

Белорусский национальный MARC-совместимый обменный формат библиографических записей BELMARC в информационном обеспечении сводного электронного каталога библиотек Беларуси

Belarusian National MARC-Compatible Exchange Format for Bibliographic Records (BELMARC) in Information Support of the Communal Electronic Catalogue of the Libraries in Belarus

Біларусьскі нацыянальны MARC-сумісны абмінны формат бібліографічных записів BELMARC в інформаційному забезпеченні зведеного електронного каталогу бібліотек Біларусі

Олейник Г.Н.

Национальная Библиотека Беларуси, Минск, Беларусь

Ильина С.В.

Белорусский университет культуры, Минск, Беларусь

Oleynik G.N.

National Library of Belarus, Minsk, Belarus

I'lina S.V.

Belarusian University of Culture, Minsk, Belarus

Олійник Г.Н.

Національна Біліотека Біларусі, Мінськ, Біларусь

Ільїна С.В.

Біларусьскі університет культури, Мінськ, Біларусь

В докладе обосновывается необходимость разработки MARC-совместимого обменного формата библиографических записей при проектировании сводного электронного каталога крупнейших библиотек Беларуси, очерчивается роль обменного формата в организации функционирования сводного каталога и в информационном обмене библиотек Беларуси с каталогизирующими центрами ближнего зарубежья, прежде всего с российскими. Излагаются принципы адаптации формата UNIMARC при разработке белорусского обменного формата. Сообщается об организации работ по проектированию формата и предварительных результатах проектных работ.

The paper explains the necessity of the working out of MARC-compatible exchange format for bibliographic records under development of the communal electronic catalogue of the largest libraries of Belarus. The role of the exchange format in the process of the organization and functioning of the communal electronic catalogue is described as well as its role in the information exchange between belarusian libraries and other cataloguing agencies abroad, especially in Russia. The principles of UNIMARC Format adaptation under development of Belarusian exchange format are discussed. The organization of the designing of the national exchange format and the preliminary results of this work are reported.

У доповіді обґрунтовано необхідність розробки MARC-сумісного обмінного формату бібліографічних записів при проектуванні зведеного електронного каталогу великих бібліотек Біларусі, обґрунтовується роль обмінного формату в організації функціонування зведеного каталогу та в інформаційному обміні бібліотек Біларусі з центрами каталогівання близького зарубіжжя, перш за все з російськими. Викладено принципи адаптації формату UNIMARC при розробці біларуського обмінного формату. Повідомлено про організацію робіт по проектуванню формату та попередніх результатах проектних робіт.

Государственная программа информатизации в Республике Беларусь предусматривает, среди прочих мероприятий, создание сводного электронного каталога крупнейших библиотек республики и телекоммуникационной инфраструктуры, обеспечивающей доступ к нему из учреждений науки и образования. Библиотеки в высшей степени заинтересованы в результативном исходе этого проекта, так как это значительно расширит доступность справочно-поискового аппарата к их фондам, увеличит использование фондов и, что также важно, повысит полезность библиотек обществу и их престиж.

Все крупнейшие библиотеки республики создают справочно-поисковый аппарат к своим фондам на машиночитаемых носителях в течение последних нескольких лет, используя для этого разные программные среды: ISIS, ISIS+ИРБИС, АСПИД+разработки фирмы ИНЕАК, МАРК, АС "Библиотека" и др. Естественно, что как сами форматы, так и требования полноты библиографических записей во всех библиотеках различны, поэтому было бы нецелесообразно проектировать организацию сводного каталога только конвертированием и слиянием подготавливаемых в локальных системах библиотек текущих массивов библиографических записей из-за низкого качества формируемых таким способом массивов, хотя от этой идеи отказались не сразу.

С другой стороны, стремление библиотек к участию в международном информационном обмене заставляет их учитывать требования международных стандартов к структуре и полноте библиографических записей.

