**Warning**: Class \_\_PHP\_Incomplete\_Class has no unserializer in **/home/content/n3pnexwpnas01\_data01/09/3283109/html/wp-content/object-cache.php** on line **894**  
  
**Warning**: session\_start(): Cannot send session cookie - headers already sent by (output started at /home/content/n3pnexwpnas01\_data01/09/3283109/html/wp-content/object-cache.php:894) in **/home/content/n3pnexwpnas01\_data01/09/3283109/html/wp-content/plugins/stats-counter/wpadm-class-wp.php** on line **9**  
  
**Warning**: session\_start(): Cannot send session cache limiter - headers already sent (output started at /home/content/n3pnexwpnas01\_data01/09/3283109/html/wp-content/object-cache.php:894) in **/home/content/n3pnexwpnas01\_data01/09/3283109/html/wp-content/plugins/stats-counter/wpadm-class-wp.php** on line **9**

[](http://excel-access-programmer.com/)

* [About Us](http://excel-access-programmer.com/about/)
* [Contact Us](http://excel-access-programmer.com/contact/)
* [Blog](http://excel-access-programmer.com/news/)

**VBA code to read text from Image**

In our day to day work at office we face a lot of challenges. We can take the challenge and learn something new or avoid the work and stick to our as usual work. I always choose to take the challenge and many occasions I learnt something new which made my day.

I am going to share one of the knowledge I recently gained by doing same. My boss asked me if he can copy some image and paste it as text where ever he wants. I knew that to read text from images we need to use OCR technology but, I did not have any idea about how to plug it with VBA. I did some digging and and found OneNote which is part of Ms-Office. Honestly speaking I never used it before. But, I learnt that it allows us to read text from images. I did some more digging and learnt a lot about the programming interface of OneNote.

As we already know that new MS-Office stands on XML structure . We need to modify the underlying XML to make any modification in OneNote.

So, We need to take care of three steps to achieve our task.

1. Save the image from clipboard to file
2. Insert the file to OneNote and read the converted text.
3. Place the text to clipboard again

**Step 1**

I added a Chat sheet to the Excel. Pasted the already copied image into it. Save the Chart to Image file. Below is the code for same.

Public Sub LoadFile()

Dim cht As Chart

Dim obj As Object

Dim shp As Shape

Set cht = ThisWorkbook.Charts("Chart1")

cht.Select

ActiveSheet.Paste

Set shp = cht.Shapes(1)

'Exporting the Chart to Image file.

cht.Export ThisWorkbook.Path & "Test1.bmp"

'Calling the Second method to read text

GetText ThisWorkbook.Path & "Test1.bmp"

'Deleting the image file

Kill ThisWorkbook.Path & "Test1.bmp"

'Deleting the image from the Chart for future use

cht.Shapes.Range(Array("chart")).Delete

End Sub

**Step 2**

There are few steps we need to perform. Convert the file to 64 bit binary data. Create a new OneNote application. Add a new page. Insert the Image file. Read the converted text. Below is the code for same.

Public Function GetFileBytes(strPath As String) As Byte()

With CreateObject("ADODB.Stream")

.Open

.Type = 1 ' adTypeBinary

.LoadFromFile strPath

GetFileBytes = .Read

.Close

End With

End Function

Function EncodeBase64(arrData() As Byte) As String

'arrData = StrConv(text, vbFromUnicode)

Dim objXML As MSXML2.DOMDocument60

Dim objNode As MSXML2.IXMLDOMElement

Set objXML = New MSXML2.DOMDocument60

Set objNode = objXML.createElement("b64")

objNode.DataType = "bin.base64"

objNode.nodeTypedValue = arrData

EncodeBase64 = objNode.text

Set objNode = Nothing

Set objXML = Nothing

End Function

Public Sub GetText(strfile As String)

Dim bytFile() As Byte

Dim base64String, imageXmlStr

Dim nCounter As Integer

Dim oneNote As OneNote14.Application

Set oneNote = New OneNote14.Application

' Get all of the Notebook nodes.

Dim nodes As MSXML2.IXMLDOMNodeList

Set nodes = GetFirstOneNoteNotebookNodes(oneNote)

Set bitmap = LoadPicture(strfile)

bytFile = GetFileBytes(strfile)

base64String = EncodeBase64(bytFile)

hh = Round(bitmap.Width / 30)

ww = Round(bitmap.Height / 30)

If Not nodes Is Nothing Then

' Get the first OneNote Notebook in the XML document.

