Content Type and Markup

Tapestry reads well-formed XML template files and renders its output as XML, with minor caveats:

- The <?xml?> XML declaration is omitted.
- Most elements render with an open and close tag, even if empty.
- Certain elements will be abbreviated to just the open tag, if empty:
    - br
    - hr
    - img
- <![CDATA[]> sections are **not** used

This is all to ensure that the markup stream, while (almost) well formed, is still properly understood by browsers expecting ordinary HTML. In fact, Tapestry may decide to render a purely XML document; it depends on the content type of the response.

When Tapestry renders a page, the output content type and charset is obtained from meta data on the page itself. Meta data is specified using the @Meta annotation.

Content Type

The response content type is obtained via meta-data key `tapestry.response-content-type`. This value defaults to "text/html", which triggers specialized XML rendering.

A page may declare its content type using the @ContentType class annotation. Content types other than "text/html" will render as well-formed XML documents, including the XML declaration, and more standard behavior for empty elements.

Character Set

The character set (aka character encoding) used when writing output and when parsing requests is normally "utf-8". UTF-8 is a version of Unicode where individual characters are encoded as one to four 8-bit bytes. Most western language characters (that is, typical ASCII characters) are encoded in a single byte. Accented characters or non-western characters (such as Japanese, Arabic, etc.) may be encoded as two or more bytes.

In Tapestry, all pages use the same encoding. The default is "UTF-8", but this can be changed using the tapestry.charset configuration symbol.