MapMate Working Note

The Data Model

Abstract:

This note provides a description of database tables and fields used by the program. It will be of use as a reference for those wishing to write their own queries in SQL and to provide a greater understanding of how your data is held and used within MapMate.

Issue 2 – 25/01/2002

1. INTRODUCTION

This document defines the 'Data Model', or database structures, used by MapMate from release 0.9.0. It is intended to give users an understanding of the way their data is stored, how it is accessed and to provide a quick reference for those writing their own queries in SQL.

1.1 Issue Status

The original issue of this document is dated 24/11/2001. It was written for users with MapMate Versions from 0.9.0 inclusive. Issue 2, dated 25/01/2002 contains minor corrections and typo revisions.

2. DEFINITIONS

2.1 The User Database

The User.mdb database is used by MapMate to access all data. 'User' contains 'links' to data tables, constant resources used by the program and all query definitions. For a default installation, this is located at c:\Program Files\MapMate\Users\User.mdb. All querying is done via the User.mdb.

2.2 Database Types and Versions

MapMate Databases are created and maintained with Microsoft 'Jet' Version 3.5 – which is the Database Engine behind Access, Office and many other Applications. You can 'open' any of the .mdb files with Microsoft Access but you should NOT change the structures to any other version. If you wish to explore these with say Access 2000 or later then do the following:
Copy the User.mdb to User2000.mdb in the same folder.
Open this new copy and choose to 'convert' this. As all other tables are connected to User - data can be accessed without changing any other structures.

2.3 Field Data Types

The basic field data types used in MapMate are restricted to the following:

Text – a text field, the maximum number of characters is given under 'size' in the descriptions
Int – a number in short integer format (approx. +/-32,000)
Long – a number in long integer format (approx. +/-2E+9)
Memo – an almost unlimited free text field (1GB max)
Y/N – a binary state field, True (or Yes) = -1, False (or No) = 0

2.4 Field Naming Conventions

All field names commencing underscore (_<field name>) indicate a 'system field' who's value is maintained by the program. These fields are used by the program but may be queried by the user. Users should not modify the contents of these fields!

All field names commencing star (*<field name>) indicates a pointer to a _guk in another table. E.g. *Site points to the _guk of a Site in the Sites table.

All field names commencing star star (**<field name>) indicate a pointer to the _guk in the current table. There is only one exception to this – the **Parent in Sites – which was changed in the 0.9.0 model to point to UKSiteParents where previously it pointed to a 'parent site' in Sites. This name was left unchanged in the data model revisions for backward compatibility.

All other fields are named to reflect their content. Note that all names are NOT unique across all Tables.

Note also that any field name containing non-alphabetic characters (including leading numbers, internal spaces or other special characters etc) should be enclosed in square braces [-] when used in any query statement. E.g. the *Site field would always be referred to as [*Site] when querying.

2.5 Replication

MapMate uses a 'native' Replication scheme written especially for the program – it does not implement the standard Microsoft Database Replication Schema.

Two special fields are present in most tables and are used for Replication and record management. These are:
_**guk** – Globally Unique Key. A field found in virtually all tables being the 'primary' key to a record in that table. Used by Replication to uniquely identify all records when shared.
_**gen** – Generation. A field used by Replication marking the position of a record in the lifetime of the dataset.

When a record is created, the [_guk] is constructed from a 'Locally Unique Key' (which is guaranteed unique on the user's machine) and this is combined with the Centre Unique Key (CUK) to make a key that will be unique on any other user's machine. This allows records to be shared between users and still uniquely identified. This scheme relies on ALL users having installations with a unique registration.

Further, when a recorded is created it is given the current Generation number (into _gen). When a record is edited it is set to minus the current Generation. _gen is used by the program to identify which records need sending to other users. Your 'Generation' increments with each synchronisation.

3. DATA TABLES

3.1 Methods

Contains Method definitions. Located in Reference.mdb or in any Dataset.

field name	type	size	comments
_guk	Text	8	GUK to this Method
Method	Text	255	Method name
Explanation	Memo	-	Short explanation of this method
**Parent	Text	8	UNUSED at 0.9.x
_gen	Long	-	Generation

Notes: Although the Methods table contains some constant and shared definitions (at installation) it can be extend with any user definitions. The **Parent field was used (prior to 0.9.0) to classify methods but is now unused.

