POSSIBLE DATA OBJECTS FOR A LIBRARY RFID SYSTEM

Introduction

Increasingly, new RFID library systems are making use of RFID tags that are compliant with ISO standards. Generally, this is ISO/IEC 18000-3 Mode 1. Sometimes the tag is referred to as being compliant with ISO/IEC 15693, which is a perfect subset of ISO/IEC 18000-3 Mode 1.

RFID library systems have often been thought of as closed systems. This is not really the case, because using ISO standards allows for a degree of interoperability of the components in the system. Furthermore, inter-library loans require a degree of interoperability of the data. Possibly of greater potential concern is the fact that if other applications use the same tag technology, there is a serious risk of system clash between library books being present in other systems, and tags from other systems being present in libraries. In contrast, the benefit that will be derived is that, as volumes increase for all applications, the cost of systems components could be reduced.

To address the question of systems clash, the ISO committee dealing with RFID for item management (which includes library books) has standardised a system whereby particular application areas can be assigned a code that will be identified at the lowest level of radio communication. This code is called the Application Family Identifier (AFI) and, currently, a request for a specific AFI for library use is being processed. Because other applications will have different AFIs, the risk of systems clash can be virtually eliminated.

This paper looks at some possible data elements that could be used in the library community. Some are based on real implementations in the United Kingdom, and systems known to be operating or proposed in the Netherlands and Denmark.

The proposals are based around the RFID Data Protocol as specified in ISO/IEC 15961 and 15962. An outline is discussed below to illustrate the flexibility of the schemes.

Excluded Components

This paper focuses completely on data associated with the item, as opposed to the user or patron. This is for two reasons. The first is concerned with standards because cards that are associated with people are within the scope of the ISO committee JTC1 SC17, whereas items are within the scope of JTC1 SC31. The second point is that, where libraries are part of a local government service, the membership card can be used (and sometimes can have a primary use) for other services provided by the local government authority.

The second excluded component is any security mechanism that deals with anti-theft. Such mechanisms can be provided as added-value components within an RFID tag that makes no use of the data encoding area of the tag. Using the data encoding area is fundamentally weak. Also, there are security systems that can be independent of the RFID technology, especially where these had been implemented in library systems that use bar code.

A third component is not so much an exclusion as a definition. Some libraries have used the AFI to distinguish between on-loan and in-stock. Procedures are in place in ISO/IEC 15961 for a specific AFI to be used for on-loan books to avoid systems clash if, for example, the book is taken into a retail store that uses the same RFID tag technology. A "closed system" AFI is available for libraries and others dealing with re-circulating items within their closed operation. This particular AFI may be used by libraries for in-stock items.

The Data Protocol

For the purposes of this paper, only an outline of the Data Protocol is necessary. An expanded version is available from Praxis Consultants.

ISO/IEC 15961 defines the application interface of the Data Protocol. This deals with defining commands and data content moved to and from the application. The commands support selective read/write to cover accessing, updating and amending of data. Data is identified using an object identifier structure, which makes it unique not only within the particular domain of the application, but also among all other domains.

ISO/IEC 15962 defines the data encoding rules and logical memory functions of the Data Protocol. This covers a standardised way to compact data for more efficient encoding within the RFID tag memory, and organising the data in the memory (hence logical memory functions), so that different data objects can be encoded, some of which can be permanently locked.

Use of the Data Protocol provides a standardised way for encoding data, but allows users the complete flexibility of which data to encoded. The advantage of this is that standardised equipment and RFID systems components, including software, can be used but users can choose which data objects are most relevant for their application to encode in the RFID tag memory.

The current memory capacity for commercially available 18000-6C tags is typically 128 bytes, or 256 bytes. As other applications make use of the same tag technology, memories of different capacities could become available. The Data Protocol has been designed to address this, so the tags with different memory capacities – either purchased at different times or for different purposes – can be accommodated in the same system.

The Object Identifier

In the RFID Data Protocol, a typical object identifier (as specified in ISO/IEC 9834-1) has a common root that refers to a particular standard or specification. For example, the following object identifier refers to a traceability code that can be used to refer to cartons and totes to move books in the supply chain:

1 0 15459 1 1

An object identifier structure for library use will need to be established as part of the process of being assigned an AFI. This matter is beyond the scope of this paper.

