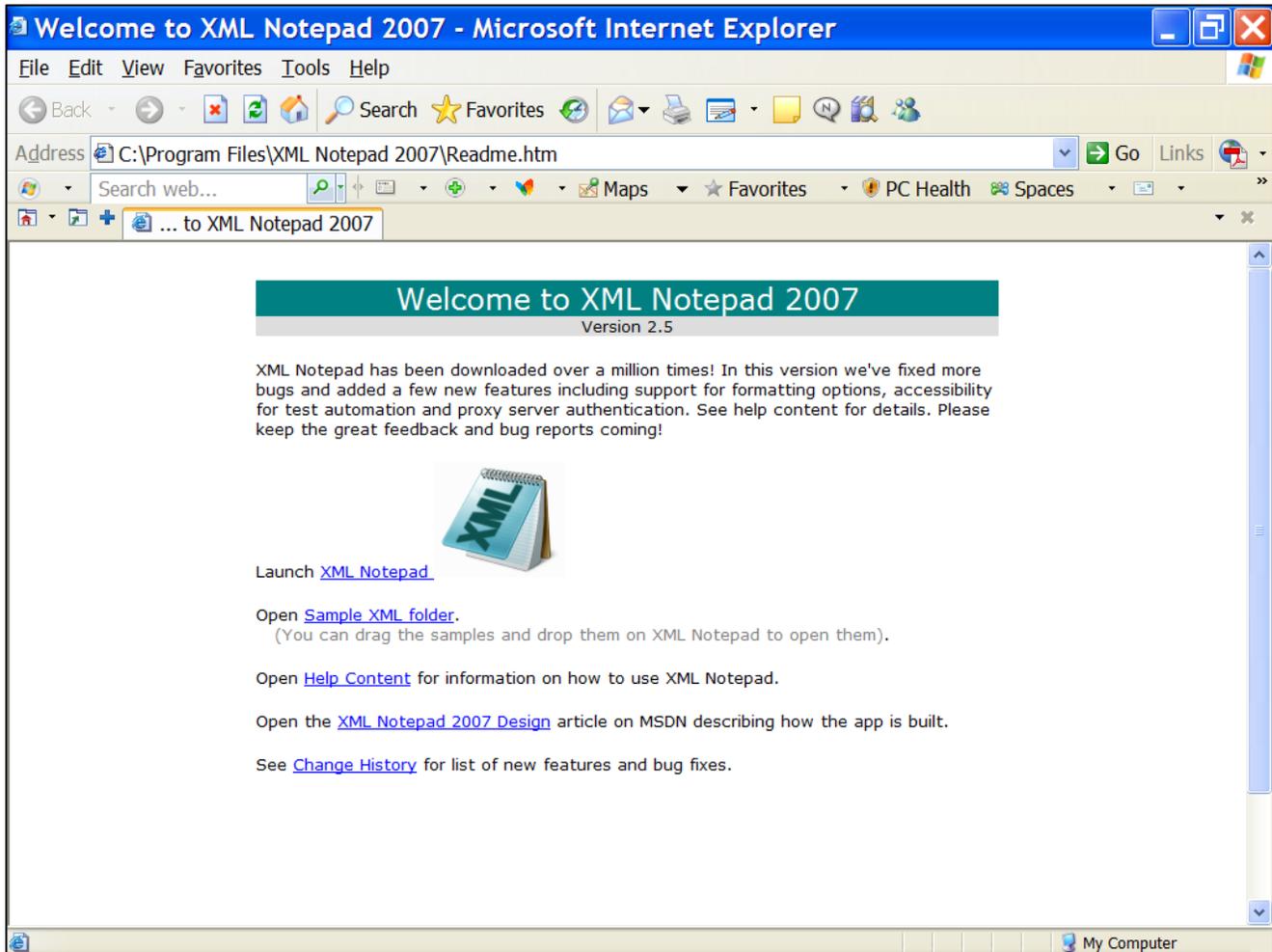
Now that your installation parameters have been defined, click "Install."



When installation is complete, a confirmation window will appear. Click "Finish" to end the installation process.

