

Text count in AnyCount and MS Office

Foreword

This document is a comparison of text count accuracy in **Microsoft® Office 2003 Professional**, **Microsoft® Office Enterprise 2007** and **AnyCount 7.0®**. The comparison has been made to determine which file and text elements are and are not included in text count results produced by **AnyCount 7.0®** and **Microsoft® Office 2003 Professional**, **Microsoft® Office Enterprise 2007** applications after each of them has been used for counting same file.

Text counts in **Microsoft® Office 2003 Professional**, **Microsoft® Office Enterprise 2007** were obtained the following way:

- **Microsoft® Word 2003**: On the Tools menu click Word Count item for all file formats which can be opened with **Microsoft® Word 2003**. These file formats are: DOC, RTF, WPD, and TXT.
- **Microsoft® Word 2007**: In the Proofing group click Review tab, then click Word Count, for all file formats which can be opened with **Microsoft® Word 2007**. These file formats are: DOC, DOCX, RTF, WPD, and TXT.
- Other **Microsoft® Office** files: Right-click the file, then in the context menu click Properties. These file formats are: XLS, XLSX, PPT, PPTX, PPS.

Note that non-office files like PDF, HLP, CHM, HTM, HTML, PUB, VSD **OpenOffice** files, like ODT, SXW, SDW, ODS, ODP, and image files BMP, JPG, PNG, GIF properties do not contain means of text count.


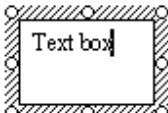
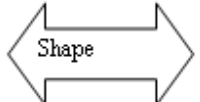

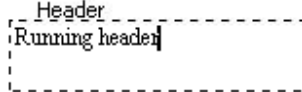
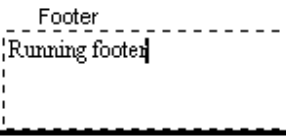


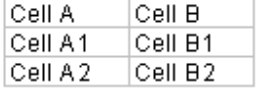
The table below displays whether certain text and file elements are included in text counts by **AnyCount 7.0®** and **Microsoft® Office 2003 Professional**, **Microsoft® Office Enterprise 2007**.

Yes – indicates that the element is included in text count results.

No – indicates that the element is not included in text count results.

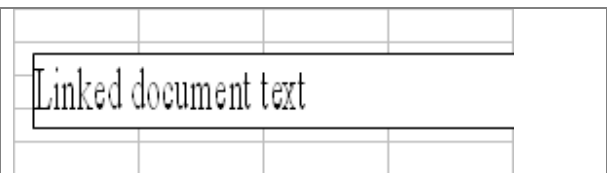
N/a – indicates that the program or program package cannot calculate amount of text in this file type.



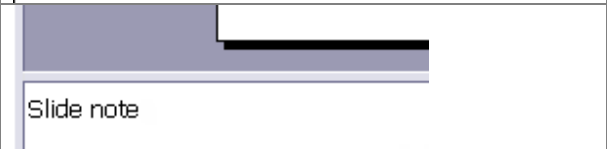


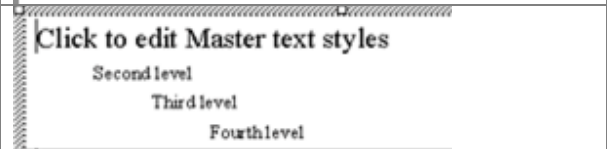


Comparison results of text counts produced by Microsoft® Office 2003 Professional, Microsoft® Office Enterprise 2007 and AnyCount 7.0®

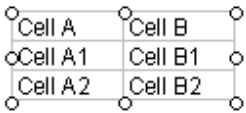
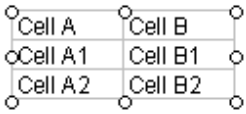
 1. DOC, DOCX, RTF, WPD – MS Word				
Object	Example	AnyCount counts	MS Word 2003 counts	MS Word 2007 counts
Text	–	Yes	Yes	Yes
Text boxes		Yes	No	Yes
Shapes		Yes	No	Yes
WordArt		Yes	No	No
Running headers		Yes	No	No
Running footers		Yes	No	No
FootNotes		Yes	Yes	Yes
End Notes		Yes	Yes	Yes
Embedded documents (i.e. XLS table embedded in DOC file)		Yes	No	No

Linked documents (i.e. XLS table linked with DOC file)		Yes	No	No
Comments		Yes	No	No
Hidden text		Yes	No	No

2. XLS, XLSX – MS Excel			
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	No
Header			
Footer			
Cell comments		Yes	No
Text Boxes		Yes	No
Shapes		Yes	No
Embedded documents (i.e. DOC document embedded in XLS file)		Yes	No

Linked documents (i.e. DOC document linked with XLS file)		Yes	No
---	--	-----	----

 3. PPT, PPTX, PPS – MS PowerPoint			
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	Yes
Shapes		Yes	Yes
Slide Notes		Yes	Yes
Slide footers		Yes	No
Presentation headers	Presentation header	Yes	No
Presentation footers	Slide footer	Yes	No
Slide master		Yes	No
Note master		Yes	No
Handout master		Yes	No
Title master		Yes	No

Embedded documents (i.e. XLS table embedded in PPT file)		Yes	No
Linked documents (i.e. XLS table linked with PPT file)		Yes	No

4. PDF – Adobe Acrobat

Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	N/a

5. TXT

Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	Yes

6. CSV – Comma-separated values

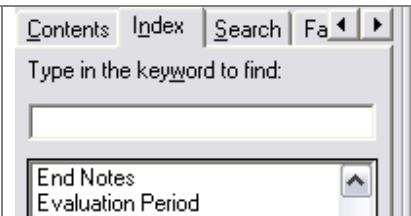
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	No

