01. CBS MARC 21 format

What is the CBS MARC 21 format?

The CBS MARC 21 format is called so, because it is a combination or synthesis of:

- MARC 21 fields
  - The MARC 21 fields of the (Library of Congress) MARC 21 format. We have implemented the bibliographic, holdings and authority format. This forms the core of the CBS MARC 21 format.
  - The additional OCLC MARC 21 fields. I.e. the fields defined by OCLC, especially for WorldCat. By OCLC several fields on the main level, but mainly on the local level have been added to the "official" MARC 21 format. This is called WorldCat MARC 21.
  - MARC 21 extensions to, i.e. these are not part of the Library of Congress and the OCLC, Dublin (WorldCat) format. These tags can be applied within the context of a CBS MARC 21 database only. "Extension" means that certain MARC 21 fields can be used on another level too. E.g. the field 500 has been defined only on main level in the "official" MARC 21 format. But in a CBS MARC 21 database, this field can also be used on copy/ LHR level.

- CBS system and non-system fields
  - CBS fields. These are non-(OCLC) MARC 21 fields. CBS fields are defined by OCLC, Leiden.

When the OCLC MARC 21 (OCLC, Dublin) format has certain limitations in comparison with the Library of Congress MARC 21 format, i.e. that certain tags, subfields or values are not allowed in the OCLC (WorldCat) MARC 21 format, the Library of Congress MARC 21 format is primarily! It implies that the CBS MARC 21 (OCLC, Leiden) format contains all MARC 21 tags in principle.

Different CBSs

CBS means Controlled Bibliographic Service.

CBS started in the eighties of the last century in the Netherlands. Later, CBSs have been implemented in e.g. Germany, France, Australia, Switzerland, United Kingdom, and Spain.

CBSs have been developed as “stand alone” regional and national systems in the past. International exchange of data and cooperation were not so important as they are now. That's why they have their own formats, like MARC 21, UNIMARC, MAB, PICA3. Although these CBSs use an (international) standard, they all have their "local flavor".

An exception is the recently developed CBS based on the CBS MARC 21 format. The starting point and purpose of this format are to be an international standard based on MARC 21 for libraries worldwide that want to use a CBS database.
Length of the field code

The tags consist of 4 characters in principle:

- The CBS MARC 21 tags in the cataloguing format, start with a capital (in principle), followed by 3 digits.
- The CBS tags in the cataloguing format consists of 4 digits (in principle).

Thus, you can see the difference between CBS MARC 21 tags and CBS tags very easily.

Levels of fields

In the CBS MARC 21 format, we distinguish three levels of the fields.

Main/master level

These CBS MARC 21 tags start with the capital letter M in the cataloguing format and the corresponding PICA+ tags begin with the digit 0. However, the use of the prefix M is optional, i.e. a library can at any moment choose to use it or not. But the syntax files need to be adapted then.

The CBS tags consist in principle of 4 digits in the cataloguing format and the corresponding PICA+ tags begin with the digit 0.

Local level

These CBS MARC 21 tags start with the capital letter L in the cataloguing format and the related PICA+ tags begin with the digit 1.

The tags, e.g. in the 59X and 69X tag ranges in the bibliographical format, are local tags.

The CBS tags consist of 4 digits in the cataloguing format and the corresponding PICA+ tags begin with the digit 1.

The local level tags, i.e. the LBD (Local Bibliographic Data) tags, contain data, which is only relevant for a specific library, i.e. the data is not universal or is not important enough or has not enough quality for putting the data on the main/master bibliographic level. The local bibliographic tags are relevant for all the copy records or Local Holdings Records (LHRs) added to a record, e.g. local subject headings or local classification codes. So hierarchically, the LBD tags are between the main/master level and the copy data/LHRs. This implies that local not only means that tags are not universal, but also that these tags are put on a local level in a record. Thus, these tags are not part of or not visible on the main level. Thus, the word local not only applies the focus of tags, but also the physical position in the record, i.e. those tags are clearly distinguished and recognizable as a group in the record.

Copy level

These CBS MARC 21 tags start with the capital letter C in the cataloguing format and the corresponding PICA+ tags begin with the digit 2.
The copy level tags\(^2\) are put in the so-called copy block(s). These records consist of tags that are only relevant for a specific copy (item) of a document. This implies that copy not only means that tags are neither universal nor local, but also that these tags are put on a copy level in a record. Thus, those tags are not part of or not visible on the main/master or local level.

The CBS tags consist of 4 digits in the cataloguing format and the corresponding PICA+ tags begin with the digit 2. The only exception is tag H\(***\) - Holding header field - in the cataloguing format for indicating the start of a copy block. E.g. tag H001 is the start of the first copy block.

In a CBS MARC 21 database, the copy blocks are added to the main/master level bibliographic record. Thus, there are no separate LHR records in a CBS MARC 21 database.

**Further details**

**Example**

CBS MARC 21 tag 500 - General Note - can be relevant on the main/master, local and copy level.