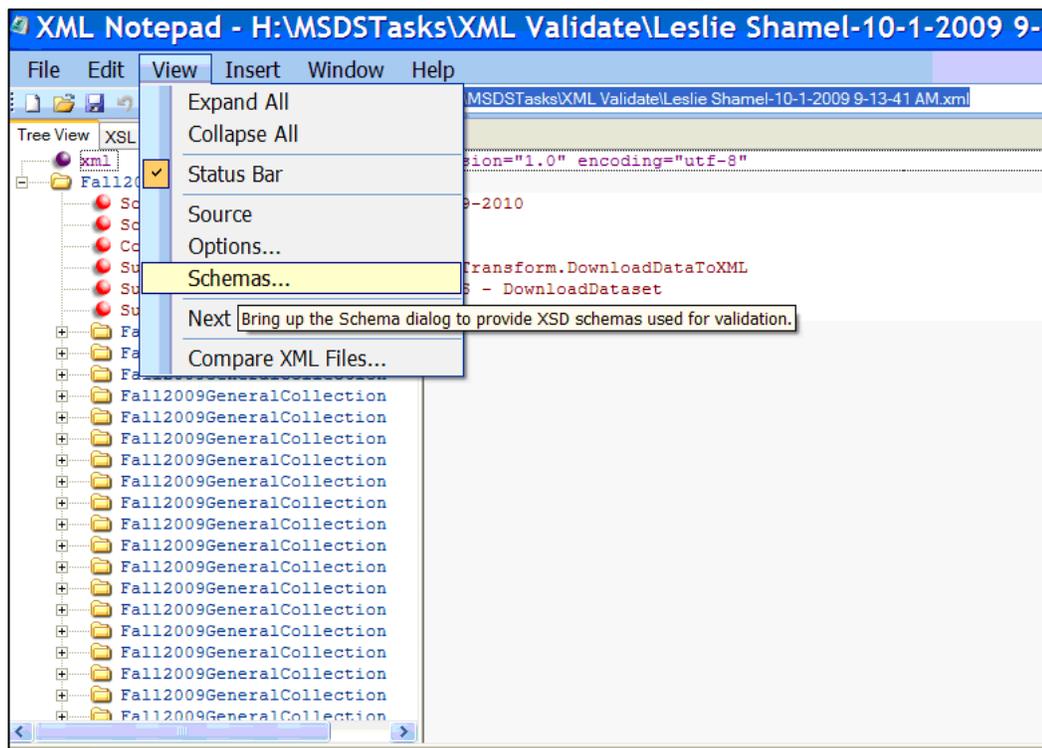


The Web page shown below will open in your browser, which contains links to additional information on the XML Notepad application.



Validating an XML File against the Schema File

1. Locate the collection schema file from CEPI's Institutions of Higher Education Web page.
2. Right click on the XML Schema and choose "Save target as." Change the file extension to .xsd and save it to your local drive.
3. Click on Start > Programs > XML Notepad 2007 > XML Notepad 2007 (this will open the XML Notepad 2007).
4. Under File Menu, open your XML file to be validated. The file will open with a vertically split window of "Tree View" on the left and the texts of values on the right.
5. Select View from the top menu and click on "Schemas..." as in the following image demonstrates.

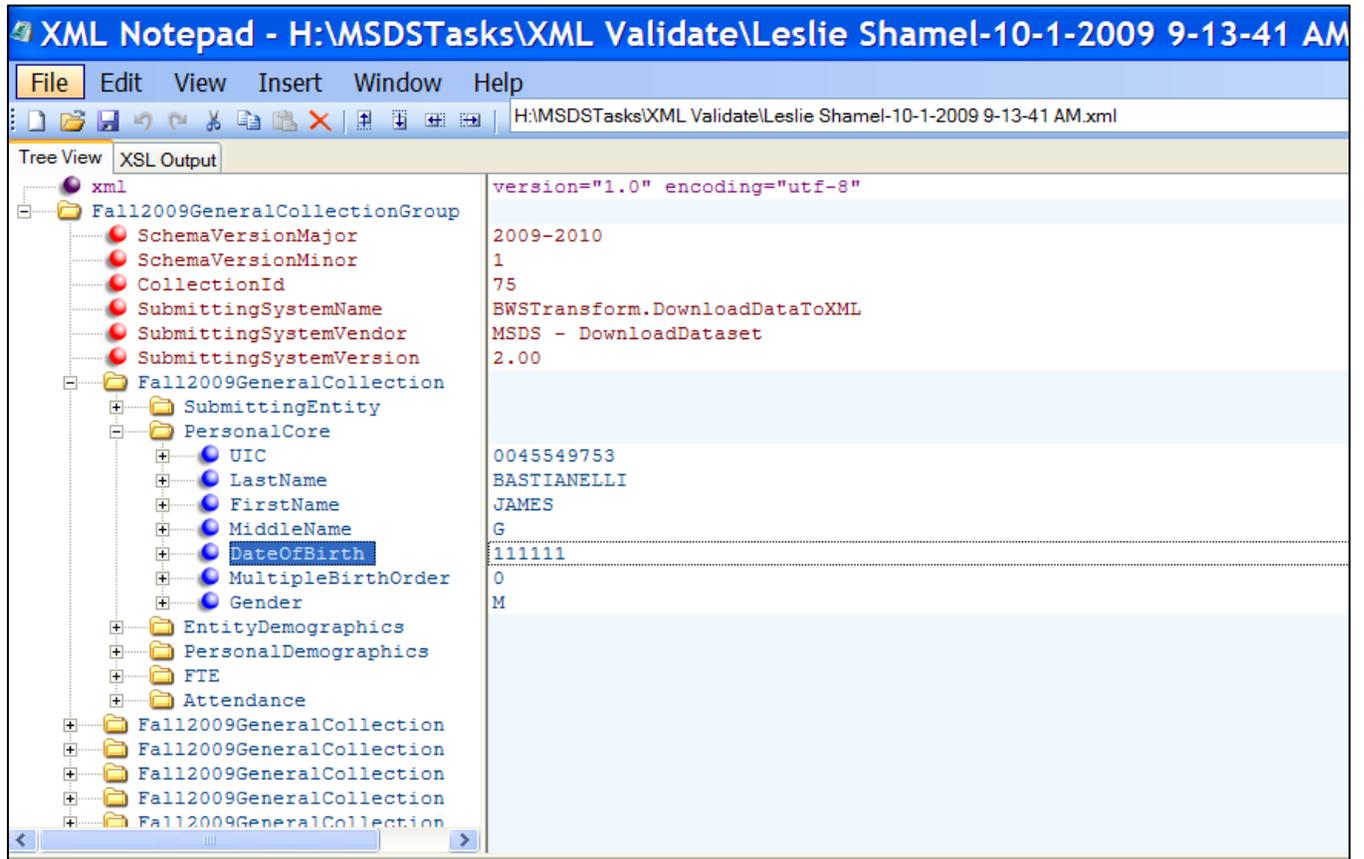


6. The XML Schemas dialog box opens.
7. If there are any schemas listed in the XML Schemas dialog box, in the File menu, click Clear. This will erase any schemas that have been previously loaded.
8. To add the node XML schema, in the File menu, click "Add schemas," type or browse to the location of the .xsd saved in step 2, and then click "Open".
9. To load the node XML schema, click "OK." If any errors occur while you are editing the node XML file, they will be reported in the Error List tab, at the bottom of the XML Notepad 2007 window.

