

Consortium of European Research Libraries

UNIMARC and Cataloguing Notes

1

UNIMARC

This document is totally new. There was no corresponding document in the former UNAD series. The text has been read by the Chair, Vice-Chair and Secretary of the Permanent UNIMARC Committee and incorporates their suggestions. Amendment: 'not' inserted into 6 (last paragraph), 1999-02-03.

1 Intellectual Rights in the Format

The International Federation of Library Associations and Institutions (IFLA) owns the copyright in UNIMARC, and also in related formats already published (UNIMARC/Authorities) and in preparation (UNIMARC/Classification, UNIMARC/Holdings, etc.)

2 Maintenance of the Format

Maintenance of the format is the responsibility of the Permanent UNIMARC Committee (PUC), which comes under the umbrella of the IFLA Universal Bibliographic Control and International MARC (UBCIM) core programme. The UBCIM Secretariat also functions as the PUC's Secretariat.

Details of the constitution, membership, work and publications of the PUC and its contact addresses and numbers can be found at its World Wide Web site (*see* Section 5, below), and also from time to time in IFLA UBCIM's quarterly *International cataloguing and bibliographic control (ICBC)*.

3 Proposals for Changes to the Format

Any interested person or body may submit suggestions for additions, corrections, etc., for consideration by the PUC. The PUC's main annual meeting at which changes are agreed takes place each spring, and proposals should reach the PUC Secretariat by the end of the previous year, in order to give PUC members time to consider them before the meeting.

Members of CERL who have suggestions for the improvement of UNIMARC are requested to communicate them to the CERL Secretariat in the first place, so that they may be considered by the Advisory Task Group, other members and the consultants.

4 Updates and Corrections

Updates to the format are published from time to time (not necessarily annually), usually several months after the PUC's spring meeting at which the changes were approved. These take the form of loose-leaf pages with additions or amendments to be inserted into the *UNIMARC manual*, together with instructions for minor amendments which can be made by hand. (*See also* the note in Section 5, Electronic Versions, below). Some corrections may occasionally be notified in *ICBC*.

5 Electronic Versions

Increasingly, electronic versions of UBCIM, and especially PUC, documents are being made available through the World Wide Web. These include the *UNIMARC manual* (complete, apart from certain appendices), and concise versions of both the bibliographic and authorities formats.

It should be noted that amendments to the format are NOT included in the electronic versions until they have been issued in print form.

For full details of these and related electronic documents, go to the UBCIM home page:

<http://ifla.inist.fr/VI/3/ubcim.htm>

(ifla.inist.fr is the European mirror site of the official IFLA site at www.ifla.org, which is the new domain name replacing the former www.nlc-bnc.ca/ifla/)

6 Editions in Various Languages

The basic language of the UNIMARC format (and the working language of the PUC) is English.

As the use of the format has spread, translations into other languages are being published. These versions can be prepared and published only with the permission of the IFLA UBCIM programme.

All approved translations are authoritative. In the event of a problem in interpretation, the English-language version is regarded as the *de facto* master version, although no formal statement to this effect appears anywhere. Occasionally, a translation may incorporate a correction not yet published for the English-language version, with the permission of the IFLA UBCIM programme.

7 National and Local Use

No person or body other than the PUC has the authority to modify the format. However, the PUC has left certain fields ('9' fields), indicators and subfields undefined and free for use in any way. These, and CERL's use of them, are described in UNICAT / 2.

8 CERL and UNIMARC

The Consortium chose UNIMARC as its preferred exchange format in 1993. Records contributed to the Hand-Press Books database may be supplied in, or converted to, UNIMARC for transmission to the Research Libraries Group. Although RLG then converts those records to its version of USMARC for the HPB database, it is part of the CERL's agreement with RLG that *all* records shall be made available for export from the database in UNIMARC, including those contributed by members in USMARC.

9 Related Formats

The *UNIMARC manual* has the subtitle *bibliographic format*. 'UNIMARC' is usually taken to mean the bibliographic format unless another is specifically named.

UNIMARC/Authorities : universal format for authorities was published in 1991; the PUC is working on a second edition of this format.

Also under preparation are UNIMARC formats for holdings and classification data.

2

UNIMARC '9' Fields : use by CERL and member libraries

This document is totally new. There was no equivalent in the former UNAD series.

1 UNIMARC

The *UNIMARC Manual* (2nd ed., 1994; Update1, 1996) states in section 4:

4.9 National and Local Use

All fields with tags containing a 9, i.e. 9--, -9-, --9, are reserved for national and local use; their definitions and indicator and subfield values remain undefined by the Permanent UNIMARC Committee. This is also true of indicator value 9 and subfield \$9.

2 Use of nationally or locally defined fields by CERL members

Several members who use UNIMARC as their national or export format have defined some fields for their own use, as they are entitled to do.

When contributing files containing '9' fields for the HPB database, members are asked to

- (a) submit full descriptions of the structure and content of those fields; and
- (b) *either* state which fields are of more than local interest and should be retained for the database and which fields contain information of purely local, internal (usually administrative) interest and should not be included, *or* themselves delete the second group of fields from the records before transmission.

3 Fields defined by CERL

3.1 A number of fields were defined by CERL for use in the HPB database. These fell into three broad groups: (a) fields to supplement the provisions of the UNIMARC format for books of the hand-press era; (b) fields to provide for holdings and other administrative details; and (c) fields to provide for cross-references from alternative forms of personal, family and corporate names (there being no general authority file available to the Consortium and the Research Libraries Group).

3.2 The fields in groups (a) and (b) have now been superseded by new fields in extensions to the format promulgated by the Permanent UNIMARC Committee; those CERL fields should no longer be used.

3.3 Because most '9' fields defined by members have been given tags in the 9-- series, all CERL fields use --9 or -9- tags.

The full list of CERL '9' fields, both current and superseded, is as follows:

009	Record identifier in source file <i>Superseded by UNIMARC 035</i>
319	General cataloguer's note <i>Superseded by UNIMARC 830</i>
349	Copy-specific note : Preservation data – Reproductions in microform, etc. <i>Not used; withdrawn.</i>

519	Title in standard modern spelling <i>Superseded by</i> UNIMARC 518
690	Personal name used as a subject – Alternative form
691	Corporate body name used as a subject – Alternative form
692	Family name used as a subject – Alternative form
790	Personal name – Alternative form
791	Corporate body name – Alternative form
792	Family name – Alternative form
899	Location

3.4 Details of the CERL '9' fields still current are available in separate documentation.

4 Use of digit '9' in indicators or subfield identifiers

Members may use the value '9' for an indicator or subfield identifier in nationally or locally defined '9' fields.

However, they should not need to use '9' as an additional indicator or subfield identifier in a *standard* UNIMARC field. Members are requested to consult the CERL Secretariat if considering this course of action for any reason.

5 Use of digit '9' in coded data

It should be noted that Section 4.9 of the *UNIMARC Manual* does **not** include the use of the digit '9' in coded data.

Some fields in the 1-- block define values for '9', either alone or in combination, for example in dates, geographical coordinates, '9999' for a serial currently being published, etc.

It follows that '9' must not be used as an optional addition in coded data fields. All permissible codes are defined by the Permanent UNIMARC Committee.

Consortium of European Research Libraries
UNIMARC and Cataloguing Notes

3
(revision 1)

Countries and Country Codes

This document is a revised version of the former UNAD / 3 (rev. 1) dated 1995-07-17.

References are made to the most recent (1997/99) edition of ISO 3166 and changes to ISO 3166-1; notes on coding former jurisdictions have been revised; some extensions and corrections have been made to the lists of UNIMARC fields and codes for European countries.

Revision 1 : Codes for Slovakia (SK, not SV) and Ukraine (UA, not UK) corrected; new code for Serbia & Montenegro (CS).

1. UNIMARC calls for the use of country codes in several places (for example, field 102), and CERL has added to these requirements by specifying a country code as the first element in fields which contain the code for a member institution.

