

## Redmine - Defect #19737

### HTML Sanitizer not working for Outlook mails

2015-04-29 12:41 - Rupesh J

<b>Status:</b> Closed	<b>Start date:</b>
<b>Priority:</b> High	<b>Due date:</b>
<b>Assignee:</b> Jean-Philippe Lang	<b>% Done:</b> 0%
<b>Category:</b> Email receiving	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b> 3.1.0	<b>Affected version:</b> 3.0.2
<b>Resolution:</b> Fixed	
<b>Description</b>	
HTML Sanitizer not working for Outlook mails. There are some unwanted tags not stripped from the mail body.	
<b>Related issues:</b>	
Related to Redmine - Feature # 16962: Better handle html-only emails	<b>Closed</b>

#### History

##### #1 - 2015-04-29 15:54 - Toshi MARUYAMA

- Target version set to 3.0.3

##### #2 - 2015-04-29 15:55 - Toshi MARUYAMA

This is Redmine 3.0 regression according to #19537#note-9.

##### #4 - 2015-04-29 15:59 - Toshi MARUYAMA

- Related to Feature #16962: Better handle html-only emails added

##### #5 - 2015-04-29 16:56 - Deoren Moor

+1

I can confirm that Outlook emails are improperly handled:

{{collapse(View details)}}

```
[if !mso]><style>v\:.* {behavior:url(#default#VML);}
o\:.* {behavior:url(#default#VML);}
w\:.* {behavior:url(#default#VML);}
.shape {behavior:url(#default#VML);}
</style><![endif]<!--
/* Font Definitions */
@font-face {font-family:Wingdings;
panose-1:5 0 0 0 0 0 0 0 0;}
@font-face {font-family:Wingdings;
panose-1:5 0 0 0 0 0 0 0 0;}
@font-face {font-family:Calibri;
panose-1:2 15 5 2 2 4 3 2 4;}
@font-face {font-family:Tahoma;
panose-1:2 11 6 4 3 5 4 4 2 4;}
@font-face {font-family:Verdana;
```

```

panose-1:2 11 6 4 3 5 4 4 2 4;}
@font-face {font-family:"Trebuchet MS";
panose-1:2 11 6 3 2 2 2 2 2 4;}
/* Style Definitions */
p.MsoNormal, li.MsoNormal, div.MsoNormal {margin:0in;
margin-bottom:.0001pt;
font-size:12.0pt;
font-family:"Times New Roman","serif";}
h1 {mso-style-priority:9;
mso-style-link:"Heading 1 Char";
margin:0in;
margin-bottom:.0001pt;
font-size:14.5pt;
font-family:"Trebuchet MS","sans-serif";
font-weight:bold;}
h2 {mso-style-priority:9;
mso-style-link:"Heading 2 Char";
margin:0in;
margin-bottom:.0001pt;
font-size:13.0pt;
font-family:"Trebuchet MS","sans-serif";
font-weight:bold;}
h3 {mso-style-priority:9;
mso-style-link:"Heading 3 Char";
margin:0in;
margin-bottom:.0001pt;
font-size:13.0pt;
font-family:"Trebuchet MS","sans-serif";
font-weight:bold;}
a:link, span.MsoHyperlink {mso-style-priority:99;
color:#2A5685;
text-decoration:underline;}
a:visited, span.MsoHyperlinkFollowed {mso-style-priority:99;
color:#2A5685;
text-decoration:underline;}
p {mso-style-priority:99;
mso-margin-top-alt:auto;
margin-right:0in;
mso-margin-bottom-alt:auto;
margin-left:0in;
font-size:12.0pt;
font-family:"Times New Roman","serif";}
span.Heading1Char {mso-style-name:"Heading 1 Char";
mso-style-priority:9;
mso-style-link:"Heading 1";
font-family:"Cambria","serif";
color:#365F91;
font-weight:bold;}
span.Heading2Char {mso-style-name:"Heading 2 Char";
mso-style-priority:9;
mso-style-link:"Heading 2";
font-family:"Cambria","serif";
color:#4F81BD;