A screenshot of the Error List tab in XML Notepad 2007. The tab is titled "Error List" and "Dynamic Help". It contains a table with the following data:

Description	File	Line	Column
The 'DateOfBirth' element is invalid - The value '111111' is invalid according to its datatype 'DateOfBirt...	Leslie%20Shamel-10-1-2...	13	14

10. When you double-click on the error, the record with the error will be opened on the right side with the value enclosed in a box.



11. Correct the value and save.

XML Schema Validation

Software Installation

In order to validate XML files, you have to download and install the tool **Notepad++** on your computer. This is a one-time installation. Once installed, this can be used to validate all XML files.

Go to the Web page: <http://notepad-plus.sourceforge.net/uk/site.htm>

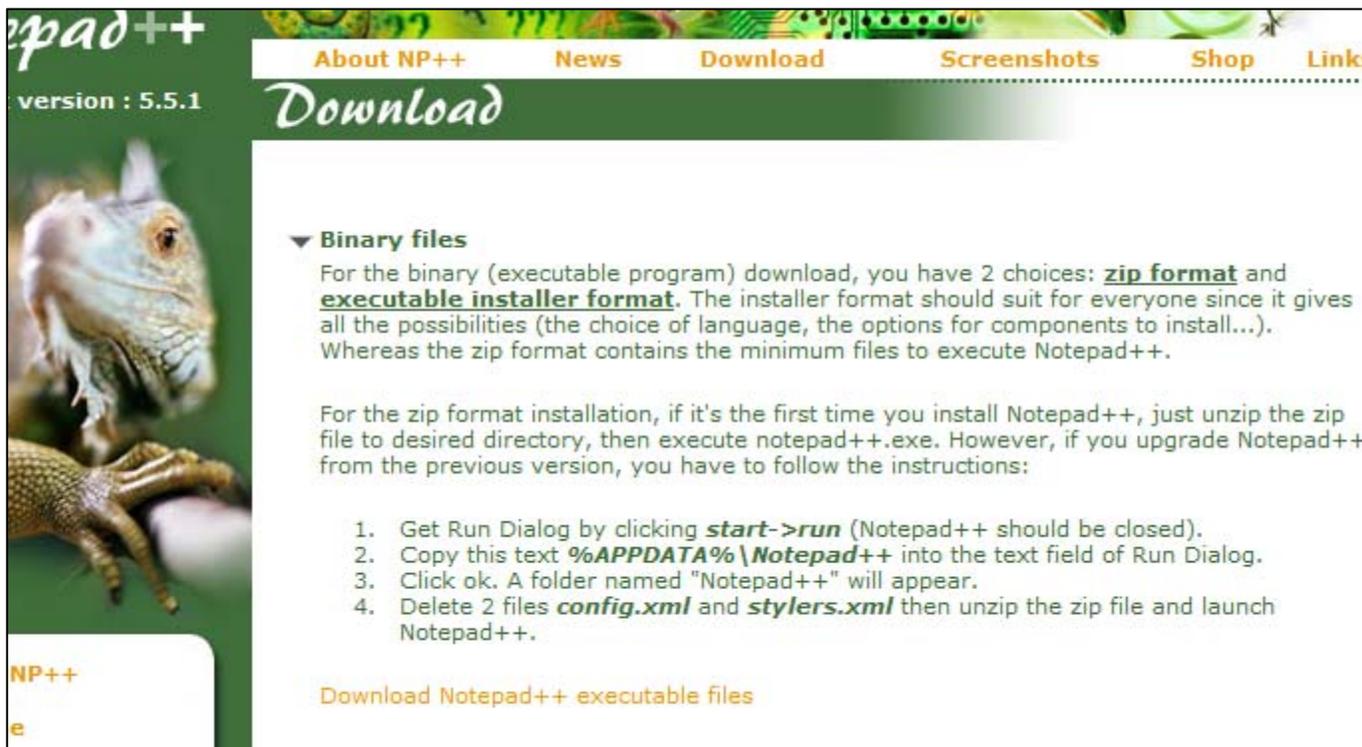
Click "Download" to continue.