A lot of memory could be consumed in encoding the object identifier for each different data element. The Data Protocol standards have a mechanism, known as the **Data Format** that enables the common root of the object identifier to be encoded once and, typically, only encoding the final node of each different data object. The data elements and data encoding discussed later in this paper only define, as examples, this final node which is known as the **Relative-OID.**

Data Elements

Each of the data elements discussed below needs to take account of some of the features of the Data Protocol, Library Management System (LMS), and data elements of message format (such as for inter-library loans) that are common in the sector. These are discussed as follows:

- Each data element needs to be defined by an object identifier. This paper simply identifies a set of Relative-OIDs that are used for illustration purposes only. A library sector standard needs to specify these more precisely. From the

perspective of the Data Protocol, Relative-OIDs valued 1 to 14 encode most efficiently.

- Each data element will be defined using one of the following three categories:
  - **Mandatory**, indicating that the object identifier must be present.
  - **Required**, indicating that this data element is mandatory under certain conditions, which will be described for each data element.
  - **optional**.

- The native character set external to the Data Protocol shall be based on ISO/IEC 8859-1, enabling support for all Latin alphabets. Particular data elements may define a subset of this, e.g. numeric digits. The encoding rules of ISO/IEC 15962 apply data compaction to achieve the most efficient encoding of a character string presented as a data object. To allow support beyond the Latin alphabet, UTF-8 transformations of ISO/IEC 10646 can be supported, with this being declared within the command structure of ISO/IEC 15961.

- The format of the data, particularly whether this is fixed or variable length. The Data Protocol itself is highly flexible, and can handle fixed length or variable length fields equally well. In a few of the data elements below, an assumption might be made on the size of the variable length data element to enable its encoding to be illustrated. Also associated with the format, is whether the structure is a subset of the full character set.

- Some data elements need to be locked to render them permanently encoded on the RF tag. This will be indicated where it is recommended.

Primary Item Identifier

This is the unique item identifier used within a particular library system. Generally, the code value is identical to the code used in any bar code system; but other code structures could be adopted at the choice of the library.

Relative-OID = 1
Mandatory
Format: variable length, supporting the full US ASCII character set
Should be locked

Coding assumption up to 16 characters

Owner Library Identifier

This data element is relevant for inter-library loans. The library identifier should be based on the ISIL as defined in ISO 15511.

Relative-OID = 2
Required for inter-library loans; otherwise optional
Format: the ISIL is a 12-character string, with particular punctuation required or allowed.

The first two characters of the ISIL are the 2-character country code defined in ISO 3166-1. The next character is the "minus/hyphen" sign, up to 9 characters follow for library identification with rules defined for each country. Some examples of ISIL codes include:

IT-RM0267

US-InU-Mu
CA-QMCBE
AU-TS:RL
DE-Tue120

This data element should be locked if the item is subject to inter-library loan.

Multi-part Indicator

This data element is required for mixed media components where a loan item can consist of more than one item that can be separated. Where only a single item is part of the loan transaction, or every item of a media set carries an RFID tag, this data element is optional. The data element enables the reading of the RFID tag to identify that other components of a media pack need to be checked, sometimes, manually.

This scheme may also be applied where items in the set might carry bar codes but do not have an RFID tag for various reasons.

Relative-OID = 3
Required, where a mixed media set does not have an RFID tag on each component; otherwise Optional
Format: fixed length of 1 byte (see encoding rule below)
Should be locked

The encoding is in the form of nn mm, with nn and mm having values 00 to 15. The following rules apply:

00 mm	A set consists of mm items with only one carrying an RFID tag
nn 00	A set has more than one item, the total number not being specified. This is item number nn
nn mm	This is item nn of a set that consists of mm items

Media Format

This data element can be used to identify the media format for bibliographic or handling purposes, including automated handling of returns. A code list of up to 256 codes (although this number is unlikely to be required) can be defined, to enable this value to be encoded as a single byte. The list covers some of the items identified from ANSI/NISO Z39.83 Part 2, and from the Danish RFID Data Model for Libraries.