Эти два фактора и определили решение пересмотреть существующую практику каталогизации в библиотеках-участницах сводного каталога, разработать требования к полноте и структуре библиографических записей с учетом международных стандартов и создать структуру входной записи сводного каталога на основе единых и согласованных всеми участниками требований. Эту задачу договорились решать путем адаптации международного стандарта UNIMARC, поддерживаемого ИФЛА, во многом учитывающего требования международного стандарта де-факто USMARC и обеспечивающего стыковку любых MARC-ориентированных форматов. Специалисты-библиографы отчетливо понимали, что приемлемая эффективность формирования сводного каталога в условиях применения структурно более сложных записей MARC-формата, чем те, что делались ими до сих пор, может быть достигнута только в условиях кооперированной каталогизации. Это дало бы возможность ввести разделение труда в машиночитаемой каталогизации, давно уже применяемое во многих других странах, а также позволило бы обеспечить более качественными машиночитаемыми записями сеть средних и мелких библиотек, как массовых, так и специальных.

Таким образом, назначение разрабатываемого в Белоруссии MARC-ориентированного (или MARC-совместимого) формата библиографических записей определяют следующие его функции:

Аккумуляция требований, основанных на международных стандартах, к структуре, полноте и качеству записей, подготавливаемых для сводного каталога библиотеками Республики Беларусь.

Создание условий для обеспечения сетей массовых и специальных библиотек республики более качественными и полными по составу записями, чем формируемые ими самостоятельно, причем с включением привычного для них поискового инструментария (например, индексов УДК и ББК), обеспечение в ближайшей перспективе получения карточек с основными и добавочными описаниями для поддержания страховых карточных каталогов в библиотеках, не отказывающихся от их ведения.

Поддержка процессов международного обмена библиографической информацией со странами ближнего и дальнего зарубежья, прежде всего с Россией, и представление на международном информационном рынке коллекций национальных белорусских документов в формате, "понимаемом" и принимаемом этим рынком.

Первая и вторая функции определяют проектируемый формат как национальный формат каталогизации для сводного каталога и как внутринациональный обменный формат. Эти две функции не запрещают иметь в его структуре элементы сверх тех, что предписаны международным стандартом. Третья функция обуславливает необходимость представления национального документа исключительно средствами международного стандарта. Роль проектируемого формата относительно этого направления его использования должна состоять в обеспечении возможности представления национального документа средствами UNIMARC без дополнительной ручной доработки записей, а только путем некоторых дополнительных алгоритмических преобразований входных записей сводного каталога.

С целью проектирования MARC-совместимого национального формата по инициативе Национальной библиотеки Беларуси и по заказу Фонда информатизации в начале 1997 г. при Национальной библиотеке была на два года открыта научная тема (руководитель – директор НББ Г.Н. Олейник) и создана рабочая группа из представителей семи библиотек г. Минска и преподавателей Белорусского университета культуры (руководитель рабочей группы – С.В. Ильина). К концу 1997 г. рабочая группа состояла из пяти функциональных подгрупп, занимавшихся структурами библиографических записей на книги, сериальные издания, аналитику, нотные документы и аудиоматериалы, а также изучением соотношения проектируемого формата с форматом USMARC. Во многих библиотеках-участницах, особенно на начальном этапе, имелись представители всех трех основных подгрупп: книжной, сериальной и аналитики. Поэтому представители одной и той же библиотеки могут разрабатывать одни и те же блоки полей формата, но с ориентацией на разные виды документов. Таким образом большинство членов рабочей группы участвует одновременно в работе двух функциональных подгрупп: подгруппы, занимающейся определенным видом документа и подгруппы, нацеленной на определенную структурную область формата. Такая организация работ позволяет вести проектирование более результативно, что особенно важно, если учесть, что все члены рабочей группы занимаются темой в свободное от основной работы время.

В основу разработки формата были положены следующие принципы:

Достаточная полнота формата, обеспечивающая наиболее полное представление национального документа на международном уровне.

