Generic CRVS Information Modelling Guide

**Contents**

1 CRVS Information Model 1

1.1 Definitions: 1

1.2 Principle CRVS Data Entities 2

1.3 Principle CRVS Data Elements 3

1.4 Entity Relationship Key and Modelling Concepts 5

# CRVS Information Model

Information modelling describes the process of analysing the data that needs to be stored by the system and the way it needs to be structured in order to provide for the immediate data processing needs of the system as well as any possible future extensions to the system. The resulting information model is a fundamental component of any digital CRVS system.

## Definitions:

**Entities** are objects or concepts that represent important data. They are typically nouns, e.g. birth event, registration event, person, mother, civil registration office, etc.

**Attributes are** characteristics of an entity; the separate data elements associated with a data entity.

**Data definitions** define the data elements themselves and the type of the data that is stored.

An information model can be defined and then elaborated at different levels:

1. **A conceptual data model** isthe simplest view and can be used as the foundation for logical data models. It :
* Includes important entities and the relationships between them.
* Does not specify attributes.
* Does not specify primary keys.
1. **A** **logical data model** defines the logical structure of the data. This model contains more detail than the conceptual data model, without regard to how information will be physically implemented in the database. Logical data models:
* Include all entities and relationships between them.
* Specify attributes for each entity.
* Specify primary key for each entity.
* Specify foreign keys, which identify the relationship between different entities.
* Involve normalization, which is the process of removing redundancy in a table so that the table is easier to modify. Normalization typically occurs by dividing an entity table into two or more tables and defining relationships between the tables.
1. **A physical data model** describes how the logical model is translated into a model that can be implemented in an electronic database management system (DBMS). It represents the process of adding information to the database. This model shows all table structures, including column name, column data type, column constraints, primary key, foreign key, and relationships between tables. The structuring of the data in the physical data model will determine the flexibility of the information model and how it can be used to satisfy data processing and reporting requirements as well as future extension, if required. Physical data models:
* Specify all tables and columns.
* Include foreign keys to identify relationships between tables.
* May include de-normalization, depending on user requirements.
* May be significantly different from the logical data model.
* Will differ depending on which DBMS (database management system) is used.

A commonly-used database management system is the relational database management system (RDBMS) that is usually modelled, at a logical level, by means of an entity-relationship model.

## Principle CRVS Data Entities

The main Vital Event and Person entities defined in the Principles and Recommendations (UNSD, 2014) are detailed in the table below.

|  |  |
| --- | --- |
| **#** | **Vital Events** |
| **1** | **Live birth** |
| 1,1 | Characteristics of the event |
| 1,2 | Characteristics of the newborn |
| 1,3 | Characteristics of the mother |
| 1,4 | Characteristics of the father |
| **2** | Death |
| 2,1 | Characteristics of the event |
| 2,2 | Characteristics of the decedent |
| **3** | Foetal death |
| 3,1 | Characteristics of the event |
| 3,2 | Characteristics of the foetus |
| 3,3 | Characteristics of the mother |
| 3,4 | Characteristics of the father |
| **4** | Marriage |
| 4,1 | Characteristics of the event |
| 4,2 | Characteristics of bride and groom (separately) |
| **5** | Divorce |
| 5,1 | Characteristics of the event |
| 5,2 | Characteristics of divorcees (husband and wife separately) |
| 5,3 | Characteristics of population at risk |

*Main Entities defined in the Principles and Recommendations (UNSD, 2014)*

## Principle CRVS Data Elements

The first order CRVS data elements as defined in the Principles and Recommendations (UNSD, 2014) are shown in Figure One, below. The model was developed in MySQL Workbench[[1]](#footnote-1), a free data modelling software tool that runs on several different operating systems. Default data types have been used for most of the data elements. This diagram is an example that shows how the notation can be used to represent the data entities and attributes, but is NOT a comprehensive and complete Entity Relationship Diagram for CRVS.



 *First Order CRVS Data Elements defined in the Principles and Recommendations (UNSD, 2014)*

## Entity Relationship Key and Modelling Concepts

Entity relationship modelling is used to model the main entities in a system as well as the main relationships between them. The notation depicts entities as tables, data elements as fields within the Tables and relationships as lines between the tables with the ends of the lines depicting the type of relationship. A simple entity relationship diagram for first order CRVS data elements is shown in Figure Two, below. The concepts and symbols are defined in Figure Three.

The relationships between the entities are shown by the connecting lines. The Cardinality and Modality are the indicators of the business rules around the relationship.

Cardinality refers to the maximum number of times an instance in one entity can be associated with instances in the related entity. Cardinality can be 1 or Many and the symbol is placed on the outside ends of the relationship line, closest to the entity. For a cardinality of 1 a straight line is drawn. For a cardinality of Many a foot with three toes is drawn.

Modality refers to the minimum number of times an instance in one entity can be associated with an instance in the related entity. Modality can be 1 or 0 and the symbol is placed on the inside, next to the cardinality symbol. For a modality of 1 a straight line is drawn. For a modality of 0 a circle is drawn.

For example, a birth event may have zero to many newborns associated to it (the baby may be stillborn, it may be a multiple birth) but a death event will only ever be associated with one decedent.



 *Normalised CRVS Entity-Relationship Model*

This illustrative example diagram has the following features:

* The model has four entities, represented by Tables in the diagram: Vital Event, Vital Event Type, Person and Role.
* Each Table has two fields (columns), an ID field and a description field.
* The Vital Event table has three additional fields that correspond to the ID field in each of the other three tables. This means that each Vital Event record has a unique person, role in the vital event (eg newborn) and the event type (eg birth event) associated with it.

This is a very simple model. In practice, the tables in a working DBMS will have many more Tables and fields within the Table as well as relationships linking the tables. The purpose is to implement best practice for database management, such as keeping only one record of a data element and normalising the relationships between tables.

Some of the ways to structure information from the Principles and recommendations (UNSD, 2014) are detailed in the document's Appendices.

Note: The Information Model will help to define whether the CRVS system is considered "vital event” or “person”-centric, depending on the organisational priorities; for example, CRVS integrated within Population Registries will tend to be more person-centric.

1. <https://www.mysql.com/products/workbench/> [↑](#footnote-ref-1)