The DTM Data Dictionary: An Overview for DTM Partners and Data Users

Background

The DTM Data Dictionary is a **centralized repository of information about data fields** used in DTM data collection tools. This repository contains a record of the indicators used in different DTM assessments worldwide as well as essential information about these indicators. This includes information such as the description, type, source and method of data collection for each data field. The key functions of the data dictionary are to ensure a standardized approach to all DTM assessments, wherever and whenever they may be executed, to reduce the time taken to establish new assessments in different contexts, and to ensure the generation of useful analysis for partners. In line with this, the Data Dictionary is used when planning and setting up DTM operations to help define the questions to be included in the DTM data collection tools and to ensure the use of indicators which facilitate compatibility across data sets from different operations. Further to operational objectives, the data dictionary also serves as a reference tool that can be used to support research partnerships with humanitarian, development, academic and private sector institutions by serving as a record of the existing data fields available for analysis or further research.

1. The Role of the Data Dictionary

The data dictionary is primarily a **data management tool**, with auxiliary functions related to analysis and use of data. It assists management, database developers and administrators, system analysts, and application programmers in planning, controlling, developing, sharing and evaluating the collection, storage and use of data. Regarding the planning and execution of data collection and storage, the data dictionary plays a role in the development of new tools, the adaptation of existing tools and in database design. It also provides DTM staff with a repository of questions that have been used in different contexts and that are recommended by DTM data users and partners.

Global clusters /AoRs and WGs are regularly consulted for inputs on data fields. From an analytic perspective, using the Data Dictionary to inform tools and database structures at country level allows for the development of systems that are not only standardized for quality purposes but are also in harmony with regional and global structures. This allows for better data analysis which can compare data across several countries or regions and allows for feedback from stakeholders and data users at all levels to be used to improve the system with global impact.

The data dictionary can also be used to regulate data sharing and to inform data governance as it defines which fields can be considered as **public** and shared with a wide range of users and which fields should be considered **restricted,** requiring defined standards of practice or data sharing agreements for access to be granted.

2. Using the Data Dictionary in DTM Operations

2.1 General Uses of the Data Dictionary

The data dictionary is used when setting up and supporting operations, guiding analysis and informing the storage, sharing and use of DTM data. The primary uses of the tool are outlined below:

a) Identify data fields for DTM assessments

The Data Dictionary must be used when identifying appropriate data fields (questions to be asked in DTM assessments). It provides all the data fields that must be included for each type of DTM assessment.

b) Developing DTM forms

The contents of the data dictionary should feed into forms, but the order, layout and format of forms should still be developed by the DTM team in the country of assessment. This process will be facilitated once the online interface tool for DTM is completed, which allows, among other functions, for the creation of tools, data collection and basic data visualization in the same application.

c) Requesting New Questions or Modifying Questions and Responses

It may be necessary to request data fields that are not found in the data dictionary or to modify questions related to context specific fields. Requesting new questions can be done through the data dictionary platform and requires approval from the DTM Global Support Team based in Geneva. Once the request has been reviewed, it is either accepted, and included as an "optional" field, or rejected and not included in the Data Dictionary. If a requested data field already exists, the country office requesting the change will be notified, and the relevant field will be coded with the existing unique identifier.

Some question texts include context-related fields that require adaptation by the mission. Some examples include the population category, time, or administrative division to which a data field applies. Any modifications to response options for core data fields can also be done within the platform.

2.2 Maintaining Operational Priorities

While the data dictionary is intended to support the standardization of DTM operations across countries, DTM ensures that **operational priorities** remain at the core of its activities in country operations. This is done by pursuing a bifurcated approach to data field classification which outlines mandatory data fields as well as optional or adaptable fields. These are defined in the data dictionary as **core** and **optional** fields. Core fields are **compulsory** and recommended **formats and labels must be followed**. A field may be classified as core due to either a current or potential interest in compiling data at regional/global level or because it has been considered a priority for humanitarian partners at regional/global level. **Optional** fields have r**ecommended formats and labels** but **variation is possible**.

In order to maintain operational priorities, DTM staff are encouraged to adapt the DTM exercises to their context and to add data fields if and when necessary. The Data Dictionary is revised with Clusters, WGs and AoRs at the global level to allow DTM to quickly collect data that are useful to humanitarian data users. At country level, it is the joint responsibility of Clusters, WGs, AoR, IOM programmes and DTM teams to contextualize the global questions so that the data can meet the specific information gaps in a response. Illustrating this further the table below outlines the different types of data field included in the Data Dictionary. Contextualization also applies to the classification of data for sharing purposes. The data dictionary provides an indication of the classification of data (public vs restricted) but this classification can vary depending on the context and should be considered carefully prior to releasing DTM products.