Разграничение минимально допустимой полноты входной записи сводного каталога и минимально допустимой полноты библиографического описания каталогизируемой единицы. Запись, не отвечающая минимально допустимому уровню полноты, не может быть принята как входная запись сводного каталога. Минимально допустимый уровень полноты описания может быть достигнут не в одной записи, а в совокупности связанных записей. Минимально допустимая полнота входной записи не может быть ниже, чем это предусмотрено в UNIMARC.

Минимально допустимый уровень полноты описания фиксируется набором обязательных полей, который несколько шире, чем набор обязательных и рекомендуемых полей UNIMARC. Как в Российском коммуникативном формате и в формате USMARC, вводится понятие условно обязательных полей/подполей, которые должны заполняться при наличии данных в каталогизируемом документе.

Ориентация наполнения разрабатываемой версии национального обменного формата на действующие стандарты оформления элементов библиографического описания, которые не полностью соответствуют требованиям ISBD, но с условием отказа от применения в записи сокращений в областях описания, оговоренных ISBD.

Принятие за основу наполнения UNIMARC, представленного его второй редакцией (1994 г.) с учетом изменений, внесенных в эту редакцию в 1996 г.

Обеспечение взаимной конвертируемости национального обменного формата с форматом USMARC.

Обеспечение максимально возможной согласованности в решениях по структуре формата с Российским коммуникативным форматом (редакция 1997 г.)

Обогащение функциональных возможностей формата только за счет полей национального использования, а не за счет функциональной дозагрузки полей/подполей UNIMARC.

Составление отдельной записи на каждую физическую единицу – объект каталогизации даже при отсутствии характерного заглавия (том многотомника, выпуск журнала).

Применение многоуровневых описаний для отражения иерархически связанных объектов каталогизации, при этом преимущественное применение восходящих связей в многоуровневом описании.

Предпочтительное использование полей блока связей (4-) для представления как вертикальных (4b-), так и горизонтальных связей бибзаписей (прочие поля блока 4-).

Использование техники вставленных полей в качестве инструмента связывания записей, причем преимущественное применение ссылок через идентификаторы записей.

Ориентация на применение технологии авторитетных файлов для представления элементов описания, подлежащих нормативному контролю.

Формирование полей примечаний только в том случае, если их нельзя получить из полей блоков 4- или 5-.

Этапность проектирования формата, состоящая в постепенном расширении репертуара видов документов, охватываемых форматом. Работы были начаты с базовых для всех библиотек видов документов: книг и сериальных изданий, описываемых на сводном, монографическом и аналитическом уровнях. Затем были добавлены диссертации, авторефераты диссертаций, нотные документы и звукозаписи. В 1998 г. должна быть разработана структура описания компьютерных файлов.

Отработка моделей записей в национальном формате для типовых случаев составления библиографических описаний (книга в авторском оформлении, том многотомника с характерным заглавием или без характерного заглавия и т.д.).

В конце 1997 г. сдан технический проект формата, подготовлен экспериментальный массив примеров, в 1998 г. ведется рабочее проектирование. Понятно, что к концу этапа рабочего проектирования некоторые из продекларированных здесь принципов могут претерпеть определенные изменения.

На стадии технического проектирования национального MARC-совместимого формата были получены три его версии для перечисленных выше видов документов:

полный национальный обменный формат, который включает все необходимые поля для возможно более полного и многоаспектного описания национального документа на национальном уровне с учетом традиций описания и представления документов в национальной библиографии, практики каталогизации ведущих библиотек республики и новых возможностей машиночитаемого представления библиографических записей (BELMARC);

национальный формат для представления национального документа на международном уровне, включающий наиболее полный репертуар только полей и подполей UNIMARC из числа представленных в национальном полном обменном формате;

национальный минимальный обменный формат, включающий только обязательные поля и подполя полного национального формата, как предлагаемые UNIMARC, так и поля национального использования.

Полный национальный обменный формат включает как поля и подполя из числа, предлагаемых UNIMARC, так и поля национального использования. Для полей UNIMARC конкретизированы условия их использования во входных записях сводного каталога, что позволяет снять функциональную вариативность формата, присущую UNIMARC как чисто обменному формату. Эти поля/подполя образуют основу национального формата, составляя подавляющее большинство его структурных единиц.