3.2 Recorders

Contains Recorder (and Determiner) definitions. Located in Reference.mdb or in any Dataset.

field name	type	size	comments
_guk	Text	8	GUK to this Recorder
Name	Text	64	Recorder (or Determiner) name(s)
*Location	Text	8	GUK to a Site (locating this Recorder) (in Sites)
Comment	Memo	-	Additional notes about this Recorder
_gen	Long	-	Generation

Notes: The Comment field should not contain any personal information about any person or persons. Recorders and Determiners (or Identifiers) are treated equally within this table. The *Location is

intended to provide a further qualification of Name to help 'locate' recorders with the same or similar names and points to a [_guk] in the Sites table.

3.3 Records

Contains your 'Records' - which are mainly pointers to other table elements. Located in Records.mdb or any Dataset.

field name	type	size	comments
_guk	Text	10	GUK to this Record
*Taxon	Text	8	GUK to a Taxon (in Taxa or [Taxa\Default])
*Sex	Text	1	GUK to a Sex (in TaxonSex)
*Stage	Text	2	GUK to a Stage (in TaxonStage)
*Status	Text	1	GUK to a Status (in RecordStatus)
*Site	Text	8	GUK to a Site (in Sites or [Sites\Default])
Date	Long	-	Record Date - in +/-days from 1=31/12/1899
DateTo	Long	-	End of Date Range (if applicable or =Date)
Quantity	Long	-	Quantity Recorded
*Method	Text	8	GUK to a Method (in Methods)
*Recorder	Text	8	GUK to a Recorder (in Recorders)
*Identifier	Text	8	GUK to a Determiner (in Recorders)
Comment	Memo	-	Any additional information
*Reference	Text	8	GUK to a Reference (in References)
_gen	Long	-	Generation

Notes: Date and DateTo are used to define arbitrary date ranges. If Date = DateTo then record is for a single day, otherwise these define the inclusive extents of a range of dates. Dates are numeric values and comply with standard Microsoft formats. Apart from Date, DateTo, Quantity and Comment, all fields are pointers to GUK's in related tables. See also section 6 for more info on [Taxa\Default] and [Site\Default] definitions.

3.4 References

Contains Reference definitions. Located in Reference.mdb or any Dataset.

field name	type	size	comments
_guk	Text	8	GUK to this Reference
Author	Text	255	Name of Author
Year	Int	-	Year of publication
Title	Text	255	Full title of this reference
PageRef	Text	32	Page number(s) (if applicable)
**SeeAlso	Text	8	UNUSED from 0.9.0
Location	Text	255	UNUSED from 0.9.0
Comment	Memo	-	Any additional information on this item
_gen	Long	-	Generation

Notes: **SeeAlso and Location are not used.

3.5 Sites

Contains Site definitions. Located in Sites.mdb or any Dataset.

field name	type	size	comments
_guk	Text	8	GUK to this Site
Name	Text	64	Site Name
**Parent	Text	8	GUK to Admin Area (in UKSiteParents)
ViceCounty	Long	-	Vice County Number (also GUK to Vice-County)
OSGridRef	Text	10	Gridreference locating this Site
Xpos	Long	-	X position in m from NG Origin (0,0)
Ypos	Long	-	Y position in m from NG Origin (0,0)
_precision	Long	-	Precision of site location (in m)
10kSquare	Text	4	Gridreference of 10km Square
2kSquare	Text	5	Gridreference of 2km Square or Tetrad
_xo	Long	-	E position of SW Corner in m from NG Origin (0,0)
_yo	Long	-	N position of SW Corner in m from NG Origin (0,0)
_xi	Long	-	E position of NE Corner in m from NG Origin (0,0)
_yi	Long	-	N position of NE Corner in m from NG Origin (0,0)
*Habitat	Text	8	GUK to a Habitat (in Habitats)
Description	Memo	-	A description of this site
_gen	Long	-	Generation

Notes: Xpos, Ypos, xo..yi and _precision are used internally by the program for locating and mapping sites - but may be called by the user in SQL. xo..yi define the smallest enclosing rectangle about the site (typically the perimeter of the Grid square locating the site). Note that [**Parent] points to UKSiteParents and NOT another site in this table - this was changed at 0.9.0 and is the only naming convention violation. When the data model was updated at 0.9.0 this name was retained for backward compatibility. Gridrefs use National Grid Format or Irish National Grid Format.

