Development of Disaster-Related Statistics Capacity in the Government of Anguilla

Funded by the European Union

Implemented by:

WORLD BANK GROUP
Urban, Disaster Risk Management, Resilience & Land

GFDRR
Global Facility for Disaster Reduction and Recovery

In partnership with:

RESEMBID

1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

Publication Design: Blue Robin Designs, LLC
Knowledge Management Lead: Kerri Cox, World Bank
Photo Credits: Lori-Rae Alleyne, Government of Anguilla

ACKNOWLEDGMENTS

This work is part of the Technical Assistance Program for Disaster Risk Financing (DRF TA) in Caribbean Overseas Countries and Territories (OCTs) to enhance long-term financial resilience to disasters in Caribbean OCTs, to the benefit of the most vulnerable. The report is the result of a collaboration between the Government of Anguilla (GoA) and the World Bank under the Caribbean DRF TA Program and is based on consultations with numerous government agencies and the private sector entities within Anguilla.

We wish to extend our appreciation and acknowledge the numerous agencies and organizations for their support and assistance in granting access to information, providing support to the activities leading to this report's development, and their availability for discussions during the assessment process. These include the Statistics Department, Department of Physical Planning, Department of Natural Resources, Department of Disaster Management, and the Department of Information Technology and E-Government Services.

This report was authored by Robert Smith (Senior Statistician), Dawnette Bryan-Lico (Senior ICT Specialist), Alan Mills (Senior Information Systems Specialist), and Michael Bordt (Senior Statistical Information Systems Specialist) under the leadership of Mary Boyer (Disaster Risk Management Specialist) with the support of Michael Fedak (Data Management Specialist) and Ivelisse Justiniano (Urban and Disaster Risk Management Specialist). We want to highlight the contribution of Michael Bordt, Ph.D., who passed away before the publication of this report, making this one of his last works. The World Bank team and the Government of Anguilla acknowledge his extensive work and valuable contributions to the development of this report.

We are grateful to the European Union (EU) for its financial support, without which the preparation of this report would not have been possible. The DRF TA in Caribbean OCTs is a partnership between the EU, the World Bank Group, and the Global Facility for Disaster Reduction and Recovery (GFDRR). The program is part of the EU-funded Caribbean OCTs Resilience, Sustainable Energy and Marine Biodiversity Program (RESEMBID), implemented by Expertise France, the World Bank Group, and the GFDRR.

Finally, the World Bank team would like to thank Ms. Lori-Rae Alleyne (Chief Statistician) and Ms. Anthea D. Ipinson (Chief Project Officer) and their dedicated teams for their data collection and technical support guidance. The report also greatly benefited from the input of key stakeholders who collaborated or assisted the World Bank team.

TABLE OF CONTENTS

TABLE INDEX

FIGURE INDEX

ABBREVIATIONS & ACRONYMS

ACOCI	Anguilla Chamber of Commerce and Industry
AHTA	Anguilla Hotel and Tourism Association
ANGLEC	Anguilla Electricity Company
ANSAC	Anguilla National Statistics Advisory Committee
ASD	Anguilla Statistics Department
ASYCUDA	Automated System for Customs Data
CARDTP	Caribbean Digital Transformation Project
CARICOM	Caribbean Community
CARPHA	Caribbean Public Health Agency
CDEMA	Caribbean Disaster Emergency Management Agency
CEO	Chief Executive Officer
CISA	Classification of International Statistical Activities
CSO	Central Statistics Office
DADS	Database for Anguilla Disaster-related Statistics
DCAT	Data Catalog Vocabulary
DCMI	Dublin Core Metadata Initiative
DDI	Data Documentation Initiative
DDM	Department of Disaster Management
DITES	Department of Information Technology and E-Government Services
DoE	Department of Education
DOHP	Department of Health Protection
DNR	Department of Natural Resources
DPP	Department of Physical Planning
DRF TA	Disaster Risk Financing Technical Assistance
DRSF	Disaster-Related Statistics Framework
DSSAT	Disaster-Related Statistics Self-Assessment Tool
E-GRIP	Electronic Government Regional Integration Project
EMIS	Education Management Information System
ESSAT	Environment Statistics Self-Assessment Tool
EU	European Union
EXCO	Executive Council
FCDO	Foreign, Commonwealth & Development Office
FDES	United Nations Framework for the Development of Environment Statistics
FGDC	Federal Geographic Data Committee
FTE	Full-Time Employee
G-WAN	Government-Wide Area Network
GDP	Gross Domestic Product
GFDRR	Global Facility for Disaster Reduction and Recovery
GIS	Geographic Information System

