

Anguilla - National Statistical System for Disaster-Related Statistics

Proposed Metadata Schema (Draft)

Prepared for: The Government of Anguilla

Submitted by: The World Bank – Urban, Disaster Risk Management, Resilience and Land (GPURL)
Alan Mills, Alan Mills Consulting Ltd.

Date: January 14, 2021

Contents

INTRODUCTION.....	3
Need for a metadata catalogue	3
Need for a metadata schema.....	3
SCHEMA RATIONALE.....	4
Adoption of international standards	4
Guiding Factors	4
Data Types:	4
Schema Logic.....	5
NEXT STEPS	7
ANGUILLA METADATA SCHEMA	8
Explanation of table	8
Element Description.....	9
Ancillary Tables	16
Contacts Table.....	16
Schema Element Mapping	18

INTRODUCTION

Need for a metadata catalogue

To ensure that a logical documentation of the data required for the multiple disaster statistics reporting required by the National Statistics Office (NSO) of the Government of Anguilla (GoA), a metadata catalogue has to be created. Metadata can be simply described as data about data; they provide information about the data which can inform an end user without the need for having access to the entire dataset itself.

The **nature** of catalogue needs to be:

1. An inventory of the existing data identifying their **custodianship** – i.e. where they are held, by whom and in what format,
2. A **mapping** of those data to the identified reporting requirements within the project, whether they be national, regional or international reporting mechanisms,
3. A **description** of the data themselves (primarily units of data held in discrete files or databases but can include if necessary other data collections or elements within a superstructure such as an nCube or larger database), covering the identification of the data, their contents, rights related to the **distribution** of such data and any **limitations** and uses of the information. They can also reference both their geographical (spatial) representation and the study design and analytical tools which created the final data.

The catalogue's **purposes** are:

1. Document the location and availability of relevant information for various reporting mechanisms,
2. Give some objective qualitative information on the status of data and potentially highlight any areas which could merit attention for improvement, i.e., serve as a component of quality assurance,
3. Help the NSO identify gaps in the required data for reporting mechanisms,
4. Assist in streamlining the cataloguing of pertinent government and other datasets to international standards to improve interoperability with international data and metadata requirements and improve the standard of data recording within the GoA,
5. Provide a simple search engine for enquiries related to disaster statistics.

Need for a metadata schema

To satisfy both the elements needed to describe the nature of the catalogue and ensure its purposes are fulfilled, a logical, systematic and standardised method of cataloguing is recommended. To that end, the first step in the design of the catalogue is to agree a schema which covers all the types of data to be documented.

The purposes of having a described schema are:

- It standardises the way data are described for the whole GoA and other disaster management related agencies in Anguilla and helps to encourage completeness and quality assurance in terms of information management,
- Using recognised terminology improves both understanding between agencies and potential interoperability with regional and national agencies using the data.
- Reduces the chance of error in recording of the data
- Encourages the adoption of international best practice within several agencies within the GoA.

SCHEMA RATIONALE

This section briefly documents the steps taken in the design of the schema and identifies its flexibility for modification as the datasets that need documenting come to light over the course of the project.

Adoption of international standards

Significant documentation exists from various international communities of interest which recommend in great detail and precision how to record metadata. These have been validated as best practice, and those who are members of initiatives have become the de facto stewards of metadata standards (e.g., the Dublin Core Metadata initiative, www.dublincore.org) and in some instances have become a recognised standard by the International Standards Organisation (ISO) in Switzerland. Metadata standards have been agreed by consensus from many experts across the globe. Adoption of metadata standards are broadly accepted by choice, but increasingly some international reporting requirements encourage the adoption of certain schema or standards. Some metadata standards cover generic top-level information about a data resource (e.g. the Dublin Core) whereas others are designed for more specific purposes (e.g. the ISO 19115 metadata standard for geographic information). It is recommended that adherence to international standards is good practice for Anguilla, however there are other guiding factors which means the Schema presented here is specific to the Anguilla Disaster Statistics reporting requirement.

Guiding Factors

For this project, the selection of which metadata elements to record were guided by four main factors:

1. Data of various types were sourced for this project and no one standard international schema would cover all the elements required,
2. Data gathered from various sources may already have metadata in a standard form and some consideration of integrating that into the design is necessary,
3. Standard schema are exhaustive and although allow for some leeway on whether an element is mandatory or not within a schema, can be unwieldy for a small organisation to maintain. Metadata cataloguing is a significant investment for an organisation and the burden of maintenance of catalogues when datasets are growing and changing has to be considered before commencing with a metadata programme.
4. Exhaustive metadata schema are designed to accurately model the highly hierarchical and interrelated nature of metadata elements but this comes at a cost in terms of database design, software choices, implementation and upkeep. For the initial tranche of data to be collected, the design of the metadatabase is kept as "flat" as possible, i.e., a rectangular table is used where one row of data contains data for one discrete dataset and one column is related to one descriptive metadata element. The only concessions to this will be in the use of selected look up tables to restrict values which can be entered, or can be used in multiple fields (e.g., agents or contacts and organisation information).