2. **ISO 3166.** The codes to be used are the two-letter codes from ISO 3166-1:1997, *Codes for the representation of the names of countries and their subdivisions – Part 1: Country codes*, known as the "alpha-2" codes to distinguish them from the three-letter and numeric codes. These codes are conventionally given in capital letters (for example, ES, FR, PT, etc.).

3. **ISO 3166 and MARC21.** The ISO codes are sometimes the same as the MARC country codes used in MARC21, UKMARC and some other derived formats (for example, BE = be, DK = dk, GR = gr, IT= it), but are often quite different (for example, Spain = ES, but *sp* in the MARC code list).

The two lists also differ in that the MARC codes include specific codes for the countries of the United Kingdom (England, Scotland, Wales and Northern Ireland), for provinces of Canada and states of the United States, while the ISO 3166-1 list does not do so. (The new ISO 3166-2 list *does* code lower-level jurisdictions. See the note to field 102, below). The MARC list has no codes for the lower-level jurisdictions of any other countries.

4. **Current and former jurisdictions.** Both lists give codes for present-day jurisdictions, and are revised from time to time as old groupings break up and new nations achieve independence. For example, the codes for the former Soviet Union are now obsolete, and new codes have been provided for Russia, Ukraine, Belarus, etc.

The implication of this for the cataloguing of materials of the hand-press period is that *there are no codes for those sovereign states which existed at the time of publication, but no longer exist today.* Conversely, there are codes for modern states (for example, Italy) which may not have existed at the time of publication, but do today.

Part 3 of the Standard (ISO 3166-3:1999) entitled *Code for formerly used names of countries* consists of codes for countries, dependencies, etc., removed from ISO 3166 since its first edition in 1974. A proposal to include codes for other jurisdictions of the nineteenth century and the first half of the twentieth was abandoned. ISO 3166-3 is of no use for CERL's purposes and should be ignored.

5. **Recommended practice.** When field 102, Country of Publication or Production, is included in a record, the rule in the *UNIMARC manual* is quite clear : ***use modern boundaries and jurisdictions to determine the code for the country of publication.*** For publications of the hand-press era, this will often result in anomalies due to changes in boundaries or the creation or dissolution of jurisdictions (for example, place of publication in field 210 \$a = Venezia and \$d = 1495, but country of publication in 102 \$a = IT, or 210 \$a = Breslau, but field 102 \$a = PL).

When field 102 is used, *ensure that help pages and other documentation draw attention to these anomalies and other pitfalls when its codes are used for searching.*

6. **UNIMARC 660.** Field 660 Geographic Area Code (GAC) contains the hierarchically-structured codes developed by the Library of Congress to denote the region covered by an item. The GAC codes are listed in Appendix D of the *UNIMARC manual*, which gives full guidance for their use. Field 660 is concerned with subject matter : it does *not* relate to place of publication of the item. ISO 3166 codes are not used in this field at all.

7. **Appendices.** There follow notes on the specific UNIMARC fields in which country and locality codes are employed (and also on field 620, which is not coded), and a list of the ISO 3166-1 codes for European countries.

UNIMARC and CERL fields and subfields requiring ISO 3166-1 alpha-2 codes

\$3 AUTHORITY RECORD NUMBERS

These subfields have the alpha-2 code as their first element, followed by a backslash (\). *See* UNICAT 9.

001 RECORD IDENTIFIER

[009 *RECORD IDENTIFIER IN SOURCE FILE (CERL field)*]

These fields have the alpha-2 code as their first element, followed by a backslash (\). *See* UNICAT 4.

[CERL 009 has been superseded by new UNIMARC field 035 published in Update 3 (2000-01-01)]

020 NATIONAL BIBLIOGRAPHY NUMBER

021 LEGAL DEPOSIT NUMBER

022 GOVERNMENT PUBLICATION NUMBER

All three fields have the alpha-2 code in subfield \$a.

035 OTHER SYSTEM CONTROL NUMBERS

\$a System Control Number

The alpha-2 code should be the first element in the subfield, followed by a backslash (\). *See* UNICAT 4.

This is a new field added to the format in Update 3 (2000-01-01); it supersedes CERL field 009.

102 COUNTRY OF PUBLICATION OR PRODUCTION

\$a **Country**

ISO 3166-1 alpha-2 code.

\$b **Locality**

The CERL Advisory Task Group recommended (March 1995) that this subfield should not be used, as having little value and lacking well-developed and internationally recognised codes. (This situation is changing : see ISO 3166-2:1998 – Country subdivision code).

620 PLACE ACCESS

\$a **Country**

\$b **State or Province, etc.**

\$c **County**

\$d **City**

This field contains a hierarchical, *non-coded* form of a place of publication, etc. The *UNIMARC manual* makes it clear that it is not necessary to include all the

subfields : subfield \$d by itself containing the name of a city is permissible, if it is unlikely to be confused with another place bearing the same name.

The *UNIMARC manual* gives no guidance concerning the language to be used. The language will usually be that of the cataloguing agency (for example, a British agency would enter **620 ##\$aItaly\$dVenice**, while a French one would enter **620 ##\$aItalie\$dVenise**).

The Advisory Task Group recommended (March 1995) that

- (a) this field should be used;
- (b) the names of countries and cities should preferably be given in the language of the country concerned, but other forms in the languages of the cataloguing agencies would also be acceptable; and
- (c) subfields \$b and \$c should *not* be used.

801 ORIGINATING SOURCE

899 LOCATION (*CERL Field*)

Both fields have the ISO 3166-1 alpha-2 code in subfield \$a. In 899 it is followed by a backslash (\).

ISO 3166-1 alpha-2 codes for European countries

AD Andorra	LI Liechtenstein
AL Albania	LT Lithuania
AT Austria	LU Luxembourg
AX Åland Islands	LV Latvia
BA Bosnia & Herzegovina	MC Monaco
BE Belgium	MD Moldova, Republic of
BG Bulgaria	MK Macedonia, the former Yugoslav republic of
BY Belarus	MT Malta
CH Switzerland	NL Netherlands
CS Serbia & Montenegro	NO Norway
CY Cyprus	PL Poland
CZ Czech Republic	PT Portugal
DE Germany	RO Romania
DK Denmark (<i>see also FO</i>)	RU Russian Federation
EE Estonia	SE Sweden
ES Spain	SI Slovenia
FI Finland (<i>see also AX</i>)	SM San Marino
FO Faroe Islands	SK Slovakia
FR France	TR Turkey
GB United Kingdom	UA Ukraine
GI Gibraltar	VC Vatican City
GR Greece	(YU Yugoslavia [= Serbia & Montenegro] <i>new code is CS</i>)
HR Croatia	
HU Hungary	
IE Ireland	
IS Iceland	
IT Italy	

Note that UK is a reserved code, to avoid confusion between United Kingdom and Ukraine. (The ICANN use of .uk in Internet addresses is irregular).

For other countries, consult the International Standard

4

Record Identifiers, and CERL Institution and File Codes

This document replaces the former UNAD / 4 dated 1995-02-17. Section 2 (Field 035) replaces the original section 2 (CERL field 009). Section 4 has a second example. Section 5 is new.

1. Field 001 : Record identifier

1.1 The *UNIMARC manual* states that the Record Identifier can take any form.

1.2 For records contributed to the CERL database the following 4-segment format is **mandatory**:

Country code\Institution code\File code\Record ID number

The segments are separated by backslashes, as shown, without any intervening spaces.

1.3 **Country codes.** These are the ISO 3166 alpha-2 codes (*see UNICAT / 3*).

1.4 **Institution codes.** Where there is a nationally-recognised set of codes used to identify libraries for interlibrary loan purposes, etc., these codes shall be used by CERL.

If there is no nationally-recognised set of codes, then the code shall be one proposed by the Member institution and agreed with the CERL Secretariat.

(It does not matter if there are identical institution codes *in other countries* -- for example, BN or KB).

