Microsoft Word to DokuWiki Converter

This is a vbScript that will do basic formatting conversion from word to dokuwiki. It's not perfect, but it will save you a lot of time if you have lots of word docs like me. Also, cells pasted into word from excel will be converted, so it's almost an excel converter.

It's poorly written ugly uncommented crappy code that has come to you as a fixed up version of an old version which supported old syntax which was created by someone else mashing together conversion code for openoffice and mediawiki. So yeh, it's crap!! Here's a readme that, just like the code, is mostly ripped off from other people. I give all the credit for this to whoever it was that did it, except for the bits I fixed up so it would actually work ;) So here's a fixed up version of their readme:

INTRO

Convert Microsoft Word document content to MediaWiki markup.

This is a Word Visual Basic macro. Usage requires a running copy of Microsoft Word that supports Visual Basic macros. (Word 97 or greater). Features

- Replaces smart quotes/double-quotes with dumb equivalents
- Escapes the following characters: * # { } [ ] ~ ^^ | '
- Converts external hyperlinks
- Converts H1-H5 headings
- Converts bold/italic/underline/strikethrough/superscript/subscript formatting
- Converts bulleted/numbered lists

CAVEATS

- No guarantees, no official support, just like the DokuWiki itself (or Word =})
- No in-document hyperlink conversion (so make Wiki links yourself after pasting)
- May not work well with documents that have highly customized styles/templates. It works best with documents written in the default “Normal” template
- Image/graphics/figures are not supported.
- Tables go a bit crazy if you have carriage returns inside the cells... I can't get around that, it's a microsoft thing.
- | or ^ characters within tables get converted to I and /\ or they'll screw it up.

INSTALLATION

```
 1. Copy the code from below, and save it to a file on your disk named
Word2Doku.bas
 2. Start Word
 3. Bring up the Visual Basic Editor (Tools->Macro->Visual Basic Editor or
Alt+F11).
 4. From the VBE, import the macro library (File->Import File...) and select
```

```
the file you downloaded.
```

USAGE

```
 1. Open a word document to convert
 2. Run the Word2DokuWiki macro by bringing up the Macros dialog
(Tools->Macro->MacrosÔÇª or Alt+F8), selecting Word2DokuWiki and clicking
Run.
 3. The macro converts the document to DokuWiki markup and places a copy of
the content on the system clipboard.
 4. Switch to a DokuWiki editor page (or new page) and paste the result in.
```

ATTRIBUTION

I did not come up with this macro, though I was looking to make something similar. The actual coding was done by http://infpro.com/projekte/wordwiki/ . Thank them for the fact you're using this. I only trivially hacked the functions to work with the slight differences inherent in DokuWiki formatting. And I rarely use Word. =}

A similar macro is available for converting OpenOffice documents to DokuWiki format at : http://homepages.paradise.net.nz/hillview/OOo/

I don't foresee any problems or need for this macro to be enhanced, but if you need to get a hold of me, the best way is by leaving a comment on my blog Across Weirdish Wild Space @ http://www.wakatara.com/blog/

Enjoy !