The data stored in this catalogue balance the need for good documentation with being accessible and easy to manage.

Data Types:

A range of Various types of data are being sourced within this project which means that several schema need to be drawn on. These include:

1. Micro statistical data, i.e., raw information down to, for example individual or household level.

2. Aggregated and time series information
3. Analysed information where modelling has reshaped the original data (e.g., through weighting),
4. Datasets with a strong geographical element, either those which are stored in a GIS ready format, or those which have some reference or database. These can be further subdivided into:
 - a. Vector GIS format commonly used for topographic, amenity, administrative areas, thematic mapping,
 - b. Raster GIS format commonly derived from satellite or aerial imagery, possibly relevant in this study given its application to land and marine cover mapping.
 - c. Tabular data with some geographic reference, either two columns which identify a point on the earth's surface using x and y coordinate (e.g., latitude and longitude), or a column which references a feature in another geographical dataset (e.g., a placename or Enumeration area).
5. Reports and other qualitative data.

Schema Logic

The following assumptions were made, based on being in keeping with the needs, use of international standards and led by the guiding principles above, in the determination of which elements to include in the schema:

Selection of schema elements

The following chain of logic was used to select elements for the schema:

- Master level of generic inventory used the Dublin Core Metadata Initiative ([DCMI](#))
- More detailed descriptive information for all datasets comes primarily from information from the [ISO 19115](#) "Geographic Information – Metadata standard", with a small number of elements derived from the Data Documentation Initiative ([DDI](#)) and Data Catalog Vocabulary ([DCAT](#)).
- Geographical representation is covered using the elements from the ISO 19115 schema
- Preliminary descriptions of statistical elements are covered using elements from the DDI schema.
- The Statistical Data and Metadata eXchange ([SDMX](#)) standards were examined in detail but primarily deal with the standard formatting of individual data packages and how they can be exchanged between systems in a platform independent manner. Much of the detail of the SDMX schema relies heavily upon agencies establishing their databases around this structure, which is beyond the scope of this project, and most of the top-level metadata elements required for describing datasets are covered already by Dublin Core, DCAT and DDI.

The current draft schema presented here is a snapshot of this logic to date. All the elements listed in the schemas mentioned above are documented in detail in a separate Excel spreadsheet and linked directly into the proposed Anguilla schema. If during the review and documentation phases of the project, there is need to revise the elements to be included those revisions can be achieved systematically and with continued reference to the relevant international standard.

Relevant Metadata - Important to the functioning of the database is to ensure only appropriate metadata for a particular data type is collected (e.g., a geographical dataset as opposed to a purely tabular information set). There are a series of gates in the catalogue where if the user chooses a type, they will be channelled down a particular line of metadata recording.

Related Resources – There is no need to document in detail the relationships between discrete datasets in the metadata record. Dublin Core has several elements which record information about related resources i.e., how a particular dataset is a subset, related to, or covers other elements, or supersedes or is superseded by other material¹. For this exercise it is proposed to exclude most of these, except successional elements which will be covered as part of the maintenance of the datasets and metadatabase under version recording and validity dates. This decision can be reviewed if the actual datasets shared show a high level of interdependence.

Dealing with organisations and people - One concession to the flat structure of the metadata catalogue is the recording of organisational information separately. This is for two reasons; first that these data are used in several parts of the schema and secondly that the additional address/contact information can be stored one time. In most schemas there is a separation between contact people and organisations (given there could be a one-to-many relationship between these two entities). To simplify the catalogue:

1. A single responsible contact within each organisation will be determined (i.e., there is a one-to-one relationship with contact name and organisation). It is expected for a small government structure as in Anguilla data are a collective responsibility within the organisations and only one named person is required who has overview of data handling and management.
2. If an organisation has departments or an affiliated agency which is supplying information, it has been given a separate record in the organisation table. A simple connection with the next tier of organisation upwards (i.e., an agency is a subsidiary of another agency) will be recorded if necessary.

Metadata schema of all types have several fields to document the agents involved with a resource, including those who create data, those who are responsible for their update and maintenance, those responsible for any distribution rights. If it appears that for all the data to be collected within this project, the number of metadata elements referring to these agents can be reduced to a single “Responsible” element.

Additional Information for Raster Data - the detailed information on the grid spatial representation, Georectified and georeferenceable is included at present, but it is expected that this level of detail will not be necessary and any relevant raster data can be documented with the generic geographical fields.