Note: The OSGridRef size was subsequently patched to '32' in order to support 10 figure grid references (i.e. SY8551283167). The latest version of the installer includes this patch.

3.6 [Taxon-Validations]

Contains user definable validation rules for records. Located in Reference.mdb or any Dataset.

field name	type	size	comments
_guk	Text	8	GUK to this Taxon Validation
*Taxon	Text	8	GUK to a Taxon (in Taxa or [Taxa\Default])
*Stage	Text	2	GUK to a Stage (in TaxonStage)
Date	Long	-	Date (in +/-days from 1=31/12/1899)
DateTo	Long	-	Extent of Date Range (if applicable)
*VC	Long	-	GUK to a VC (in Vice-County)
*PlacedBy	Text	8	GUK to a Recorder (in Recorders)
Warning	Text	255	Warning message for this validation rule
Comment	Text	255	Additional comments if applicable
_gen	Long	-	Generation

Notes: Date and DateTo are the date on which the rule was applied. If a [*Taxon] is found in [*Stage] and in [*VC] then 'Warning' message is shown when a user tries to save this record.

4. CONSTANT LOOKUP TABLES

The following tables are used to provide constant lookup definitions. These definitions are controlled centrally, shared by all users and are resources used by the program and your data tables.

4.1 Habitats

Contains Habitat classifications. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Habitat
Habitat	Text	64	Habitat name
HabitatCode	Text	16	Habitat classification key
NBNKey	Text	16	NBN Code number for this Habitat
_gen	Long	-	Generation

4.2 RecordStatus

Defines standard terms used for a Record 'Status'. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	1	GUK to this Record Status
Status	Text	64	Status term
Description	Text	255	Short description of this term
Breeding	Y/N	-	Yes/No state indicating if 'Breeding'
UsedBy	Text	255	List of Taxa groups that use this term
_gen	Long	-	Generation

Notes: The 'Breeding' flag (Yes=-1/No=0) indicates if this status term is indicative of a breeding state. The UsedBy field defines Taxa groups, listed and separated by a comma (,) that use a particular term. This is to allow sub-sets of Status terms to be pre-selected according to a recorder's interests. The default is 'All' - indicating that the term is always used.

4.3 Taxa

Contains Taxa definitions. Located in Taxa.mdb or any Dataset. Configured by users via 'My Configuration'. Managed by the system. The Taxa table maintains a 'short list' of taxa from the Taxa Library (TaxaLib) according to current user's interests.

field name	type	size	comments
_guk	Text	8	GUK to this Taxon
Taxon	Text	64	Scientific Name of Taxon
*TaxonType	Text	8	GUK to a Type (in TaxonType)
**Parent	Text	8	GUK to a Taxon in this table
Vernacular	Text	64	Vernacular or Common name of this Taxon
Authority	Text	64	Authority and Date for this definition
*Status	Text	8	GUK to a Status (in TaxonStatus)
Comment	Text	255	Any additional info on this Taxon
*Reference	Text	8	GUK to a Reference (in TaxonSources)
Code	Text	64	Most popular coding system for this Taxon
**RefersTo	Text	8	GUK to a Taxon in this table
_specificindex	Text	64	Index to specific names
_gen	Long	-	Generation

Notes: [**Parent] is a pointer to the next taxon up the hierarchy (e.g. for a species this might be the 'Family' name). The [*Reference] points to a [TaxonSource] and not a Reference. [**RefersTo] points to

the currently accepted definition for this taxon if a synonym or itself if not. I.e. if [_guk]=[**RefersTo] then this is the current name, or if [_guk]<>[**RefersTo] then [**RefersTo] points to the current name.

[_specificindex] is a fast index, used by the program, to locate binomials by their specific name only.

