

Process

The process of turning data from the MacDougall access database created by the KVRF staff into data loadable into luna is fraught.

- 1) Please record the export options, and procedure, so as to make the process reproduceable.
- 2) The character set encoding should be carefully treated. If iso-8859-1 output is created, mark that as the encoding in the output xml file. If utf-8 output is created, then mark that as the encoding in the xml file.
- 3) A valid xml file should be created. Always validate the file against the vra (BETA) schema.

here is an error message, given by one validator, when the default encoding does not match the actual encoding.

```
vra_20080918.xml:233: parser error : Input is not proper UTF-8, indicate encoding Bytes: 0xFC 0x64 0x69 0x6E
Jaenicke, Anselm; Mamallapuram und die Welt der Südindischen
Kun
```

here is an error message given when an invalid character (a windows mdash) is used in a date:

```
vra_20080918-iso.xml:65535: element date: Schemas validity error : Element '{http://www.vraweb.org/vracore4.htm}
date', attribute 'start': \[facet 'pattern'\] The value '1913Â27' is not accepted by the pattern '(-)*\[0-9\]
{1,12}(-\[0-9\]{2}(-\[0-9\]{2}))*'. vra_20080918-iso.xml:65535: element date: Schemas validity error : Element
'{http://www.vraweb.org/vracore4.htm}date', attribute 'start': '1913Â27' is not a valid value of the atomic
type '{http://www.vraweb.org/vracore4.htm}dateValueType'.
```