Statistical Aggregation Levels and statistical methods– At present there are no statistical aggregation levels that are recorded in the metadata. No specific schema elements have identified the ability to record this, although within DDI there are various elements covering the study methodology which can include descriptive text related to these concepts. There are several potential dimensions to deal with:

1. Aggregation from individual records to another unit (Household, neighbourhood, EA, PSU, Country),
2. Statistical method used for aggregation (SUM, MEAN, MODE, MEDIAN,)
3. Models or analyses used to create indicator or indices.

¹ E.g. hasPart, hasFormat, isFormatof , isVersionOf

At this stage it is proposed that these issues be discussed with the working group and when data are collected, they can be analysed to see to what extent this level of metadata documentation is required.

DDI does allow for various levels of documentation not just of datasets (files) and collections but descriptions of hyperdimensional tables (nCubes), nested data collections, as well as variables used in forms, questionnaires and other ancillary instruments used in the collection of statistics. It is expected the disaster related statistics metadata will predominantly use data files, but the provision is there to include more metadata elements if that description is needed.

NEXT STEPS

The design of the metadata schema has made several assumptions about the nature of the reporting required, the nature of the data themselves and the extent of existing metadata recording. Inevitably the current schema is a draft, and it is proposed that the working group review the metadata schema and attempt to respond to the questions below. The schema will then be translated into a working Excel spreadsheet for the cataloguing of information and shared with those identified within NSO to be supporting the metadata process. Then once data are identified and started to be collected a trial entry period to test both schema and catalogue and any adaptations made.

Underlying the simple schema detailed in this document here is a separate Excel document which meticulously records all the elements required for all standard schemas. This is connected into the Anguilla schema so that each element can be traced back to its standard source, and it is expected updates can be made in a systematic and logical form.

ANGUILLA METADATA SCHEMA

Explanation of table

The following table has these columns:

Major Theme	Groups the metadata elements into a small series of top level themes related to Description, Access, Content, Spatial Representation, and statistical study description
Sub Theme	Each of these major themes may be further subdivided into sections
Label	Simple label to be used for each element which is machine readable
Item Name	A descriptive name for the element
Schema Element Source	The standard schema from which the element is as exactly matched as possible (Dublin Core, DDI, DCAT or ISO 19115)
Obligation	The obligation on any cataloguer to document this element (Note this is open to change once it has been tested with real data). M – Mandatory – must be entered O – Optional – this is an optional field C – Conditional – if other fields are filled in this element must have data. H – Hierarchical, a useful metadata placeholder which relates to several more detailed elements but may also be a useful element in its own right for documentation.
Condition terms	If Obligation is C the terms of those conditions are specified here.
Description	A description of the element. Where possible these descriptions come from the official schema description

Element Description

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
MAIN LIST							
Description	Identification	identifier	Unique Identifier	Dublin Core	M		An unambiguous reference to the resource within a given context.
Description	Identification	title	Title	Dublin Core	M		A name given to the resource.
Description	Identification	alternative	Alternative Title	Dublin Core	C	If alternative exists	An alternative name for the resource
Description	Identification	abstract	Abstract	Dublin Core	M		A summary of the resource.
Description	Identification	description	description	Dublin Core	C	If a lengthier explanation of the resource is needed than placed in the abstract	An account of the resource.
Description	Citation	bibliographicCitation	Bibliographic Citation	Dublin Core	C	If related to a published document or material	A bibliographic reference for the resource.
Description	Responsibility	creator	Creator	Dublin Core	M		An entity responsible for making the resource.
Description	Responsibility	contributor	Contributor	Dublin Core	C	If different from the creator	An entity responsible for making contributions to the resource.
Description	Responsibility	publisher	Publisher	Dublin Core	M		An entity responsible for making the resource available.
	Responsibility	Source	Source	Dublin Core	M		A related resource from which the described resource is derived.
Description	Responsibility	rights	Rights	Dublin Core	M		Information about rights held in and over the resource.
Description	Responsibility	rightsHolder	Rights Holder	Dublin Core	M		A person or organization owning or managing rights over the resource.
Description	Responsibility	provenance	Provenance History	Dublin Core	M		A statement of any changes in ownership and custody of the resource since its creation that are

