

NAME

HTML::FormatText – Format HTML as plaintext

VERSION

version 2.12

SYNOPSIS

```
use HTML::TreeBuilder;
$tree = HTML::TreeBuilder->new->parse_file("test.html");

use HTML::FormatText;
$formatter = HTML::FormatText->new(leftmargin => 0, rightmargin => 50);
print $formatter->format($tree);
```

or, more simply:

```
use HTML::FormatText;
my $string = HTML::FormatText->format_file(
    'test.html',
    leftmargin => 0, rightmargin => 50
);
```

DESCRIPTION

HTML::FormatText is a formatter that outputs plain latin1 text. All character attributes (bold/italic/underline) are ignored. Formatting of HTML tables and forms is not implemented.

HTML::FormatText is built on HTML::Formatter and documentation for that module applies to this – especially “new” in HTML::Formatter, “format_file” in HTML::Formatter and “format_string” in HTML::Formatter.

You might specify the following parameters when constructing the formatter:

leftmargin (alias *lm*)

The column of the left margin. The default is 3.

rightmargin (alias *rm*)

The column of the right margin. The default is 72.

SEE ALSO

HTML::Formatter

INSTALLATION

See perlmodinstall for information and options on installing Perl modules.

BUGS AND LIMITATIONS

You can make new bug reports, and view existing ones, through the web interface at <http://rt.cpan.org/Public/Dist/Display.html?Name=HTML-Format>.

AVAILABILITY

The project homepage is <https://metacpan.org/release/HTML-Format>.

The latest version of this module is available from the Comprehensive Perl Archive Network (CPAN). Visit <http://www.perl.com/CPAN/> to find a CPAN site near you, or see <https://metacpan.org/module/HTML::Format/>.

AUTHORS

- Nigel Metherringham <nigel@cpan.org>
- Sean M Burke <sburke@cpan.org>
- Gisle Aas <gisle@ActiveState.com>

COPYRIGHT AND LICENSE

This software is copyright (c) 2015 by Nigel Metherringham, 2002–2005 Sean M Burke, 1999–2002 Gisle Aas.

This is free software; you can redistribute it and/or modify it under the same terms as the Perl 5 programming language system itself.