Changes in the model for meta database 2 -> 2.1

Changes are listed in the same order as the proposal in "Remaining proposals for changes in the data model", and after each change there is a number of the corresponding proposal.

In order for a change in the data model to be fully successful within the database area, many of the changes must also be carried out in *PC-Axis format*, so that the metadata in the SQL database can be transferred to the PC Axis family and affect handling there.

All tables and columns described below are included in the universal model. If a column is described as optional, this means it does not have to be filled in with any values and can be NULL.

**The list is broken down into the following sections:**
- **A.** General
- **B.** Subject area - main table
- **C.** Contents, sub table, variable
- **D.** Value pool, value set, value, point in time, group
- **E.** Footnotes and links to further information
- **F.** Persons and organisation
- **G.** Administrative tables

**A. General**

"*NULL, prikk, missing*" (NPM) (A2)
NPM affects several parts of the data model and is described in each part:
- New code in MainTable. SpecCharExists is found in section B.
- Supplement to PresMissingLine in Contents is found in section C.
- Changes in MetaAdm are found in section G.
- Changes in SpecialCharacter are found in section G.
B. Subject area - main table

Other languages (B4)

Decision: To allow the possibility of an unlimited number of languages, MetaAdm is supplemented with information on languages and two status columns are moved from MainTable to the table of the respective language. For each extra language, except the basic language and English, tables are then created in the meta database with the language as a suffix. The new tables will have the same appearance and content as the English tables. So that different interfaces can present the name of the language clearly, the text is inserted in the Text catalogue.

Changes in the data model:
1) In the MetaAdm table, an obligatory line is inserted with the number of languages - NoOfLanguage. For each language, an obligatory line is inserted - Language1, Language2 etc.

Example of a database with the languages of Swedish (basic language), English and Spanish:

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language1</td>
<td>SVE</td>
</tr>
<tr>
<td>Language2</td>
<td>ENG</td>
</tr>
<tr>
<td>Language3</td>
<td>ESP</td>
</tr>
</tbody>
</table>

Language1 shall always be the basic language. For other languages, extra tables are created in the meta database with the suffix from MetaAdm, e.g. SubTable_ENG and SubTable_ESP

2) In the table TextCatalogue, an obligatory line is inserted for each language, with the name of the language clearly stated.

Example:

<table>
<thead>
<tr>
<th>TextTyp</th>
<th>PresText</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language1</td>
<td>Svenska</td>
</tr>
<tr>
<td>Language2</td>
<td>English</td>
</tr>
<tr>
<td>Language3</td>
<td>Espanol</td>
</tr>
</tbody>
</table>

3) MainTable.StatusEng and MainTable.PublishedEng is moved to MainTable_Eng and is placed directly after MainTable_Eng. MainTable and the suffix of the columns Eng is removed.

If a new table is set up for a new language, the status information is placed in the same way in this table.
The table MainTable for different languages will thus always contain the same columns, and the columns have the same names. The only difference is the suffix in the table name.

*PC-Axis format is affected: No*

**Changes format in SubjectCode in MainTable (B6)**

*Decision:* Code for subject area is found in several places in the meta database, but does not have the same format today. In MenuSelection.Menu and MenuSelection.Selection, varchar (20) is used, and in MainTable.SubjectCode char (2). To get the same format in all places, MainTable.Subject.Code is changed.

Limitations: For the time being, a maximum of 5 characters is used, since PC-Axis cannot handle any more.

*Changes in the data model:*
MainTable.SubjectCode is changed from char(2) to varchar(20).

*PC-Axis format is affected: No*

**New code for SpecCharExists in MainTable (A2)**

*Decision:* Sometimes a data table can contain NPM columns - empty or with codes - but these will not be used for retrievals. To be able to describe this, SpecCharExist is supplemented with a new code.

*Changes in the data model:*
1) Definition of the column SpecCharExists has been changed to:
"Inform if column for special characters exists in the data table(s), and if these shall be used for retrievals."

2) A new code E = "Yes, special character column(s) exist, but will not be used for retrievals" is added in the MainTable.SpecCharExists. The same code is used in the English data model.

3) The current codes are clarified:
Y = "Yes, special character column(s) exist and are used for retrievals"
N = "No, special character column(s) are missing"

*PC-Axis format is affected: No*
C. Contents, sub table, variable

NPM - Different symbols for presentation of data that has not been stored (C5, A2)

Decision: Today the presentation of data that has not been stored is steered by the contents in PresCellsZero. If Y, the missing data cell is presented as 0 (zero), otherwise as .. (two dots). To be able to present data cells that have not been stored in different ways for different tables and contents, and to use several different symbols, a new column is added to Contents. Later on, the new column can replace PresCellsZero, but we will keep it for now to allow for a gradual implementation.

Changes in the data model:
A new column PresMissingLine varchar(2) is added after PresCellsZero in the Contents. The column is called PresMissingLine in the English data model.