```

font-weight:bold;}  
span.Heading3Char {mso-style-name:"Heading 3 Char";  
mso-style-priority:9;  
mso-style-link:"Heading 3";  
font-family:"Cambria","serif";  
color:#4F81BD;  
font-weight:bold;}  
span.footer {mso-style-name:footer;  
font-style:italic;}  
span.header {mso-style-name:header;}  
span.EmailStyle24 {mso-style-type:personal-reply;  
font-family:"Calibri","sans-serif";  
color:#1F497D;}  
.MsoChpDefault {mso-style-type:export-only;  
font-size:10.0pt;}  
@page WordSection1 {size:8.5in 11.0in;  
margin:1.0in 1.0in 1.0in 1.0in;}  
div.WordSection1 {page:WordSection1;}  
/\* List Definitions \*/  
@list l0 {mso-list-id:688027544;  
mso-list-template-ids:-2060393820;}  
@list l0:level1 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:.5in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Symbol;}  
@list l0:level2 {mso-level-number-format:bullet;  
mso-level-text:o;  
mso-level-tab-stop:1.0in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:"Courier New";  
mso-bidi-font-family:"Times New Roman";}  
@list l0:level3 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:1.5in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level4 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:2.0in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level5 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:2.5in;

mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level6 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:3.0in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level7 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:3.5in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level8 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:4.0in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l0:level9 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:4.5in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Wingdings;}  
@list l1 {mso-list-id:960186181;  
mso-list-template-ids:-1253122092;}  
@list l1:level1 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:.5in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:Symbol;}  
@list l1:level2 {mso-level-number-format:bullet;  
mso-level-text:o;  
mso-level-tab-stop:1.0in;  
mso-level-number-position:left;  
text-indent:-.25in;  
mso-ansi-font-size:10.0pt;  
font-family:"Courier New";  
mso-bidi-font-family:"Times New Roman";}  
@list l1:level3 {mso-level-number-format:bullet;  
mso-level-text:□;  
mso-level-tab-stop:1.5in;  
mso-level-number-position:left;

```

text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level4 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:2.0in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level5 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:2.5in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level6 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:3.0in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level7 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:3.5in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level8 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:4.0in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
@list l1:level9 {mso-level-number-format:bullet;
mso-level-text:□;
mso-level-tab-stop:4.5in;
mso-level-number-position:left;
text-indent:-.25in;
mso-ansi-font-size:10.0pt;
font-family:Wingdings;}
ol {margin-bottom:0in;}
ul {margin-bottom:0in;}
-->[if gte mso 9]><xml>
<o:shapedefaults v:ext="edit" spidmax="1026" />
</xml><![endif][if gte mso 9]><xml>
<o:shapelayout v:ext="edit">
<o:idmap v:ext="edit" data="1" />
</o:shapelayout></xml><![endif]

```

}}

The actual human-readable text is three lines. However a similar issue is also happening with iOS (v7.x) emails as I reported on #15716.

Another issue I noticed along with this one is that the **Truncate emails after one of these lines** setting does not appear to be working. I've opened #19740 for that.

Additional notes:

- I tested sending from Outlook and converting to Plaintext when replying and the email was scraped properly. The truncate option also seemed to work properly

- Same (good, expected) results when responding via Exchange web mail and first converting format to Plain text before replying.

- I tested using Exchange web mail and got similar results (see below)

- I tested replying using Outlook and leaving the email as-is and the email was rejected with **Validation failed: Documented cannot be blank**.

Presumably this is because of the Truncate option not working properly because the custom **Documented** field is listed well below the Truncate phrase that we're using.

Reply from Exchange webmail, left in HTML format

```
<!--
body {font-family:Verdana,sans-serif;
font-size:0.8em;
color:#484848}
h1, h2, h3 {font-family:"Trebuchet MS",Verdana,sans-serif;
margin:0px}
h1 {font-size:1.2em}
h2, h3 {font-size:1.1em}
a, a:link, a:visited {color:#2A5685}
a:hover, a:active {color:#c61a1a}
fieldset.attachments {border-width:1px 0 0 0}
hr {width:100%;
height:1px;
background:#ccc;
border:0}
span.footer {font-size:0.8em;
font-style:italic}
-->
P {margin-top:0;margin-bottom:0;}
```

Responding via web interface. I'm leaving the format as HTML and not manually pruning existing text in the reply.