1.5 **File codes.** These shall (preferably) consist of two characters, alphabetic or numeric or mixed.

The codes may be sequential (01, 02, 03 ... or A1, A2, A3 ... *etc.*) or mnemonic (for example, OC for "Old Catalogue", *etc.*). The choice is left to the Member institution.

Alphabetic characters are *not* case sensitive : AA, Aa, aA and aa will be regarded as identical.

1.6 **Record ID numbers.** These may take any form and be of any length.

1.7 EXAMPLES:

001 SE\S\01\820907g01
A record from Kungliga Biblioteket, Stockholm. "S" is the code nationally used for the library. The arbitrary file number 01 has been assigned.

001 DE\GyMuBSB\AK\8013379054
A record from Bayerische Staatsbibliothek, München. There are no library codes commonly used throughout Germany, so the style used by Die Deutsche Bibliothek has been accepted for the time being. The mnemonic code AK (=Alphabetischer Katalog) has been used for the file.

2. Field 035 : Other System Control Numbers This field, published in Update 2 to the *UNIMARC manual*, supersedes CERL field 009, Record Identifier in Source File. (It is believed that 009 has not been used so far).

If 035 is used, then the provisions of the *Manual* must be followed : \$a (non-repeatable) holds the System Control Number and \$z (repeatable) holds cancelled or invalid control numbers. The code or name of the institution is given first, in parentheses, followed by the control number.

If the code for the name of the institution is not internationally recognized, use the CERL pattern – Country code\Institution code – inside the parentheses:

```
001      XY\AbCd\01\234567
035 ##$a(ISO country code\Institution code)Original
        record's ID
```

If a distinct file identifier is required to make the record ID unique, it should precede the ID, outside the parentheses:

```
035 ##$a(Institution code)File code\Original record's ID
```

Note that, unlike field 001, field 035 *does* have indicators (two blanks) and subfield identifiers.

3. Record IDs in other fields

The CERL Record ID must appear *in full* not only in field 001, but also in all other fields in which it appears. In particular, this applies to every 4-- linking field which contains a Record ID.

This is essential in order to ensure that fields containing otherwise identical record numbers from various source files are not loaded into the CERL database without being differentiated by their Institution and File Codes.

(Some examples in the CERL *UNIMARC format and descriptive cataloguing* specification are oversimplified and do not show the full format -- for example, the multivolume formats illustrated in Section 10).

EXAMPLE:

```
461 #0$1001DE\GyMuBSB\AK\8013387049
```

4. Institution Codes in other fields

In those fields where the code for the Institution is required (but not the File code or Record ID number), the code for the Institution must still be preceded by the Country code.

EXAMPLES:

```
801 #0$aSE$bS ... etc.
      Originating source. The country and institution are in separate subfields.

899 ##$aSE\S$b ... etc.
      CERL field for locations.
```

Note that this rule also applies to all occurrences of subfield \$5, *Institution to which the field applies*.

5. Authority File Record IDs

These follow the same pattern. For further details consult UNICAT 9

5

Fill Characters, Blanks and Standard Characters

This document supersedes the former UNAD / 5 (revision 1) dated 1996-02-06. Minor changes have been made to the presentation throughout, but the information and rules remain unaltered.

1 The *UNIMARC manual* (Section 4.5) states the rules for the use of fill characters, and contrasts them with blanks and certain other characters often used as codes with standard meanings. The essential features of these rules are restated here, together with examples showing their use in various contexts.

2. Fill character

2.1 The fill character is a graphic character in ISO 646, Basic Latin Character Set (the default G0 set for UNIMARC records). It is the vertical line, (|), position 7/12 (7C) in the 7-bit code table.

[Note that this is an actual character, unlike the \$, @, % and # signs which are only conventional printed representations of the control functions IS1, IS2 and IS3 and the blank (2/0) from ISO 646 : *see* Section 2.2 in the *UNIMARC manual*].

2.2 The fill character is used

- (a) when it is not possible to assign the correct value to a coded data item (for example, when there is no corresponding value in the source format); and
- (b) when the agency never assigns values to a particular category of data in a coded field or to the indicators in particular types of fields.

2.3 Fill characters may *not* be used in the Record Label, Directory, or tags, or to indicate missing or uncertain characters, etc., in textual data (variable data fields).

3. Blank

The blank (ISO 646 position 2/0) is conventionally represented by a hash sign (#) in the *UNIMARC manual*, 2nd ed., 1994 (a crossed lower-case letter ' b ' was used formerly).

The blank is mainly used

- (a) for indicators when no values have been defined in the format (*see* 8.2 below); and
- (b) as "padding" to fill data elements in fixed-length coded subfields in specified circumstances (*see* 10.2 below).

4. Standard codes

Five other characters have been used as *general codes with standard meanings* in many (but not all) coded subfields. These are:

- u Unknown.** The cataloguer would have assigned a definite code if it had been possible to discover the necessary information.
- v Combination.** A combination of several individually coded characteristics appears in the item [but only one coded position is available].

For example, in field 140, Coded data for antiquarian materials, \$a character position 8 = Illustration techniques. An item which uses two or more techniques is coded **v**.

- x Not applicable.** The characteristic being coded does not apply to the *type* of material being catalogued.

For example, in field 126, Coded data for sound recordings, \$a character position 3 = Groove width. If the recording is a CD, which has no grooves, the appropriate code is **x**.

- y Not present.** The characteristic being coded is not present in the item.

[Note: this must be distinguished from the use of the fill character, which would mean that the cataloguing agency *never* codes the characteristic in question, even if it is present in the item].

- z Other.** This code is used when the characteristics of the item are known but none of the other available codes is appropriate. **z** is also occasionally used for "Unknown".

- 5 **Record Label.** This must contain the specified numeric values, alphabetic codes or blanks. The fill character must never be used in the Record Label.

- 6 **Directory.** This contains only numeric values (computed automatically), followed by the field separator IS2 (ISO 646 position 1/14 (1E), conventionally represented by @).

- 7 **Tags.** Tags must be used as specified. No digit in a tag can be replaced by a blank or fill character.

8 Indicators.

- 8.1 Blanks are specific indicator values. A blank normally means that no definition has been given to the indicator.

Exceptions occur in fields 321 (External Indexes/Abstracts/References Note) and 606 (Topical Name used as Subject), where several values are defined for Indicator 1, and # = *No information given* and *No information available*, respectively.

- 8.2 Fill characters may be used in place of indicators when EITHER the institution never assigns values which would be represented by indicators OR the source format does not hold values which can be converted to indicators.

- 9 **Subfield Identifiers.** Blanks or fill characters can NEVER be used to replace subfield identifiers, which consist of IS1 (ISO 646 position 1/15 (1F), conventionally represented by \$), followed by one of the designated alphabetic or numeric characters.

10 Fixed-length coded data subfields. (*These are discussed in more detail in UNICAT / 6*)

- 10.1 Fill characters may be used as stated above (paragraph 2.2):

- (a) when the agency never supplies value for particular data elements; or
- (b) when the source format has no values which can be converted to UNIMARC.

- 10.2 Blanks are used to fill the remaining character positions when more than one position is allotted to a particular data element, but not all are needed.

11 '9' fields and subfields.

The digit '9' in tags, indicators and subfield identifiers is reserved for national and local use throughout the format. (*See Section 4.9 in the UNIMARC Manual*).

Use of '9' fields, etc., by CERL and its members is discussed in UNICAT / 2.

6 (revision 1)

Fixed-length Coded Data Subfields

This document was a reprint of the former UNAD / 6 (rev. 1) dated 1995-11-02, with minor changes which did not affect the instructions.

Revision 1 : Section 9 (Mandatory coded data fields) has been completely revised in line with a ruling by the Permanent UNIMARC Committee (March 1999).

1 The majority of these subfields occur in the 1-- CODED INFORMATION BLOCK. Some of the fields consist of a single subfield, but several contain two or more subfields.