Type of Data Field	Description
Core	*Core* designates the fields that are compulsory and **recommended formats, answer choices as applicable and labels must be followed**. A field may be classified as core due to either a current or potential interest in compiling data

	at regional/global level or because it has been considered a priority for humanitarian partners at the regional/global level.
Core CDW	*Core_cdw* designates the core fields that will also be consolidated and stored in specific tables (L2 tables) in the Central Data Warehouse (CDW), allowing for quick comparison of certain data fields across all countries with DTM operations. As with core, these are compulsory, and the recommended formats and labels must be followed.
Optional	*Optional* fields are those commonly used at country level and aren't expected to be consolidated across countries. Optional fields have recommended formats and labels, but variation is possible.
Recommended by Global Experts	*Recommended by cluster* designates the fields that are recommended to provide sector/cluster-specific information. These fields have been reviewed and approved by the relevant clusters at the global level.
Recommended by DTM	*Recommended by DTM* designates the information that is not required (core), but more encouraged than optional. These include questions on the host community, so proportions of the observed population group can be compared to the total population in an area or location.
Required for Severity	*Required by clusters to conduct severity analysis* indicates the fields that are required to conduct a severity analysis on a given cluster/sector.
Requested by mission	*Requested by mission* are indicators that a mission/region requests because they couldn't find what suited their data collection exercise, more like optional but specific to a country or region

3. Contents of the Data Dictionary

A look at the in-depth profile view of each data field shows how the Data Dictionary provides information to support operations, guide analysis and inform the storage, sharing and use of the data field.

- **Unique ID:** A number in ordinal sequence to give a count of the number of questions and to be able to identify the same question across different DTM exercises and databases.

- **Dissemination Category:** The Data Dictionary indicates how you can share data resulting from each question. The categories follow the IOM Migration Data Governance Policy (IN/253). There are three categories in the Data Dictionary:
  I. Confidential Data, non-personal but sensitive data, require the organization/institution receiving the data to fill a DTM Data Access Request form.
  II. Secret data (personal data) require an official IOM personal data sharing/transfer agreement approved by IOM Legal Department
  III. Public Data: non-personal data that may be made publicly available, after a Do No Harm analysis is conducted by DTM team, Management and Protection colleagues.

  These categories refer to the sharing of data at the level they were collected. For example, confidential or secret data collected at individual and household level cannot be shared publicly at individual and household level. They may be shared publicly, however, once they are aggregated at location or at district or at national level, if a Do No Harm analysis is conducted, using the DTM Do No Harm checklist, and confirms that sharing such aggregated data does not create harm. As another example, confidential data collected at location level cannot be publicly shared at location level, while they may be aggregated at higher administrative levels and publicly shared, if the team considers that appropriate

after the Do No Harm analysis."Question Constraint: Provides stipulations that apply to the answer – for example, the response must be a positive number, select one, select multiple, or open text.

- **Type of Question:** This defines which questions must be included in the assessment and which are optional. Consult the table above for complete information.

- **Question Text:** Shows the question as it will appear on the form. Some question texts can be modified depending on the unit of observation, the population group assessed and the time-frame of the assessment.

- **Data Collection Method:** The approach of data collection is different for different fields of study, depending on the required information. This is to guide on the best method to use to get the most credible information

- **Question Answer Type:** Specifies the data type of the answer response expected to the question e.g. Integer, text, select one etc.

- **Response Options:** Shows the options available for each data field. This could be a list of choices, a text box, yes/no or integers. For the core data field, the proposed answers are minimum compulsory categories.

- **Preconditions for Data Collection:** Identifies questions were previous steps must be taken before asking that question. (e.g. signed consent forms, referral mechanism in place, etc.)

- **Recommended Source of Information:** Person or entities from whom the information needed will ideally come from: local authorities, community leaders, camp managers, etc.

- **Example of Visualization:** Shows which visualizations could be used by DTM to show a data field. This is currently only available for data fields in the DTM Field Companion.[1]

- **Example of Descriptive Analysis:** Gives examples of a possible descriptive analysis of the data field. i.e. the type of wording that would be used in a report.

- **Example of what can be done by users:** Gives examples of potential uses of the data field by IOM or partners.

- **To whom the data field is useful:** Humanitarian response sectors/clusters that could use the specific data field for analysis and action.

- **Instructions:** A set of directions at the beginning of a survey tool that provide important information for participants regarding the purpose of the data collection and how to correctly complete the data assessment.
- **Data Field Name:** To be used in generating KOBO forms and for database column headers in order to enforce standard data management across the different missions and regions

4. Governance of the Data Dictionary

The Data Dictionary database is maintained by the DTM Global Team in Geneva. The contents of the data dictionary are sensitive because they form the foundation of DTM operations. Modifications to

[1] https://displacement.iom.int/dtm-toolkit/dtm-partners-toolkit

the database can impact the entire DTM data infrastructure. As a result of the sensitive nature of the data dictionary, access has been defined based on specific roles. These are outlined in the subsequent section. Protocols for sharing elements of the data dictionary, such as indicators, with external actors have not yet been defined. As a result, access to the data dictionary is currently strictly internal to IOM.

5.1 Accessing the Data Dictionary

Different roles have been assigned to the Data Dictionary to address different needs in access. These are outlined in the document below

Role	Access	Position
Administrator	Access to the platform including approving and rejecting users and modifying the interface.	System Administrator in Geneva
Custodian	Rights to modify the contents of the Data Dictionary – such as add new questions, create forms for missions, edit components and re-code question types (core, optional, etc.)	System Custodians in Geneva
Region Admin	Rights to view contents, download questions, request new questions and eventually create forms.	Regional DTM coordinators or other designated regional DTM staff
User	Rights to view and download questions.	Access granted upon request.
Country Admin	Rights to view contents, download questions, request new questions and eventually create forms for specified country.	Country DTM coordinators or other designated regional DTM staff

Access to the Data Dictionary may be requested through the http://dtmsupport/datadictionary or by contacting dtmsupportcore@iom.int. Access is granted on a case-by-case basis.