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>TAGS</th>
<th>CBS MARC 21</th>
<th>CBS MARC 21 PICA+</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main/master</td>
<td>M500</td>
<td>060A</td>
<td></td>
<td>General Note – Main/master level</td>
</tr>
<tr>
<td>Local</td>
<td>L500</td>
<td>160A</td>
<td></td>
<td>General Note – Local level</td>
</tr>
<tr>
<td>Copy</td>
<td>C500</td>
<td>260A</td>
<td></td>
<td>General Note – Copy level</td>
</tr>
</tbody>
</table>

Advantages of the above three levels structure:

- It is immediately clear to what level a tag applies to.
- Furthermore, you can see immediately which CBS MARC 21 tag is involved. The meaning of a tag is clear, because only the prefixes M, L and C differ. So, the prefix defines the level.\(^4\) We need the prefix, because the combination of the prefix and the three digits number is necessary to make a one to one relation between the CBS MARC 21 and the internal PICA+ format. A big advance of this approach is that fields like 69X - Local Subject Access Fields - are not necessary anymore!
- There is a clear distinction between the CBS MARC 21 tags and the CBS tags. The latter consist of 4 digits in principle, so again you can see immediately what kind of tag it is.
- Both the order and numbering of the CBS MARC 21 tags and the corresponding PICA+ tags are logical.
- The syntax of a tag on different levels is the same in most cases. This makes the maintenance easier and has great advantages for all kinds of topics related to configuration like validation, displays (including expansions and click effects), indexes, import and export of data, etc.

However, the fact that the subfield structure of several of the copy tags differs from that of the equivalents of the...
bibliographic format on the local and/or main/master level is not essential, but is only a fact to be aware of.

So, the structure of the format is very logical, easy to maintain, easy to extend, etc.

**Cataloging format and database/internal format PICA+ in a CBS**

In the context of a CBS, we make a distinction between software and configuration. The CBS software forms the core of a CBS around which configuration is built. Because the CBS software must be able to function well independent from the cataloguing format used by a certain library or a group of libraries, the internal or database format PICA+ has developed. So, in a CBS a distinction is made between the **cataloging format** on the one hand and the **database or internal format** on the other hand. In general, both the CBS software and configuration use the internal or database format PICA+, not the cataloguing format like CBS MARC 21. E.g., the data is stored in a CBS database in the PICA+ format, not in the cataloguing format.

Examples of cataloguing formats are MARC 21, UNIMARC, PICA3, MAB, CBS MARC 21. These are called cataloging formats, because catalogers use them to catalog, i.e. to create and update the records in their CBS.

A CBS has one cataloguing format. But data can be displayed in different formats, like the cataloging format PICA3 and the export format MARC 21, or the cataloguing format CBS MARC 21 and the export format MARC 21 for WorldCat. In the cases of export formats, the data in the database is converted to and displayed in those formats.

The internal or database format is always PICA+ and is **CBS dependent** and exists in several “dialects” only. In other words, the PICA+ related to PICA3 (Dutch CBS), UNIMARC (Abes and Bibliothèque Nationale (France)), MARC 21 for NLA (Australia), MAB (Germany) or in a CBS MARC 21 database are not the same!

So, **PICA+ is not a format as such! A PICA+ tag as such doesn’t mean anything!** When you talk about PICA+, the first question should be “Which PICA+?”. There are big differences between the several CBSs, mainly because the cataloging formats differ, the local, regional or national wishes and requirements for the specific CBS and the fact that the PICA+ is also a reflection of the state of the art of the time in which a CBS was implemented. Of course, there are similarities between the several PICA+ formats, but these are mainly related to "system" tags, i.e. tags with a close relation to the CBS software. Thus, PICA+ is the "translation" of a specific cataloguing format into a structure that can be used by the CBS configuration and software.

**Why is the internal or database format called PICA+?**

PICA is the former name of OCLC EMEA. The organization PICA started in the Netherlands at the end of the seventies.

**PICA** means literally (in English): **Project of Integrated Cataloguing Automation**.

**Characteristics of PICA+ fields in general**

Independent from the cataloguing format, the **structure** of a PICA+ is always the same. A PICA+ tag consists of three (or four) elements.

- **Position 1** defines the level: 0 = main; 1 = local; 2 = copy/LHR.
- **Position 2 and 3**: two digits defining the “range”. The word range indicates which tags are grouped together.
• Position 4: a capital letter or at sign, indicating the specific tag within a range.
• Position 5-7 (8) are optional: a tag may have an occurrence, consisting of 2 (3) digits, preceded by a slash.

1. The LBD tags, i.e. the Local Bibliographic Data, in WorldCat look like the local tags in a CBS but are not (exactly) the same.
2. The LHR tags, i.e. the Local Holdings Records, in WorldCat look like the copy tags in a CBS but are not (exactly) the same.
4. When in a web client the level of a tag is related to the part of the cataloguing template, the prefixes M, L and C are implicit.