There are also some fixed length subfields which form part of fields which otherwise contain variable-length data. Examples are subfield \$z *Language of Parallel Title Proper* in field 200 and the 51- fields, the 3-digit relator codes following \$4 in the 7-- fields, and any subfield consisting of the ISO 3166 alpha-2 country code (described in UNICAT / 3).

2 **Character position.** In the fixed-length subfields the meanings of the codes are strictly dependent upon their positions in the string. **All** the character positions must therefore be present and occupied by specified or permissible alphabetic or numeric characters, fill characters or blanks.

This rule applies even if the cataloguing agency only assigns values to the first *n* character positions. It cannot be assumed that the remaining positions are unused : they must be occupied by fill characters as a default minimum.

For example, to encode a regularly-published weekly newspaper,

```
110 ##$acca@           is invalid; but
110 ##$acca|||||||@    is correct, because 110 $a is an 11-character subfield.
```

3 **Data elements** may be represented by 1-, 2-, 3- or 4-character alphabetic or numeric codes. A subfield may consist solely of one type, or may contain a mixture. For example, field 101 consists of subfields which hold only 3-letter language codes. By way of contrast, field 100 \$a contains single-character alphabetic and numeric codes, 2-figure numeric codes, 4-figure dates, 2-letter script (alphabet) codes and 3-letter language codes.

4 **Fill character.** The fill character (|) is used to replace codes

- (a) when it is not possible to assign the correct value to a coded data item (for example, when there is no corresponding value in the source format); and
- (b) when the agency never assigns values to a particular category of data in a coded field or to the indicators in particular types of fields.

For example, if the Target Audience (coded in field 100 \$a, character positions 17-19) is never designated by the cataloguing agency, fill characters are inserted:

```
100 ##$a19950303d1974####| | |y0engy01#####ba@
                ^^^
```

5 Unused positions. If two or more character positions (c/ps) are allocated to a data element, and some but not all of them are needed, **blanks**, *not* fill characters are used to fill the remaining unused positions.

For example, field 100 has room for up to four 2-digit codes to denote the character sets used in the record, in c/ps 26-33. If only one, two or three character sets are specified, the remaining positions are filled by blanks:

```
100 ##$a19950303d1974####|||y0engy01#####ba@
100 ##$a19990329d1998####|||y0gery0103#####ba@    etc.
```

6.1 If a **data element is not present** in the item being catalogued (but would be coded by the agency if it were present), the standard code **y** is normally used, followed by blanks if necessary.

For example, in field 105 \$a, c/ps 0-3 record the types of illustrations present in the item, and c/p 12 indicates the kind of biography, if applicable. A work which has no illustrations and is not biographical is therefore coded:

```
105 ##$ay###.....y@
```

The equivalents in field 140, Coded Data Field: Antiquarian - General, are in c/ps 0-3 and 19:

```
140 ##$ay###.....y.....@
```

6.2 At other times, especially if the data element is simply defined as **present/not present**, without any other qualification, the numeric values **1** and **0** respectively are used.

For example, in field 105 \$a, c/p 9 is the Festschrift indicator:

```
105 ##$a.....1...@    = the item is a Festschrift;
105 ##$a.....0...@    = the item is not a Festschrift.
```

(Note that in field 140, however, a specific *two*-letter code must be entered in one of the available c/ps 9-16:

```
140 ##$a.....ja.....@ = the item is a Festschrift).
```

7 Irregular coding. Be aware of special instructions in the *UNIMARC manual* which may require "irregular" coding, contrary to normal expectation.

For example, in field 126, Coded Data Field: Sound Recordings – Physical Attributes, \$a c/ps 7-12 provide for up to six single-character codes for textual material accompanying recordings. If there is no accompanying material, the cataloguer should apparently enter six blanks, *not* **y#####**, as code **y** is not listed. (Compare the example in 6.1 above).

Field 105 \$a uses c/p 11 for the Type of Literature code. If the item is a combination of types (for example, Humorous (**d**) poetry (**g**)), then the code **z**, 'multiple or other literary forms' is to be used, *not* **v** for 'mixed'. Code **v** is not listed.

8 Alternatives. Exceptionally, field 106, Textual material - Physical attributes, states that "Where the textual material is regular print," [that is, not Braille, large print, microprint, etc.] "the field may contain code '**r**' or be omitted altogether".

CERL recommendation: omit field 106 for normal printed materials.

9 Mandatory coded data fields. The *UNIMARC manual*, 2nd ed., 1994, specifies that a valid UNIMARC record must contain fields 001 Record Identifier, 100 General Processing Data, 101 Language of the Item (if the work has language), 200 \$a Title and Statement of responsibility and 801 Originating Source, plus three other fields which apply solely to cartographic materials.

The Permanent UNIMARC Committee decided at its meeting in March 1999 that if the source file bears no explicit coded or textual information about the language of the items in its records which can be converted directly or by algorithm, then the cataloguing agency need not regard field 101 as mandatory.

CERL RECOMMENDATION (with immediate effect): do not enter the default values

101 |# \$a | | | [Indicator 1 = fill character; \$a code = 3 fill characters]

but omit the field altogether. (The default values are not *incorrect* in these circumstances, but serve no useful purpose, other than to draw attention to the fact that the field could not be completed). The fact that the field has been omitted should be noted in the accompanying file documentation. For **original** (not retroconverted) cataloguing, the field remains mandatory.

Fields 001, 100 and 801 are always mandatory.

These rules apply equally to the fields for cartographic materials, and fields which may be defined in the future as mandatory for other kinds of materials.

10 No values available for non-mandatory fields and subfields. Similarly, if no values at all can be entered in a non-mandatory coded subfield, *omit it*: do not fill it with blanks or fill characters. It follows that if the field consists of a single \$a subfield, or of two or more subfields, none of which contains data, the entire field is to be omitted.

7

Dates of Publication, etc.

This document is a radical revision of the former UNAD / 7 dated 1995-03-04.

Several changes have been made in line with decisions taken by the Permanent UNIMARC Committee at its meetings in 1997-99, and published in Update 3 to the UNIMARC Manual.

*CERL members' attention is drawn in particular to the amended provisions in field 100, character position 8, for code **f**, code **h** when there is no publication date in the item, and the new code **u** for retroconverted records when there is no publication date information in the source file.*

The sole subfield in field 100 is \$a : this has been omitted below, so 100/8-16 = 100 ##\$a/8-16.

1. Bibliographic description

The dates recorded as part of the bibliographic description of the item (an element in the ISBD Publication, Distribution, etc., Area 4), are transcribed or supplied according to the provisions of the cataloguing rules used by individual organizations.

2. UNIMARC**2.1 Dates in the bibliographic description**

The standard practice is to record the date(s) in field 210, subfield \$d. (It is recognised that, when cataloguing old books, some libraries transcribe the titlepage, including imprint data, in 200 \$a). Note that if 210 \$d records a date in the Julian calendar, the date to be coded in field 100 is the year in the Gregorian calendar, which may differ.

These topics are not discussed further here, except for consideration of probable, uncertain and multiple dates. The examples are illustrative of most types.

2.2 Coded dates

The *type* of publication date is coded in field 100, General Processing Data, character position (c/p) 8, followed by either one or two dates in c/ps 9-12 and 13-16. The significance of these dates is dependent upon the code entered in c/p 8.

Codes **a**, **b** and **c** apply to records for serials; codes **d**, **f**, **g**, **h**, and **j** apply to records for monographs; and codes **e** and **u** apply to both monographs and serials. (Code **i** is used for films and music when there is a difference between the production date and the release/issue date: it will not appear in CERL records and is not listed in section 4 below).

3 Special types**3.1 Certain and probable dates**

Note that for the purpose of coding field 100, no distinction is made between *established* years, decades, etc., shown variously in field 210, subfield \$d as 1718, [1718], [171-], etc., and *probable* years, decades, etc., shown as [1718?], [171-?], and so on.