7. HLP – Windows Help

Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	N/a

8. CHM

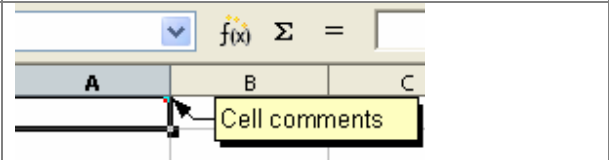
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	N/a
Alt texts		Yes	N/a
Form controls	Name: <input type="text"/> <input type="button" value="Subscribe"/> E-mail: <input type="text"/>	Yes	N/a
Contents		Yes	N/a

Index keywords			Yes	N/a
----------------	---	--	-----	-----

8. ODT, SXW, SDW – Open Office

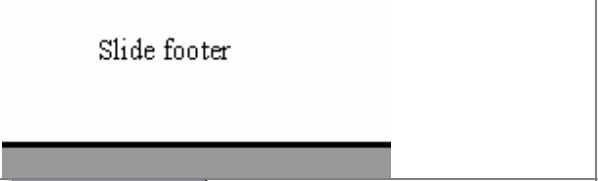

Object	Example	AnyCount counts	MS Office counts
Text	-	Yes	N/a
Running headers		Yes	N/a
Running footers		Yes	N/a
End Notes		Yes	N/a
FootNotes		Yes	N/a
Text boxes		Yes	N/a


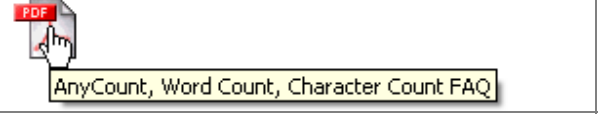

9. ODS – Open Office

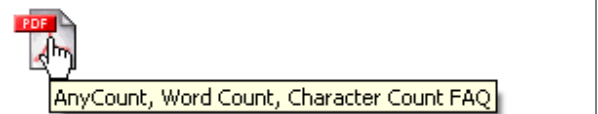
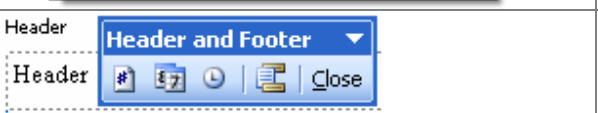

Object	Example	AnyCount counts	MS Office counts
Text	-	Yes	N/a
Cell comments		Yes	N/a

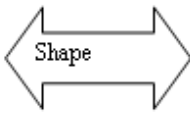
10. ODP – Open office


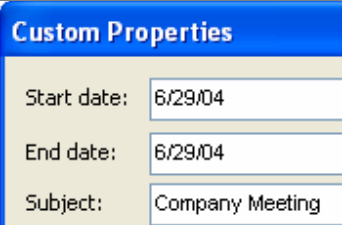



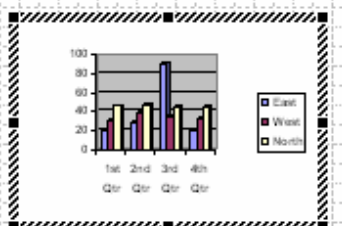
Object	Example	AnyCount counts	MS Office counts
Text	-	Yes	N/a


Slide footer		Yes	N/a
Slide Note		Yes	N/a

10. HTM, HTML			
Object	Example	AnyCount counts	MS Office counts
Text (without code elements)		Yes	No
Meta description	<pre><meta name="description" content="Word count software produces"></pre>	Yes	No
Meta keywords	<pre><meta name="keywords" content="Word Count, Line Count, Page Count"></pre>	Yes	No
Alt texts		Yes	No
Form controls (i.e. text in HTML forms, buttons, and so on)		Yes	No

11. PUB – MS Publisher			
Object	Example	AnyCount counts	MS Office counts
Alt Text		Yes	No
Header		Yes	No
Footer		Yes	No

Shapes		Yes	No						
Embedded object (i.e. XLS table embedded in PUB file)	<table border="1" data-bbox="614 443 865 533"> <tr><td>Cell A</td><td>Cell B</td></tr> <tr><td>Cell A1</td><td>Cell B1</td></tr> <tr><td>Cell A2</td><td>Cell B2</td></tr> </table>	Cell A	Cell B	Cell A1	Cell B1	Cell A2	Cell B2	Yes	No
Cell A	Cell B								
Cell A1	Cell B1								
Cell A2	Cell B2								
Linked object (i.e. XLS table linked with PUB file)	<table border="1" data-bbox="614 622 865 712"> <tr><td>Cell A</td><td>Cell B</td></tr> <tr><td>Cell A1</td><td>Cell B1</td></tr> <tr><td>Cell A2</td><td>Cell B2</td></tr> </table>	Cell A	Cell B	Cell A1	Cell B1	Cell A2	Cell B2	Yes	No
Cell A	Cell B								
Cell A1	Cell B1								
Cell A2	Cell B2								

12. VSD – MS Visio			
Object	Example	AnyCount counts	MS Office counts
Markup		Yes	No
Custom properties		Yes	No
Header		Yes	No
Footer		Yes	No
Shapes		Yes	No
Embedded object (i.e. XLS diagram embedded in VSD file)		Yes	No

Linked object (i.e. DOC text linked with VSD file)		Yes	No
--	---	-----	----

 13. BMP, JPG, PNG, GIF – Image file formats			
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	N/a

 14. SLP – Sisulizer: http://www.sisulizer.com/			
Object	Example	AnyCount counts	MS Office counts
Text	–	Yes	N/a