Werk

Titel: Describing Geomatic Data Sets with ISBD and UNIMARC: Problems and Possible Soluti...

Autor: Smits, Jan

Ort: Graz

Jahr: 1995

PURL: https://resolver.sub.uni-goettingen.de/purl?514854804_0005|log74

European Research Libraries Cooperation:
The LIBER Quarterly, 5 (1995), 292-311.

Describing Geomatic Data Sets with ISBD and UNIMARC: Problems and Possible Solutions

JAN SMITS
Koninklijke Bibliotheek, The Hague

Identification and retrieval

"O brave new world, that has such people in it. Let's start at once", exclaimed the Savage in Aldous Huxley's famous novel when he met the first representatives of the new society. However, his new acquaintance retorted: *"You have a most peculiar way of talking sometimes. ... And, anyhow, hadn't you better wait till you actually see the new world?"*[1] Before we fall for the magic of a new possibility we have to check if and how we can control it to make practical use of it.

Since the first appearance of the ISBDs in the early 1970s many have been published for specific materials: monographs, serials, non-book materials, cartographic materials, music, antiquarian items and, only recently, computer files[2]. All of them fit well into the ISBD(G) in which only area 3 (material, or type of publication, specific area) is not defined specifically. All these ISBDs were primarily developed "to aid the international exchange of bibliographic records between national bibliographic agencies and throughout the international library and information community"[3]. This was a laudable goal except that the exchange of bibliographic records was intended to be realised electronically. For

1 Huxley, Aldous. Brave new world, Chatto & Windus, 1932. The implication of this quote might be that map curators, like the Savage, could revert to old ways because they are daunted by the new technologies.

2 The first draft appeared in 1986. As there was not then much experience with computer files in libraries areas 3 (type and extent of file) and 5 (physical description) were still causing many problems.

3 ISBD(CM) : international standard bibliographic description for cartographic materials / recommended by the ISBD Review Committee. - Rev. ed. - London : IFLA Universal Bibliographic Control and International MARC Programme, 1987. P. 1: purpose.

this purpose MARC-formats were created from the late 1960s onwards. Unfortunately, as this depended on available hard- and software many different MARCs were developed, albeit on the same basis and with the same functions. IFLA tried (and still tries) to alleviate the problem of exchange by creating UNIMARC with the primary purpose of putting an exchange format at the disposal of the library and information community.

However, the MARCs have more functions than only supporting the exchange of bibliographic records. Indeed they code information within a bibliographic record in such a way that electronic retrieval is made possible. This function is even more important than the exchange function as it makes our bibliographic records machine-processable and allows us to query the contents of a bibliographic database in such a way that we can trace information according to most of our requirements. Here lies the added value of MARC compared to ISBD. Where ISBD is primarily developed to identify specific publications, the MARC is primarily developed to store and process information contained within these ISBD-records and beyond. To enlarge this function of retrieving information most MARCs have been complemented with a "coded information block" and an "intellectual responsibility block". In this respect, UNIMARC is possibly one of the most complete standards so far available[4]. But it also creates many problems. As information needs to be standardised, especially in the coded information block, and be acceptable in the user community when new codes are added, the revision cycles are too long for practical use. I shall try to propose a possible solution to this problem in the course of this paper.

Published and unpublished materials

Up to now we have described cartographic materials which have been published as finished products, e.g. maps, globes, aerial photographs, CD-ROMs, etc. However, in our field there are many remote-access cartographic databases which are dynamic and available to the public. If we do not possess a certain cartographic item we usually refer our client to the organisation which has it at its disposal. The same is true for digital cartographic databases which are only available online or from which one can order a part or the whole on request. It is highly unlikely that we shall ever possess the bigger databases ourselves (I am

4 UNIMARC manual / ed. by Brian Holt with the assistance of Sally H. McCallum & A.B. Long. - [London] : IFLA Universal Bibliographic Control and International MARC Programme, 1987. Coded information block: p. 52-134; Subject analysis block and intellectual responsibility block: p. 289-347.

A second edition of the UNIMARC manual has been issued in 1994: UNIMARC manual : bibliographic format. -2nd ed. - München [etc.] : K.G. Saur, 1994. - (UBCIM publications - New series ; Vol. 14).

referring mainly to base maps, etc., of official organisations) because we cannot afford them financially[5]. There is a real danger that some of these base maps may no longer be published in hard copy in the future or only on request[6]. Will that make our holdings into collections of old maps or will they evolve into cartographic information centres? As far as I can glean from developments in the U.S.A. and Canada and within the Koninklijke Bibliotheek I think the latter is the more likely[7]. If so, this means we have to know the contents and potential of these databases and integrate them as metadata descriptions into our bibliographic apparatus[8]. This may make our field of work more diffuse than it is, at least for the time being. So far we have occupied ourselves with

5 The total update of the "Topographic subjectmap of The Netherlands" of the topographic survey will cost appr. NLG 420,000.-. The "Photographic map of The Netherlands" of the private firm ROBAS in hardcopy costs now NLG 518,000.-, and will be flown every two years; when the colour aerial photographs will be digitized the price will not be less. The Netherlands is a small country, so I presume that larger countries have to think in millions.