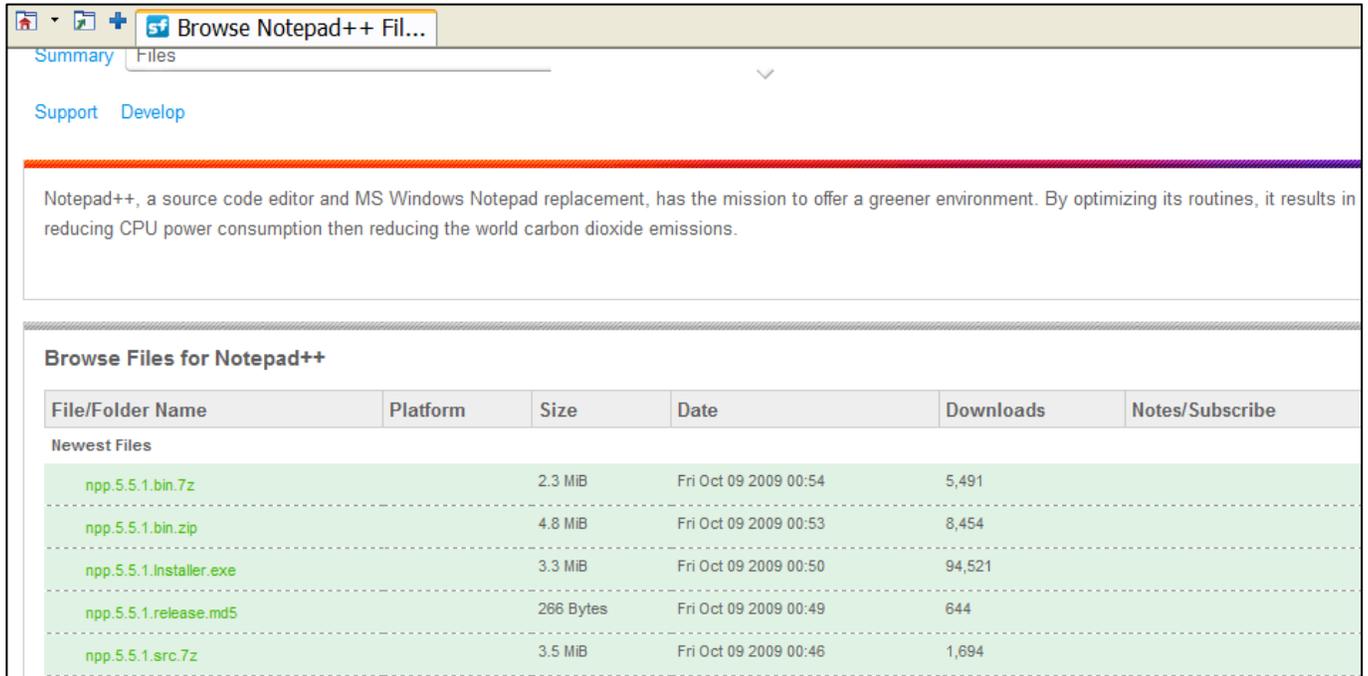
On the Download page, select the link for binary files.



Select the link to download Notepad++ executable files.



On the next screen, scroll down to see the sub-header Browse Files for Notepad++, and click the link titled, "npp.5.5.1.Installer.exe."



The screenshot shows the Notepad++ website's file browser. A red box highlights the 'Browse Files for Notepad++' section. Below it is a table with columns: File/Folder Name, Platform, Size, Date, Downloads, and Notes/Subscribe. The table lists several files, with 'npp.5.5.1.Installer.exe' highlighted in green.

File/Folder Name	Platform	Size	Date	Downloads	Notes/Subscribe
Newest Files					
npp.5.5.1.bin.7z		2.3 MiB	Fri Oct 09 2009 00:54	5,491	
npp.5.5.1.bin.zip		4.8 MiB	Fri Oct 09 2009 00:53	8,454	
npp.5.5.1.Installer.exe		3.3 MiB	Fri Oct 09 2009 00:50	94,521	
npp.5.5.1.release.md5		266 Bytes	Fri Oct 09 2009 00:49	644	
npp.5.5.1.src.7z		3.5 MiB	Fri Oct 09 2009 00:46	1,694	

To install the application upon the completion of the download, click the "Run" button. If you wish to save the file for later use click "Save."



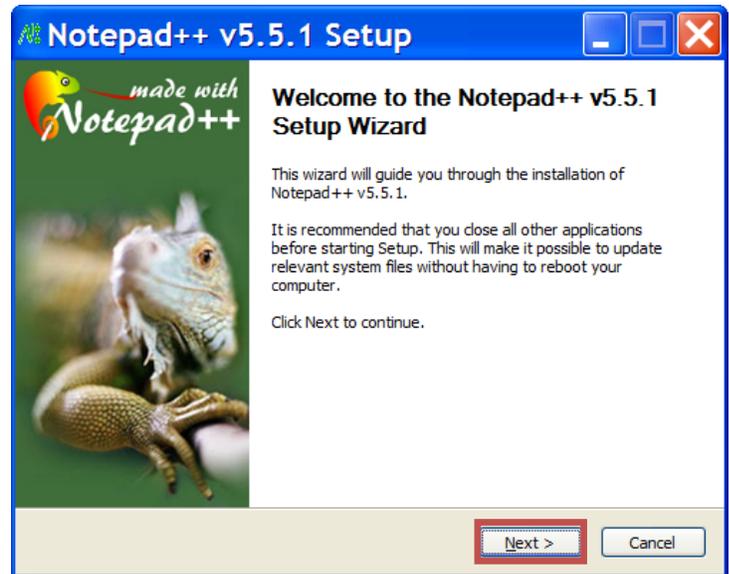
You may be prompted by a security warning when you being the installation process. Click "Run" to continue.



