

Text format guesser unable to identify the correct input document format

This technical note provides a workaround for situations in which the text format guesser is unable to identify the input document format as text.

Platforms affected

All platforms

Problem

HP Output Server 3.5 identifies some special characters as literal instead of text.

Solution

The problem is due to inappropriate locale settings of the operating system. When a text document format is identified as literal, check the input document for the presence of special characters and set an appropriate locale.

For example, `en_US.iso88591` supports special characters, such as ñ, ó, é.

You can add the locale settings to the `setup_env.sh/csh` file. Alternatively, you can set them on UNIX and Windows as described in the following sections.

Changing locale settings on UNIX

1. Identify the current locale settings by entering the following command:

```
locale
```

The output on HP-UX 11.11 is as follows (this may differ on other platforms):

```
LANG=C
```

```
LC_CTYPE=C
```

```
LC_COLLATE=C
```

```
LC_MONETARY=C
```

```
LC_NUMERIC=C
```

Solution

```
LC_TIME=C  
LC_MESSAGES=C  
LC_ALL=C
```

2. Enter the following command to generate a list of supported locales:

```
locale -a
```

3. Modify locale settings, as required.
 - Check for the locale of the characters that are not supported in the current locale.
 - Set all the variables noted in step 1 to the new locale.
4. Restart HP Output Server for the changes to take effect.

NOTE: Changing the locale may cause changes in log messages and other output messages.

NOTE: Locale settings are applicable only for the session in which they are made. Within a separate session, you must make the same locale changes.

Changing locale settings on Windows

1. From the Control Panel, select **Regional and Language Options**.
2. On the **Languages** tab, click **Details**.

The **Text Services and Input Languages** dialog box is displayed.
3. Select your preferred language from the **Default input language** list. If the language is not listed, click **Add** to specify the language.
4. Click **OK**.