All codes (except **e**) distinguish between single years and ranges of years, both established and probable in both instances. However, note in particular that [ca. 1750], for example, is coded as a *single* year, even though there is an implied range of possible years.

3.2 No date

All the ISBDs, together with many cataloguing codes, assume that *some* date can be given, even if only an uncertain year or a broad range of possible dates. They make no provision for "No date" (n.d., or equivalents in other languages, s.a., u.å., etc.).

Cataloguers using rules which do permit the use of "n.d.", etc., should enter the appropriate term or abbreviation in 210 \$d. The usage, and the abbreviations employed, should be noted in the documentation describing the file. The earliest and latest possible dates should be given in field 100 following code **f** (for works known to have been published in a single year) or **g** (for those published over more than one year).

The new code **u** will be most used in retroconverted records when there is no data in the source file which can be used to supply a date or dates for 100 c/ps 9-16. It should never appear in new cataloguing.

4. Type-of-Publication-Date codes, and examples.

Examples of dates in 210 \$d and the matching codes and dates in 100 c/ps 8-16 are given in the following paragraphs.

Note! Character positions (c/ps) 0-7 are represented by leading dots in these examples; c/ps 17-35 are not shown at all.

a - c Codes for serials (see also code **e** below)

The text in the *UNIMARC Manual* for Publication Date 1 following each of these codes reads "*the beginning year of publication or coverage if coverage differs from publication.*" This implies that the date of coverage is to be preferred to the actual date of publication. A similar text is not, however, used for Publication Date 2 following codes **b** and **e**: it is reasonable to assume that the same criteria are valid. See second example to code **b**.

Blanks (#) may replace digits if the specific years following these codes are uncertain. However, the *Manual* offers no guidance for the rare (?) occasions when a possible range of dates overlaps the turn of a century. **1###**, for example, would be logical but of little value for retrieval!

CERL recommendation : If the range of dates given in 210 \$d is approximately evenly spread on either side of the century year (e.g., [between 1796 and 1804]), use the earlier decade for the coded date (179#). Otherwise code the decade with the greater number of possible years (e.g., for [between 1798 and 1805], use 180#).

```
100 ##$a.....b179#1837
210 ##$d[between 1798 and 1802]-1837
                        [decade of beginning date uncertain]
```

a = currently published serial.

Date 1 = beginning year of publication (or coverage, if different). If the beginning date is uncertain, digits may be replaced by blanks.

Date 2 = **9999**.

```
100 ##$a.....a18359999
210 ##$d1835-

100 ##$a.....a181#9999
210 ##$d[between 1811 and 1813]-    [precise year uncertain]

100 ##$a.....a18##9999
210 ##$d[182-? or 183-?]-        [decade uncertain]
```

b = serial no longer published.

Date 1 = beginning year of publication (or coverage, if different).

Date 2 = year publication ceased. Either or both dates may contain blanks if uncertain.

```

100 ##$a.....b17361757
210 ##$d1736-1757

100 ##$a.....b17731787
200 0#$aA digest of the most notable events which occurred in
    the year ...
207 #0$a1773-1787
210 ##$d1774-1788                                [years of publication differ
                                                    from those of coverage]

100 ##$a.....b182#1853
210 ##$d[between 1821 and 1823]-1853

100 ##$a.....b171#173#
210 ##$d[171-?]-[173-]
    
```

c = serial of unknown status; publication may be continuing or have ceased.

Date 1 = beginning year of publication (or coverage, if different). If the beginning date is uncertain, digits may be replaced by blanks.

Date 2 = #####

```

100 ##$a.....c1825####
210 ##$d1825-
306 ##$aLast issue received dated 1982; current publication
    status unknown.
    
```

d, f - h, j Codes for monographs (see also codes **e** and **u**)

d = monograph complete when issued, or all volumes/parts issued in one calendar year.

Date 1 = year of publication.

Date 2 = #####

```

100 ##$a.....d1695####
210 ##$d1695

100 ##$a.....d1752####
210 ##$d1752
215 ##$a3 Bd.                                [work complete in 3 vols, all issued in the
    the same year, even if not simultaneously]

100 ##$a.....d1825####
210 ##$d[1825]
306 ##$aPreface dated 12 February 1825.
    
```

Use code **d** when 210 \$d contains a privilege or copyright date, but lacks a date of publication:

```

100 ##$a.....d1675####
210 ##$dpriv. 1675                                [See EX 8 in UNIMARC Manual]
    
```

Use code **d** also when 210 \$d contains a single uncertain (or probable) year expressed as 4 digits, even though a range of possible years may be implied:

```

100 ##$a.....d1687####
210 ##$d[1687?]                                [See EX 6 in UNIMARC Manual]

100 ##$a.....d1690####
210 ##$d[ca. 1690]
    
```

(If the uncertain date in 210 \$d contains *less* than 4 digits, use code **f**. If the monograph was published over a number of years, whether the dates are certain or not, use code **g**).

e = reproduction of a document (reprint, facsimile, reissue, etc., but *not* a new edition). This code is used for both monographs and serials.

Date 1 = year of publication of the reproduction. Date 2 = date of publication of the original. Either or both dates may contain blanks if uncertain.

```
100 ##$a.....e19771610
210 ##$d1977
324 ##$aFacsimile reprint of: 1610 ed.

100 ##$a.....e17701769
210 ##$d1769, reprinted 1770
```

If either the original or the reproduction (or both) was (or were) issued over a span of years, the *beginning* years of reproduction and publication are used for Date 1 and Date 2:

```
100 ##$a.....e18101976
210 ##$d1976-1978
324 ##$aFacsimile reprint. Serial first published 1810-1843.
```

f = monograph, date of publication uncertain. Date 1 = earliest possible year of publication. Date 2 = latest possible year of publication. Blanks are permitted: see final example.

```
100 ##$a.....f16981703
210 ##$d[between 1698 and 1703]

100 ##$a.....f17001701
210 ##$d[1700 or 1701?]

100 ##$a.....f17001709
210 ##$d[170-]

100 ##$a.....f17001799
210 ##$d[17--]

100 ##$a.....f####1489
210 ##$d[not later than 1489] [See EX 16 in the UNIMARC Manual. It should normally be possible to give an earlier date]
```

Use code **f** for monographs published as a single physical item, and also for works in several volumes or parts known to have been issued in the same calendar year, although the precise year is not known. If the uncertain date in 210 \$d is expressed as a single year in 4 digits (for example, [**ca. 1715**]), use code **d**.

Use code **g** if the monograph consists of several volumes/parts issued over a period of years, *whether the dates are certain or not*.

g = monograph issued over a number of years.

Date 1 = beginning year of publication.

Date 2 = final year of publication, or **9999** if still in progress. Either or both dates may contain blanks if uncertain.

```
100 ##$a.....g17891801
210 ##$d1789-1801

100 ##$a.....g19479999
210 ##$d1947-

100 ##$a.....g165#166#
210 ##$d[165-?-166-?]

100 ##$a.....g183#1849
210 ##$d[183-]-1849
306 ##$aVol.1 published not earlier than September 1832 and
not later than March 1835.
```

CERL recommendation : Use code 'g' also for made-up collections whose contents were created or issued over a number of years (Record label / 7 = 'c').

```
100 ##$a.....g17551770
200 ##$a[Shropshire handbills]
210 ##$a[Various places]$c[Various publishers]d1755-1770
300 ##$aA collection of handbills published in or relating to
    Shropshire 1755-1770. Not catalogued individually.
```

h = monograph with both actual and copyright/privilege date. The date of publication differs from the copyright/privilege date.

Date 1 = year of publication.

Date 2 = copyright/ privilege date.

```
100 ##$a.....h17221716
210 ##$d1722, priv. 1716
```

If there is a copyright/privilege date but *no* date of publication, treat the copyright/privilege date as if it were the date of publication, and use code **d**:

```
100 ##$a.....d1745####
210 ##$dpriv. 1745
```

CERL recommendation : Use code 'g', *not* 'h', for items with both copyright/privilege dates and publication dates (or copyright/privilege dates but *no* publication dates), issued over a period of years. For the former, code the publication dates.

```
100 ##$a.....g1658-1663
210 ##$ad1658-1663, priv. 1655-1660.

100 ##$a.....g1585-1587
210 ##$adpriv. 1585-1587.
```

j = document with detailed date of publication.