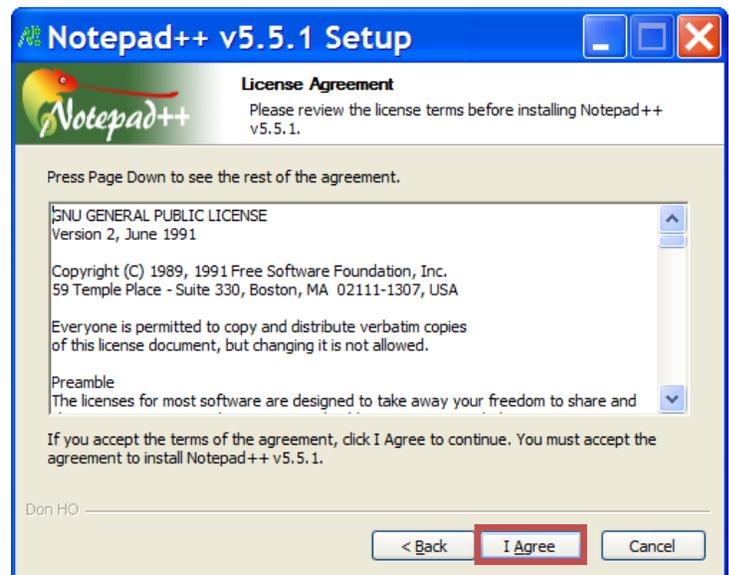
Select a language for installation.



Click "Next" to proceed with the installation process.



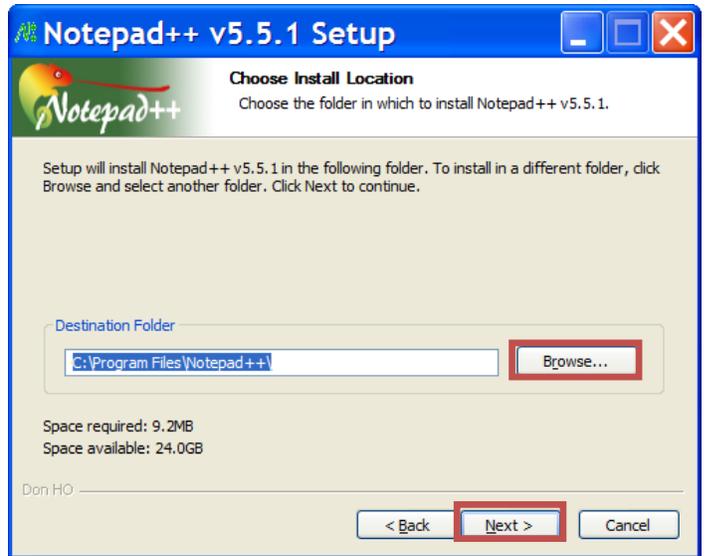
Review the terms for the application and click "I Agree" to continue.



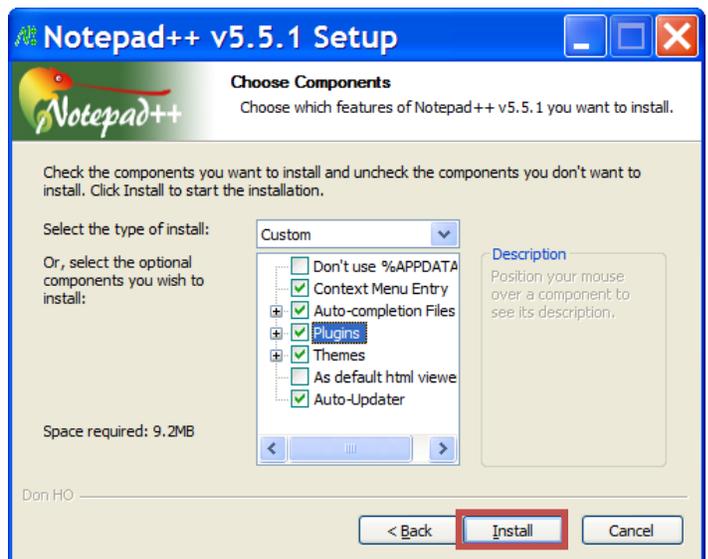
The default destination folder should be:

C:\Program Files\Notepad++\

If you wish to change that location, click "Browse" to choose a location. Click "Next" when you are ready to continue.



Advanced users may wish to customize their installation from this menu. To continue with the default installation, click "Install."