The following is a list of possible code values:

0	undefined
1	book
2	audio tape
3	music tape
4	music CD
5	DVD
6	video tape
7	CD-ROM
8	Bound journal
9	Magazine
10	Book with audio tape
11	Book with CD
12	Book with diskette
13	Book with CD
14	Diskette

15 Microform
::::
98 Item requiring special handling
99 Small item requiring careful handling

100)
255) Internal codes for libraries

Relative-OID = 4
Optional data element
Format
Fixed length: single byte encoding integer value
Locking optional

Type of Usage

A Type of Usage code could be of benefit, particularly where there are different degrees of freedom on how long an item can be borrowed. ANSI/NISO Z39.83 Part 2 defines a data element **Item Use Restriction Type**, which could form the basis for some of the code points. These include:

In library use only
Limited circulation, long term period
Limited circulation, normal loan period
Limited circulation, short loan period
No reproduction
Not for loan
Overnight only
Renewal not permitted
Supervision required
Term loan
Use only in controlled access
User signature required

One of the issues to be decided is whether such a list should be standardised or left to each library to define by itself.

Relative-OID = 5
Optional data element
Format: fixed length of 1 byte encoding integer code value
Locking optional

EAN-UCC Product Identifier

The ISBN-13 is being introduced this year. This is based on the EAN-13 code with a prefix 978 and 979. The EAN-13 code structure (and its associated North American UCC-12 structure) is also used on other media products. This code can provide direct access to various bibliographic and catalogue information. Other code structures are supported (see data element Other Produce Identifier Codes).

Relative-OID = 6
Optional
Format: fixed length of 13 numeric digits
Locking optional

Current 10-character ISBN codes should be converted to the ISBN-13 format, because there is a one-to-one mapping between these codes. If software to carry out this conversion is not available then, in the last resort, data element Other Produce Identifier Codes may be used.

Other Product Identifier Codes

Various other types of item could be identified including serial publications, music, government publications and so on. Some of these have well-defined code structures and others do not. Where a code structure exists, we propose that the coding be done in two parts:

- 2-digits to identify a particular identification code scheme (the two digit code needs to be specified in an appropriate library standard).
- The specific code structure of the coding scheme.

This allows the code structure to self-declare, and retain the code structure of the legacy system.

Relative-OID = 7
Optional
Format: variable length, generally US ASCII character set, but UTF-8 should be supported for non-Latin languages
Locking optional

Other Candidate Data Elements

Nothing has been included in the list of data elements above to deal specifically with inter-library loans. A number of data elements could be considered, with the prime candidate being the **ISIL of the receiving library**. Other data could include any agreed or contracted **return date** to the owner library.

To enable internal inventory systems to operate even with the LMS temporarily out of action, additional codes could be considered for the RFID tag to enable it to be a direct means, for example, of relocating an item to its **storage position**.

The **book title** is already being encoded in some applications.

The significant advantage of using object identifiers is that the dictionary of data elements can be expanded as and when required, without changing the basic system.

The Danish proposal has supply chain data. If this is considered relevant, the data elements could be assigned Relative-OID codes. The data could be encoded on initial supply and over-written for library loan purposes.

Given that the majority of data elements are defined as optional, each library is able to use standard RFID tags and, within the constraints of the memory capacity, choose which data elements are most appropriate for the library operation. This does not even have to be consistent for all items, and some classes of item could have different encoded data elements to others.

Encoding Examples

The encoding and formatting rules of ISO/IEC 15962 are sufficiently sophisticated to encode any combination of Object Identifier, each having a fixed length or variable length data object made up of different characters. Instructions can be applied, using command arguments, to selectively lock data. Different RFID tag memory formats are supported, both in terms of

total memory and block size (a block being the smallest unit of memory that can be written to, read from, or locked).

Although it is not possible to show a complete encoded example, given that a number of features have been assumed. The following encoding guidelines are representative of the rules:

- If it is assumed that the library community has a single standard, then one object identifier structure can be applied to all data elements.

- Assume that the root-OID is recognised in the data protocol as one justifying a data format. This means that no user memory is used for encoding this.
  NOTE: The worst case is that approximately 6 bytes would be required to encode the root-OID.