GoA	Government of Anguilla
GSBPM	Generic Statistical Business Process Model
HAA	Health Authority of Anguilla
IATI	International Aid Transparency Initiative
ICT	Information and Communication Technology
IFRC	International Federation of Red Cross and Red Crescent Societies
INSPIRE	Infrastructure for Spatial Information in the European Community
MoU	Memorandum of Understanding
NDMC	National Disaster Management Committee
NSDS	National Strategy for the Development of Statistics
NSO	National Statistical Office
OCTs	Overseas Countries and Territories
OECS	Organisation of Eastern Caribbean States
PAHO	Pan-American Health Organization
PDNA	Post-Disaster Needs Assessment
PUC	Public Utilities Commission
PY	Person Year
RESEMBID	Resilience, Sustainable Energy and Marine Biodiversity Program
SDGs	Sustainable Development Goals
SDMX	Statistical Data and Metadata eXchange
SIAP	Statistical Institute for Asia and the Pacific
SIDS	Small Island Developing States
SOP	Standard Operating Procedure
UNECE	United Nations Economic Commission for Europe
UNECLAC	United Nations Economic Commission for Latin America and the Caribbean
UNESCAP	United Nations Economic and Social Commission for Asia and the Pacific
UNOSAT	United Nations Satellite Centre
UNSD	United Nations Statistics Division
WHO	World Health Organization

EXECUTIVE SUMMARY

Caribbean countries are exposed to high levels of risk from meteorological hazards, which have a negative impact on their economic and fiscal stability. These natural hazards are being exacerbated by the adverse impacts of climate change—intensifying hazard patterns and increasing stress on water availability, coastal investments, and livelihoods. The high costs of recovery and reconstruction have resulted in increased debt, unsustainable budgetary deficits, and unreliable funding streams for many countries in the Caribbean. The absence of macroeconomic stability makes it harder for Caribbean countries to implement poverty reduction policies.

Anguilla is a microstate, one of the five UK Overseas Territories in the Caribbean. Lying at the northern end of the eastern Caribbean Island chain, it covers an area of 90 km2. Anguilla has a gross domestic product (GDP) of US$281.1 million and a population of 17,422. While tourism and construction are the main economic drivers and employment suppliers in Anguilla, the economy is fueled by luxury tourism, offshore banking, fishing, and remittances. Services, primarily tourism, account for 76.3 percent of GDP, followed by industry, which represents 21.3 percent, and agriculture, 2.4 percent of GDP.

This report provides detailed technical guidance to the Government of Anguilla (GoA) on improving its capability to assess the macroeconomic, fiscal, and socioeconomic impacts of disasters. The impetus for this project arose from a request for technical assistance from the GoA and the Anguilla Statistics Department (ASD) for assistance in compiling statistics required for assessing the impacts of disasters. The major challenge faced by the ASD was obtaining the required information from private and public sector entities in Anguilla and then compiling this information using statistically rigorous methods which meet user requirements. The specific questions asked in the preparation of this project were the following:

- What is the local enabling environment, concerning legislation and policy, for data governance?
- What capacity exists in the GoA, and the Anguillan private sector, for statistics development and full life cycle data management?
- What strategies and plans exist for information and communication technology (ICT) and statistics development within the GoA, and to what extent are these actively implemented?
- What is the ICT infrastructure and capacity, in terms of hardware and software, available within the GoA and the Anguillan private sector?
- What statistics are required by key stakeholders within the GoA, and how can these statistics be assembled according to international best practices?

A series of assessments, using interviews, surveys, training sessions, and workshops, were conducted to answer these questions. The answers informed the development of tangible and actionable technical products which the ASD can use to assess the socioeconomic impacts of disasters. These included the following.

- **A complete set of disaster-related statistics for Anguilla** based on the United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP) Disaster-Related Statistics Framework (DRSF). These statistics were determined from the application of international best practices tailored to the unique conditions of Anguilla based on consultations with local stakeholders through a technical working group. These statistics capture a full spectrum of socioeconomic impacts of disasters and can be used to inform post-disaster impact assessments.