6 See: James D. Elliot: Digital map data: archiving and legal deposit implications for U.K. copyright map libraries. In: ERLC The LIBER Quarterly, Vol. 2(1992), No. 2, pp. 119-127.

The municipalities of The Hague, Utrecht and Amsterdam have now ceased to publish large scale maps (1:1,000 and 1:2,000). Also The State Service for Road- and Waterways Management has ceased to publish the 'Waterstaatskaart van Nederland 1:50,000' (Water Management Map). My experience up till now has shown that hard copy of remote-access dynamic digital maps on request is far more expensive than traditional analoguous material, because there is no large print-run anymore and because production has to be more cost-effective. Whereas an A0 analoguous colour map would cost something like NLG 15.-, the same map, but now digital, will cost at least some NLG 35.- for an A4 hard opy, which is a differential rate of almost 20.

7 Chris Perkins of Manchester University believes that the differences in the level of service will increase. On the one hand there will be a few well-equipped map collections which can offer a large array of conventional and digital services (the elite groups in society), on the other hand a large majority of map collections will continue to offer only analogue products to their clients.

See: Chris Perkins: De kwaliteit van kaartbeheer en kaartdocumentatie in het GIS-tijdperk. In: Kartografie in het GIS-tijdperk / red.: P.G.M. Mekenkamp. - [Amersfoort : NVK, 1994]. - (NVK publikatiereeks ; nummer 11). Pp. 59-68.

8 I do not agree with Chris Perkins who suggests in his paper [see above] that the only short term solution would be to set up separate bibliographic databases for digital and analoguous cartographic materials. Especially since there are as yet no descriptive standards for digital material, this suggestion may even widen the gap between the elite and the poorer map collections. In view of the continuity in time of cartographic (geo-referenced) information I feel that his proposal shows the same kind of bias as that of people wishing to create separate bibliographic databases for old and modern cartographic materials, merely because present use seems to indicate such a course (which I doubt). If this is realised, inevitable problems will arise when modern materials become old materials, as separate databases tend to diverge in standards and contents.

cartographic materials. But when it comes to databases, our scope will be enlarged to the whole field of geospatial data[9]. Many databases do not contain maps. They contain geo-referenced data and can also contain programmes with which one can create maps. If not, a separate programme has to be added to create maps. Though it is in visualisations of spatial data where our main qualities as map curators lie, we have to upgrade our knowledge in order to handle the underlying unprocessed data to be able to (help) create such visualisations. This in its turn will probably show itself in the descriptions we will add to our cataloguing apparatus[10].

ISBD

Let us first try to see whether the present ISBDs can be used for describing electronic documents and what problems they might pose.

9 Part of the title of this paper derives from the the Canadian publication Geomatic data sets : cataloguing rules. In this publication 'geomatics' is defined as follows: "The scientific and technical domain concerned with methods, procedures and technologies associated with computer systems for the collection, manipulation, display and dissemination of geographically referenced data". This may be the field with which map curators in future will occupy themselves, thereby broadening their field of work to include production and use.

See further: Velma Parker (ed.): Geomatic data sets : cataloguing rules / prepared by the Canadian General Standards Board ; approved by the Standards Council of Canada. - Ottawa : Canadian general Standards Board ; Canadian Library Association, 1994.

10 It would be a great help if throughout the world initiatives were taken as are now taken in the United States of America. In President Clinton's Executive Order of April 11, 1993, titled: "Coordinating geographic data acquisition and access: the National Spatial Data Infrastructure" he called upon the Federal Geographic Data Committee (FGDC) "... [to] adopt a schedule ... for documenting, to the extent practicable, geospatial data previously collected or produced, either directly or indirectly, and making that data documentation electronically accessible ...". This has resulted in the following standard which is meant for producers: "Content standards for digital geospatial metadata (June 8). Washington D.C., Federal Geographic Data Committee, 1994". [I could not ascertain why extra-terrestrial spatial data are for the time being excluded.] These standards describe in minute detail how a producer should create a metadata record of a certain digital set and what kind of information the elements should contain (including as to fitness of use). However, the standards are not meant in the first instance to provide for ISBD- or MARC-descriptions. But certain elements are analogous to or can be used in MARC formats as was pointed out by Gary Fitzpatrick of the Library of Congress. I think it worthwhile to consider whether European producers can match this initiative in cooperation with supranational bodies, maybe also at the instigation of map curators?

The standards and related documents are available from anonymous FTP server fgdc.er.usgs.gov in directory GDC\METADATA or on the Internet by electronic mail on gdc@usgs.gov.

When describing cartographic materials we can work with ISBD(CM), ISBD(A) and ISBD(CF). Of course it is possible to use all three for one cartographic data package[11], as the ISBDs are not mutually exclusive. Though the ISBDs are created for the description and identification of certain kinds of materials, where the material is more specified by its form than by its contents, future developments can be incorporated. Should one cataloguing agency decide to use only one of the ISBDs the same material might be described by another cataloguing agency with a different ISBD. In this way arbitrariness is introduced, and that is one of the possibilities we want to prevent by using the same kind of rules for the same kind of material. It would be wise for the ISBD Review Committee of IFLA to put more emphasis on the possible integrated use of the ISBDs with reference to these problems in the next revision cycle of the ISBDs.