Date 1 = year of publication.

Date 2 = detailed date in form "MMDD". Months 1-9 and days 1-9 are given in the form 01, 02, *etc.* If the day positions are not used, they are filled with blanks.

```
100 ##$a.....j17660707
210 ##$d7 July 1766

100 ##$a.....j14981225
210 ##$ddie Natalis Christi 1498

100 ##$a.....j166608##
210 ##$dAugust 1666
```

The examples in the *UNIMARC manual* are of modern report literature. It is evident that the technique is equally applicable to early printed books, broadsheets, *etc.*, if wanted.

Note that information retrieval systems will commonly search on the *year* recorded in Date 1. However, it may be possible to make a supplementary search for the detailed date as a string; this could be useful if the exact form recorded in 210 \$d is unknown.

u = date(s) of publication unknown.

Date 1 = ####. Date 2 = ####.

```
100 ##$a.....u#####  
210 ##$d[n.d.]
```

This code should be used for converted records when the source file contains "[n.d.]", "[s.a.]", etc., (or no data at all) in the equivalent of 210 \$d, and no coded or other data anywhere else which can be used to produce meaningful coded date information for 100.

CERL recommendation: Code 'u' should only be used as a last resort. Converted records should be edited, whenever possible, so that they contain more precise coding of publication dates. Original cataloguing should not use code 'u' at all.

8

Coded Data Fields 105: Textual Material, Monographic, and 140: Antiquarian, General

This document is a reprint of the former UNAD / 8 dated 1995-03-15 with a few minor amendments.

1. Introductory

Coded Data Field (CDF) 105 was (and still is) the standard field for recording in coded form the types of illustrations and the form of contents – arrangement of material, literary and biographical form, etc. – found in printed monographic texts.

Field 140 has been developed as a more specialised CDF for use with antiquarian printed monographic texts, and has a fuller range of codes for types of illustrations, the kinds of literature encountered in older materials, the type of paper or other material used for text and illustrations, and so on. Conversely, it does not contain codes for categories never or exceedingly rarely found in publications of this period, for example, Programmed text books, Project descriptions, and Standards. Early examples of categories which have not been allotted specific codes can be given the code for 'other'.

Field 110 is the CDF for serials which corresponds to 105 and 140. *There is no separate CDF for antiquarian serials.* These three fields are mutually exclusive : a UNIMARC record can only contain one of them.

2. Technical

CDF 105. Indicators : 1 = # 2 = #
Subfields : \$a Monograph coded data.
13 character positions, conventionally numbered 0-12.

CDF 140. Indicators : 1 = # 2 = #
Subfields : \$a Antiquarian coded data – General.
28 character positions, conventionally numbered 0-27.
C/p 8 (illustration technique) and 20-25 (type of paper, presence of watermarks, printers' and publishers' devices) have no equivalents in 105. C/p 26-27 are at present unassigned and contain two blanks, ##.

3. Conversion from 105 to 140

The table on the following pages is an attempt to provide a convenient guide for those libraries which may wish to change records previously coded using field 105 to bring them into line with other records using field 140.

It does not reproduce all the detail to be found in either 105 or 140 and must certainly not be regarded as a substitute for the official pages in the *UNIMARC manual*, which are definitive.

UNIMARC coded data fields for monographs : conversion table from 105 to 140

105 \$a		140 \$a	
c/p 0-3	Illustration codes	0-3	Illustration codes
	<i>(105 and 140 : 4 single-character code positions; unused positions filled by blanks, #)</i>		4-7
			Plates
a	illustrations	a	a
b	maps	j	j
c	portraits	h	h
d	charts	k	k
e	plans	l	l
f	plates	[use c/p 4-7]	
g	music	m	m
h	facsimiles
i	coats of arms	n	n
j	genealogical tables	o	o
k	forms
l	samples
m	sound recordings
n	transparencies
o	illuminations	b	...
y	no illustrations	y	y
...		z	other
	<i>No code for 'other' in 105, c/p 0-3.</i>		<i>140 contains codes for several types of illustrations, for example, ornamental letters, rubrics, miniatures, not listed in 105.</i>
#	value position not needed	#	#
c/p 4-7	Form of contents codes	9-16	Form of contents codes
	<i>(4 one-character code positions; unused positions filled by #)</i>		<i>(4 two-character code positions; unused positions filled by ##)</i>
a	bibliography	fc	<i>specifically; or include in</i> fa = Reference work
b	catalogue	fb	= <u>library</u> catalogue <i>specifically; or include in fa</i>
c	index	fe	<i>specifically; or include in fa</i>
d	abstract or summary	zz	
e	dictionary	ff	<i>specifically; or include in fa</i>
f	encyclopaedia	fg	<i>specifically; or include in fa</i>
g	directory	fa	
h	project description	zz	
i	statistics	zz	
j	programmed textbook	...	
k	patent	zz	
l	standard	zz	
m	dissertation or thesis	bb	
n	laws and legislation	da	
o	numeric table	zz	
p	technical report	zz	
q	examination paper	zz	

r	literature survey, reviews	zz	
s	treaties	da	
t	cartoons or comic strips	zz	
z	other	zz	<i>Note that many of the forms considered as 'other' in 105 have their own codes in 140, for example, sermons, ephemera.</i>
#	value position not needed	##	
c/p 8	<i>Conference or meeting code</i>	9-16	
0	not a conference	...	
1	conference publication	zz	
c/p 9	<i>Festschrift indicator</i>		
0	not a festschrift	...	
1	festschrift	ja	
c/p 10	<i>Index indicator</i>		
0	does not contain an index	...	
1	contains an index	...	<i>Note that c/p 9-16 'fa' = the work <u>is</u> an index</i>
c/p 11	<i>Literature code</i>	17-18	<i>Literature code</i>
a	fiction	ea	<i>or use more specific forms eb-ej</i>
b	drama	ca	<i>libretto = da</i>
c	essays	fa	
d	humour, satire	ga	
e	letters	ha	
f	short stories	ej	<i>specifically; or include in ea</i>
g	poetry	aa	<i>romance, specifically; gesta, pastoral romance, etc. = ab</i>
h	speeches, oratory	ma	
y	not a literary text	yy	
z	multiple or other literary forms	zz	<i>Note that many of the forms considered as 'other' in 105 have their own codes in 140, for example, proverbs, mystical literature.</i>
c/p 12	<i>Biography code</i>	19	<i>Biography code</i>
			<i>Note that these are used <u>in addition to</u> the relevant form codes in c/p 17-18 (diaries, etc. = lc-lf) for which there are no equivalents in 105 \$a, c/p 11.</i>
a	autobiography	a	
b	individual biography	b	
c	collective biography	c	
d	contains biographical information	d	
y	not biographical	y	
...		z	multiple or other
	<i>No code for multiple or other in 105.</i>		

9

Authority file record numbers and cross-references

This document is a partial revision of the former UNAD / 9 dated 1995-08-09.

1. CERL recommendation.

Some members' bibliographic files have records which contain references to authority files, or fields which hold alternative forms of names.

Members are *strongly recommended* to leave these authority file numbers and alternative forms in place when they are preparing files for contribution to the CERL database, on the grounds that

- other members may have access to those authority files, or may be able to make direct use of the alternative forms of names in the records;
- in the longer term, the data will be a valuable resource when the CERL authority file is created.

2. Subfield \$3.

The *UNIMARC manual* specifies the use of subfield \$3 in fields 600-602, 605, 606, 700-702, 710-712 and 720-722 to hold IDs linking to records in authority files using the *UNIMARC/Authorities* format. By extension, subfield \$3 may also be used in CERL fields 69x and 79x.

