

使用java将网页保存为mht格式(2)计算机二级考试 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/597/2021_2022__E4_BD_BF_E7_94_A8java_c97_597125.htm //设置网页正文

```
MimeTypeName = "text/html";
MimeBodyPart bp = new MimeBodyPart(). bp.setText(content, strEncoding).
bp.setHeader("Content-Type", "text/html.charset=" strEncoding).
bp.setHeader("Content-Location", strWeb.toString()).
mp.addBodyPart(bp). int urlCount = urlScriptList.size(). for (int i =
0. i lt. urlCount. i ) { bp = new MimeBodyPart(). ArrayList urlInfo =
(ArrayList) urlImageList.get(i). // String url =
urlInfo.get(0).toString(). String absoluteURL =
urlInfo.get(1).toString(). bp .addHeader("Content-Location",
javax.mail.internet.MimeUtility
.encodeWord(java.net.URLDecoder .decode(absoluteURL,
strEncoding))). DataSource source = new
AttachmentDataSource(absoluteURL, "image").
bp.setDataHandler(new DataHandler(source)).
mp.addBodyPart(bp). } msg.setContent(mp). // write the mime
multi part message to a file msg.writeTo(new
FileOutputStream(strFileName)). } /** *方法说明：mht转html *
输入参数：strMht mht文件路径. strHtml html文件路径 *返回类
型： */ public static void mht2html(String strMht, String strHtml) {
try { //TODO readEmfFile InputStream fis = new
FileInputStream(strMht). Session mailSession =
Session.getDefaultInstance(System.getProperties(), null).
MimeMessage msg = new MimeMessage(mailSession, fis). Object
```

```

content = msg.getContent(). if (content instanceof Multipart) {
MimeMultipart mp = (MimeMultipart)content. MimeBodyPart bp1
= (MimeBodyPart)mp.getBodyPart(0). String strEncoding =
getEncoding(bp1). String strText = getHtmlText(bp1, strEncoding).
if (strText == null) return. File parent = null. if (mp.getCount() lt.
mp.getCount(). i) { MimeBodyPart bp =
(MimeBodyPart)mp.getBodyPart(i). String strUrl =
getResourcesUrl(bp). if (strUrl == null) continue. DataHandler
dataHandler = bp.getDataHandler(). MimePartDataSource source =
(MimePartDataSource)dataHandler.getDataSource(). File resources
= new File(parent.getAbsolutePath() File.separator getName(strUrl,
i)). if (saveResourcesFile(resources, bp.getInputStream())) strText =
JHtmlClear.replace(strText, strUrl, resources.getAbsolutePath()). }
saveHtml(strText, strHtml). } } catch (Exception e) { // TODO
Auto-generated catch block e.printStackTrace(). } } /** *方法说明
: 得到资源文件的name *输入参数 : strName 资源文件链接,
ID 资源文件的序号 *返回类型 : 资源文件的本地临时文件名
*/ public static String getName(String strName, int ID) { char
separator = ' / '. System.out.println(strName).
System.out.println(separator). if( strName.lastIndexOf(separator)
>= 0) return
format(strName.substring(strName.lastIndexOf(separator) 1)).
return "temp" ID. } /** *方法说明 : 得到网页编码 *输入参数
: bp MimeBodyPart类型的网页内容 *返回类型
: MimeBodyPart里的网页内容的编码 */ private static String
getEncoding(MimeBodyPart bp) { if (bp != null) { try {

```

```
Enumeration list = bp.getAllHeaders(). while
(list.hasMoreElements()) { javax.mail.Header head =
(javax.mail.Header)list.nextElement(). if
(head.getName().compareTo("Content-Type") == 0) { String
strType = head.getValue(). int pos = strType.indexOf("charset="). if
(pos != -1) { String strEncoding = strType.substring(pos 8,
strType.length()). if
(strEncoding.toLowerCase().compareTo("gb2312") == 0) {
strEncoding = "gbk". } return strEncoding. } } } catch
(MessagingException e) { // TODO Auto-generated catch block
e.printStackTrace(). } } return null. } /** *方法说明：得到资源文
件url *输入参数：bp MimeBodyPart类型的网页内容 *返回类型
：资源文件url */ 100Test 下载频道开通，各类考试题目直接下
载。详细请访问 www.100test.com
```