**#6 - 2015-05-08 09:11 - Jean-Philippe Lang**

- Status changed from New to Needs feedback

| *HTML Sanitizer not working for Outlook mails.*

Please attach such an email (with dummy text), so we can start working on this.

**#7 - 2015-05-08 16:16 - Rupesh J**

- File *sample2.htm* added

I saved the outlook msg file as HTML,  
let me know if this attachment is ok?

Thanks.

**#8 - 2015-05-09 10:41 - Jean-Philippe Lang**

A .eml file that contains the full email source would be better, thanks.

**#9 - 2015-05-10 09:23 - Jean-Philippe Lang**

- Target version changed from 3.0.3 to 3.0.4

**#10 - 2015-05-11 18:03 - Deoren Moor**

- File Test email.msg added

Jean-Philippe Lang wrote:

| A .eml file that contains the full email source would be better, thanks.

From the attached msg file the following was scraped and inserted into the OP of a new issue on our Redmine v3.0.2 installation (raw text, grabbed by choosing to "edit" the OP so I could copy/paste here):

{{collapse(View details)

```
[if gte mso 9]&gt;&lt;xml&gt;
&lt;o:OfficeDocumentSettings&gt;
&lt;o:AllowPNG/&gt;
&lt;o:OfficeDocumentSettings&gt;
&lt;/xml&gt;&lt;![endif][if gte mso 9]&gt;&lt;xml&gt;
&lt;w:WordDocument&gt;
&lt;w:Zoom&gt;140&lt;/w:Zoom&gt;
&lt;w:SpellingState&gt;Clean&lt;/w:SpellingState&gt;
&lt;w:TrackMoves/&gt;
&lt;w:TrackFormatting/&gt;
&lt;w:EnvelopeVis/&gt;
&lt;w:PunctuationKerning/&gt;
&lt;w:ValidateAgainstSchemas/&gt;
&lt;w:SaveIfXMLInvalid&gt;false&lt;/w:SaveIfXMLInvalid&gt;
&lt;w:IgnoreMixedContent&gt;false&lt;/w:IgnoreMixedContent&gt;
&lt;w:AlwaysShowPlaceholderText&gt;false&lt;/w:AlwaysShowPlaceholderText&gt;
&lt;w:DoNotPromoteQF/&gt;
&lt;w:LidThemeOther&gt;EN-US&lt;/w:LidThemeOther&gt;
&lt;w:LidThemeAsian&gt;X-NONE&lt;/w:LidThemeAsian&gt;
&lt;w:LidThemeComplexScript&gt;X-NONE&lt;/w:LidThemeComplexScript&gt;
&lt;w:Compatibility&gt;
&lt;w:BreakWrappedTables/&gt;
&lt;w:SnapToGridInCell/&gt;
&lt;w:WrapTextWithPunct/&gt;
&lt;w:UseAsianBreakRules/&gt;
&lt;w:DontGrowAutofit/&gt;
&lt;w:SplitPgBreakAndParaMark/&gt;
&lt;w:EnableOpenTypeKerning/&gt;
```