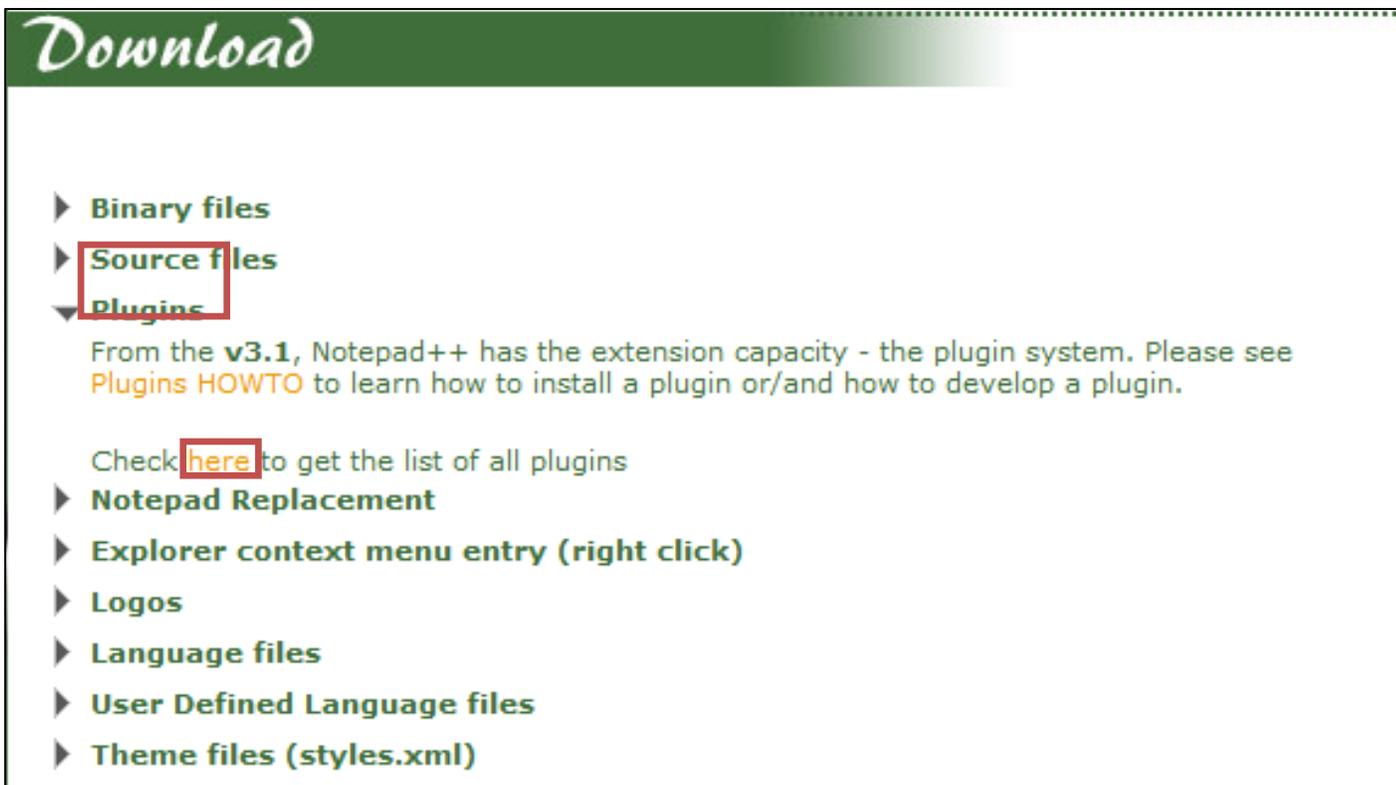


When installation is complete, a confirmation window will appear. Click "Finish" to end the installation process.



Plug-in Installation

Notepad++ is installed, but now we have to install the plug-in tools for Notepad++ to validate XML files. Return to the Notepad++ Web site (<http://notepad-plus.sourceforge.net/uk/site.htm>) and click "Download" on the top menu. From the Download screen, select "Plugins" and click the "here" link:



Download

- ▶ Binary files
- ▶ **Source files**
- ▶ **Plugins**

From the **v3.1**, Notepad++ has the extension capacity - the plugin system. Please see [Plugins HOWTO](#) to learn how to install a plugin or/and how to develop a plugin.

Check [here](#) to get the list of all plugins

- ▶ Notepad Replacement
- ▶ Explorer context menu entry (right click)
- ▶ Logos
- ▶ Language files
- ▶ User Defined Language files
- ▶ Theme files (styles.xml)

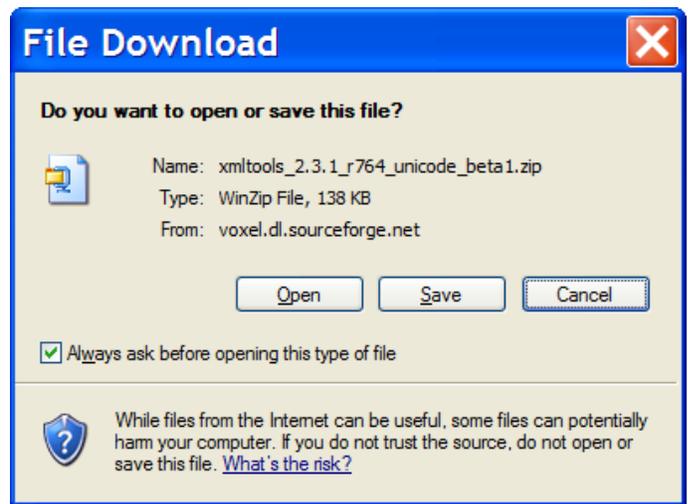
You will be taken to the SourceForge.net Plugin Central Web site. Scroll down the page until you reach the section of the page with the line "XML Tools" – click this link to continue.

<p>Version 1.1 (ANSI / Unicode) XBrackets Lite (Unicode) XBrackets Lite (ANSI)</p>	<p>XBrackets Lite allows to autocomplete brackets ({{""}}) i.e. it inserts corresponding right bracket when the left bracket is typed. The plugin uses "smart" autocompletion:<p></p> <ul style="list-style-type: none"> ▪ next character is analysed for {{{ brackets; ▪ next & previous characters are analysed for " quote. 	<p>Dovgan Vitaliy</p>
<p>Version 2.3.1β1 r764 (Unicode) XML Tools</p>	<p>This plugin is a small set of useful tools for editing XML with Notepad++. The plugin is libXML2-based.</p> <p>The plugin features are:</p> <ul style="list-style-type: none"> ▪ XML syntax Check ▪ XML Schema (XSD) + DTD Validation ▪ XML tag autoclose ▪ Pretty print ▪ Linearize XML ▪ Current XML Path ▪ Conversion XML <-> Text ▪ Comment / Uncomment ▪ XPath expression evaluation 	<p>Nicolas Crittin</p>

Your download will begin on the next screen in a few seconds. If for some reason it does not begin, click the "direct link" to manually download the file.



