

Converting Millennium ILS bibliographic records into Dublin-Core XML format for DSpace

Alan NG
Hong Kong University, Hong Kong
lbalan@hku.hk

The University of Hong Kong Libraries have been working on its IR (Institutional Repository) database, namely "The HKU Scholars Hub", which is based on the DSpace system. Part of the DSpace records originate from our Millennium ILS (integrated library system), such as student thesis and other HKU publications. Since these publications already have detailed and high quality bibliographic MARC record inside the ILS, it will be a duplicate effort to re-entry these information into DSpace. In an attempt to save the effort and automate the process, a Perl program was developed to directly retrieve the MARC record and convert it into the Dublin-Core XML format required by the DSpace system.

Given a list of bibliographic record numbers, the program will retrieve the MARC record directly via the HTTP protocol from the Millennium Web-OPAC, instantly one by one. The retrieved MARC record is in plaintext format which is a suitable format for furthering processing. Different profiles for different DSpace community subset are made to control from which MARC fields the program choose extract. Having the content of the MARC fields, the program will generate a Dublin-Core XML file for each of the bibliographic record.

In this presentation, examples on the IR database will be shown. The details of the programming effort will also be covered, such as the computer environment for running the program, the connection between the program and the Millennium ILS ... etc. Having the appropriate Internet connection, a live demonstration can be arranged.