Fennica linkitettynä datana tilannekatsaus

Osma Suominen

Tietomalliryhmä 11.4.2017

BIBFRAME2-muunnin otettu käyttöön

Subject: First impressions of new MARC to BIBFRAME 2.0 converter
From: Osma Suominen <[log in to unmask]>
Reply-To: Bibliographic Framework Transition Initiative Forum <[log in to unmask]>
Date: Thu, 23 Mar 2017 15:11:19 +0200
Content-Type: text/plain
Parts/Attachments: text/plain (115 lines)

Hi!

I'm developing a conversion pipeline [1] for converting MARC records
from the Fennica bibliographic database into Linked Data (using mainly
Schema.org as the target data model). Until now, I've used the LOC
marc2bibframe [2] conversion tool as one important step in that
pipeline. The new MARC to BIBFRAME 2.0 converter marc2bibframe2 [3] was
released last week and I've now switched to that. What follows are some
impressions and experiences from the switchover that I hope people on
this list will be interested in.

1. Technical architecture

The converter is based on XSL stylesheets, a well-established XML
processing technology. This gives it a level of technical independence
as there are many implementations of XSL available. In practice, the
README file suggests using the xsltproc XSL processor, so I've done that.

The code is split into around 20 XSL files, each corresponding to a
conversion specification published as Excel sheets on the LOC website.
In addition, there are corresponding unit tests written using XSpec that
verify that the conversion produces the desired result.

I like the architecture a lot. XSL performs much better than XQuery,
which the old marc2bibframe used. Also the chain from specification
(Excel) to implementation (XSL) and unit test (XSpec) makes it very
clear what the requirements for the conversion are and how it is
implemented and verified. In the old converter, the code itself served
as the specification and there were no unit tests.

COMHIS-yhteistyö: vuosilukujen putsaaja