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
							significant for its authenticity, integrity, and interpretation.
Description	Temporal	valid	Period of Validity	Dublin Core	M		Date (often a range) of validity of a resource.
Description	Temporal	created	Date Created	Dublin Core	M		Date of creation of the resource
Description	Temporal	dateAccepted	Date Accepted	Dublin Core	M		Date of acceptance of the resource.
Description	Temporal	dateCopyrighted	Date Copyrighted	Dublin Core	M		Date of copyright of the resource.
Description	Temporal	dateSubmitted	Date Submitted	Dublin Core	M		Date of submission of the resource.
Description	Temporal	issued	Date Issued	Dublin Core	M		Date of formal issuance of the resource.
Description	Temporal	modified	Date Modified	Dublin Core	M		Date on which the resource was changed.
Description	Identification	language	Language	Dublin Core	M		A language of the resource.
Description	Identification	subject	Subject	Dublin Core	M		A topic of the resource.
Description	Identification	keyword	Keyword (s)	ISO 19115	M		Words or phrases which describe the resource
Description	Identification	type	Format Type	Dublin Core	M		The nature or genre of the resource.
Description	Identification	license	Official License	Dublin Core	M		A legal document giving official permission to do something with the resource.
Description	Identification	conformsTo	Conforms To	Dublin Core	M		An established standard to which the described resource conforms.
Access	Distribution	format	Format	Dublin Core	H		The file format, physical medium, or dimensions of the resource.
Access	Distribution	medium	Format Medium	Dublin Core	C	subset of Format	The material or physical carrier of the resource.
Access	Distribution	onLine	Online Location	ISO 19115	O		information about the online sources from which the resource can be obtained
Access	Distribution	offLine	Offline Location	ISO 19115	O		information about the offline sources from which the medium can be obtained
Access	Distribution	available	Date Available	Dublin Core	O		Date that the resource became or will become available.
Access	Distribution	transferSize	Transfer Size		M		estimated size of a unit in specified transfer format in Mb
Access	Usage	specificUsage	Specific Usage	ISO 19115	O		Brief description of the resource and/or resource series usage
Access	Usage	audience	Audience	Dublin Core	M		A class of agents for whom the resource is intended or useful.

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
Access	Usage	accessRights	Access Rights	Dublin Core	M		Information about who access the resource or an indication of its security status.
Access	Constraints	useLimitations	Use Limitations	ISO 19115	M		limitations affecting fitness for use of the resource or metadata
Access	Constraints	constraintApplicationScope	Spatial Temporal Extent Constraints	ISO 19115	O		spatial or temporal extents/levels for application of the constraint
Access	Constraints	graphic	Constraint Graphic	ISO 19115	O		graphic symbol indicating the constraint
Access	Constraints	reference	Constraint Reference	ISO 19115	O		citation for the limitation or constraint
Access	Constraints	releasability	Constraint Releasability	ISO 19115	O		information concerning parties to whom the resources can or cannot be released.
Access	Constraints	useConstraints	Use Protection Constraints	ISO 19115	M		access constraints applied to assure protection of privacy or intellectual property, and any special restrictions of limitations on obtaining the resource or metadata
Access	Constraints	otherConstraints	Other Constraints	ISO 19115	O		other restrictions and legal prerequisites for accessing and using the metadata or resource
Access	Constraints	responsibleParty	Constraint Responsibility	ISO 19115	M		party responsible for the resource constraints
Content	Accrual	accrualMethod	Accrual Method	Dublin Core	M		The method by which items are added to a collection
Content	Accrual	accrualPeriodicity	Accrual Periodicity	Dublin Core	M		The frequency with which items are added to a collection.
Content	Accrual	accrualPolicy	Accrual Policy	Dublin Core	M		The policy governing the addition of items to a collection.
Content	Coverage	coverage	Coverage	Dublin Core	H		The spatial or temporal topic of the resource, spatial applicability of the resource, or jurisdiction under which the resource is relevant.
Content	Coverage	extent	Extent	Dublin Core	O		The size or duration of the resource.
Content	Coverage	temporal	Temporal Characteristics	Dublin Core	H		Temporal characteristics of the resource.
Content	Coverage	temporalResolution	Temporal Resolution	DCAT	M		Minimum time period resolvable in the dataset

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
Content	Coverage	geogUnit	Geographic Unit of Observation	DDI	M		Lowest Level of geographic aggregation covered by the data
Content	Coverage	Unit of Analysis	Unit of Analysis	DDI	M		Basic unit of analysis or observation that the file describes: individuals, families/households, groups, institutions/organizations, administrative units, etc. The "unit" attribute is included to permit the development of a controlled vocabulary for this element.
SPATIAL							
SpatialRepresentation	GeneralSpatial	spatial	Spatial Characteristics	Dublin Core	H		Spatial characteristics of the resource.
SpatialRepresentation	GeneralSpatial	spatialRepresentation	Spatial Representations	ISO 19115	M		method used to spatially represent geographic information
SpatialRepresentation	GeneralSpatial	RepresentativeFraction	Scale (1: Denominator)	ISO 19115	M		Scale - number below the line in vulgar fraction
SpatialRepresentation	GeneralSpatial	sourceSpatialResolution	Source Resolution	Dublin Core	H		spatial resolution expressed as a scale factor, a distance, and angle or level of detail.
SpatialRepresentation	BoundaryBox	westBoundLongitude	westernmost coordinate	ISO 19115	C		westernmost coordinate
SpatialRepresentation	BoundaryBox	eastBoundLongitude	easternmost coordinate	ISO 19115	C		easternmost coordinate
SpatialRepresentation	BoundaryBox	northBoundLatitude	northernmost coordinate	ISO 19115	C		northernmost coordinate
SpatialRepresentation	BoundaryBox	southBoundLatitude	southernmost coordinate	ISO 19115	C		southernmost coordinate
SpatialRepresentation	BoundaryBox	Polygon	set of points defining bounding box	ISO 19115	C	if bounding box coordinates not submitted	set of points defining bounding box
SpatialRepresentation	GridSpatialRepresentation	numberOfDimensions	Number of dimensions	ISO 19115	O		number of independent grid-spatial temporal axes
SpatialRepresentation	GridSpatialRepresentation	transformationParameterAvailability	Transformation Parameters Available	ISO 19115	O		indication whether geographic transformation parameters exist