When describing cartographic materials I usually emphasise that I do not describe a certain document for identification purposes[12] but that we are mainly interested in analysing its contents and putting them it in a form which helps us meet the demands of our clients. For me the ISBD is a vehicle to give such form to a description that it is internationally understandable. Which does not mean that ISBD has no inherent value: it provides us with a structure.

As soon as we are able to describe digital maps and databases there are two more or less defined types that we will have to handle. There are the products which are finished and there are dynamic databases. Finished products have a certain lay-out and have identifying data like more traditional materials, like a title page, credits, physical data, edition statement etc[13]. They can be treated in roughly the same way as traditional maps. But dynamic databases lack most of these features, so we have to do some creative cataloguing to incorporate them.

11 It is possible that one has to describe an old map (CM and A), a modern map (CM), a computer-map (CM and CF), and a scanned and computerized old map (CM, A and CF).

12 If cartographic materials are described for identification purposes it is mainly for historical reasons and usually concerns old or antiquarian maps. My practical experience is that hardly any client demands a modern document by title or other discriminatory bibliographic data. Not to mention the fact that, for instance, titles in many modern cartographic materials are hardly relevant because of their generic contents. I assume that even if map producers had been more creative with titles it would not have made much difference to the demands of our clients.

See for further information: Jan Smits: Report on the 'Inquiry into map-use and user-habits in Europe'. In: ERLC The LIBER Quarterly, Vol. 1 (1991), No. 3, pp. 283-310.

13 ISBD(CF), paragraph 0.5.1 Order of preference of sources: "Sources internal to the computer file shall be preferred to all other sources. Such information must be formally presented and can usually be found in title screens, in the main menu or prominently in the listing of the file's programme statement".

Form

As the form in which information can be published[14] proliferates through time -currently more quickly than ever before- and as the amount of information seems to grow exponentially, the form becomes less meaningful to those who are seeking information. At a later stage form might be important if they can choose between different formats of the information, though they might also prefer to opt for more than one form. But first and foremost they are probably interested in the content of the information. This may lead to the question whether form should still be the decisive aspect of the ISBDs. I can imagine that the ISBDs will be remade in ISBDs for contents and that they will have a special field for form-attributes. To begin with I would look to area 5 (physical description area) to fulfil this function. Area 3 (material specific area) will still be reserved to distinguish the different kinds of information[15]. This would mean that the ISBD(CF) first of all would be reserved for describing pure computer programmes, etc. Fortunately the ISBDs include the following remark: "Each ISBD is intended to embody a coherent set of provisions for its own type of publication, but there has been no attempt to make any ISBD exclusive. Users will, on occasion, need to refer to several ISBDs when, for example, the item for description exhibits the characteristics described in other ISBDs, such as a computer-readable item published as a [map], or with an accompanying monograph"[16]

ISBD-Description

When we combine ISBD(CM) and ISBD(CF) the description of a finished dynamic digital map may look as follows[17]:

14 "Published" should be read here as information that has been made available to the public in one way or another. Even if this has some restrictions, as may be the case with copyrighted or trademarked commodities, as to purpose or group of people for which it is intended, I have defined it as 'published'. Thus it includes analogue information, computer files, audio and video packages. In contradistinction to former definitions, which usually refer to finished articles, it can also include intermediate forms of publication.

15 In this I disagree with the Canadian rules (see note 9). Although this element is repeatable I would opt, if possible, for only one GMD (General Material Designation), which is also in field 3 of the ISBD. This to discriminate between form and contents of the described material.

16 ISBD(CF): international standard bibliographic description for computer files. - London : IFLA International Office for UBC, [1986]. - Draft for worldwide review. Paragraph 0.1.1. Scope.

17 Most data have been translated for this article.

Transport atlas of the southern North Sea : display programme / Rijkswaterstaat ; Delft Hydraulics. - various scales (W 005-E 012/N 063-N 050). - [The Hague] : Rijkswaterstaat, Service for Tidal Waters ; [Nijmegen : Mooren, dist., 1987]. - Computer data (1 file, 260 Kb) and programmes (6 files, 93 Kb) on 1 computer floppy disk : 5¼ inch, DS,DD ; 14 cm + 1 atlas.
Optional auto-scaling: a. logarithmic; b. reversed logarithmic.
Maps can be presented in black and white as well as in color.
By means of the programme one can create maps in an infinite number of scenarios by manipulating influx as well as concentration-factor (of substances); the programme offers the possibility of creating detailed and general maps in which are depicted how the effluents of rivers entering the North Sea and the substances dissolved therein will, on average, be distributed over the North Sea.
System demands: IBM (-compatible), MS-DOS, colour/graphic adaptercard; Olivetti, MS-DOS, colour/graphic adaptercard
ANNEX: Transport atlas of the southern North Sea / W.P.M. de Ruiter ... [et al.] ; graphic design and production: Mooren. Scales vary from [ca. 1:2,750,000 to ca. 1:8,000,000]. Contains maps of the North Sea with the influx of respectively The Channel, Firth of Forth, Tyne, Tees, Humber, Thames, Schelde, Rijn/Maas, IJsselmeer, Ems, Weser and Elbe; the maps show the distribution of various watermasses over the North Sea, with predominantly southwesterly winds of 3.5 m/s and 4.5 m/s. The atlas contains technical information about how to start the programme.
Classification: <3.113>; 543.54
Sign. DNP: KC3.113 -0000/001/00000/00/1987/1 D870000

Title and credits are derived from the title screen.