Следует отметить, что не весь репертуар полей UNIMARC для охваченных данной версией формата видов документов был включен в состав полного национального обменного формата. Это обусловлено, в основном, двумя причинами. Во-первых, стремлением разработчиков некоторых функциональных подгрупп к определенному "упрощению" формата для "своего" вида документа, нежеланием излишне, по их мнению, детализировать семантику формата. Поэтому если в структуре UNIMARC имелось поле с обобщающим значением для группы полей, то иногда выбиралось именно это обобщающее поле (например, подгруппой сериальных изданий был сокращен набор полей для предшествующих и последующих описаний). Второй и очень весомой причиной было стремление исключить из структуры формата поля с дублирующими функциями. Например, поле 520 ("Прежнее заглавие сериального издания") не было включено из-за явного дублирования этим полем ряда полей связи блока 4 - с функцией "предшествующие описания".

Отбор полей/подполей UNIMARC в структуру национального формата сопровождался, в ряде случаев, уточнением их функций и условий использования. Ограничение функций полей/подполей проводилось только на основе вариантов, предлагаемых UNIMARC. Для функций, не предусмотренных UNIMARC, вводились новые поля для национального использования с оговоренными правилами их перевода в поля UNIMARC при международном обмене.

Поля национального использования дают возможность более полно, чем позволяет UNIMARC, описать на национальном уровне какой-то аспект документа или более удобно описать этот аспект во входном формате с последующим преобразованием соответствующих UNIMARC элементов данных при международном обмене.

Поля национального использования были введены в блоки 0-, 4-, 6-, 9- Идеология UNIMARC позволяет использовать для полей национального использования не только блок 9-, специально для этого предназначенный, но и другие блоки. Коды полей национального использования в других блоках должны при этом кодироваться с включением цифры "9". Рассматривается также возможность ввода индикатора со значением "9" в некоторые поля из схемы UNIMARC, но только для внутринационального использования.

В процессе технического проектирования были также накоплены материалы для последующей разработки национального формата ввода библиографических записей, который должен включать только поля и подполя, подлежащие ручному вводу (путем набора символов на клавиатуре или выбора соответствующего пункта меню), т.е. те поля/подполя, элементы данных для которых не могут быть получены программно ни на основе элементов данных, введенных вручную, ни иным способом.

Хотелось бы также остановиться на наиболее сложных проблемах, с которыми столкнулась рабочая группа в ходе проектных работ.

Как отмечали и российские разработчики [2, 4], при работе с UNIMARC существует проблема адаптации и толкования понятийного аппарата, представленного в Руководстве к формату. Трудность освоения некоторых новых для нас понятий связана как с более обобщенным представлением привычных для каталогизаторов-практиков реалий, так и с некоторой неопределенностью и многозначностью некоторых важнейших терминов в самом Руководстве в силу его принципиальной лояльности к любым правилам каталогизации. Например, в понятийном содержании термина "точка доступа" применительно к нашим правилам каталогизации основными являются признаки "поисковый элемент" и "возможность выступать в качестве заголовка основной или добавочной записи". В отношении традиционного каталога эти два признака присутствуют в содержании термина всегда. В электронном каталоге пометки относительно заголовков записей имеют смысл только для выходных форм, а поисковыми могут быть и элементы, совершенно не предназначенные быть заголовками записей, в библиографических пособиях. Поэтому может быть целесообразнее в наших условиях использовать термины "поисковый элемент" и "заголовок записи", а не термин "точка доступа". Большой конкретизации требует также понятие "общее унифицированное заглавие", что напрямую связано с уточнением условий применения этого типа заглавия.