So yeh, ignore that stuff above, if you're a MIDIBoxer, just hit me up if you need support. Cheers!

```
Attribute VB_Name = "Word2DokuWikiv3"

Sub Word2DokuWiki()
    Application.ScreenUpdating = False
    ReplaceQuotes
    DokuWikiEscapeChars
    DokuWikiConvertHyperlinks
    DokuWikiConvertH1
    DokuWikiConvertH2
    DokuWikiConvertH3
    DokuWikiConvertH4
    DokuWikiConvertH5
    DokuWikiConvertItalic
    DokuWikiConvertBold
    DokuWikiConvertUnderline
    DokuWikiConvertStrikeThrough
    DokuWikiConvertSuperscript
    DokuWikiConvertSubscript
```

```
    DokuWikiConvertLists
    DokuWikiConvertTable
    UndoDokuWikiEscapeChars
    ' Copy to clipboard
   ActiveDocument.Content.Copy
   Application.ScreenUpdating = True
End Sub

Private Sub DokuWikiConvertH1()
    ReplaceHeading wdStyleHeading1, "======"
End Sub

Private Sub DokuWikiConvertH2()
    ReplaceHeading wdStyleHeading2, "====="
End Sub

Private Sub DokuWikiConvertH3()
    ReplaceHeading wdStyleHeading3, "===="
End Sub

Private Sub DokuWikiConvertH4()
        ReplaceHeading wdStyleHeading4, "==="
End Sub

Private Sub DokuWikiConvertH5()
    ReplaceHeading wdStyleHeading5, "=="
End Sub

Private Sub DokuWikiConvertH6()
    ReplaceHeading wdStyleHeading5, "="
End Sub

Private Sub DokuWikiConvertBold()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Font.Bold = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
```

```
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "**" Then
                    .InsertBefore "**"
                    End If
                    If Not Right(.Text, 2) = "**" Then
                    .InsertAfter "**"
                    End If
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.Bold = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertItalic()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Font.Italic = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "//" Then
                    .InsertBefore "//"
                    End If
                    If Not Right(.Text, 2) = "//" Then
                    .InsertAfter "//"
                    End If
```

```
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.Italic = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertUnderline()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Font.Underline = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "__" Then
                    .InsertBefore "__"
                    End If
                    If Not Right(.Text, 2) = "__" Then
                    .InsertAfter "__"
                    End If
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.Underline = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertStrikeThrough()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
```

```
        .Font.StrikeThrough = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "<del>" Then
                    .InsertBefore "<del>"
                    End If
                    If Not Right(.Text, 2) = "</del>" Then
                    .InsertAfter "</del>"
                    End If
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.StrikeThrough = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertSuperscript()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Font.Superscript = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
```

```
                .Text = Trim(.Text)
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "<sup>" Then
                    .InsertBefore "<sup>"
                    End If
                    If Not Right(.Text, 2) = "</sup>" Then
                    .InsertAfter "</sup>"
                    End If
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.Superscript = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertSubscript()
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Font.Subscript = True
        .Text = ""
        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                .Text = Trim(.Text)
                If Len(.Text) > 1 And InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
                If Not .Text = vbCr Then
                    If Not Left(.Text, 2) = "<sub>" Then
```

```
                    .InsertBefore "<sub>"
                    End If
                    If Not Right(.Text, 2) = "</sub>" Then
                    .InsertAfter "</sub>"
                    End If
                End If
                .Style = ActiveDocument.Styles("Default Paragraph Font")
                .Font.Subscript = False
            End With
        Loop
    End With
End Sub

Private Sub DokuWikiConvertLists()
    Dim para As Paragraph
    For Each para In ActiveDocument.ListParagraphs
        With para.Range
            .InsertBefore "  "
             If .ListFormat.ListType = wdListBullet Then
                 .InsertBefore "*"
             Else
                  .InsertBefore "-"
              End If
            For i = 1 To .ListFormat.ListLevelNumber
                   .InsertBefore "  "
           Next i
            .ListFormat.RemoveNumbers
        End With
    Next para
End Sub

Private Sub DokuWikiConvertHyperlinks()
    Dim hyperCount As Integer
    hyperCount = ActiveDocument.Hyperlinks.Count
    For i = 1 To hyperCount
        With ActiveDocument.Hyperlinks(1)
            Dim addr As String
            addr = .Address
            .Delete
            .Range.InsertBefore "["