As accuracy is probably not crucial for small scales it is not given in this publication.

The physical description area is a combination of areas 3 (type and extent of file) and 5 (physical description) of ISBD(CF). As certain drives do not accept certain floppy disks it is advisable to specify the kind of floppy disk (DS, DD, HD or other).

To be able to interpret the data on the screen well, it is necessary to use the atlas as the distribution of the water masses depends on the wind speed and wind direction.

In 1992 we have made some experimental ISBD descriptions, describing the publication as an atlas with a floppy attached (CM), and as a digital publication with a hard copy atlas (CF) and (CM) + (CF). My present position is that it should be described as shown above (CM) with computer-related data in area 5 (physical description) and area 7 (notes).

A description of a remote-access dynamic cartographic database could look as follows[18]:

[Digital topographic subject map of The Netherlands] / Topografische Dienst Nederland. - Situation on 31-12-1992. - Representational scale 1:5,000 to 1:25,000, standard deviation 1.8 m ; with coordinates of the shifted Dutch triangulation system (E 3 20-E 7 15/N 53 35-N 50 45). - Emmen : Topografische Dienst Nederland, 1991-... - ca. 1,980 Mb data.
Mapping scale 1:10,000.
Update: depending on area 4, 6 or 8 years; the update programme is drawn up according to the present requirements of the Ministry of Defence.
Extent of database: urban area ca. 1.6 Kb per hectare; rural area ca. 0.3 Kb per hectare.
Format of delivery: SUF2, DGN (microstation format), DXF, DWG (autocad format); 9 inch magnetic reeltape, Exabite (8 mm cassette); small files on diskette.
Deliverable on 31-12-1992: Sheet 9W, 9O, 14W, 14O, 15W, 19W, 19O, 20W (excl. Texel); delivery per sheet or part thereof.
Language: English
Contents: Coded vector-database with all topographic point-, line- and area-information, as is represented on the present maps of 1:10,000 and 1:25,000; classification of several topographic features ('layers') such as roads, water, buildings, etc.; within each theme there is subcoding.
Structure: All line-segments are noded and, when they have an identity of their own, fitted with their own coding; the polygons (objects) which thus are created are coded with a centroid.
Literature: Van basisbestand naar kernbestand : de Topografische Dienst als producent van een kernbestand 1:10.000 / P.W. Geudeke. In: Kartografisch tijdschrift, 1993.XIX.2, pp. 24-28.
Classification: <4.210>; 273
Sign. DNP:KC4.210 -0000/001/00000/00/1991/1 D9400001

There are some differences with a traditional description. The first one is the title. Almost always this title will not appear on the screen, which means that the bibliographic agency has to create one, if possible in cooperation with the producer. To make the title relevant it should include a statement of both subject and area and, when applicable, a contents date. A lot of other information could

18 Data derived from: De Kaarten van tafel : aanbod en gebruik van digitale kaartbestanden en andere ruimtelijke gegevens bij de waterschappen. - [Den Haag] : Unie van Waterschappen, 1993. See Appendix 1 for a description of this map. All data have been translated for this article.

be put between square brackets, but one may also use accompanying texts as source (see note 3: sources). I guess that most producers will publish instructions explaining the uses of their databases. For argument's sake I shall assume that this has happened. As with traditional material we can always put in a note to indicate the source from which the title has been derived.

A second problem is the continuous update of these databases. One can, of course, make an open description which would only show in area 4 (publication, distribution etc.). However, in that case the history of the database can only been shown by continuously adapting the description. I would prefer to include a description of the database annually in the national bibliography and to show this in area 2 (edition). If necessary I even can make a three-level description, describing on the first level the database in general, on the second level a specific sheet and on the third level the situation at the end of each year. I agree that this is arbitrary, but it need not confuse the user of the bibliography or bibliographic database. This may become even more obvious when we decide to archive a copy of the database (though that is many Gigabytes for each year) and have to justify this bibliographically.

In area 3 (mathematical data) the representational scale is given, as this is the scale clients would be confronted with when they want to view the map[19]. I agree with Chris Perkins that accuracy is even more important than scale and I have given this as a specification[20].

In area 4 (publication, distribution etc.) the data concerning the producer are stated as well as the date(s) when the database is (was) in active use.

As it concerns a dynamic database, only an approximation of size is given in area 5 (physical description). In the ISBD(CF) it is usual to put this data in area 3 (type and extent of file) but for obvious reasons I have put this in area 5 and reserved area 3 for mathematical data.