Dim node As MSXML2.IXMLDOMNode

Set node = nodes(0)

Dim noteBookName As String

noteBookName = node.Attributes.getNamedItem("name").text

' Get the ID for the Notebook so the code can retrieve

' the list of sections.

Dim notebookID As String

notebookID = node.Attributes.getNamedItem("ID").text

' Load the XML for the Sections for the Notebook requested.

Dim sectionsXml As String

oneNote.GetHierarchy notebookID, hsSections, sectionsXml, xs2010

Dim secDoc As MSXML2.DOMDocument60

Set secDoc = New MSXML2.DOMDocument60

If secDoc.LoadXML(sectionsXml) Then

' select the Section nodes

Dim secNodes As MSXML2.IXMLDOMNodeList

Set secNodes = secDoc.DocumentElement.getElementsByTagName("one:Section")

If Not secNodes Is Nothing Then

' Get the first section.

Dim secNode As MSXML2.IXMLDOMNode

Set secNode = secNodes(0)

Dim sectionName As String

sectionName = secNode.Attributes.getNamedItem("name").text

Dim sectionID As String

sectionID = secNode.Attributes.getNamedItem("ID").text

' Create a new blank Page in the first Section

' using the default format.

Dim newPageID As String

oneNote.CreateNewPage sectionID, newPageID, npsDefault

' Get the contents of the page.

Dim outXML As String

oneNote.GetPageContent newPageID, outXML, piAll, xs2010

Dim doc As MSXML2.DOMDocument60

Set doc = New MSXML2.DOMDocument60

' Load Page's XML into a MSXML2.DOMDocument60 object.

If doc.LoadXML(outXML) Then

' Get Page Node.

Dim pageNode As MSXML2.IXMLDOMNode

Set pageNode = doc.getElementsByTagName("one:Page")(0)

Dim newElement As MSXML2.IXMLDOMElement

Dim newNode As MSXML2.IXMLDOMNode

' Create Outline node.

Set newElement = doc.createElement("one:Outline")

newElement.setAttribute "lang", "en-US"

Set newNode = pageNode.appendChild(newElement)

' Create OEChildren.

Set newElement = doc.createElement("one:OEChildren")

Set newNode = newNode.appendChild(newElement)

' Create OE.

Set newElement = doc.createElement("one:OE")

newElement.setAttribute "lang", "en-US"

Set newNode = newNode.appendChild(newElement)

' Create Image.

Set newElement = doc.createElement("one:Image")

'newElement.setAttribute "format", "bmp"

Set newNode = newNode.appendChild(newElement)

' Create Size.

Set newElement = doc.createElement("one:Size")

newElement.setAttribute "width", ww

newElement.setAttribute "height", hh

newElement.setAttribute "isSetByUser", "true"

newNode.appendChild newElement

'Push the image bnary data

Set newElement = doc.createElement("one:Data")

newElement.text = base64String

newNode.appendChild newElement

' Update OneNote with the new content.

oneNote.UpdatePageContent doc.XML, , , True

Dim strxml As String

'Get the contnt back from OneNote Page

oneNote.GetPageContent newPageID, strxml

doc.LoadXML strxml

Set nodes = doc.DocumentElement.getElementsByTagName("one:OCRText")

nCounter = 1

Do While nodes.Length = 0

oneNote.GetPageContent newPageID, strxml

doc.LoadXML strxml

Set nodes = doc.DocumentElement.getElementsByTagName("one:OCRText")

nCounter = nCounter + 1

If nCounter = 100 Then

CopyToClipboard "Image is not readable"

Exit Sub

End If

Loop

Set nodes = doc.DocumentElement.getElementsByTagName("one:OCRText")

CopyToClipboard nodes(0).text

Debug.Print doc.XML

End If

Else

MsgBox "OneNote 2010 Section nodes not found."

End If

Else

MsgBox "OneNote 2010 Section XML Data failed to load."

End If

Else

MsgBox "OneNote 2010 XML Data failed to load."

End If

End Sub

Private Function GetAttributeValueFromNode(node As MSXML2.IXMLDOMNode, attributeName As String) As String

If node.Attributes.getNamedItem(attributeName) Is Nothing Then

GetAttributeValueFromNode = "Not found."

Else

GetAttributeValueFromNode = node.Attributes.getNamedItem(attributeName).text

End If

End Function

Private Function GetFirstOneNoteNotebookNodes(oneNote As OneNote14.Application) As MSXML2.IXMLDOMNodeList

' Get the XML that represents the OneNote notebooks available.

