



链滴

sax解析xml报错SAXParseException

作者: [714593351](#)

原文链接: <https://ld246.com/article/1401759442408>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

最近偶尔在外部订单导入的时候会报

java.lang.Exception: javax.xml.bind.UnmarshalException: unexpected element (uri: " address"). Expected elements are <{}inventoryTranferDtIDto>,<{}inventoryTrnsferDto>;这种错误, 后来发现是由于参数中包含有不可见的非法字符, 在重调的时候对参数做下理就好了:

```
private static void processUnicode(Object serviceDispatchDto) {
    Field[] fields = serviceDispatchDto.getClass().getDeclaredFields();
    for (Field field : fields) {
        try {
            field.setAccessible(true);
            Object value = field.get(serviceDispatchDto);
            if (value instanceof Collection) {
                Collection<?> col = (Collection<?>) value;
                for (Object item : col) {
                    processUnicode(item);
                }
            } else if (value instanceof String) {
                if (value != null) {
                    value = ((String)value).replaceAll( "[\\x00-\\x08\\x0b-\\x0c\\x0e-\\x1f] ", "[X] ");
                    field.set(serviceDispatchDto, value);
                }
            }
        } catch (Exception e) {
            log.error( "Error while processUnicode" , e);
        }
    }
}
```

这些错误的发生是由于一些不可见的特殊字符的存在, 而这些字符对于XML文件来说又是非法的所以XML解析器在解析时会发生异常, 官方定义了XML的无效字符分为三段:

0x00 – 0x08
0x0b – 0x0c
0x0e – 0x1f

解决方法是: 在解析之前先把字符串中的这些非法字符过滤掉:

string.replaceAll("[\\x00-\\x08\\x0b-\\x0c\\x0e-\\x1f] ", "")