Some of the more important information has been relegated to area 7 (notes). The ISBD does not compel us to structure the notes in terms of the areas of the ISBD, though for practical use this is advisable. As most of the notes refer to ISBD areas (except mapping scale, update and extent of database) the sequence here is accidental. However, it would be wise if a future ISBD could try

19 This would probably never be a single scale denominator, but give a range of scale denominators, as this is somewhat dependent on the screen used. Some might want to put the mapping or input scale here. As scale and standard deviation provides an indication as to the fitness of use of the data I think that producers will give a scale range which gives the user optimum possibility of using the particular dataset. In this I also differ from the Canadian rules (see note 9) as they prescribe 'input scale', which I relegate to the notes as 'mapping scale'.

20 "Though digital maps can be presented on any scale, in practice this is not very significant. Accuracy and level of detail determine on which 'scale' a representation has fitness of use." From: De kaart van tafel. See also note 18.

to structure these to make them more transparent. The notes should at least include a statement about the 'format of delivery', 'contents' and 'structure'.

Should the database include a specific application system then this should be described in a note which also contains information about its size (number of programme files and bytes).

Contents

If one were able to create a very practical (national) bibliography with ISBD-descriptions and with enough indeces it might even serve certain research purposes. However, in this instance one could only access the material by title, area, subject and name or a combination thereof and if finances permitted, also include scale and year of publication. This may seem quite a lot of possibilities to find the information one needs, if it were not for the fact that many modern cartographic information packages seem rather generic, e.g. aerial photographs, remote sensing images, etc. To increase retrieval possibilities several MARCs have included a 'coded information block'. The information is defined in terms of classified values at fixed positions within a specific field. The positions are filled with codes which are defined in the MARC manuals. Though this coded information is not yet used for retrieval purposes, the more than 50,000 descriptions which the Koninklijke Bibliotheek has processed in the CCK (Dutch Union Map Catalogue) have all been supplemented with all the coded information which can be provided by UNIMARC. For a description of the coded information I refer to the UNIMARC manual (see note 4). As a lot of these codes refer to the specific contents of the cartographic item, it sometimes presumes a wider knowledge of techniques than is normally required for describing these materials.

Below I shall try to point out in what way the contents of these coded fields might be used and/or amended to include digital geospatial information.

tag 120: general

The codes for index, narrative text, relief, map projection and prime meridian are all applicable.

The codes for colour should be extended, as some digital maps can be shown in black and white as well as in colour. The following codes would then be needed:

a = one colour
b = multi-colour
c = one/multi-colour, optional

The CCK has added a code for triangulation system, as many base maps use this as reference. More codes may be added for relief and projection as computer cartography creates more possibilities, but we can also use tag 131 for this kind of information.

tag 121: physical attributes

The codes for physical dimension, primary cartographic image and geodetic adjustment in subfield $a and all codes in subfield $b (aerial photograph and remote sensing) are applicable. We can add to the codes for physical medium, creation technique, form of reproduction and physical form of publication in subfield $a, so that digital information can be accommodated. However we can also create a subfield $c specifically for digital information. This should then contain coded information on the physical presentation of these files. It could contain the following elements:

(1) kind of presentation system used:
a = standard
b = commercial
u = unknown

(2) kind of programme used:
aa = Arc/info
ab = Atlas GIS
ac = Smallworld
ad = etc.
uu = unknown

tag 122: time period of item content

This field contains a formatted indication of the period [down to the hour] covered by the item. For research purposes it is a very valuable field as it provides a better specification of the possible use of the item. I believe this information to be more important than the year of publication and advise including it in the description independent of the kind of material one describes. I think it will come in especially handy with remote sensing images and aerial photographs and any material which depicts a historical situation. The manual gives as possibilities:

single date
multiple single dates
range of dates.

The CCK has altered this into:

single date
two dates
three dates
period between two dates
estimated period between two dates

Of course all these are also applicable to digital items

tag 123: scale and coordinates

In this field there are subfields for all kind of scales and coordinates. For scales there are the subfields:

a = linear scale
b = angular scale
c = other type of scale

Though it is possible to put it under 'linear scale' or 'other type of scale' I would like to add:

d = Representational scale (digital maps)

This could be complemented with "standard deviation" with the contents:

(a) = number (0 to 9)
(b) = metric unit codes
c = centimetres
i = decimeters
m = metres
d = decameters
h = hectameters
k = kilometres
x = not applicable
z = other[21]

tag 124: specific material designation analysis

"This field contains fixed length coded data relating to the characteristics of photographic, non-photographic and remote sensing image types of cartographic materials". This quote from the UNIMARC manual does not exclude digital materials and a lot of the subfields are applicable. However, I would like to see them being amended with the following subfields:

structure:	a = open vectors
	b = polygons
	c = points
	d = etc.
	u = unknown
	z = other
encoding:	a = standard 1
	b = standard 2

21 There are standard deviations, especially with large scale maps, which read like '0.2 times mapping scale'.

c = etc.
y = not encoded
z = other (company encoded)[22]

tag 131: geodetic, grid and vertical measurement

For specific collections this coded data field is created, which has not yet been implemented in the CCK. Presumably that is because some of its contents are already implemented in a more general way in tag 120. But I can imagine that some institutions with large scale base maps might want to use this for specific purposes.

tag 135: computer files (provisional)[23]

This tag is described very briefly, probably because in 1987 (the publication date of the UNIMARC manual) there was not enough practice and knowledge about the kind of coded information needed by the library and information field. In the manual the code consists only of 1 position which can be filled in as:

a = numeric
b = computer programmes
c = representational
d = text
u = unknown
v = combination
z = other

As we may describe a multi-media package which contains computer programmes, audiovisual and text data the number of positions in this subfield could be enlarged to 4 (u and z can be one position with three blanks, while v can be deleted). Should any of the coded information for digital publications not be incorporated in fields 120, 121, 123 and 124, this should be incorporated in field 135.