&lt;w:DontFlipMirrorIndents/&gt;  
&lt;w:OverrideTableStyleHps/&gt;  
&lt;/w:Compatibility&gt;  
&lt;m:mathPr&gt;  
&lt;m:mathFont m:val=&quot;Cambria Math&quot;/&gt;  
&lt;m:brkBin m:val=&quot;before&quot;/&gt;  
&lt;m:brkBinSub m:val=&quot;&#45;-&quot;/&gt;  
&lt;m:smallFrac m:val=&quot;off&quot;/&gt;  
&lt;m:dispDef/&gt;  
&lt;m:lMargin m:val=&quot;0&quot;/&gt;  
&lt;m:rMargin m:val=&quot;0&quot;/&gt;  
&lt;m:defJc m:val=&quot;centerGroup&quot;/&gt;  
&lt;m:wrapIndent m:val=&quot;1440&quot;/&gt;  
&lt;m:intLim m:val=&quot;subSup&quot;/&gt;  
&lt;m:naryLim m:val=&quot;undOvr&quot;/&gt;  
&lt;/m:mathPr&gt;&lt;/w:WordDocument&gt;  
&lt;/xml&gt;&lt;![endif][if gte mso 9]&gt;&lt;/xml&gt;  
&lt;w:LatentStyles DefLockedState=&quot>false&quot; DefUnhideWhenUsed=&quot>true&quot; DefSemiHidden=&quot>true&quot;  
DefQFormat=&quot>false&quot; DefPriority=&quot;99&quot; LatentStyleCount=&quot;267&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;0&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;Normal&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;heading 1&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 2&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 3&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 4&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 5&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 6&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 7&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 8&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;9&quot; QFormat=&quot>true&quot; Name=&quot;heading 9&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 1&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 2&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 3&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 4&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 5&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 6&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 7&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 8&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;39&quot; Name=&quot;toc 9&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;35&quot; QFormat=&quot>true&quot; Name=&quot;caption&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;10&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;Title&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;1&quot; Name=&quot;Default Paragraph Font&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;11&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;Subtitle&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;22&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;Strong&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;20&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
QFormat=&quot>true&quot; Name=&quot;Emphasis&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; Priority=&quot;59&quot; SemiHidden=&quot>false&quot; UnhideWhenUsed=&quot>false&quot;  
Name=&quot;Table Grid&quot;/&gt;  
&lt;w:LsdException Locked=&quot>false&quot; UnhideWhenUsed=&quot>false&quot; Name=&quot;Placeholder Text&quot;/&gt;











```

QFormat=&quot;true&quot; Name=&quot;Intense Reference&quot;/&gt;
&lt;/w:LsdException Locked=&quot;false&quot; Priority=&quot;33&quot; SemiHidden=&quot;false&quot; UnhideWhenUsed=&quot;false&quot;
QFormat=&quot;true&quot; Name=&quot;Book Title&quot;/&gt;
&lt;/w:LsdException Locked=&quot;false&quot; Priority=&quot;37&quot; Name=&quot;Bibliography&quot;/&gt;
&lt;/w:LsdException Locked=&quot;false&quot; Priority=&quot;39&quot; QFormat=&quot;true&quot; Name=&quot;TOC Heading&quot;/&gt;
&lt;/w:LatentStyles&gt;
&lt;/xml&gt;&lt;![endif]&lt;!--
/* Font Definitions */
@font-face
    {font-family:Calibri;
    panose-1:2 15 5 2 2 2 4 3 2 4;
    mso-font-charset:0;
    mso-generic-font-family:swiss;
    mso-font-pitch:variable;
    mso-font-signature:-536870145 1073786111 1 0 415 0;}
/* Style Definitions */
p.MsoNormal, li.MsoNormal, div.MsoNormal
    {mso-style-unhide:no;
    mso-style-qformat:yes;
    mso-style-parent:&quot;&quot;;
    margin:0in;
    margin-bottom:.0001pt;
    mso-pagination:widow-orphan;
    font-size:11.0pt;
    font-family:&quot;Calibri&quot;,&quot;sans-serif&quot;;
    mso-ascii-font-family:Calibri;
    mso-ascii-theme-font:minor-latin;
    mso-fareast-font-family:Calibri;
    mso-fareast-theme-font:minor-latin;
    mso-hansi-font-family:Calibri;
    mso-hansi-theme-font:minor-latin;
    mso-bidi-font-family:&quot;Times New Roman&quot;;
    mso-bidi-theme-font:minor-bidi;}
a:link, span.MsoHyperlink
    {mso-style-noshow:yes;
    mso-style-priority:99;
    color:blue;
    mso-themecolor:hyperlink;
    text-decoration:underline;
    text-underline:single;}
a:visited, span.MsoHyperlinkFollowed
    {mso-style-noshow:yes;
    mso-style-priority:99;
    color:purple;
    mso-themecolor:followedhyperlink;
    text-decoration:underline;
    text-underline:single;}
p.MsoPlainText, li.MsoPlainText, div.MsoPlainText
    {mso-style-priority:99;
    mso-style-link:&quot;Plain Text Char&quot;;
    margin:0in;
    margin-bottom:.0001pt;
    mso-pagination:widow-orphan;