- **The design and prototype of a database for disaster-related statistics** for the ASD are based on disaster-related statistics. The design and prototype are constructed in a way that allows the GoA to further develop the database depending on needs and available resources. Significant guidance has also been provided with regard to operating procedures, standards, and the level of effort required for the implementation and maintenance of the system. Datasets from within Anguilla as well as from regional and international sources have also been identified for possible inclusion in the database.

- **The design and implementation of a metadata profile and catalogue** to accompany the database. This catalogue is based on the application of widely used international metadata standards to the specific needs of Anguilla, determined through a series of consultations with data managers and users within the GoA. In addition to the catalogue, standard operating procedures (SOPs) were developed, and training was offered on these procedures and the catalogue to GoA staff. The catalogue has already been used to document datasets for inclusion in the DADS.

The tools and instruments identified within this project support the Chief Statistician in their drive to improve how statistics can be obtained for disaster management, encourage interdepartmental cooperation and liaison, and smooth data exchange. Proper application and further development of these tools would be extremely beneficial to improving the capacity of the GoA in the assessment of disaster impacts. However, there are obstacles to implementation, including the following:

1. Limited capacity within the GoA to implement and maintain this, and other, systems. This includes available personnel and the level of training required for personnel.

2. Fundamental legislation and policies for data governance should be developed to better enable agencies to collaborate and enable the public and international organizations to formally access GoA data.

3. A lack of a strategy for ICT which addresses the development of the national statistical system and the integration of systems. The GoA has made progress toward establishing a robust government wide area network (G-WAN), but systems being established on this network are not being established in a way that would facilitate automatic data sharing with the ASD.

Overcoming these obstacles will take significant strategic effort on the part of the GoA. The technical working group established as part of this project can assist with improving coordination among agencies and in spreading awareness of the need for improved statistics. There are also international and regional opportunities for technical assistance and knowledge sharing from which Anguilla can benefit.

1. Introduction

In 2017, Hurricane Irma caused severe damages and losses in Anguilla. One life was lost, and homes and businesses across the island were severely damaged.

The lead agency for this initiative is the Anguilla Statistics Department (ASD), which is the Central Statistics Office (CSO) of the country. The mandate of the Statistics Department is set out in Section 2 of the Statistics Act of 2000. Fulfilling their mandate requires collaboration with many agencies within the Government of Anguilla (GoA). Due to a lack of official policies around data sharing within the GoA, the CSO relies on ad hoc arrangements with other agencies around the sharing of data. This has resulted in a lack of interoperability between various information systems and has required significant amounts of extra staff time to transform datasets into a form appropriate for the creation of statistics. In the wake of a disaster, these ad hoc arrangements prove even more difficult.

The GoA has a centralized information and communication technology (ICT) unit (Department of Information Technology and E-Government Services [DITES]) tasked with operating the government-wide area network (G-WAN), computer servers, and voice over internet protocol phones, as well as developing and maintaining e-government systems. However, there is no national ICT strategy that would clearly define the priorities and initiatives of the GoA with regard to ICT. The lack of a guiding strategy with associated policies and standards has led to agencies developing information systems that are not designed to be interoperable. According to the 2019 budget, the GoA has a number of in-progress information systems projects, including an upgrade to the Financial Management System (Smartstream), a Revenue Information System, a beneficial ownership system, a tax administration information system, and a data portal (the Anguilla Data Gateway). The need for guiding policies is appreciated by DITES and the CSO. DITES has produced a proposal for the development of an ICT strategy, and the CSO has been working on the development of the national statistical system since 2014. With limited resources available to complete the system, progress has been slow.

To build resilience, access to good data—the right data to inform decisions—is crucial. To produce statistics in the aftermath of a disaster, the process for data sharing needs to be as seamless as possible, and interoperable systems are key to this. However, the CSO does not have a clear picture of the available information systems and applications in-house, which would be useful for this purpose, and the degree to which they are interoperable from a technical, policy, or enterprise culture point of view. The desired outcome of this project would be to create a springboard from which a response can be launched by the GoA to dynamically respond to the need for data to make decisions on time, with the knowledge and comfort that the data used are of high quality, mainly coherent and consistent after a disaster.

The main objectives were the following:

(a) To build the capacity of the ASD to be able to produce statistics required to assess the impact of disasters, including damages, losses, and financial expenditures.

(b) To investigate the feasibility of integrating systems, and data sources, the key to the supply of disaster impact data.