Exceptionally, CERL records may use \$3 to link to authority files which, although not using *UNIMARC/Authorities* itself, are very closely aligned with it (accommodating UNIMARC heading structures, etc.).

3. Authority File Record IDs.

- 3.1 To avoid the possibility of confusion, authority record numbers must be prefixed by the ISO 3166 Alpha-2 country code, the code for the organization creating and maintaining the authority file and a code for the file. The country code, organization code, file code and record numbers must be separated by backslashes. This is identical to the format used for CERL record IDs in field 001, etc., as explained in *UNICAT 4*.
- 3.2 Note that the prefixed codes may not necessarily be those of the organization creating the bibliographic record : it is quite possible for a library to make use of an authority file created by a different cataloguing agency – even one in a different country.
- 3.3 Order of subfields. The statement in the *UNIMARC manual* (Section 4.3) means that whenever a \$3 Authority Record Number appears in a 6-- or 7-- field, it should be the *first* subfield in the field.

4. Records with fields containing alternative names or forms of names**4.1 UNIMARC**

The UNIMARC bibliographic format makes no provision for fields containing alternative names or forms of names (unlike UKMARC and some other formats). The assumption is that these will be found in an associated authority file.

4.2 CERL fields

CERL has provided fields 690-692 and 790-792 for alternative names and forms of names found in fields 600-602 and 700-722 respectively. The details of these fields are given in the *CERL Supplementary Fields* series.

CERL 69X and 79X fields do not carry pointers to the specific fields to which they refer. Unless explanatory phrases or authority record IDs are added, the connections between them can only be inferred if there are several 60X/69X or 7XX/79X fields in the record. This is a serious disadvantage, and CERL members are **recommended** to include explanatory phrases and/or authority record IDs whenever possible.

The examples show (1) a heading without explicit connections – there could well be other 701 and 702 headings in the record, with associated 790 fields; (2) a heading with an embedded \$3 subfield pointing to the authority file – the reference forms are not shown in the bibliographic record; (3) a 790 field with an added explanatory phrase showing which 70X field it refers to; (4) three 70X headings and their 79X non-preferred forms, with \$3 pointers to the associated authority file as well.

5. Documentation.

If files contain authority record numbers, members must include information about the authority file (its format, and the maintaining agency) in the file documentation.

EXAMPLES:

EX 1

```
700 #0$aFelio$cPapa, V, Antipapa
790 #0$aAmedeo$cSavoyen, Herzog, VIII
790 #0$aAmedeus$cSavoyen, Herzog, VIII
```

700 Author heading and CERL 790 fields from Bayerische Staatsbibliothek "AK" file. There are no links between them, although they do, in fact, refer to the same person. There could be other 70X fields present in the record.

EX 2

```
702 #1$3HR\NSK\1\950123071$aCovarrubias$bDiego de
```

Author heading from Nacionalna i Sveucilisna Knjiznica, Zagreb, with embedded ID for that heading in the library's authority file. Alternative forms are not shown in the bibliographic file.

EX 3

```
700 #1$aGrick$bFriedrich
790 #1$aAgnostus$bIrenaeus$cpsseud. [i.e. Friedrich Grick]
```

Author heading from British Library "K17" file, with explanatory phrase in the CERL 790 field.

EX 4

```
700 #0$3IT\ICCU\SBLV\218266$aDino $b: del Mugello
702 #1$3IT\ICCU\BVEV\020845$aDescousu$b, Celse Hugues $f<1480ca.-1540>
702 #1$3IT\ICCU\RMLV\020396$aBohier$b, Nicolas $f<1469-1539>
790 #0$3IT\ICCU\SBLV\218268$aDinus $b: de Muxello
790 #1$3IT\ICCU\BVEV\027885$aHugo$b, Celsus
790 #1$3IT\ICCU\RMLV\020947$aBoerius$b, Nicolaus
790 #1$3IT\ICCU\SBLV\218267$aRossono$b, Dino
```

Part of record from ICCU file showing three 70X fields, all with associated 79X fields *and* \$3 IDs. In this system the authority records do not use a single number for all forms of a name : the connections are made internally.

10

The UNIMARC Record Label

1. General features

The Record Label is a mandatory, non-repeatable element of UNIMARC and all other formats which are implementations of ISO 2709.

It is found at the beginning of every record, preceding the Directory, and consists of a fixed field of 24 characters whose positions are conventionally numbered 0-23. It has no field tag, indicators or subfield identifiers. (The tag “000” or “LAB”, etc., seen in some diagnostic displays is the product of the particular software used by the system).

The data elements in the Record Label are a mixture of

- fixed values which are the same for all UNIMARC records;
- variable values which may be converted from a source format or, where UNIMARC is itself the source format, assigned by the cataloguer individually or sometimes as default values; and
- numerical data calculated by the system when the record is complete and formatted.

The following sections will deal with each of these three groups. Problems are encountered only in interpretation of the variable values.

2. Valid characters

Characters valid in specified positions in the Record Label are the digits 0-9, lower-case letters a-z and the blank (represented by the sign '#' in the examples below). The fill character '|' may **not** be used in any position (*see* section 4.5 in *UNIMARC Manual* and UNICAT 5).

3. Fixed values

The UNIMARC Record Label has invariant values in the following character positions:

Pos Value Definition

9	#	Undefined
10	2	Indicator length
11	2	Subfield identifier length
19	#	Undefined (unused position in Additional Record Definition group of elements)
20-23		<i>Directory map</i>
20	4	Length of “Length of field” part of each Directory entry
21	5	Length of “Starting character position” part of each Directory entry
22	0	Length of implementation-defined portion of each Directory entry
23	#	Undefined (unused position in Directory map)

These values may therefore be input automatically as constants for every UNIMARC record:

```

0           1           2
012345678901234567890123
.....#22.....#450#
    
```

4. Automatically calculated numeric values

Character positions 0-4 give the total number of characters in the entire record (right justified, with left-hand zero fill if necessary, e.g. 00953). It is this element which determines that the theoretical maximum size of a UNIMARC (or any ISO 2709) record is 99999 characters.

Character positions 12-16 give the *Base Address* of the data relative to the beginning of the record, that is, the starting character position of the first data character immediately following the Record Label and Directory (including the End-of-Field character which terminates the Directory). The

Base Address gives the starting point from which the position of each field in the record is calculated in the Directory. This element is similarly justified. Note that "data" in this context include indicators and subfield identifiers.

Both these elements will normally be calculated and supplied automatically by the system and no action is needed on the part of the cataloguer. Only in extremely rare cases of software errors, etc., would it be necessary to check the accuracy of these calculations manually.

5. Variable values

The values in many of the remaining character positions may often be pre-set as defaults when cataloguing is organised so that all the records within each batch are of the same type. For example, if all the records are *new*, describe *printed language materials* and are *monographic*, and the library *does not use hierarchical relationships* (46- fields), is cataloguing the items *fully, book-in-hand*, using *some, but not all, ISBD elements*, then the following may be pre-set if the system allows this (they will otherwise have to be entered manually into each record):

```

0           1           2
012345678901234567890123
.....nam#.....#i.....

```

Similarly, a batch of records for *[printed] serials, corrected and upgraded to full level, with the items in hand*, using *hierarchical relationships* and *ISBD punctuation*, could have the following pre-set coding:

```

0           1           2
012345678901234567890123
.....cas.....##.....

```

In this example, position 8 is not pre-set because the hierarchical coding could be '0', '1' or '2'.

5 Record status

Value Definition

c Corrected record

A record to which changes have been made – correcting errors
– bringing it up to date
– deleting fields

The earlier record may have the value 'n', 'o' or 'p'.

