# template extraction

Many documents will have boilerplate headings and text that we aren’t particularly interested in.

## <<field:DocumentTitle>>

It is likely that pieces of metadata are likely to be contained in headings, like the one above. There may also be text, such as dates, mentioned in the body, like the following:

<<field:DocumentDate >>

### Tabular data

We may also find text of interest in table form, such as this:<<record:foo:begin>>

|  |  |
| --- | --- |
| Title | <<field:PersonTitle>> |
| Full Name: | <<field:PersonFullName>> |
| Address: | <<field:PersonAddress>> |
| DoB | <<field:PersonDoB>> |

**Description:**

<<field:Description>><<record:foo:end>><<record:bar:begin>>

|  |  |
| --- | --- |
| Title | <<field:PersonTitle2>> |
| Full Name: | <<field:PersonFullName2>> |
| Address: | <<field:PersonAddress2>> |
| DoB | <<field:PersonDoB2>> |

**Description:**

<<field:Description2>><<record:bar:end>>

Finally, there may be information in the footer, as at the bottom of this page.