

Printservice v2.3 - HTML nach PDF Fehlerbehebung

Diese Ergänzung beschreibt den Fehler "Markup im Dokument vor dem Root-Element muss ordnungsgemäß formatiert sein" und die neue Normalisierung im HTML->PDF-Druckpfad.

1. Ursache

Der PDF-Renderer openhtmltopdf arbeitet deutlich strenger als ein Browser. Er erwartet wohlgeformtes XHTML. HTML, das im Browser noch sichtbar ist, kann beim PDF-Rendering fehlschlagen, wenn Tags nicht geschlossen sind, ein UTF-8-BOM vor dem Dokument steht oder versehentlich JSON/PDF-Inhalt statt HTML gespeichert wurde.

2. Typische Auslöser

Auslöser	Beispiel	Lösung
HTML-Fragment ohne Root	<table>...</table>	v2.3 bettet Fragmente automatisch ein
Void Tags nicht XML-konform	 , 	v2.3 normalisiert zu XHTML
BOM/Steuerzeichen	■<html>	v2.3 entfernt BOM und NUL
Falsche Datei gespeichert	document.html enthält JSON	v2.3 meldet klaren Fehler
PDF statt HTML	document.html beginnt mit %PDF	templateType=PDF_TEMPLATE verwenden

3. Neuer Ablauf ab v2.3

```
HTML-Vorlage + JSON-Daten
-> Platzhalter und {{#each:items}} Tabellen rendern
-> BOM und NUL entfernen
-> HTML-Fragment bei Bedarf in <html><body>...</body></html> einbetten
-> jsoup normalisiert Browser-HTML zu XHTML
-> openhtmltopdf erzeugt PDF
-> PDF wird gedruckt oder heruntergeladen
```

4. Validierungsendpunkt

Vor dem Drucken kann eine HTML-Vorlage gegen den HTML->PDF-Pfad geprüft werden:

```
POST /api/html/validate

curl -X POST http://localhost:8080/api/html/validate ^
-H "Content-Type: application/json" ^
-H "X-Printservice-Token: YOUR_TOKEN" ^
-d @src/main/resources/examples/rest-validate-html-pdf-request.json
```

5. Beispiel für eine gültige HTML-Tabelle

```
<html>
<head>
  <meta charset="UTF-8" />
  <style>table { width:100%; border-collapse:collapse; } td,th { border:1px solid #444; }</style>
</head>
<body>
<table>
  <tr><th>Pos.</th><th>Artikel</th><th>Barcode</th></tr>
  {{#each:items}}
  <tr>
    <td>{{text:pos}}</td>
    <td>{{text:partNumber}}</td>
    <td></td>
  </tr>
  {{/each}}
</table>
</body>
</html>
```

6. Hinweise

Für neue Vorlagen weiterhin möglichst vollständiges HTML verwenden: , , , . Die automatische Normalisierung macht den Druckpfad robuster, ersetzt aber keine fachliche Prüfung des Layouts.

Wenn der Fehler weiterhin auftritt, die Antwort enthält nun `htmlPreview` und `xhtmlPreview`. Damit kann geprüft werden, welcher Inhalt tatsächlich an den PDF-Renderer übergeben wurde.