As the data transfer format is a specific computer- or programme-related item it might be advisable to include it in this tag as a subfield[24]. However, I am

22 Code 'a' to 'w' should be used for standards. However it is possible that many official organizations and private companies use their own encoding system, especially when it takes many years before accepted standards are in common use. Though 'z' in UNIMARC is used for unforeseen categories which do not fit the coding, here it seems wise to use this also for company encoding.

23 The Permanent UNIMARC Committee will reconsider this field with its revision of the ISBD(CF).

24 The FGDC metadata standards (see note 9) gives a list of 35 data transfer formats. Another requirement is that a number must be added which signifies the version of the format.

not sufficiently knowledgeable to give an indication how this subfield might be formatted. Some of our GIS-colleagues might be able to help here.

One could also think about the possibility of including a subfield that contains codes which stand for the minimum hardware needed to use the digital item, but here again I lack the specific expertise. I would include at least the following:

system demands:
position 1-2: platform (kind of PC/microcomputer, computer etc.)
position 3: operating system
position 4-8: internal memory: 3-7: number
8: kind (K, M, G, T etc.[25])
position 9-12: external memory: 9-11: number
12: kind (K, M, G, T etc.)
position 13-14: special additions

All other data which cannot be incorporated in the ISBD-tags (tag 200-225) can be written in tags 3xx (note block = ISBD note area[26]) or 4xx (linking entry block = ISBD note area)

UNIMARC-description

The UNIMARC-description of the "Digital topographic subject map of The Netherlands" might look something like this:

TAG	*1 2 SF*	*TEXT*[27]
001		D9400001
020	$a	NL$D9400001
100	* * $a	19940101i19919999u**a*engy01******BA
101	1 * $a	eng$cdut

25 K = kilobyte, M = megabyte, G = gigabyte, T = terabyte.

26 I think it would be advisable to create special note-tags for structure and encoding of data and system-demands and the like. In this way the note area structure would be improved. One could follow the structure given in the Canadian rules (see note 9) which state under rule 7B1 'Nature and scope and system requirements': a) nature and scope; b) System requirements; c) Mode of access; under rule 7B16 'Other formats'. But this still is not very satisfying.

27 1 = indicator 1; 2 = indicator 2; SF = subfield. * = blank. Italics in the TEXT means additions to the present Unimarc format. For explanation see 'Unimarc manual' (note 4) and the preceding chapter.

Should I have made any mistakes in coding certain kinds of information, this would be due to the fact that the CCK format is more elaborate [i.e. the structure of tags, indicators and subfields have a more logical construction and sequence in the cataloguing-module than the format the computer works with or that which is used in exchange] than the Unimarc format and thus that I have to translate my everyday practical experience into Unimarc.

102	* * $a	NL
120	* * $a	byygk**afaa***ba*[28]
121	* * $a	a*ahyyxz
122	2 * $a	d1986******$ad9999******
123	3 * $a	*D*$b00005000.000025000$de00320**$ee00715**$fn5335**$gn 5045**[29]*$o2M*
124	* * $a	abdcacai*hbciz*
131	* * $a-$l	?[30]
135	* * $a	c
200	1 * $a	[Digital topographic subject map of The Netherlands][31]$b[Cartographic Material]

28 Grid/reference system could be incorporated here (as it is in the CCK), when field 131 is not used.

29 For retrieval purposes these will not be geographical coordinates, as these have positive and negative values. Behind the screen a conversion program can work which converts the geographic coordinates of tag 206 (mathematical data) into vector coordinates. See further: E.H. van der Waal: The application of geographical co-ordinates for retrieval of maps in a computerized map-catalogue, In: International yearbook of cartography, XIV, 1974, pp. 166-173.

For the sake of clarity the subfields for coordinates have been enumerated.

30 I can imagine that this field is widely used where base maps and e.g. DEMs (Digital elevation model) are concerned, as these data are of interest for those who want to use these maps for specific purposes.

Appendix F of the UNIMARC Manual (p. 396-412) gives very extensive code-lists for spheroid ($a), horizontal datum ($b), grid and referencing systems ($c), vertical datum ($f), and unit of measuring height ($g). As far as I know even the Content standards for digital geospatial metadata (see note 10) does not incorporate such extensive listings.

For less specific purposes fields 120 and 121 can be used.