```

font-size:11.0pt;  
mso-bidi-font-size:10.5pt;  
font-family:"Calibri";"sans-serif";  
mso-fareast-font-family:Calibri;  
mso-fareast-theme-font:minor-latin;  
mso-bidi-font-family:"Times New Roman";  
mso-bidi-theme-font:minor-bidi;}

span.PlainTextChar

{mso-style-name:"Plain Text Char";  
mso-style-priority:99;  
mso-style-unhide:no;  
mso-style-locked:yes;  
mso-style-link:"Plain Text";  
mso-bidi-font-size:10.5pt;  
font-family:"Calibri";"sans-serif";  
mso-ascii-font-family:Calibri;  
mso-hansi-font-family:Calibri;}

span.Spelle

{mso-style-name:"";  
mso-spl-e:yes;}

.MsoChpDefault

{mso-style-type:export-only;  
mso-default-props:yes;  
font-family:"Calibri";"sans-serif";  
mso-ascii-font-family:Calibri;  
mso-ascii-theme-font:minor-latin;  
mso-fareast-font-family:Calibri;  
mso-fareast-theme-font:minor-latin;  
mso-hansi-font-family:Calibri;  
mso-hansi-theme-font:minor-latin;  
mso-bidi-font-family:"Times New Roman";  
mso-bidi-theme-font:minor-bidi;}

@page WordSection1

{size:8.5in 11.0in;  
margin:1.0in 1.0in 1.0in 1.0in;  
mso-header-margin:.5in;  
mso-footer-margin:.5in;  
mso-paper-source:0;}

div.WordSection1

{page:WordSection1;}

-->[if gte mso 10]><style> /\* Style Definitions \*/

table.MsoNormalTable

{mso-style-name:"Table Normal";  
mso-tstyle-rowband-size:0;  
mso-tstyle-colband-size:0;  
mso-style-noshow:yes;  
mso-style-priority:99;  
mso-style-parent:"";  
mso-padding-alt:0in 5.4pt 0in 5.4pt;  
mso-para-margin:0in;  
mso-para-margin-bottom:.0001pt;  
mso-pagination:widow-orphan;  
font-size:11.0pt;

```
font-family:&quot;Calibri&quot;,&quot;sans-serif&quot;;
mso-ascii-font-family:Calibri;
mso-ascii-theme-font:minor-latin;
mso-hansi-font-family:Calibri;
mso-hansi-theme-font:minor-latin;}
</style>&lt;![endif][if gte mso 9]&gt;&lt;xml>
<o:shapedefaults v:ext=&quot;edit&quot; spidmax=&quot;1026&quot; /&gt;
</xml>&lt;![endif][if gte mso 9]&gt;&lt;xml>
<o:shapelayout v:ext=&quot;edit&quot;&gt;
<o:imap v:ext=&quot;edit&quot; data=&quot;1&quot; /&gt;
</o:shapelayout>&lt;/xml>&lt;![endif]
```

Simple, unadorned test email generated by Outlook 2010. It is in HTML format, but no special formatting has been chosen. I'm going to save this as a draft and then manually drop it into the Inbox for scraping by Redmine 3.0.2.

```
}}
```

Thanks for your work on this!

#### #11 - 2015-05-19 16:43 - Deoren Moor

If I can provide any further information please let me know. Thanks.

#### #12 - 2015-05-19 22:51 - Geoff Maciolek

- File Redmine-Mail-HTML-Mess-via-OutlookWebAccess2010.eml added

For another datapoint, I've uploaded another .eml file (and included the original content below), pulled via IMAP right out of the mailbox; I've sanitized the company info, but it's otherwise unmodified. I believe you'll find this easier to work with than Deoren Moor's message, as that seems to have some outlook-specific content. This one is plane jane ASCII. (With HTML embedded, of course).