If PresCellsZero=N, the column is either empty or contains a character type that is shown in SpecialCharacter.CharacterType. If the column is empty, DefaultCodeMissingLine in MetaAdm is used. Otherwise SpecialCharacter.PresCharacter is used.

If PresCellsZero=Y the column is empty.

PC-Axis format is affected: No

D. Value pool, value set, value, point in time, group

Presentation text for value set (D7)

Decision: Add a column for presentation text in value set. In different contexts, this could instead be used for the presentation text of the variable, e.g. in selection boxes in the retrieval interface. NOTE that if you would like the presentation text of the value set to be shown in the table heading, it will also be the name of the variable in the PX file. For example, instead of having two PX files with the variable "region" that can be combined, you have one file with the variable "county", and one with the variable "municipality", which cannot be combined.

Changes in the data model:
The column PresText varchar (80) is added after the column ValueSet in the table ValueSet and ValueSet_Eng. Filling in the column is optional. (Compare with PresText for Grouping.)
The column has the same name in the English data model.

*PC-Axis format is affected: No*

**Short and long value texts (D8)**

*Decision:* To allow use of both short and long value texts, certain additions of new codes in the value pool and value set have been made to handle the short text. Space for the short text is already found in the data model, but has not been used yet. The codes that describe the short text have existed earlier, but were removed when the short text was not used.

At the same time, the short text is extended to 250 characters to allow full flexibility. Each installation can then have special limitations on length, but general software should be able to handle texts consisting of 250 characters.

*Changes in the data model:*

1) The column `ValueTextS` is changed from `varchar(35)` to `varchar(250)`

2) Two new codes are added in `ValueTextExists` in `ValuePool`:
   - S = Short value text exists
   - B = Both short and long value text exists

3) Two new codes are added in `Valuepres` in `ValuePool` and `ValueSet`:
   - A = Both code and short text is presented
   - S = Short value text is presented

Also, the description of code B is changed from "Both code and text" to "Both code and long text" and the code T is changed from "Value text is presented" to Long value text is presented".

*PC-Axis format is affected: No*
**E. Footnotes and links to further information**

**Footnote for table and value (E6)**

*Decision:* To avoid linking a value footnote on every contents column, a new linking table is added, enabling value footnotes to be directly linked to a main table.

*Changes in the data model:*
1) The table FootnoteMaintValue is added. The table contains the following columns:

- **MainTable**  varchar (20)
- **Variable**  varchar(20)
- **Value pool**  varchar(20)
- **Value code**  varchar(20)
- **FootnoteNo**  numeric(6.0)
- **Logoncode**  varchar(20)
- **Date**  smalldatetime

All columns are obligatory.

Footnotes shall have FootnoteType in the Footnote = 9 (footnote to value + main table).

2) **Contents.FootnoteValue**

The definition is changed to "State if there is a footnote linked to a certain value that belongs in the contents column (FootnoteType = 4) or main table (FootnoteType = 9)."

*PC-Axis format is affected:* No

**F. Person and organisation**

**New role (F2)**

*Decision:* In order for us to see data that has not yet been published, we are today required to be contact persons or persons responsible for updating. In order for a person to have the role of examiner of data that has not been published, a new role is created.

*Changes in the data model:*

A new code V for the role - verify data that has not been published - is added to the MainTablePerson.

*PC-Axis format is affected:* No
G. Administrative tables

NPM changes in SpecialCharacter table (A2)

Decision: A new column is added so that the retrieval programs will know how to present a data cell with a certain number of special characters - with data and special characters or only with special characters. To avoid different codes in the Swedish and English data models, the codes are adjusted in DataCellIfylld. A new code is also added.

Changes in the data model:
1) The obligatory column DataCellPres char(1) is added after Aggregable in the table SpecialCharacter. The column has the same name in the English data model.

Y = data cell is presented together with the special character.
N = Only the special character is presented.
It is obligatory to fill in the column.

2) DataCellIfylld is used to check indata and to inform how a data cell with a certain special character is filled in.

The codes below replace the present codes in DataCellIfylld:
V = Value must be given.
N = No, the data cell should not be filled in but should be NULL.
V = Any, i.e. the data cell can be filled in or can be NULL.
0 = The data cell is only allowed to contain 0 (zero).
Codes are the same in English and Swedish.

PC-Axis format is affected: No
Descriptions and names of properties in English in the table MetaAdm (G1)

Decision: MetaAdm is supplemented with a column that describes the property. To simplify retrieval programs, all properties will in the future only have an English name and not both a Swedish and an English one. The exception to this rule applies to already existing properties.

Changes in the data model:
The column Description varchar(200) is added after Value in MetaAdm. Filling in the column is optional.

Existing properties:
1) SenasteFotnotNr
Describes the highest footnote number that has been given a footnote in the meta database. Since we assess that this property is only used by Statistics Sweden, we have kept the Swedish name and there is no English name.

2) MenuLevels
Gives the number at the lowest level in the hierarchy (= MainTable) in MenuSelection. Both the Swedish name and the English name will be used.