4.4 TaxaLib

Contains ALL taxonomic names. Located in TaxaLibrary.mdb. Managed by the system. Note that this is a constant reference table that may be queried by the user. Working taxa definitions are always taken from Taxa or [Taxa\Default].

field name	type	size	comments
_guk	Text	8	GUK to this Taxon
Taxon	Text	64	Scientific Name of Taxon
*TaxonType	Text	8	GUK to a Type (in TaxonType)
**Parent	Text	8	GUK to a Taxon in this table
Vernacular	Text	64	Vernacular or Common name of this Taxon
Authority	Text	64	Authority and Date for this definition
*Status	Text	8	GUK to a Status (in TaxonStatus)
Comment	Text	255	Any additional info on this Taxon
*Reference	Text	8	GUK to a Reference (in TaxonSources)
Code	Text	64	Most popular coding system for this taxon
**RefersTo	Text	8	GUK to a Taxon in this table
_specificindex	Text	64	Index to specific names
_gen	Long	-	Generation

Notes: See Notes for Taxa Table above. This table is called 'Taxa' within the Taxa Library.

4.5 TaxonSex

Defines basic sex terms used by the program. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	1	GUK to this Sex term
Sex	Text	64	Sex term
Description	Text	255	Description of this term
UsedBy	Text	255	List of Taxa groups that use this term
_gen	Long	-	Generation

Notes: UsedBy lists high level taxa groups, separated by commas, using this term. The default value is All' indicating 'used by all groups'.

4.6 TaxonSources

Defines the References for Taxonomic definitions. Located in TaxaLibrary.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Reference
Author	Text	255	Name of Author
Year	3	2	Year of Publication
Title	Text	255	Full title of this Reference
PageRef	Text	32	Page number(s) (if applicable)
**SeeAlso	Text	8	UNUSED from 0.9.0
Location	Text	255	UNUSED from 0.9.0
Comment	Memo	-	Any additional information on this item

field name	type	size	comments
_gen	Long	-	Generation

Notes: This table is centrally controlled alongside the taxa definitions and is identical in structure to the References table - see notes for that.

4.7 TaxonStage

Defines Stage terms used in Taxa and TaxaLib. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	1	GUK to this Stage term
Stage	Text	64	Stage Term
Description	Text	255	Description of this term
Mature	Y/N	1	Yes/No flag indicating a 'mature' stage
UsedBy	Text	255	List of Taxa groups that use this term
_gen	Long	-	Generation

Notes: UsedBy lists high level taxa groups separated by commas using this term. The default value is All' indicating 'used by all groups'.

4.8 TaxonStatus

Defines Status terms used by Taxa. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Status term
Status	Text	64	Status Term
Description	Text	255	Description of this term
_gen	Long	-	Generation

4.9 TaxonType

Defines taxon classification 'types'. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Taxon Type
Classification	Text	64	Classification
Level	Int	-	Hierarchy (7=highest..0=lowest)
_gen	Long	-	Generation

Notes: Level is a numeric value used by the program to order classification levels, 7=Domain .. 0=Specific or Sub-specific.

4.10 UKSiteParents

Defines Admin Area names. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Area
Name	Text	64	Name of Admin Area
Type	Text	64	Type of Area
_gen	Long	-	Generation

Notes: UKSiteParents are used by Site.[**Parents] to assign an Admin Area to a site definition. Type is a literal classification for this area with applicable dates.

4.11 [Vice-County]

Contains Vice County definitions. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Long	-	GUK to this Vice County
Name	Text	64	Vice County name
VC	Text	3	Vice County number
_gen	Long	-	Generation

Notes: Irish VC's are re-numbered here from 201 - so H1:=201 etc. The [_guk] to this table is an exception in that it is a numeric value and not a string. The main use for this is for linking to Sites.[ViceCounty] which is the VC number. In this table [_guk] always equals [VC]. In most queries the VC is used directly from Sites.[ViceCounty] – you can, however, join in this table to get VC literal values.

5. ADDITIONAL TABLES

5.1 [_ver]

Contains User database version info. Located in User but also present in other databases. Managed by the system.

field name	type	size	comments
_issue	Long	-	issue number
_notes	Text	255	any notes on this issue

Notes: This [_ver] structure is also used in other tables to list version history. Only the User.mdb [_ver] info however is visible without directly accessing another database. You can report say the TaxaLib ver using:
```
SELECT *
FROM [_ver] in 'c:\Program Files\MapMate\Taxa\TaxaLibrary.mdb'
```

Some Databases (if inherited from older program versions) may not have this version info.