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
SpatialRepresentation	Georectified	checkPointAvailability	Check Points available	ISO 19115	O		accuracy points available
SpatialRepresentation	Georectified	transformationDimensionDescription	Transformation Description	ISO 19115	O		general description of transformation
SpatialRepresentation	Georectified	transformationDimensionMapping	Transformation Information	ISO 19115	O		information about which grid axes are spatial map axes
SpatialRepresentation	Georeferenceable	controlPointAvailability	Control Points Available	ISO 19115	O		indication whether control points exist
SpatialRepresentation	VectorSpatialRepresentation	topologyLevel	Topology Level	ISO 19115	O		code identifying degree of complexity of spatial relationships
STATISTICAL REPRESENTATION							
Study Description	Other Study Description Materials	relation	Related Resources	Dublin Core	M		A related resource.
Study Description	Study Scope	universe	Universe	DDI	O		The group of persons or other elements that are the object of research and to which any analytic results refer. Age, nationality, and residence commonly help to delineate a given universe, but any of a number of factors may be involved, such as sex, race, income, veteran status, criminal convictions, etc. The universe may consist of elements other than persons, such as housing units, court cases, deaths, countries, etc.
Study Description	Study Scope	respRate	Response Rate	DDI	O		The percentage of sample members who provided information.
Study Description	Study Scope	sampProc	Sampling Procedure		O		The type of sample and sample design used to select the survey respondents to represent the population. May include reference to the target sample size and the sampling fraction.
Study Description	Study Scope	dataKind	Kind of Data	DDI	O		The type of data included in the file: survey data, census/enumeration data, aggregate data, clinical data, event/transaction data, program source code, machine-readable text,

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
							administrative records data, experimental data, psychological test, textual data, coded textual, coded documents, time budget diaries, observation data/ratings, process-produced data, etc.
Study Description	Study Scope	frequency	Frequency of Data Collection	DDI	O		For data collected at more than one point in time, the frequency with which the data were collected. The "freq" attribute is included to permit the development of a controlled vocabulary for this element.
Study Description	Study Scope	resInstru	Type of Research Instrument	DDI	O		The type of data collection instrument used. "Structured" indicates an instrument in which all respondents are asked the same questions/tests, possibly with precoded answers. If a small portion of such a questionnaire includes open-ended questions, provide appropriate comments. "Semi-structured" indicates that the research instrument contains mainly open-ended questions. "Unstructured" indicates that in-depth interviews were conducted. The "type" attribute is included to permit the development of a controlled vocabulary for this element.
Study Description	Study Scope	weight	Weighting	DDI	O		The use of sampling procedures may make it necessary to apply weights to produce accurate statistical results. Describe here the criteria for using weights in analysis of a collection. If a weighting formula or coefficient was developed, provide this formula, define its elements, and indicate how the formula is applied to data.
Study Description	Study Scope	estSmpErr	Estimates of Sampling Error	DDI	O		Measure of how precisely one can estimate a population value from a given sample.
Study Description	Study Scope	DefExist	Definitions Exist	Local	M		Is there documentation about the Definitions used in the dataset