This report contains the results of diagnostics and recommendations based on these objectives to assist the GoA in increasing its ability to create and manage the data and statistics necessary for assessing the impacts of disasters.

2. Approach

The approach for meeting these objectives involved two activity streams: data and systems and statistics development. Within these, assessments took place through interviews, workshops, and meetings of a technical working group of GoA officers created for disaster-related statistics development. Based on the assessment activities, a design of a database for disaster statistics in Anguilla was produced along with a metadata profile for GoA data and a data inventory for populating

the database. This assessment and the accompanying products will be used by the GoA to develop its capacity in determining the financial impact of disasters.

2.1 Data and Systems

Effective and sustainable development of statistics requires a reliable and consistent data pathway supported by necessary infrastructure, qualified personnel, and policies. Attempting to assess each of these aspects in isolation would potentially result in missing obstacles and gaps blocking the creation of disaster-related statistics. The investigation of governance sought to determine the existence of fundamental enabling policies and legislation concerning data sharing and management as well as the capacity within the GoA and private sector for data management. The assessment was structured along the four main sectors (Infrastructure, Social, Environmental, and Economic) used by the United Nations Economic Commission for Latin America and the Caribbean (UNECLAC) Post-Disaster Needs Assessment (PDNA) methodology. A list of interview questions was drafted to assess the digital data management capabilities of the identified agencies in the following five areas:

1. Data governance
2. Data management protocols (based on the data life cycle process)
3. Personnel capacity
4. Data management infrastructure
5. Metadata standards.

Knowledge of relevant data that support disaster-related statistics, that is, metadata, was pointed out as vital in building the capacity of the ASD. Production of a simple catalogue was identified as an appropriate vessel for that information within the time frame of the current project—what data are available, where they reside, their quality and extent, and how they support the provision of disaster-related statistics. A metadata catalogue and profile were developed through the following activities:

1. Development of a schema built on international good practice that describes the various types of data needed to supply disaster-related statistics and be appropriate and relevant to Anguilla.
2. Design of a metadata catalogue of the documented data that conforms to this schema, so it is ready to be migrated to other schema and formats with minor modifications.
3. Identification of the environment in which it is necessary for having effective metadata documentation. While the policy-level interventions are inextricably linked with the general environment for data sharing described elsewhere, a system for recommended protocols for how to manage and steer the metadata process and individual procedures (serving as a manual for operation) was outlined in an SOP Manual.
4. Raising of awareness and training of the need and operation of the metadata catalogue. Wider awareness-raising was given during the general training sessions; then, a more focused group from the ASD and the Department of Physical Planning (DPP) were given more in-depth training on metadata, the catalogue, and tools/documentation provided.
5. Identification of the parameters needed to transfer this into a web-based solution. While the request was focused on the technological solution and design of any server framework, wider considerations were needed covering both the technical capacity of the GoA and operational matters.

2.2 Disaster-Related Statistics

Having statistics on disaster vulnerability and post-disaster impacts is a priority for the ASD, particularly in the wake of Hurricane Irma, where the department was expected to provide immediate information on the hurricane impacts. The

department found it did not have information readily available to produce the necessary statistics and did not have an understanding of the quality of the information available from other agencies. This series of activities focused on the ASD and its ability to assemble statistics based on the UNESCAP DRSF. This framework is relatively new, first published in 2018, and outlines a generic framework for the development of disaster-related statistics. This framework has not been tried in a Caribbean country until now. To do this, the capabilities and operations of the ASD were reviewed through two activities. The first was a generic assessment of the department's operations compared to international best practices for statistical business practices. The second was an assessment of its operations specifically as they relate directly to disaster-related statistics.

Assessment of the ASD's operations with respect to disaster-related statistics was undertaken using a modified version of a tool first developed by the United Nations Statistics Division for assessing environment statistics. This tool—known as the Environment Statistics Self-assessment Tool (ESSAT)—comprises two parts, an institutional-level assessment, and a statistical-level assessment

The ESSAT was chosen as the starting point for the assessment of disaster-related statistics in Anguilla for two reasons. First, the tool is well established and has a track record of application in assessing statistics around the world.[1] Second, the ESSAT was designed specifically to assist countries in building their statistical programs following the United Nations Framework for the Development of Environment Statistics (UNSD 2017). The FDES includes disaster-related statistics as one of its six components, so the ESSAT provided a logical starting point for developing a self-assessment tool for disaster-related statistics. It required only minor modifications to serve this purpose. The modified version of the ESSAT, labelled the Disaster-related Statistics Self-Assessment Tool (DSSAT), can be found in Annex 4.

