

TET Migration Guide

Latest version covered in this document: TET 5.2

The TET API generally remains compatible among major releases. Sometimes API methods are phased out and replaced with an improved successor. In these situations a newer PDFlib version declares the old method as deprecated without changing the functionality in any way. In most cases the new method is more general or more powerful than the old one which it replaces. The new method is always preferable. New code should not use deprecated API methods, and existing code should switch to the new replacement method. Deprecated methods may be removed in a future release.

This document contains recommendations for users who migrate existing TET application code which has been developed with an older release. The migration guide explains how to identify deprecated API features which are used in application code. Once identified, the deprecated features should be replaced with the recommended newer ones.

Note The document compatibility.txt in the distribution package contains additional version-by-version information regarding API compatibility.

1 Identify deprecated TET API Features

1.1 Identify deprecated TET Method Calls at Compile Time

Since TET 5.2 does not contain any API methods which are marked as deprecated this section does not apply.

1.2 Identify deprecated API Method Calls at Runtime

Since TET 5.2 does not contain any API methods which are marked as deprecated this section does not apply.

1.3 Identify deprecated Options

All language bindings. Deprecated options can be identified in the logging output with the logging class `api=2` which can be set as follows:

```
tet.set_option("logging={filename=deprecated.log classes={api=2}}");
```

With this setting the log file contains a message for all API methods which are called with deprecated options, e.g.

```
TET_open_page(tet_0x7ffffba9c0e70, 0, 3, "imageanalysis={smallimages={maxheight=50}}")  
[Option "smallimages" is deprecated]
```

These logging entries are only warnings. TET execution continues and PDF contents are extracted as usual. However, it is strongly recommended to adjust the application code to avoid deprecated options.

2 Deprecated TET API Features

2.1 Deprecated PDFlib TET Methods

Table 2.1 *Deprecated TET API functions/methods and recommended replacements*

deprecated API method	deprecated since	removed in	recommended replacement method
<code>TET_document_mem()</code>	TET 3.0	TET 5.1	<code>TET_create_pvf()</code> and <code>TET_open_document()</code>
<code>TET_get_xml_data()</code>	TET 5.0	TET 5.2	<code>TET_get_tetml()</code>
<code>TET_utf16_to_utf8()</code>	TET 4.1	TET 5.2	<code>TET_convert_to_unicode()</code>
<code>TET_utf8_to_utf16()</code>			
<code>TET_utf32_to_utf8()</code>			
<code>TET_utf8_to_utf32()</code>			
<code>TET_utf16_to_utf32()</code>			
<code>TET_utf32_to_utf16()</code>			

2.2 Deprecated Options

Table 2.2 *Deprecated TET API options and recommended replacements*

deprecated option or suboption	deprecated since	removed in	recommended replacement option
<code>TET_open_document()</code>:			
<code>tetml</code> , <code>suboption version</code>	TET 4.0	TET 4.4	n/a (TET 3 XML output is no longer supported)
<code>tetml</code> , <code>suboption elements</code> , <code>suboption docxmp</code>	TET 5.0		<code>suboption metadata</code>
<code>keepppua</code>	TET 4.0		<code>TET_open_document()</code> , option <code>fold</code> with one the following foldings: <code>fold={{[:Private_Use:] preserve}}</code> or <code>fold={{[:Private_Use:] unknownchar}}</code>
<code>TET_open_page()</code> and <code>TET_process_page()</code>:			
<code>imageanalysis</code> , <code>suboption smallimages</code>	TET 5.2		<code>suboptions</code> <code>heightrange</code> , <code>sizerange</code> , <code>widthrange</code>
<code>skipengines</code>	TET 5.0		<code>engines</code>
<code>contentanalysis</code> , <code>suboption ideographic</code>	TET 5.0		n/a (the default is now the reasonable behavior of <code>ideographic=keep</code>)
<code>contentanalysis</code> , <code>suboptions lineseparator</code> , <code>paraseparator</code> , <code>wordseparator</code>	TET 4.0		<code>options</code> <code>lineseparator</code> , <code>paraseparator</code> , <code>wordseparator</code> of <code>TET_open_document()</code>
<code>contentanalysis</code> , <code>suboption zoneseparator</code>	TET 3.0	TET 5.2	n/a (option was never supported and didn't have any effect since TET 5.0)
<code>TET_write_image_file()</code> and <code>TET_get_image_data()</code>:			
<code>compression</code> , <code>preferredtiffcompression</code>	TET 5.2		n/a
<code>format</code>	TET 4.0	TET 5.2	n/a (didn't have any effect since TET 4.0)
<code>smallimages</code> , <code>suboption maxcount</code>	TET 5.0		n/a

2.3 TET History

Table 2.3 TET release history

<i>TET version</i>	<i>release date</i>
<i>TET 3</i>	2009
<i>TET 4</i>	2010
<i>TET 4.1</i>	2012
<i>TET 4.2</i>	2013
<i>TET 4.3</i>	2014
<i>TET 4.4</i>	2015
<i>TET 5</i>	2015
<i>TET 5.1</i>	2017
<i>TET 5.2</i>	2019