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
Study Description	Study Scope	DSD	Data Structure Definition	SDMX	C	If DefExist = Y	Set of structural metadata associated to a Data Set, which includes information about how Concepts are associated with the Measures, Dimensions, and Attributes of a data cube, along with information about the Representation of data and related descriptive metadata
Study Description	Study Scope	classesExist	Number of Classification Systems	Local	M		Number of classifications systems used in the dataset for which there are descriptions
Study Description	Study Scope	CLASS_SYSTEM	Classification system	SDMX	C	If classesExist = Y	Metadata element used to a) list the classification(s) being used for a given Data Set or set of Data Sets, and b) describe how these conform to internationally agreed standards, guidelines, or good practices.
Study Description	Variables	var.name	Variable Name	DDI	O		usually contains the so-called "short label" for the variable, limited to eight characters in many statistical analysis systems such as SAS or SPSS.
Study Description	Variables	var.wgt	Variable Weight	DDI	O		The attribute "wgt" indicates whether the variable is a weight. But this can also be used as reference to "wgt-var" which identifies and references the weight variable(s) for this variable.
Study Description	Variables	var.aggrMethod	Variable Aggregation Method	DDI	O		indicates the type of aggregation method used, for example 'sum', 'average', 'count'.
Study Description	Variables	var.nature	Variable Nature	DDI	O		records the nature of the variable, whether it is 'nominal', 'ordinal', 'interval', or 'ratio'.
METADATA LOG							
Metadata Log		Respondent	Respondent	Internal	M		Who reported the data and metadata
Metadata Log		Agency	Agency	Internal	M		From which agency the respondent is part
Metadata Log		Reporter	Reporter	Internal	M		Who entered the data in the catalogue
Metadata Log		Date of Original Report	Date of Original Report	Internal	M		Date the dataset was first catalogued

Major Theme	Sub Theme	Label	Item Name	Schema Element Source	Obligation (recommended)	Condition terms	Description
Metadata Log		Last Revision Date	Last Revision Date	Internal	M		The most recent time metadata has been reviewed and/or revised.
Metadata Log		Next Review Date	Next Review Date	Internal	M		Date (Determined by respondent) when metadata for this record should be reviewed
Metadata Log		Filename	Filename	Internal	O		If data is a discrete file - can be shared here.
Metadata Log		Comments	Comments	Internal	O		Comments on the data or metadata (e.g. required data covered by another catalogued dataset).
Metadata Log		Recommendations	Recommendations	Internal	O		Recommendations on gap

Ancillary Tables

The following ancillary table is proposed which will be an inventory of all personnel/agencies connected with the sharing of disaster related statistics. It will form a drop-down list that is used in the master table to identify, amongst others, data creators, publishers, custodians and main users. It is proposed to have a simple contacts table rather than splitting between contacts and organisations. The specification is based on the ISO 19115 details but map to both DDI and SDMX schema elements.

Contacts Table

ID	Unique identifier code.
name	Contact name
contactInfo	Type of contact information gathered (e.g., Office, Consultant, online repository – free text allowed)
positionName	Position in organisation
logo	link to logo image
deliveryPoint	Physical Address
city	Settlement name
administrativeArea	Administrative area (if applicable)
postalCode	Post Code or Zip Code (if applicable)
country	Country (Anguilla or other)

electronicMailAddress	email address
phone	main contact number
Address	Postal Address (if different from Physical Address)
onlineResource	URL - website address
hoursOfService	Opening hours (if applicable)

At this time, it is not proposed to have any other ancillary tables. However a series of look up ranges will be created within the Excel file for many textual fields which require standard values to be entered.

Schema Element Mapping

The following table details how each element used in the proposed Anguilla metadata schema maps to elements in other standard elements, if at all. The comparison is made between Dublin Core, DCAT, DDI, SDMX and ISO 19115 standard schemas. For completeness the sets identified in the Terms of Reference for the project are also included; all have been included in the schema. In some cases, the mapping is fuzzy, i.e., the descriptions in each schema do not entirely match; these are shown in italics in this table. In some cases, a single element in one schema may relate to *MULTIPLE* elements in others.

Anguilla Schema Label	Anguilla Schema Item Name	Schema Element Source	TOR Requirements	Dublin Core	ISO19115	DCAT	SDMX	DDI
identifier	Unique Identifier	Dublin Core		identifier		identifier	ID	2.1.1.5 Identification Number
title	Title	Dublin Core	Dataset Name	title	title	title	Name	2.1.1.1 Title
alternative	Alternative Title	Dublin Core		alternative	alternateTitle	-	-	2.1.1.3 Alternative Title
abstract	Abstract	Dublin Core	Abstract	abstract	abstract	-	-	2.2.2 Abstract
description	description	Dublin Core	Description	description	<i>purpose</i>	description	Description	<i>Summary</i>
bibliographicCitation	Bibliographic Citation	Dublin Core		bibliographicCitation	<i>citation</i>	-	-	2.1.7 Bibliographic Citation
creator	Creator	Dublin Core	Creator	creator	<i>contact or pointOfContact</i>	-	<i>DataProviders</i>	2.1.2.1 Authoring Entity/Primary Investigator
contributor	Contributor	Dublin Core		contributor	credit	qualifiedAttribution	DataProvider	2.1.2.2 Other Identifications/Acknowledgements
publisher	Publisher	Dublin Core		publisher	Distributor	publisher	<i>contact</i>	2.1.3.1 Producer
Source	Source	Dublin Core			sourceCitation			2.3.1.8 Sources Statement
rights	Rights	Dublin Core		rights	<i>various under Constraints</i>	rights		2.1.3.2 Copyright
rightsHolder	Rights Holder	Dublin Core		rightsHolder	responsibleParty	-		<i>2.1.3.1 Producer</i>
provenance	Provenance History	Dublin Core		provenance	-	-		-
valid	Period of Validity	Dublin Core		valid	extent or maintenanceAndUpdateFrequency	-		-
created	Date Created	Dublin Core	Creation Date	created	date	<i>modified</i>	<i>Prepared</i>	2.1.3.4 Place of Production