            .Range.InsertAfter "-" & addr & "]"
        End With
    Next i
End Sub

' Replace all smart quotes with their dumb equivalents
Private Sub ReplaceQuotes()
    Dim quotes As Boolean
    quotes = Options.AutoFormatAsYouTypeReplaceQuotes
    Options.AutoFormatAsYouTypeReplaceQuotes = False
    ReplaceString ChrW(8220), """"
```

```
    ReplaceString ChrW(8221), """"
    ReplaceString "ë", "'"
    ReplaceString "í", "'"
    Options.AutoFormatAsYouTypeReplaceQuotes = quotes
End Sub

Private Sub DokuWikiEscapeChars()
    EscapeCharacter "*"
    EscapeCharacter "#"
    EscapeCharacter "_"
    EscapeCharacter "-"
    EscapeCharacter "+"
    EscapeCharacter "{"
    EscapeCharacter "}"
    EscapeCharacter "["
    EscapeCharacter "]"
    EscapeCharacter "~"
    EscapeCharacter "^^"
    EscapeCharacter "|"
    EscapeCharacter "'"
End Sub

Private Function ReplaceHeading(styleHeading As String, headerPrefix As
String)
    Dim normalStyle As Style
    Set normalStyle = ActiveDocument.Styles(wdStyleNormal)
    ActiveDocument.Select
    With Selection.Find
        .ClearFormatting
        .Style = ActiveDocument.Styles(styleHeading)
        .Text = ""

        .Format = True
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
        .Forward = True
        .Wrap = wdFindContinue
        Do While .Execute
            With Selection
                If InStr(1, .Text, vbCr) Then
                    ' Just process the chunk before any newline characters
                    ' We'll pick-up the rest with the next search
                    .Collapse
                    .MoveEndUntil vbCr
                End If
                ' Don't bother to markup newline characters (prevents a
loop, as well)
               If Not .Text = vbCr Then
```

```
                    .InsertBefore headerPrefix
                    .InsertBefore vbCr
                    .InsertAfter headerPrefix
                End If
                .Style = normalStyle
            End With
        Loop
    End With
End Function

Private Sub DokuWikiConvertTable()
Dim TotTables As Long
Do While ActiveDocument.Tables.Count() > 0
ActiveDocument.Tables(1).Range.Select
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = " $s$|$s$ "
.Replacement.Text = "I"
.Forward = True
.Wrap = wdFindContinue
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = " $s$^^$s$ "
.Replacement.Text = "/\"
.Forward = True
.Wrap = wdFindContinue
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Find.ClearFormatting
Application.DefaultTableSeparator = "|"
Selection.Rows.ConvertToText Separator:=wdSeparateByDefaultListSeparator,
NestedTables:=True
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
```

```
.Text = "^p"
.Replacement.Text = "|^p|"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.InsertBefore ("|")
Selection.InsertParagraphAfter
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = "^p|^p"
.Replacement.Text = "^p"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = "$s$blank$s$"
.Replacement.Text = ""
.Forward = True
.Wrap = wdFindContinue
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = "||"
.Replacement.Text = "|  |"
.Forward = True
.Wrap = wdFindStop
```

```
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
With Selection.Find
.Text = "||"
.Replacement.Text = "|  |"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = "| |"
.Replacement.Text = "|  |"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
With Selection.Find
.Text = "| |"
.Replacement.Text = "|  |"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Selection.Paragraphs(1).Range.Select
Selection.Find.ClearFormatting
```

```
Selection.Find.Replacement.ClearFormatting
With Selection.Find
.Text = "|"
.Replacement.Text = "^^"
.Forward = True
.Wrap = wdFindStop
.Format = False
.MatchCase = False
.MatchWholeWord = False
.MatchWildcards = False
.MatchSoundsLike = False
.MatchAllWordForms = False
End With
Selection.Find.Execute Replace:=wdReplaceAll
Loop
End Sub
Private Sub UndoDokuWikiEscapeChars()

    UndoEscapeCharacter "*"
    UndoEscapeCharacter "#"
    UndoEscapeCharacter "_"
    UndoEscapeCharacter "-"
    UndoEscapeCharacter "+"
    UndoEscapeCharacter "{"
    UndoEscapeCharacter "}"
    UndoEscapeCharacter "["
    UndoEscapeCharacter "]"
    UndoEscapeCharacter "~"
    UndoEscapeCharacter "^^"
    UndoEscapeCharacter "|"
    UndoEscapeCharacter "'"

End Sub

Private Function EscapeCharacter(char As String)
    ReplaceString char, " $s$" & char & "$s$ "
End Function

Private Function UndoEscapeCharacter(char As String)
    ReplaceString " $s$" & char & "$s$ ", char
End Function

Private Function ReplaceString(findStr As String, replacementStr As String)
    Selection.Find.ClearFormatting
    Selection.Find.Replacement.ClearFormatting
    With Selection.Find
        .Text = findStr
        .Replacement.Text = replacementStr
        .Forward = True
        .Wrap = wdFindContinue
        .Format = False
```

```
        .MatchCase = False
        .MatchWholeWord = False
        .MatchWildcards = False
        .MatchSoundsLike = False
        .MatchAllWordForms = False
    End With
    Selection.Find.Execute Replace:=wdReplaceAll
End Function
```

From:
https://wiki.midibox.org/ - **MIDIbox**

Permanent link:
https://wiki.midibox.org/doku.php?id=stryd_one_word2doku&rev=1160904901

Last update: **2008/06/25 10:01**