The DSSAT, like the ESSAT, is intended to be applied through a multi-stakeholder consultation process led by the ASD, involving all the main stakeholders in the production, dissemination, and use of statistics. This multi-stakeholder approach facilitates agreement on a common assessment of the situation in the country in a clear and transparent manner. Thus, with the agreement of the ASD, a multi-stakeholder group—the Anguilla Disaster-statistics Data Assessment Working Group (the Working Group from here on)—was constituted with the mandate to apply the DSSAT in Anguilla and report on its findings (see Annex 5 for the terms of reference of the Working Group). The Working Group met three times for the purposes of completing the DSSAT. The group was chaired by the Chief Statistician, Ms. Alleyne-Franklin. Members of the Working Group are listed in Annex 6.

3. Data and Systems

This chapter presents an assessment of the current state of play in Anguilla with regard to the enabling elements required, including essential legislation, physical computing, and networking hardware and software, institutional policies governing data sharing, and institutional capacity in terms of personnel and roles. Before the commencement of this assessment, metadata were identified as a key tool for enabling data exchange. As part of this assessment, a metadata profile and database were developed based on the data holdings identified from the assessed agencies.

3.1 Data Governance

Data governance is the set of principles that sets out who has the authority and control over data assets within any organization and determines how those assets can be used. The amount and complexity of data have increased exponentially with both digitization and the ability to share over the internet. Organizations and members of cross-cutting networks (for example, departmental government) who lack data governance are both frustrated by not having visibility of useful and available data for their operations and spending unnecessary time, money, and resources on data gathering and discovery which could be used elsewhere. What is worse is that much valuable data are lost or non-transferable, and routinely collected data change format so they cannot be compared easily over time or have been lost when individuals have moved in their career. Anguilla currently does not have legislation governing data sharing, use, or security. Citizens and organizations have no legal right or pathway to accessing government data, and there is no legislation or policy within government defining data security or privacy rights.

1 https://unstats.un.org/unsd/envstats/meetings/-2019Caricom Region/documents/Session 1.3.1 UNSD - Implementation of the FDES 2013.pdf.

Table 1: Enabling Legislation of Data Producing Agencies

Title of Legislative Act	Date Adopted/Revised
Disaster Management Act	May 1, 2008
Disclosure in the Statement of Account Regulations	December 15, 2006
Education Act	April 1, 2009
Electricity Act	February 6, 2003
Financial and Accounting Regulations	December 15, 2002
Fisheries Protection Act	December 15, 2000
Health Authority of Anguilla Act	January 1, 2004
Land Surveyors Act	November 1, 2010
Telecommunication Act	June 10, 2003
Telecommunications Administrations Account Regulations	December 15, 2006

3.1.1 Social Sector

The social sector agencies assessed for data governance were the Ministry of Infrastructure, Health Authority of Anguilla (HAA), Department of Education (DoE), and Department of Disaster Management (DDM). In assessing their data governance legislation, none of the agencies in the social sector has legislation related to data governance. However, all these agencies do have acts that have regulations that guide what data should be collected.

When assessing the data policies, most social sector agencies again indicated that they have no data policies or official steering committees that deal with data governance. The HAA has a Health Authority Board that performs some data governing functions. The board determines what data are collected and how they should be used.

The HAA is also the only agency in this sector that has formal agreements signed. This agency has memorandums of understanding (MoUs) with various agencies with which they share essential health-related data. These agencies are the Ministry of Health, the Pan-American Health Organization (PAHO), and the Caribbean Public Health Agency (CARPHA). These MoUs were created with the input of the Chief Executive Officer (CEO), Ministry of Health, and the Attorney General. All MoUs for the HAA are signed by the Ministry of Health in collaboration with the HAA. All MoUs that require a legislative change must be approved by the Executive Council (EXCO), a ministerial body that includes all the ministers of government and the Governor.

3.1.2 Infrastructure Sector

The infrastructure sectors that were assessed for the data governance were Public Utilities Commission (PUC), Anguilla Electricity Company (ANGLEC), Flow, and DITES. Under this sector, PUC, ANGLEC, and FLOW all indicated that they have legislated regulations around data governance.