Anguilla Schema Label	Anguilla Schema Item Name	Schema Element Source	TOR Requirements	Dublin Core	ISO19115	DCAT	SDMX	DDI
dateAccepted	Date Accepted	Dublin Core		dateAccepted	usageDateTime	<i>issued</i>	<i>datePrepared</i>	
dateCopyrighted	Date Copyrighted	Dublin Core		dateCopyrighted	<i>editionDate</i>	-	<i>validFrom</i>	
dateSubmitted	Date Submitted	Dublin Core		dateSubmitted	<i>date</i>	-	-	2.1.4.4 Date of Deposit
issued	Date Issued	Dublin Core		issued	<i>usageDateTime</i>	-		2.1.4.5 Date of Distribution
modified	Date Modified	Dublin Core		modified	<i>maintenanceDate</i>	modified		1.1.6.1 Version
language	Language	Dublin Core		language	language	language	-	-
subject	Subject	Dublin Core		subject	topicCategory	-	-	2.2.1.2 Topic Classification
keyword	Keyword (s)	ISO 19115	Keyword	-	keyword	Keyword	-	2.2.1.1 Keywords
type	Format Type	Dublin Core	Format	type	(Various)	type	(Various)	2.2.3.10 Kind of Data
license	Official License	Dublin Core	Applicable Licensing Terms	license	-	license	-	-
conformsTo	Conforms To	Dublin Core		conformsTo	-	-		-
format	Format	Dublin Core	Format	format	(Various)	-	(Various)	3.1.5 Type of File
medium	Format Medium	Dublin Core		medium	medium	mediaType	<i>distribution</i>	3.1.6 Data Format
onLine	Online Location	ISO 19115	Dataset Location - URL	-	onLine	<i>accessURL or download URL</i>	urn	5.5.4.1 Distributor
offLine	Offline Location	ISO 19115	Dataset Location - Address	-	offLine	-		5.5.3.4 Place of Production
available	Date Available	Dublin Core		available	<i>usageDateTime</i>	-		5.5.4.5 Date of Distribution
transferSize	Transfer Size			<i>extent</i>	<i>transferSize</i>			-
specificUsage	Specific Usage	ISO 19115		-	<i>specificUsage</i>	<i>wasGenerated By</i>	-	-
audience	Audience	Dublin Core		audience	userContactInfo	-	<i>Receiver</i>	
accessRights	Access Rights	Dublin Core		accessRights	<i>accessConstraints</i>	-	-	
useLimitations	Use Limitations	ISO 19115		-	useLimitations	-	-	2.4.2.3 Restrictions
constraintApplicationScope	Spatial Temporal Extent Constraints	ISO 19115		-	constraintApplicationScope	-	-	2.4.2.3 Restrictions
graphic	Constraint Graphic	ISO 19115		-	graphic	-	-	-

Anguilla Schema Label	Anguilla Schema Item Name	Schema Element Source	TOR Requirements	Dublin Core	ISO19115	DCAT	SDMX	DDI
reference	Constraint Reference	ISO 19115		-	reference	-	-	2.4.2.5 Citation Requirement
releasability	Constraint Releasability	ISO 19115		-	releasability	-	-	2.4.1.3 Availability Status
useConstraints	Use Protection Constraints	ISO 19115		-	useConstraints	-	-	2.4.2.3 Restrictions
otherConstraints	Other Constraints	ISO 19115		-	otherConstraints	-	-	2.4.2.3 Restrictions
responsibleParty	Constraint Responsibility	ISO 19115	ResponsibleAgency	-	responsibleParty	contactPoint	-	2.4.2.4 Contact Persons
accrualMethod	Accrual Method	Dublin Core		-	maintenanceScope	-	-	-
accrualPeriodicity	Accrual Periodicity	Dublin Core		accrualPeriodicity	maintenanceAndUpdateFrequency	accrualPeriodicity	-	-
accrualPolicy	Accrual Policy	Dublin Core		accrualPolicy	maintenanceNote	-	-	-
coverage	Coverage	Dublin Core		coverage	MULTIPLE	-		multiple
extent	Extent	Dublin Core		extent	Multiple under Ex_Extent	-		Geographic Coverage
temporal	Temporal Characteristics	Dublin Core		temporal	TM_Duration (MULTIPLE)	temporal		Time Period
temporalResolution	Temporal Resolution	DCAT						
geogUnit	Geographic Unit of Observation	DDI			-			2.2.3.5 Geographic Unit
Unit of Analysis	Unit of Analysis	DDI						2.2.3.8 AnlyUnit
spatial	Spatial Characteristics	Dublin Core			Ex_GeographicExtent (MULTIPLE)	Spatial		2.2.3.4 Geographic Coverage
spatialRepresentation	Spatial Representations	ISO 19115		spatial	Spatial Representations	-		
RepresentativeFraction	Scale (1: Denominator)	ISO 19115			RepresentativeFraction			
sourceSpatialResolution	Source Resolution	Dublin Core			spatialResolution	spatialResolution inMeters		
westBoundLongitude	westernmost coordinate	ISO 19115		-	westBoundLongitude	-		2.2.3.6.1 West Bounding Longitude
eastBoundLongitude	easternmost coordinate	ISO 19115		-	easternmost coordinate	-		2.2.3.6.2 East Bounding Longitude
northBoundLatitude	northernmost coordinate	ISO 19115			northernmost coordinate	-		2.2.3.6.3 South Bounding Latitude