This table is not replicable and exists in User.mdb (see also notes above) – it should reflect the issue at Installation or last Update.

5.2 Filters

Contains definitions for recording defaults and replica filters. Both Taxa and Site defaults are defined here. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to this Filter
_filterClass	Text	16	Filter class: value is currently either 'Taxa' or 'Sites'
_filterName	Text	50	Filter name (as it appears in default lists)
_filterSql	Memo	-	SQL for this filter definition
_gen	Long	-	Generation

Notes: This table contains all of the available default prototypes. They are currently classified (by _filterClsss) as either 'Taxa' or 'Sites'. When a filter or default is selected the [_filterSql] is loaded into

[Sites\Default] or [Taxa\Default] at run time. This table, although replicable, has [_gen] set to 2 for user's current preferences via 'My Configuration'. See also Section 6. Run-time Tables.

5.3 UKMatrix

Contains a list of all UK 10km squares with details of their Admin Area and VC number. Located in User.mdb. Managed by the system.

field name	type	size	comments
_guk	Text	8	GUK to a [UK Matrix] element
OS10k	Text	4	10km Square
County	Text	20	Admin Area
VC	Long	-	Vice County number
_gen	Long	-	Generation

Notes: Where an OS10k square has parts in multiple Admin Areas or VC's then a record exists for each combination. This table is used internally by the program for quick determination of locations. This table could be used however for functions like:
To list all OS10km squares for a VC:

```
SELECT Distinct OS10k
FROM UKMatrix
WHERE VC = [Enter a VC number]
```


5.4 Other System Tables

A number of other tables exist in User. These contain system or program related information and are not documented here.

6. 'RUN-TIME' QUERY BASED TABLES

The program uses two important query based tables. These are [Sites\Default] and [Taxa\Default] and are used to return site and species definitions under the user's currently selected recording defaults. Although these are not strictly database 'tables' they are used as such and form an integral part of the working data model.

6.1 [Sites\Default]
This returns an identical set of fields as found in the Sites table and is defined as:

```
SELECT Sites.*
FROM Sites
WHERE <current default site restriction>
```


6.2 [Taxa\Default]
This returns an identical set of fields as found in the Taxa table and is defined as:

```
SELECT Taxa.*
FROM Taxa
WHERE <current default taxa restriction>
```

In both cases, the <current default -- restriction> defines a subset of sites or taxa according to rules obtained from the [_filterSql] definition in the Filters table for a particular [_filterName] definition. When a user (or the program) selects a particular default, the [_filterSql] is loaded into the respective […\Default] query automatically.

The default 'All' settings are defined as

```
SELECT Sites.*
FROM Sites
```

and

```
SELECT Taxa.*
FROM Taxa
```

Note that these have no restrictions and that the Taxa default is always based on the Taxa table and not from the Taxa Library (TaxaLib) so it can only ever include items in 'My Checklist'. Equally, they could use TaxaLib.

See Filters table for more details of default definitions and their use.

6.3 Run-time Data Model

Data tables are all 'attached' to User at run-time. As the program starts, it selects your default dataset 'Records' for use by making the following tables available: Records, Sites, Methods, Recorders, References, Taxon-Validations, Habitats and Taxa. Equally, as any 'Dataset' is opened, any tables this contains are selected into the program from its .mdb file. End users should be aware of this 'dynamic' nature of their User database. When the program is quit, User is left in whatever state of configuration was in place at the point of termination.

7. PHYSICAL MODEL

7.1 File locations

For a normal (default) installation all files are contained within a MapMate folder located at:

c:\Program Files\MapMate

Folders under this are structured as follows (folder names in **bold** contain your databases):

\Cache – used for temporary files in Replication (also has Up and Download sub folders)
\Maps – used for mapping and map reference data
\Output – default folder for program output
\Records – contains Records.mdb, Dataset.mdb (dataset prototype) and <any dataset>.mdb
\Reference – contains Reference.mdb
\Sites – contains Sites.mdb
\Taxa – contains Taxa.mdb (your checklist) and TaxaLibrary.mdb
\Users – contains User.mdb (program interface and resource database)