31 It is a pity that in UNIMARC subfields $c, $d cannot be exchanged for a second indicator as is done with the CCK:

0 No specific bibliographic relation with preceding subfield

1 Parallel title proper or statement of responsibility (ISBD '=')

2 Title proper by another author (ISBD ".")

3 Second or further title proper by same author (ISBD ';')

4 Second part of title when making a one-level description of a series (ISBD '.')

For further information concerning the CCK format see: G.J.K.M. van der Velden et al., CCK: making cartographic materials accessible. In: ERLC the LIBER Quarterly, Vol.2 (1992), No. 2, pp. 192-208.

As far as I have understood UNIMARC only recognises the function of the data in a given subfield, but thereby does not structure the whole tag. It is my experience that titles and statements of responsibility in complicated maps need structuring to make sense to users of the bibliographic record and to make this structure computer-processable. If not, the construction of title keys may be a hard job.

		$fTopografische Dienst Nederland
205	* * $a	Situation on 31-12-1992[32] *$r19921231*
206	* * $a	Representational scale 1:5,000 to 1:25,000, *standard deviation 1.8 m*, stereographic proj., with coordinates of the shifted Dutch triangulation system (E 3 20-E 7 15/N 53 35-N 50 45)[33]
210	* * $a	Emmen$cTopografische Dienst Nederland$d1991-...
215	* * $a	Ca. 1,980 Mb data
305	* * $a	Update: depending on area: 4, 6 or 8 years; the update programme is drawn up according to the current requirements of the Ministry of Defence
307	* * $a	Extent of database: urban area ca. 1.6 Kb per hectare, rural area ca. 0.3 Kb per hectare
310	* * $a	Deliverable on 31-12-1992: Sheet 9W, 9O, 14W, 14O, 15W, 19W, 19O, 20W (excl. Texel); delivery per sheet or part thereof
315	* * $a	Mapping scale 1:10,000
327	1 * $a	Contents: coded vector-database with all topographic point-, line and area-information, as is represented on the present maps of 1:10,000 and 1:25,000; classification of several topographic features ('layers') such as roads, water, buildings, etc.; within each theme there is subcoding; structure: all line-segments are noded and, when they have an identity of their own, fitted with their own coding; the polygons (objects) which are thus created are coded with a centroid
337	* * $a	Format of delivery: SUF2, DGN (microstation format), DXF, DWG (autocad format); classic 9 inch magnetic tape, Exabite (8 mm cassette); small files on diskette
488	* * $a	REFERENCE LITERATURE[34]: Van basisbestand naar kernbestand: de Topografische Dienst als producent van een

32 For second indicator see comment in note 31. There should be a subfield for sorting purposes, e.g.:

205 * * R 19921231

The text could either be an alphanumerical code chosen at random or a date which is encoded as in field 122.

33 It would be better for each information unit to have its own subfield so as to facilitate identification of its contents. Coordinates would then get a fixed length with precursor zeros when necessary, e.g.

206 * * $b E 003 20$cE 007 15$dN 053 35$eN 050 45

As the coordinates are in a fixed position with a fixed length it is easier to have a programme running behind the screen which converts these coordinates to vector-coordinates. Then the cataloguing agency does not have to put these data in coded form in field 123, subfield $d to $g for the second time. See further note 29.

34 This would be a computer-generated text linked to this specific field when showing an ISBD.

		kernbestand 1:10.000 / P.W. Geudeke. In: Kartografisch tijdschrift, 1993,XIX.2, pp. 24-28
675	* * $a	273$v3$zDUT
686	* * $a	<4.210>[35]$dcck
710	0 2 $a	NA098714$aNederland (state)$bTopografische Dienst Nederland$c(civil service)$4180
801	* 0 $a	NL$bCCK$c19940901$gISBD(CM)

Manual update

One of the problems in an ever changing world is that the contents of certain codes become obsolete or that codes have to be added for new developments. A standard like UNIMARC still takes too long to incorporate changes which take place in the actual world. This is mainly due to the evaluation and consultation stage being a very arduous one. I can understand that this is necessary when new fields or subfields are added as they may be hard to incorporate in existing programmes. However, when the coded contents of a position or positions have been defined it should be possible to add to the range of contents more quickly. Examples are subfields $f (name of satellite for remote sensing image) and $g (recording technique for remote sensing image) in field 124 (specific material designation analysis). Subfield $f has a two-character code where these codes signify a certain platform, e.g. for earth resource satellites:

ga	=	ERTS
gb	=	Landsat I
gc	=	Landsat II
gd	=	Landsat III
ge	=	Seasat
gf	=	Skylab
gg	=	Spacelab

But by now we have many images of Landsat IV and V (where has Landsat VI gone to?), SPOT (of which subsequent platforms will be launched) and several Russian platforms. The coding allows for 1,296 alphanumerical codes (if I am right: aa-zz, 00-99, a0-z9, 0a-9z) which should be enough for some time to come. The same is true for subfield $g, the codes for map projection (tag 120, subfield $a), physical medium (tag 121, subfield $a), and presentation techniques (tag 124, subfield c) as well as a new subfield $c in tag 121 as proposed by me. With the current and future possibilities of electronic communications it should be possible to create new codes fairly quickly. I can imagine the UBCIM office[36] to be in touch with launching bodies (like NASA, ESA and Russian or other organisations) and even without consultation to designate a specific code to a specific platform when this is functioning or to a new recording technique. The