Anguilla Schema Label	Anguilla Schema Item Name	Schema Element Source	TOR Requirements	Dublin Core	ISO19115	DCAT	SDMX	DDI
southBoundLatitude	southernmost coordinate	ISO 19115			southernmost coordinate	-		2.2.3.6.4 North Bounding Latitude
Polygon	set of points defining bounding box	ISO 19115			set of points defining bounding box	-		2.2.3.7 Geographic Bounding Polygon
numberOfDimensions	Number of dimensions	ISO 19115			Number of dimensions	-		-
axisDimensionProperties	Axes dimension properties	ISO 19115			Axes dimension properties	-		-
cellGeometry	Cell geometry	ISO 19115			Cell geometry	-		-
transformationParameterAvailability	Transformation Parameters Available	ISO 19115			Transformation Parameters Available	-		-
checkPointAvailability	Check Points available	ISO 19115			Check Points available	-		-
checkPointDescription	Check Points Description	ISO 19115			Check Points Description	-		-
cornerPoints	Bounding Box Locations	ISO 19115			Bounding Box Locations	-		-
centrePoint	Centre Point Location	ISO 19115			Centre Point Location	-		-
pointInPixel	Point in pixel definition	ISO 19115			Point in pixel definition	-		-
transformationDimensionDescription	Transformation Description	ISO 19115			Transformation Description	-		-
transformationDimensionMapping	Transformation Information	ISO 19115			Transformation Information	-		-
controlPointAvailability	Control Points Available	ISO 19115			Control Points Available	-		-
orientationParameterAvailability	Orientation Parameters Available	ISO 19115			Orientation Parameters Available	-		-
orientationParameterDescription	Orientation Parameters Description	ISO 19115			Orientation Parameters Description	-		-
georeferencedParameters	Georeferenced Parameters	ISO 19115			Georeferenced Parameters	-		-
parameterCitation	Parameter Citation	ISO 19115			Parameter Citation	-		-
topologyLevel	Topology Level	ISO 19115			Topology Level	-		-

Anguilla Schema Label	Anguilla Schema Item Name	Schema Element Source	TOR Requirements	Dublin Core	ISO19115	DCAT	SDMX	DDI
geometricObjects	Information about geometric Objects	ISO 19115			Information about geometric Objects	-		-
relation	Related Resources	Dublin Core						2.5.1 Related Materials
universe	Universe	DDI						2.2.3.9 Universe
respRate	Response Rate	DDI						2.3.3.1 Response Rate
sampProc	Sampling Procedure	DDI						2.3.1.4 Sampling Procedure
dataKind	Kind of Data	DDI						2.2.3.10 Kind of Data
frequency	Frequency of Data Collection	DDI						2.3.1.3 Frequency of Data Collection
resInstru	Type of Research Instrument	DDI						2.3.1.7 Type of Research Instrument
weight	Weighting	DDI						2.3.1.12 Weighting
DefExist	Definitions Exist	Local						
DSD	Data Structure Definition	SDMX					DSD	
classesExist	Number of Classification Systems	Local						
CLASS_SYSTEM	Classification system	SDMX					CLASS_SYSTEM	
estSmpErr	Estimates of Sampling Error	DDI						2.3.3.2 Estimates of Sampling Error
var.name	Variable Name	DDI						4.3 Variable
var.wgt	Variable Weight	DDI						4.3 Variable
var.aggrMethod	Variable Aggregation Method	DDI						4.3 Variable
var.nature	Variable Nature	DDI						4.3 Variable