- Assume that all the Relative-OIDs are in the range 1-14. The metadata or "overhead" for each encoded data element then requires 2 bytes.
  NOTE: It would be wise to ensure that little used data elements have a higher value Relative-OID to ensure that popular data elements encode as efficiently as possible.

- Each data object is encoded based on its character string, taking into account the range and variety and number of characters. This process is called **compaction** and is an integral part of ISO/IEC 15962. The compaction process achieves efficient use of the RFID tag memory and reduces the transmission across the air interface. The data protocol decodes (de-compacts) the data object to present it as a normal character string to display or interface with a computer system.

- Individual data sets (Object Identifier, data object, metadata and syntax) may be locked. This is not covered in the examples below. Locking requires the data set to be aligned on block boundaries to enable adjacent unlocked data sets to be changed. ISO/IEC 15962 provides all the processes, based on a user decision, to lock or not lock. Typically, locking for an ISO/IEC 18000-3 Mode 1 requires 2-3 additional bytes if all the locked data sets are adjacent to one another.

The following examples of the data elements, previously discussed, show the number of bytes required for the compacted data object. The overhead of 2 bytes has not been included.

- **Primary Item Identifier**
  Assume 16 characters
  If all-numeric: 7 bytes
  If alphanumeric: 12 bytes

  Shorter codes could result in a shorter encoding.

- **Owner Library Identifier (based on ISIL)**
  The encoding of the examples previously shown is as follows:
  IT-RM0267 7 bytes
  US-InU-Mu 8 bytes
  CA-QMCBE 6 bytes
  AU-TS:RL 6 bytes
  DE-Tue120 8 bytes

- **Multi-part Indicator**
  Irrespective of value: 1 byte

- **Media Format**
  Irrespective of value: 1 byte

- **Type of Usage**
  Irrespective of value: 1 byte

- **EAN-UCC Product Identifier (e.g. ISBN-13)**
  6 or 7 bytes, depending on whether the first digit is a 0. The ISBN-13 always encodes in 6 bytes.

- **Other Product Identifier Code**
  The ISRC is 12 characters long and, like the ISBN, uses hyphens to separate the constituent parts for display purposes only. An example is:

  FR-Z03-98-00212

  Leaving out the hyphens, and adding a 2-digit prefix to indicate that what follows is an ISRC, the encoding of the ISRC would always require 11 bytes.

- **Title**
  Using, as the example **Does Anything Eat Wasps?**, this 24 character string would require 21 bytes encoded upper and lower case. Changing to all upper case DOES ANYTHING EAT WASPS? reduces the encoding to 18 bytes.

using the longest examples for all the above data, 78 bytes would be required for encoding to the rules of ISO/IEC 15962 in an RFID tag. This is within the capacity of many currently available RFID tags used by libraries.

Although the encoding illustration results in 78 bytes of memory, the entire memory does not have to be read on each occasion. It is advisable to encode the Primary Item Identifier in the first primary positions. Using the *Read First Object* command of ISO/IEC 15961 it is possible to only read this across the air interface. Other commands allow for selected or complete reading when required.

Summary

This paper has considered a number of data elements that have been identified as potential candidates for encoding on RFID tags used in library systems. The encoding has been based on the ISO Data Protocol for RFID for Item Management, which controls the AFI that needs to be used on library books to ensure no systems clash with other applications using the same RFID technology.

The Data Protocol enables complete flexibility by encoding data objects based on their structure, and allows the user to select the data elements that are most relevant for the application.
The paper has described a number of candidate data elements and shown the implications of encoding these on an RFID tag. The candidate data elements that have been identified can all be encoded within the memory capacity of commercially available ISO/IEC 18000-3 Mode 1 tags.

The Data Protocol also supports selective locking of data, whilst still retaining the flexibility of the encoding. It does this by aligning data to block boundaries, which then allows the air interface command of **lock block** to be implemented.

The paper now needs to be considered by those in the library community, with a view of moving forward to developing an application standard. This application standard will define, more precisely, the data elements and their basic formatting rules. The advantage of the Data Protocol is that the standard can be revised by adding or removing data elements and their associated object identifier, without needing to change any of the RFID equipment or supporting RFID software.

Paul Chartier
Principal, Praxis Consultants
Project Editor ISO/IEC 15961 & 15962