

This concise book is for library programmers who want to learn to use MARC4J to process bibliographic data. MARC4J is an open source software library for working with MARC, MARCXML and related bibliographic standards in Java. The library is designed to bridge the gap between MARC and XML. MARC4J has built-in support for reading and writing MARC and MARCXML data, thus providing a programming environment for crosswalking: take records in one bibliographic format as input and output these records in a different format.

Crosswalking explains the MARC4J API in detail. It covers how to use the API to write programs to read, update and write MARC and MARCXML data. You will also learn how MARC4J integrates with XML environments and how the software library can be used to process other bibliographic standards like MODS and Dublin Core. It also teaches you how to index and search MARC data using Lucene, an open source text search engine library written in Java, and how to develop a servlet-based SRU server with MARC4J and Lucene. Other topics include how to work with character encodings, how to perform character conversions and how to use XSLT to pre- or post-process MARC data.

This book provides useful information for both developers learning about MARC4J for the first time and developers returning for reference and more advanced material. The chapters provide many reusable examples, while appendixes provide a reference to the API and the command-line utilities.

Bas Peters is the project owner and core developer of MARC4J. He works for Capgemini in the Netherlands, where he specializes in Enterprise Content Management.



Crosswalking: Processing MARC in XML Environments with MARC4J

Crosswalking

Processing MARC in XML Environments with MARC4J

Bas Peters