Depending on your browser's configuration you will be prompted to open or save the file, or one of those options will occur automatically.



Once your download is complete, the file will open with your system's ZIP default application or you will open it manually. Once you have opened it, the following files will be seen. Right click on the XMLTools.dll file and click "Copy."

Name ^	Size	Type	Date Modified
Closed Tasks		File Folder	10/6/2009 7...
XML Validate		File Folder	10/20/2009 ...
xmltools_2.3.1_r764_unicode_beta1.zip	139 KB	WinZip File	10/20/2009 ...
XMLTools.dll	304 KB	Application E...	10/16/2009 ...

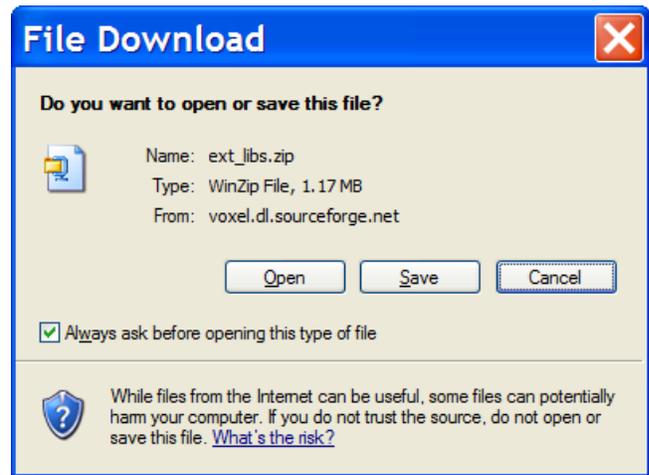
In your Windows Explorer, go to the folder listed below and paste the file:
C:\Program Files\Notepad++\plugins (this location may be different if you did not use the default installation location)

Go to the Web page: <http://sourceforge.net/projects/npp-plugins/files/XML%20Tools/External%20libs/>
Click the link "ext_libs.zip."

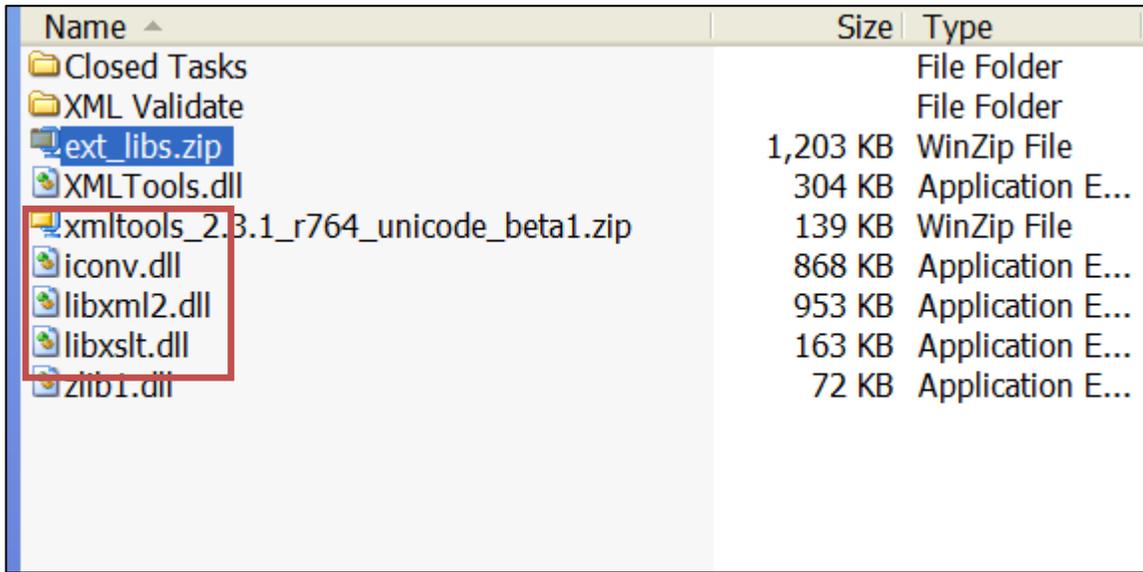
The screenshot shows the SourceForge project page for 'Notepad++ Plugins' by 'jenslorenz'. The page includes a breadcrumb trail, a share button, and a description of the project. Below the description is a table titled 'Browse Files for Notepad++ Plugins' with columns for File/Folder Name, Platform, Size, Date, Downloads, and Notes/Su. The file 'ext_libs.zip' is listed under the 'External libs' subdirectory and is highlighted with a red box.

File/Folder Name	Platform	Size	Date	Downloads	Notes/Su
Subdirectory (view all files)					
XML Tools					
External libs					
ext_libs.zip		1.2 MiB	Mon Oct 27 2008 18:06	83,574	

Depending on your browser's configuration you will be prompted to open or save the file, or one of those options will occur automatically.



Once your download is complete, the file will open with your system's ZIP default application or you will open it manually. Once you have opened it, the following files will be seen. Select the four DLL files shown below, right click, and select "Copy."