(A full record replacing a pre-publication (e.g., CIP) record is coded 'p', *not* 'c' – but CERL files should not contain any records with the code 'p').

d Deleted record

A record indicating that an earlier record bearing the same control number is no longer valid. The record may contain only the Record label, Directory and field 001 (Record Control Number) *or* all the fields in the record as originally issued. In either case field 300 General Note may be added, giving the reason(s) for the deletion.

n New record

A new record, including a pre-publication (e.g. CIP) record, before it is upgraded to a full(er) record. If code 'o' for a new record below the highest level in a hierarchy applies, *prefer 'o' to 'n'*.

CERL ruling: use code 'n' for new UNIMARC records, even if they have been converted from records previously existing in a different machine-readable source format.

o Previously issued higher level record [= New lower level record]

A new record at any level in a hierarchy below the highest when a higher level record has already been issued. When character position 5 = 'o', c/p 8 = '2'.

CERL ruling: If both the highest and lower level records are created and issued as part of the same cataloguing operation, use code 'n' for all of them. If a higher level record has been issued previously, and new records are created at more than one level below it in the same hierarchy, use code 'o' for all these new records.

p Previously issued as an incomplete, pre-publication record.

E.g., an upgraded CIP record. **This code is invalid for CERL records.**

6-9 Implementation codes

6 Type of record

These codes indicate the actual type of material being catalogued rather than its secondary physical format – e.g., photographic (including microphotographic) and digital reproductions of printed books are coded 'a', not 'g' or 'l'. The secondary physical format is indicated by the presence of the appropriate Coded Data Field (105-141), the General and Specific Material Designations (200 \$b and 215 \$a) and Material Specific Area fields (206-208 and 230).

Records in the CERL database are restricted to the description of original items created by letterpress printing, although they may carry information about the availability of reproductions, etc.

The code 'a' is always valid for CERL records and can normally be set as a default; 'c', 'e' and 'k' are occasionally valid; the remainder are not.

Value Definition

a	language materials, printed	[Normal value for CERL records]
b	language materials, manuscript	
c	music scores, printed	[Scores printed by engraving or lithography are outside CERL's scope, but scores printed from movable types may perhaps be included]
d	music scores, manuscript	
e	cartographic materials, printed	[Normally excluded; a few items which also contain letterpress may be included]
f	cartographic materials, manuscript	
g	projected and video material (motion pictures, filmstrips, slides, transparencies, video recordings)	
i	sound recordings, non-musical performance	
j	sound recordings, musical performance	
k	two-dimensional graphics (pictures, designs, etc.)	[Normally excluded; a few items which also contain letterpress may be included]
l	computer media	
m	multimedia	
r	three-dimensional artefacts and realia	

7 Bibliographic level

Value Definition

a	analytic (component part).
m	monographic = bibliographic item complete in one physical part or intended to be completed in a finite number of parts. This code, the most common in CERL records, may often be set as a default.
s	serial = bibliographic item issued in successive parts and intended to be continued indefinitely.
c	collection = bibliographic item that is a made-up collection. Code 'c' is for use when a collection is described as a whole, for example, a box containing a group of pamphlets and other ephemera not catalogued individually. When an "artificial" collection is described in separate records with linking 481/482 fields, each constituent record in the set should be given its appropriate code (usually 'm'), not code 'c'.

8 Hierarchical level code

Value Definition

- # Hierarchical relationship undefined.
The cataloguing agency *never* uses hierarchical relationships.
**If c/p 8 = '#', then all the records in the file must have 8 = '#'
and no record in the file may contain field(s) 46–**
- 0 No hierarchical relationship [associated with this record]
The cataloguing agency *does* use hierarchical relationships for multivolume, etc., records, but not for this particular item.
**If c/p 8 = '0', then no record in the file may have 8 = '#'
and this record must not contain field(s) 46–**
- 1 Highest level record
- 2 Record below highest level (all levels below)
**If c/p 8 = '1' or '2', then no record in the file may have 8 = '#'
and this record must contain field(s) 46–**
Although records for all levels are commonly created at the same time, it is quite possible for a higher or highest level record to be created on one occasion and further lower level records to be created for addition to the file at a later date. If this occurs, then c/p 5 = 'o' for the later additions.

9 [Undefined]

- # (blank)
A constant value. See section 3 above.

17-19 *Additional record definition*

17 Encoding level

"A one-character code indicating in general the degree of completeness of the machine record, and whether or not the item was examined when the record was created". – *UNIMARC Manual*.

This code has caused more problems of interpretation than any other, because (a) it fails to make it absolutely clear whether "completeness of the machine-readable record" is primarily concerned with completeness of the *record structure* in terms of provision of tags, indicators and subfield identifiers or *data content* (fullness of cataloguing), and (b) it confuses the two quite distinct questions of *fullness of the record* – however defined – with whether or not the *item was (re)examined* when the record was prepared for inclusion in the machine-readable database. Ostensibly, the code is concerned with fullness, but the explanatory text to the first two levels seems to define this by reference to examination of the item.

Whether the UNIMARC record has been created by original cataloguing or by conversion may be shown by the absence or presence of subfield \$2 in Field 801, Originating Source (see Section 6 below).

Value Definition

- # Full level. The item was examined when the record was prepared for inclusion in the machine-readable database.
CERL rulings: Use this code for the library's fullest, modern records, whether created in UNIMARC or converted subsequently. The records must have been created item-in-hand, and be fairly full in data content, with tagging, etc., appropriate to that content.
It is recognised that interpretations of "fullness" of data content will vary from library to library.
In cases of doubt whether a record qualifies for Full level, use Sublevel 1 or 3.

- 1 Sublevel 1. "The item was *not* examined when the record was prepared for inclusion in the machine-readable database" – *UNIMARC Manual*.

Typically, this may apply to retrospective conversion from card or printed catalogues when tags, identifiers and subfield indicators may not have been added with the same degree of reliability as can be achieved with the item in hand.

Records coded at Sublevel 1 may be full in terms of data content – sometimes as full as or fuller than records coded at Full level – but are not guaranteed to be fully and accurately tagged, etc.; coded data may also be lacking.

- 2 Sublevel 2. Pre-publication (CIP) record.

This code is not valid for CERL records.

- 3 Sublevel 3. Less than full cataloguing which may or may not be subsequently upgraded to full level.

Use Sublevel 3 for any records which are recognised as being well below current cataloguing standards, whether or not the items were (re)examined at the time of record creation

18 Descriptive cataloguing form

Code indicating whether the fields in the 2-- block use the provisions of ISBD.

Some other fields may also use ISBD conventions for cited works (for example, 324 Original Version Note, 327 Contents Note), but these should not be taken into account in determining the correct code here.

Value Definition

- # All the ISBD data elements conform to the provisions of ISBD.
- i Incomplete or partial use of ISBD : some but not all 2-- fields conform to ISBD.
- n Non-ISBD : none of the descriptive data elements necessarily conforms to ISBD.

(See also section 7 below)

19 [Undefined]

- # (blank) A constant value. See section 3 above.

6. Converted formats

If a record has been converted from another format, this is not indicated in the Record Label, but in field 801 Originating Source, subfield \$2 System Code. The codes are listed in Appendix H. (This subfield was added to the format in Update 2, 1998).

CERL ruling: If no code is listed for a source format, use "ISO2709" for any format which conforms to that standard, and "Non-ISO2709" for any other, until specific codes have been approved by the Permanent UNIMARC Committee. Draw attention to the usage in the documentation accompanying the file.

7. Cataloguing rules

The *specific* cataloguing code being used (which may or may not incorporate ISBD) is identified in field 801 Originating Source, subfield \$g Cataloguing Rules (Descriptive Conventions). The codes are listed in Appendix H.

EXAMPLES

***** = *Machine-generated numeric values*

EX 1

0 1 2
012345678901234567890123
*****nam##22*****###450#

New, printed, monograph; hierarchical relationships not used; full cataloguing; full ISBD.

EX 2

0 1 2
012345678901234567890123
*****oas2#22*****3i#450#

New lower-level record [= record above it in the hierarchy created and issued on a previous occasion], printed, serial; low-level cataloguing; partial ISBD.