PC-Axis format is affected: No
NPM changes in the table MetaAdm (A2)

*Decision:* A number of new properties are added to MetaAdm.

*New properties:*

NOTE: Names of properties are only in English.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Obligatory?</th>
<th>Affects PX format</th>
</tr>
</thead>
<tbody>
<tr>
<td>DataNotAvailable</td>
<td>The value that is presented if the data cell contains NULL and there is no NPM character. If the value exists in the SpecialCharacter table, this will be used. Otherwise the character is used in DataNotAvailable. For example .. (two dots)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>DataNoteSum</td>
<td>This value is presented after the sum if data cells with different NPM markings are summed up. For example * 1A + 2B = 3*</td>
<td>Yes</td>
<td>Yes, but not obligatory</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>New key word = DataNoteSum</td>
</tr>
<tr>
<td>DataSymbolSum</td>
<td>This value is presented after the sum if data cells with different NPM markings are summed up and no sum can be created. For example N.A. . + .. = N.A.</td>
<td>Yes</td>
<td>Yes, but not obligatory</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>New key word = DataSymbolSum</td>
</tr>
<tr>
<td>DataSymbolNil</td>
<td>This value is presented for absolute 0 (zero) This is needed if there are several different values for 0 (zero in the SpecialCharacter table. For example: -</td>
<td>Yes</td>
<td>Yes, but not obligatory</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>New key word = DataSymbolNil</td>
</tr>
<tr>
<td>PxdataFormat</td>
<td>Matrix = all retrievals are stored in matrix format Keysnn = retrievals with keys are made again nn &gt; locked data cells *100 / maximum number of data cells in the complete matrix For example 40 (Matrix is default)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td>KeysUpperlimit</td>
<td>Maximum number of data cells that the complete matrix is allowed to have if the retrieval is able to be done with Keys If larger, the retrieval is done in matrix format. For example 1000000</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>DefaultCodeMissingLine</td>
<td>The value that is presented in data cells that are not stored. Is used if neither presentation with 0 nor special character is specified.</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

*PC-Axis format is affected:* See the table above
Compilation of properties in MetaAdm in the data model version 2.1
The following properties shall/can be found in MetaAdm in data model 2.1:

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Obligatory?</th>
</tr>
</thead>
<tbody>
<tr>
<td>SenasteFotnotNr</td>
<td>Describes the highest footnote number that has been given a footnote in the meta database. (Used by Statistics Sweden only) For example 5768</td>
<td>No</td>
</tr>
<tr>
<td>MenuLevels</td>
<td>Gives the number at the lowest level in the hierarchy (= MainTable) in MenuSelection. For example 5</td>
<td>Yes</td>
</tr>
<tr>
<td>NoOfLanguage</td>
<td>Number of languages For example 3</td>
<td>Yes</td>
</tr>
<tr>
<td>Language1</td>
<td>Basic language For example SVE</td>
<td>Yes</td>
</tr>
<tr>
<td>Language2</td>
<td>Short text that is used as a suffix in the tables in the meta database. For example ENG SubTable_Eng</td>
<td>Yes, for each extra language in NoOfLanguage.</td>
</tr>
<tr>
<td>Codepage1 (one for each language)</td>
<td>Describes the characters that can be used and how they are presented. Is used for creating the key word Codepage in the px file. Is there used at converting to XML. Three examples: iso-8859-1, windows-1251, big5.</td>
<td>Yes</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td>Example</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>DataNotAvailable</td>
<td>The value that is presented if the data cell contains NULL and there is no NPM character. If the value exists in the SpecialCharacter table, this will be used. Otherwise the character is used in DataNotAvailable. For example .. (two dots)</td>
<td></td>
</tr>
<tr>
<td>DataNoteSum</td>
<td>The value that is presented after the sum if data cells with different NPM markings are summed up. For example * 1A + 2B = 3*</td>
<td></td>
</tr>
<tr>
<td>DataSymbolSum</td>
<td>The value that is presented if data cells with different NPM markings are summed up and no sum can be created. For example N.A. . + .. = N.A.</td>
<td></td>
</tr>
<tr>
<td>DataSymbolNil</td>
<td>This value is presented for absolute 0 (zero) This is needed if there are several different values for 0 (zero in the SpecialCharacter table. For example -</td>
<td></td>
</tr>
</tbody>
</table>
| PxdataFormat          | Matrix = all retrievals are stored in matrix format  
Keysnn = retrievals with keys are made again  
nn > locked data cells *100 / maximum number of data cells in the complete matrix  
For example 40 (Matrix is default)                                                      |         | Yes      |
### KeysUpperlimit
- **Definition**: Maximum number of data cells that the complete matrix is allowed to have if the retrieval is able to be done with Keys. If larger, the retrieval is done in matrix format.
- **Example**: For example 1000000
- **Requirement**: Yes

### DefaultCodeMissingLine
- **Definition**: The value that should be presented in data cells that are not stored. Is used if neither presentation with 0 nor special character is specified.
- **Requirement**: Yes