These agencies all have data governance policies created by their Board of Directors, which guide them on how they should collect, store, and disseminate their data. These agencies' boards also function as their data governance bodies. They determine what data should be collected and how they should be used.

DITES does not have any defined data policies; however, it does have network user policies on how users should operate on the GoA network.

The PUC and the telecommunication agency of FLOW have a formal sharing agreement. All the telecommunication companies must sign a Licensing and Frequency Agreement, which is part of an Operations Contract, with the PUC. This agreement mandates telecommunication companies to supply essential data for regulatory purposes.

3.1.3 Economic Sector

The agencies assessed included the Ministry of Finance, Agriculture Unit, Economic Development Unit, Anguilla Hotel and Tourism Association (AHTA), and Anguilla Chamber of Commerce and Industry (ACOCI). These agencies neither have any data policies related to data governance nor have an official data governance body. However, some of the agencies do have governing bodies that perform data governing duties. The Ministry of Finance and the Economic Development Unit have a Financial Economic Planning Committee that uses their financial and fiscal data to make financial projections.

When assessing the formal agreements in the economic sector, it was observed that there are limited formal sharing arrangements between the agencies in the economic sector with other public or private agencies. The Ministry of Finance did not indicate that they have any formal agreements with any agency; however, they are expected to supply financial information monthly to the Eastern Caribbean Central Bank, and upon request, to the Caribbean Development Bank, International Monetary Fund, and UK Foreign, Commonwealth & Development Office (FCDO). The AHTA, despite having 185 members, does not have any official sharing arrangements with its members.

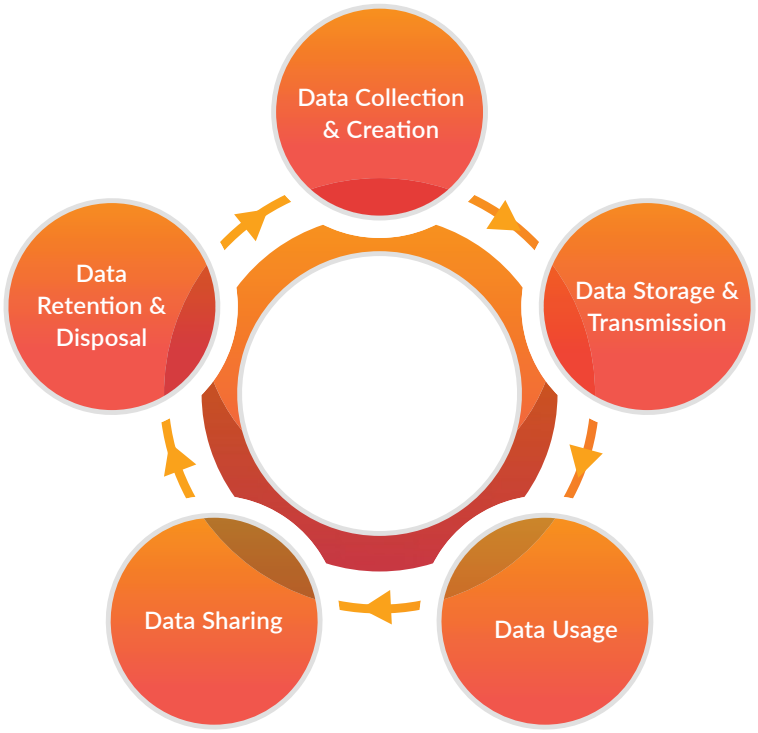
3.1.4 Environmental Sector

Like most other sectors, agencies under the environmental sector do not have any data governance legislation. In assessing the data governance policies, the DPP, Department of Lands and Survey, and the Environmental Unit do have data policies in relation to how data are collected and stored. The Department of Lands and Survey has a Lands Act that guides its data governance regulations. The DPP did not provide any of the policies it has related to data governance. The Department of Health Protection (DOHP) does not have any data policies; however, it follows the World Health Organization (WHO) policy guidelines when collecting data.

The Environmental Unit, an agency under the Department of Natural Resources (DNR), was the only environment agency that indicated it has a formal information-sharing agreement. This agreement was signed as part of the Anguilla Natural Capital Accounting Project that was implemented with assistance from the Economics for the Environment Consultancy Ltd., a UK-based agency. All the data from this project were shared with the sponsoring agency through a formal agreement. The Environmental Unit does not have any formal agreements with government agencies, and data are only shared upon request.