Name	Size	Type
Closed Tasks		File Folder
XML Validate		File Folder
ext_libs.zip	1,203 KB	WinZip File
XMLTools.dll	304 KB	Application E...
xmltools_2.3.1_r764_unicode_beta1.zip	139 KB	WinZip File
iconv.dll	868 KB	Application E...
libxml2.dll	953 KB	Application E...
libxslt.dll	163 KB	Application E...
zlib1.dll	72 KB	Application E...

In your Windows Explorer go to the folder listed below and paste the files:

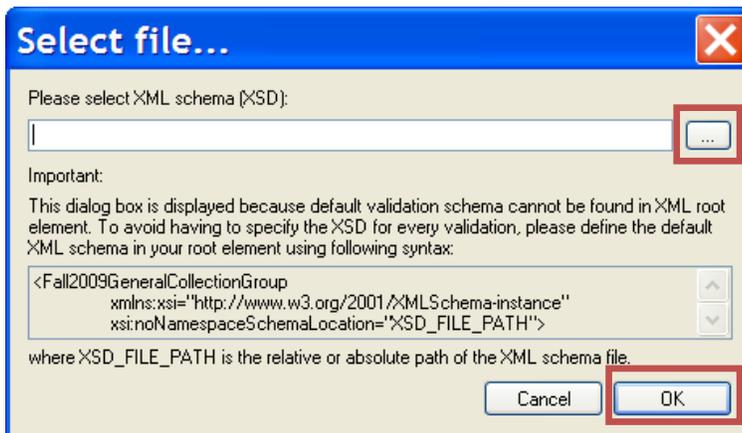
C:\Program Files\Notepad++\plugins (this location may be different if you did not use the default installation location).

Validating a Schema File

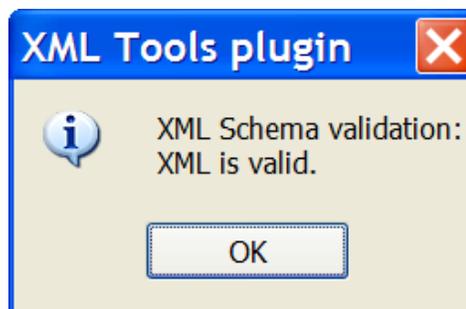
1. Locate the collection schema file from CEPI's Institutions of Higher Education Web page.
2. Right click on the XML Schema and choose "Save target as" and save it to your local drive as .xsd.
3. Open Notepad++ (click on Start > Programs > Notepad++ > Notepad++).
4. Under File Menu, Open your XML file to be validated.
5. When the file is opened, press Ctrl+Alt+Shift+M

Click on the browse button (labeled "...") next to XML Schema to select the .xsd file saved in the second step.

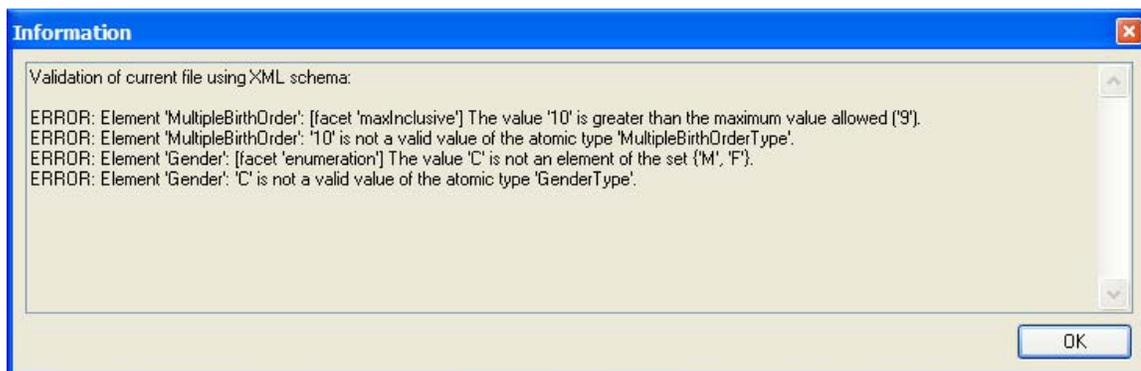
Click "OK" to continue.



If the XML file is valid, the following window will open. Click on "OK" to close.



If the file is not valid, then the following Information window will open.



How to Read the Error Message

Each error will have two lines of code in the information window. The first line is the value from the Schema (XSD) file and the second line is the error that has occurred in the XML file.

In the illustrated example above, in the first error, the first line indicates that the maximum value allowed for MultipleBirthOrder is 9. The second line indicates that a record contains number 10 as the value of Multiple Birth Order.

In the second error below, the first line indicates that only M or F are allowed in the Gender Type. The second line indicates there is a record in the XML file with the value of C instead of M or F.

In order to fix the errors, select the Search menu and click on Find. In the Find What text box, type "<elementname> value" and click "Find Next" (e.g., [<Gender>C]).

The search will point to the error record.

```
<SubmittingEntityTypeCode>D</SubmittingEntityTypeCode>
<SubmittingEntityCode>25260</SubmittingEntityCode>
</SubmittingEntity>
<PersonalCore>
  <UIC>0019191919</UIC>
  <LastName>DOE</LastName>
  <FirstName>JAMES</FirstName>
  <MiddleName>G</MiddleName>
  <DateOfBirth>2008-09-05</DateOfBirth>
  <MultipleBirthOrder>0</MultipleBirthOrder>
  <Gender>C</Gender>
</PersonalCore>
<EntityDemographics>
  <OperatingISDESANumber>25</OperatingISDESANumber>
```

Fix the error and perform the validation again.