35 CCK area code for The Netherlands

36 Office for the IFLA Universal Bibliographic Control and International MARC Programme.

information would then be made available to users through some Internet-function like the IFLA-gopher[37]. For coded information regarding specific cartographic items (like tag 120, tag 121, subfield $b, and $c of tag 124) I can imagine the UBCIM Office to be in touch with an ICA[38]-commission (or ICA-IFLA commission) responsible for defining the new techniques and applications and communicating them in the same manner. Of course the other MARC-offices could offer the same service to their clients.[39]

Conclusion

When discussing the problems of how to catalogue digital material with colleagues, so far we have not gone much further than looking at the ISBDs. In order to be able to use the descriptions for cataloguing purposes and in OPAC environments, however, we have to adapt the MARC-formats at the same time. To keep up with developments in the practical field we will have to adapt, to use the electronic communication networks in order to be able to amend part of the MARC-formats more quickly.

I would like to thank Professor Dr. Fer-Jan Ormeling jun. for checking the part of the draft concerned with digital cartography. He prompted me to rethink some of my statements[40]. I would also like to express my appreciation to Brian Holt (National Bibliographic Service, The British Library) for his comments and advice on the part concerned with the changes I propose in UNIMARC. He was so kind as to edit the UNIMARC-description to bring it in accordance with the new edition of the manual. From his comments I understand that I may also thank Jim Elliot of the same department for the thoughts he has given to the problems I have discussed. I would also like to express my thanks to Govert van

37 The IFLA-gopher is available through: gopher.konbib.nl

It contains currently information pertaining to the different bodies and the functions of IFLA and is managed by IFLA Headquarters in The Hague, The Netherlands. My proposal would add to its possible contents.

See further: IFLA Headquarters and the Internet. In: IFLA journal, Vol. 20 (1994) No. 3, pp. 369-370.

38 International Cartographic Association

39 UKMARC and USMARC already have a kind of facility like this. When there are sufficient changes new pages to the loose-leaf manual would replace the old ones with its manuscript additions. Now they only need to go electronically.

40 He posed the interesting question as to how I would describe electronic atlases like the National Atlas of Sweden. Would I mention all statistical subjects, aggregate possibilities and combinations of the two? And would I mention all analytical tools and, if necessary, GIS-functions? In the absence (yet!) of a Content standards for digital geospatial metadata I think I would try to dissect this electronic atlas as we did with the second edition of the analogue National Atlas of The Netherlands, creating multi-level records. See further: Jan Smits: Automation and multi-part description. In: ERLC the LIBER Quarterly, 2.1992.2, p. 128-136.

der Velden (former Manager CCK) for rousing my interest in the kind of problems addressed in this paper and for the way he has constantly guided me through these difficult matters. It feels good to have been working with somebody who truly knows about cartography and about the problems of automating bibliographic databases for cartographic materials. The English text has been edited by Mrs. Lysbeth Croiset van Ughelen-Brouwer.

Based on a paper read during the 9th Conference of the Groupe des Cartothécaires de LIBER in Zürich, Switzerland, 26-29 September 1994.

APPENDIX 1

Information concerning a remote-access dynamic digital geospatial database (see note 15)

Supplier:	Topografische Dienst Nederland (Topographical Survey)
Product:	Topographic subject map
Description:	Coded vector-database with all topographic point-, line- and area-information, as is represented on the present maps of 1:10,000 and 1:25,000. Classification of several themes ('levels') such as roads, water, buildings, etc., and within each theme there is subcoding.
Mapping scale:	1:10,000
Representation scale:	1:5,000 to 1:25,000
Standard deviation:	1.8 m
Structure:	All line-segments are noded and, when they have an identity of themselves, fitted with their own coding. The polygons (objects) which are thus created are coded with a centroid.
Encoding:	Company encoding

Relation RD-system[41]:	Yes
Extent of database:	- built-up area: *c.* 1.6 Kb per hectare; - rural area: *c.* 0.3 Kb per hectare; - average 6 Mb per sheet (c. 650 sheets)
Supply:	Per sheet (rectangular cut possible, price per km2)
Price-indication:	Depending on area 3 tariff-areas: price per sheet per area: 1. NLG 1,050 annually 2. NLG 700 annually 3. NLG 525 annually On average NLG 0.12 per hectare annually. One-time costs NLG 100 per sheet; parts are NLG 100 extra per sheet.
Update:	Depending on area 4, 6 or 8 years
Format of delivery:	SUF 2; DGN (microstation format); DXF; DWG (autocad format); magnetic tape (9 inch), Exabite (8 mm cassette) and small parts of sheets on diskette
Continuity:	Guaranteed
Country-wide coverage:	25%
Remarks:	The Topographical Survey strives to have 100% coverage in 1997
Address	Topografische Dienst Nederland Postbus 115, 7800 AC Emmen, The Netherlands, Tel. +31 5910 96911

41 This is the official Dutch grid and reference system.