I had attempted to include some friendly expanding code blocks to show what arrived, but when I did so I got blocked as a spammer, so... I'm trying a bit less:

Note that the below "What I get in Redmine as source" is also what I get if I "edit" the comment.

```
{{collapse(What I get in Redmine as source [expand])
```

```
&lt;!--
body
  {font-family:Verdana,sans-serif;
  font-size:0.8em;
  color:#484848}
h1, h2, h3
  {font-family:&quot;Trebuchet MS&quot;,Verdana,sans-serif;
  margin:0px}
h1
  {font-size:1.2em}
h2, h3
```

```
{font-size:1.1em}
a, a:link, a:visited
  {color:#2A5685}
a:hover, a:active
  {color:#c61a1a}
fieldset.attachments
  {border-width:1px 0 0 0}
hr
  {width:100%;
  height:1px;
  background:#ccc;
  border:0}
span.footer
  {font-size:0.8em;
  font-style:italic}
--&gt;
P {margin-top:0;margin-bottom:0;}
```

A mess.

--Geoff Maciolek  
MYCOMPANYNAME, LLC

```
}}
```

```
{{collapse(What I get in Redmine visibly [expand])
```

```
<!--
```

```
body {font-family:Verdana,sans-serif;
font-size:0.8em;
color:#484848}
h1, h2, h3 {font-family:"Trebuchet MS",Verdana,sans-serif;
margin:0px}
h1 {font-size:1.2em}
h2, h3 {font-size:1.1em}
a, a:link, a:visited {color:#2A5685}
a:hover, a:active {color:#c61a1a}
fieldset.attachments {border-width:1px 0 0 0}
hr {width:100%;
height:1px;
background:#ccc;
border:0}
span.footer {font-size:0.8em;
font-style:italic}
-->
P {margin-top:0;margin-bottom:0;}
```

A mess.

--Geoff Maciolek  
MYCOMANYNAME, LLC

```
}}
```



#13 - 2015-05-19 22:59 - Geoff Maciolek

```
{{collapse(The apparently problematic HTML [expandable])
<html dir=3D"ltr">
<head>
<meta http-equiv=3D"Content-Type" content=3D"text/html; charset=3Diso-8859-
1">
<style>=0A=
<!--=0A=
body=0A=
  {font-family:Verdana,sans-serif;=0A=
  font-size:0.8em;=0A=
  color:#484848}=0A=
h1, h2, h3=0A=
  {font-family:"Trebuchet MS",Verdana,sans-serif;=0A=
  margin:0px}=0A=
h1=0A=
  {font-size:1.2em}=0A=
h2, h3=0A=
  {font-size:1.1em}=0A=
a, a:link, a:visited=0A=
  {color:#2A5685}=0A=
a:hover, a:active=0A=
  {color:#c61a1a}=0A=
fieldset.attachments=0A=
  {border-width:1px 0 0 0}=0A=
hr=0A=
  {width:100%;=0A=
  height:1px;=0A=
  background:#ccc;=0A=
  border:0}=0A=
span.footer=0A=
  {font-size:0.8em;=0A=
  font-style:italic}=0A=
-->=0A=
</style><style id=3D"owaParaStyle" type=3D"text/css">P {margin-top:0;margin=
-bottom:0;}</style>
</head>
<body ocsi=3D"0" fpstyle=3D"1">
<div style=3D"direction: ltr;font-family: Tahoma;color: #000000;font-size: =
10pt;">A mess.<br>
}}
```

As you can see, it looks like it's not parsing the HTML comment inside HEAD as a comment, but then something else weird happens that's leading to the P {...} section showing up as well. It looks to me like this would be resolved by having everything inside &lt;HEAD&gt;...&lt;/HEAD&gt; get ignored - maybe easier said than done?

#14 - 2015-06-11 20:16 - Deoren Moor

Is there any further information we can provide?

**#15 - 2015-06-15 23:21 - Jean-Philippe Lang**

- Status changed from Needs feedback to Closed
- Assignee set to Jean-Philippe Lang
- Target version changed from 3.0.4 to 3.1.0
- Resolution set to Fixed

Thanks for the examples, these 2 are now fixed as part of #16962 and added as tests.

**Files**

---

sample2.htm	129 KB	2015-05-08	Rupesh J
Test email.msg	51 KB	2015-05-11	Deoren Moor
Redmine-Mail-HTML-Mess-via-OutlookWebAccess2010.eml	4.57 KB	2015-05-19	Geoff Maciolek