Наибольшую сложность вызывает, конечно, дифференциация уровней описания, поскольку решено было применять многоуровневое описание. Кроме того существует максималистское намерение создавать записи на составную часть документа путем формальной отсылки к документу-хозяину при росписи тех документов-хозяев, которые имеются или должны быть в сводном каталоге. Это вынуждает составлять отдельную запись на каждую физическую единицу-объект каталогизации, включая документы без характерных заглавий. Вследствие такого подхода в некоторых сложных случаях возникают громоздкие иерархические цепочки связанных библиографических записей и появляется соблазн их упростить за счет применения подполей 200Sh и 200Si не только в тех случаях, где без этих подполей

мы не считаем возможным обойтись (например, в составных наименованиях и при обозначении части документа, выпущенной в виде физически отдельной единицы). Наиболее, пожалуй, сложным вопросом является формирование правил построения многоуровневой записи для сериальных изданий, а также произведений, помещенных в нескольких физически отдельных единицах, относящихся к разным томам и/или годовым комплектам сериального издания.

Особой заботой для нас является согласование принимаемых нами решений с решениями Российского коммуникативного формата, т.к. именно российская литература составляет большую часть объемов поступлений в фонды библиотек Белоруссии и мы надеемся на активное сотрудничество в будущем с российскими каталогизирующими центрами. Все логичные и приемлемые для нас решения российских разработчиков мы стараемся применить и в своей версии адаптации UNIMARC. Так мы практически целиком приняли решения россиян по блоку примечаний [1, 3]. Россияне ранее нас начали адаптацию UNIMARC, и проработанность решений по книгам и сериальным изданиям в российской версии формата, безусловно, выше, что в немалой степени помогает и нам. Однако проектирование двух форматов проходит почти параллельно, поэтому, как бывает во всякой новой работе, мы не согласны с частью решений российских разработчиков по адаптации полей/подполей UNIMARC, а они, вероятно, кое в чем не согласятся с нами. Разница в подходах к некоторым вопросам может быть объяснена и не вполне совпадающим назначением наших форматов. Кроме того, для принятия окончательных решений нужна практика использования нового обменного формата, которой пока у нас нет.

В целом работа по созданию белорусского национального MARC-совместимого обменного формата, помимо своей практической важности, позволила библиографам взглянуть на процесс каталогизации с несколько иной точки зрения, более общей, и это безусловно ценно для теории и практики каталогизации.

Список литературы :

Российский коммуникативный формат представления библиографических записей в машиночитаемой форме: (российская версия UNIMARC): (книги и сериальные издания) / Российская Библиотечная Ассоциация. – 1997.

Бахтурина Т.А. Терминологические проблемы разработки Российского коммуникативного формата // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы конф., г. Судак, Респ. Крым, Украина, июнь 1997 г. Т. 2. – С. 465 - 467.

Скворцов В.В. Альтернативы в формате UNIMARC. Решения для Российского коммуникативного формата // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы конф., г. Судак, Респ. Крым, Украина, июнь 1997 г. Т. 2. – С. 472 - 475.

Фролова Е.Е. Проблемы внедрения международного формата UNIMARC в национальную среду России // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы конф., г. Судак, Респ. Крым, Украина, июнь 1997 г. Т. 2. – С. 446 – 449.

Структурная и семантическая совместимость форматов BELMARC и USMARC

Structural and Semantic Compatibility of BELMARC and USMARC Formats Устименко

Структурна та семантична сумісність форматів BELMARC та USMARC

*Устименко А.И., Федорина А.И., Козлова Т.В.
Белорусский университет культуры, Минск, Беларусь*

*Ustimenko A.I., Fedorina A.I., Kozlova T.V.
Belarus University of Culture, Minsk, Belarus*

*Устименко А.И., Федорина А.И., Козлова Т.В.
Біларуський університет культури, Мінська, Біларусь*

В докладе анализируется возможность структурной и семантической совместимости национальных форматов BELMARC и USMARC с целью преобразования библиографических записей, представленных в одном формате в другой.

Анализ проводился на уровне полей и подполей форматов с ориентацией на монографические издания (книги). Проанализированы такие структурные элементы, как код, имя поля (подполя), индикаторы, повторяемость поля и статус обозначаемого элемента. Выявлена структурная неадекватность форматов, которая, практически, не отражается на семантической совместимости.