

## Re: Structure of Office XP Word Document

**Source:**

<http://www.tech-archive.net/Archive/Office/microsoft.public.office.developer.automation/2005-03/0051.html>

---

**From:** ITJRW (*justice\_at\_community.nospam*)

**Date:** 03/06/05

Date: Sat, 5 Mar 2005 18:19:03 -0600

Thanks Cindy, I've just finished e-mailing officeff@microsoft.com to obtain the BIFF. You were a big help. By the way, I had an painting teacher when I was 9-11 yrs. old (about 50 years ago) in Houston by the name of Daphne Meister. Don't know if that name is common but the lady sure could teach art.

Dick

"Cindy M -WordMVP-" <C.Meister-C@hispeed.ch> wrote in message news:VA.0000a84b.0044de6e@speedy...

> *Hi Dick,*

>

> *In order to examine the "real" information in the binary file format you'd have to request the BIFF from Microsoft. There's a Knowledge Base article that*

> *tells how to go about doing it, although I don't have the number at hand.*

>

> *Since you have Office XP, an alternate approach might be to compare what you*

> *get when you save the document as a webpage (to round-trip HTML format).*

> *OTOH,*

> *this often results in throwing out the damaged part of the document (much*

> *as*

> *copying without the last paragraph mark does).*

>

> *FWIW, it's a known fact that Word documents are very susceptible to*

> *"hiccups"*

> *during any kind of transmission process, be it over a network or via*

> *Email.*

> *The only safeguard I'm aware of is to ZIP the file (Winzip, for example)*

> *before transmitting. That, or fix the faulty network connection.*

>

>> *My office is required to use a UNIX (AIX) Universe (IBM) database*

>> *application to host and access our data through a telnet connection. Our*

>> *office uses mailmerge in two different implemenations, and as part of the*

>> *process, through host processes and VBA macros, Word is opened and a Word*

>> *files stored on this UNIX server are trasmitted to our Windows 2000*

>> *client*  
>> *machines. The master document is transmitted and a .dat merge data file*  
>> *is*  
>> *created on the host and also transmitted to the client. This process is*  
>> *failing from time to time and requires numerous retries to complete the*  
>> *transmission process. Somehow the documents are corrupted during this*  
>> *failure and the only way to "fix" the documents is for someone who has*  
>> *never merged the document to open the document for editing, copy and save*  
>> *to*  
>> *another document all content up to the last paragraph symbol in the*  
>> *document*  
>> *(the final character showing with All checked to show tabs, spaces,*  
>> *paragraphs, etc.). Pasting this new content into a new document and*  
>> *saving*  
>> *it on the host remedies the problem. However I want to be able to*  
>> *compare*  
>> *the contents of the "header" document information that is found at or*  
>> *after*  
>> *that last paragraph symbol (at least as far as what I can see being*  
>> *transmitted) in both a corrupted document and the newly created document*  
>> *to*  
>> *try to identify how this corruption is taking place during access and*  
>> *transmission.*  
>>  
>> *My question is where can I obtain a the contents of the document*  
>> *information*  
>> *stored at or after the last paragraph symbol at the end of every Word*  
>> *document? What would be the best way to examine this in each document?*  
>>  
>  
> *Cindy Meister*  
> *INTER-Solutions, Switzerland*  
> *<http://homepage.swissonline.ch/cindymeister> (last update Jun 8 2004)*  
> *<http://www.word.mvps.org>*  
>  
> *This reply is posted in the Newsgroup; please post any follow question or*  
> *reply in the newsgroup and not by e-mail :-)*  
>