

Java_J2EE中文问题终极解决之道 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/144/2021_2022_Java_J2EE_E4_c104_144645.htm

Java中文问题一直困扰着很多初学者，如果了解了Java系统的中文问题原理，我们就可以对中文问题能够采取根本的解决之道。最古老的解决方案是使用String的字节码转换，这种方案问题是不方便，我们需要破坏对象封装性，进行字节码转换。还有一种方式是对J2EE容器进行编码设置，如果J2EE应用系统脱离该容器，则会发生乱码，而且指定容器配置不符合J2EE应用和容器分离的原则。在Java内部运算中，涉及到的所有字符串都会被转化为UTF-8编码来进行运算。那么，在被Java转化之前，字符串是什么样的字符集？Java总是根据操作系统的默认编码字符集来决定字符串的初始编码，而且Java系统的输入和输出的都是采取操作系统的默认编码。因此，如果能统一Java系统的输入、输出和操作系统3者的编码字符集合，将能够使Java系统正确处理和显示汉字。这是处理Java系统汉字的一个原则，但是在实际项目中，能够正确抓住和控制住Java系统的输入和输出部分是比较难的。J2EE中，由于涉及到外部浏览器和数据库等，所以中文问题乱码显得非常突出。J2EE应用程序是运行在J2EE容器中。在这个系统中，输入途径有很多种：一种是通过页面表单打包成请求（request）发往服务器的；第二种是通过数据库读入；还有第3种输入比较复杂，JSP在第一次运行时总是被编译成Servlet，JSP中常常包含中文字符，那么编译使用javac时，Java将根据默认的操作系统编码作为初始编码。除非特别指定，如在Jbuilder/eclipse中可以指定默认的字符集。输出途

径也有几种：第一种是JSP页面的输出。由于JSP页面已经被编译成Servlet，那么在输出时，也将根据操作系统的默认编码来选择输出编码，除非指定输出编码方式；还有输出途径是数据库，将字符串输出到数据库。由此看来，一个J2EE系统的输入输出是非常复杂，而且是动态变化的，而Java是跨平台运行的，在实际编译和运行中，都可能涉及到不同的操作系统，如果任由Java自由根据操作系统来决定输入输出的编码字符集，这将不可控制地出现乱码。正是由于Java的跨平台特性，使得字符集问题必须由具体系统来统一解决，所以在一个Java应用系统中，解决中文乱码的根本办法是明确指定整个应用系统统一字符集。指定统一字符集时，到底是指定ISO8859_1、GBK还是UTF-8呢？（1）如统一指定为ISO8859_1，因为目前大多数软件都是西方人编制的，他们默认的字符集就是ISO8859_1，包括操作系统Linux和数据库MySQL等。这样，如果指定Java统一编码为ISO8859_1，那么就有下面3个环节必须把握：开发和编译代码时指定字符集为ISO8859_1。运行操作系统的默认编码必须是ISO8859_1，如Linux。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