Dim notebookXml As String

' OneNote fills notebookXml with an XML document providing information

' about what OneNote notebooks are available.

' You want all the data and thus are providing an empty string

' for the bstrStartNodeID parameter.

oneNote.GetHierarchy "", hsNotebooks, notebookXml, xs2010

' Use the MSXML Library to parse the XML.

Dim doc As MSXML2.DOMDocument60

Set doc = New MSXML2.DOMDocument60

If doc.LoadXML(notebookXml) Then

Set GetFirstOneNoteNotebookNodes = doc.DocumentElement.getElementsByTagName("one:Notebook")

Else

Set GetFirstOneNoteNotebookNodes = Nothing

End If

End Function

**Step 3**

Now, we have the converted text. We need to put that text into clipboard. Microsoft provides Microsoft Forms DataObject to do some basic clipboard operations. We will use it to put the text to clipboard. Below is the code to achieve the same.

Public Sub CopyToClipboard(strText As String)

Dim objDo As DataObject

Set objDo = New DataObject

objDo.SetText strText

objDo.PutInClipboard

Set objDo = Nothing

End Sub

Here you can [download](https://drive.google.com/file/d/0B12rLMDd7HDPT2swaHg5dHZNejA/view?usp=sharing) a working copy of the program.

(Copy the image and press ctrl+shift+T on Chart Sheet. Then paste anywhere the extracted text)

Tagged on: [How to read text from image in excel](http://excel-access-programmer.com/tag/how-to-read-text-from-image-in-excel/)    [VBA code for OCR](http://excel-access-programmer.com/tag/vba-code-for-ocr/)    [VBA code to add new page in OneNote](http://excel-access-programmer.com/tag/vba-code-to-add-new-page-in-onenote/)    [VBA code to insert image in OneNote page](http://excel-access-programmer.com/tag/vba-code-to-insert-image-in-onenote-page/)    [VBA code to modify OneNote page](http://excel-access-programmer.com/tag/vba-code-to-modify-onenote-page/)    [VBA code to read text from image](http://excel-access-programmer.com/tag/vba-code-to-read-text-from-image/)

[OfficeAccelerators](http://excel-access-programmer.com/author/officeaccelerators/) [March 20, 2016](http://excel-access-programmer.com/vba-code-to-read-text-from-image/) [Uncategorized](http://excel-access-programmer.com/category/uncategorized/), [VBA code to read text from image](http://excel-access-programmer.com/category/vba-code-to-read-text-from-image/) [No Comments](http://excel-access-programmer.com/vba-code-to-read-text-from-image/#respond)

* [← How to show Google Map on Excel sheet](http://excel-access-programmer.com/how-to-show-google-map-on-excel-sheet/)
* [VBA code to upload/download file from SharePoint using credentials →](http://excel-access-programmer.com/vba-code-to-uploaddownload-file-from-sharepoint-using-credentials/)

**Leave a Reply**

Top of Form

Your email address will not be published. Required fields are marked \*

Comment 

Name \* 

Email \*

Website

Bottom of Form

Top of Form

Bottom of Form

Top of Form



Bottom of Form

**Topics**

* [July 2016](http://excel-access-programmer.com/2016/07/) (1)
* [April 2016](http://excel-access-programmer.com/2016/04/) (7)
* [March 2016](http://excel-access-programmer.com/2016/03/) (3)
* [December 2015](http://excel-access-programmer.com/2015/12/) (1)
* [July 2015](http://excel-access-programmer.com/2015/07/) (1)
* [June 2015](http://excel-access-programmer.com/2015/06/) (1)
* [March 2015](http://excel-access-programmer.com/2015/03/) (3)
* [February 2015](http://excel-access-programmer.com/2015/02/) (4)
* [January 2015](http://excel-access-programmer.com/2015/01/) (1)
* [November 2014](http://excel-access-programmer.com/2014/11/) (1)
* [September 2014](http://excel-access-programmer.com/2014/09/) (1)
* [May 2014](http://excel-access-programmer.com/2014/05/) (2)
* [January 2014](http://excel-access-programmer.com/2014/01/) (1)
* [December 2013](http://excel-access-programmer.com/2013/12/) (4)
* [November 2013](http://excel-access-programmer.com/2013/11/) (14)

Copyright © 2017 . Powered by [WordPress](http://wordpress.org). Theme: Spacious by [ThemeGrill](http://themegrill.com/themes/spacious).