3.2 Data Management

Figure 1: Data Lifecycle Management Process



This part of the assessment investigated whether data life cycle management (Figure 1) is known and applied by agencies. For each step of the process, the following were investigated:

(a) For data collection and creation, defined roles and responsibilities for data collection as well as any standard procedures

(b) For data storage and transmission, access restrictions and security measures used around the storage and transmission of information, such as encryption, and the existence of data backup and recovery procedures

(c) For data usage, existence of procedures to monitor data usage and gather feedback from end-users

(d) For data sharing, all data-sharing arrangements, both formal and informal

(e) For data retention and disposal, how frequently datasets were erased and whether there are any defined disposal techniques (that is, overwriting deleted datasets).

Generally speaking, data management in Anguilla can be subdivided into themes: data holding, data sharing, data quality assurance, and strategic data management.

Data holding. This implies that individual agencies hold data. The way data are held varies from agency to agency and varies from paper records through individual data forms and worksheets (for example, in Excel) to geographic information system (GIS) files. Integrated database holdings and those shared across agencies are few and far between. In many agencies, the method of data holding is more connected to how the data are collected and not how they can be analyzed or disseminated. Only the ASD and DPP demonstrated any datasets which gave more consideration to end-use or presentation/dissemination.

Data sharing. The majority of data sharing is done through the bespoke release of data upon written requests, which have to go through formal channels, often involving at least heads of department and probably ministry (Permanent Secretary or

ministerial level). The ASD shares some data through its website, well presented and with useful minimal metadata, but it has to go through rigorous processes to display these, given they are considered official statistics. However, even with this facility, some of the links to the data are broken, or downloadable data are not available for download.

Data quality assurance. Only the ASD demonstrates care and attention to upholding data quality assurance. Its role as the provider of official statistics which are used both by the public and for top-level governmental decision-making has placed an onus on the ASD to deliver data that can be validated, and the sources and assumptions used to create the data are documented.

Strategic data management. There is no agency within the GoA which is mandated to actively manage data across the whole government. While DITES puts forward the system for intranet data holding, it has no responsibility for the software, format, or structuring of data within each department's holdings. Nobody looks across government to identify data programs, collection targets, dissemination portals, and quality assurance. The Anguilla National Statistics Advisory Committee (ANSAC) plays a role in identifying the data needed but at present does not follow through with a complete set of guidelines to manage these other strategic needs.

The following sections present sector-specific findings of this assessment based on the data life cycle process.

3.2.1 Social Sector

In the social sector, the following agencies were assessed; the Ministry of Infrastructure, HAA, DoE, and the DDM. The DoE indicated that it is familiar with the data life cycle model and its processes. However, the DoE could not confirm that it is using any of the processes from the data life cycle. The HAA is not familiar with the process; however, it is internally drafting data collection roles, responsibilities, and procedures to assist with data management. These policies are being drafted by an internal body with the assistance of the Ministry of Health.

All agencies were found to have implemented some level of data management security measures. These agencies are on the GoA network, which has a centralized data management system. All their data are stored on a centralized server, and presently none of the GoA data are in the cloud. All the agencies access the GoA network through fiber optics except the HAA clinics that access the network through wireless links. Data are exchanged using email. The data are archived after five years, and no deletion of data takes place.

The social sector agencies have all digitized their data except the HAA and DoE. The HAA is presently implementing a new health administrative system software called Evident, developed by CPSI.[2] This new system will allow the agency to digitize all its health information data and generate management reports. The DoE is implementing an Education Management Information System (EMIS)[3] for the collection, integration, processing, maintenance, and dissemination of data. The EMIS is being implemented with assistance from the European Union.

3.2.2 Infrastructure Sector

In the infrastructure sector, the ANGLEC and FLOW indicated that they know about the data life cycle and apply its data management processes. Both of these infrastructure agencies have defined roles, responsibilities, and procedures for data collection. The Water Corporation is familiar with the data life cycle but does not apply any of its processes for data management.

FLOW and ANGLEC are agencies that have private secured networks. All the data collected by these agencies are stored in the cloud. The semi-autonomous agencies such as Water Authority and the PUC are on the GoA network, where the data are stored on a centralized server. Most of the agencies in this sector share their electronic data by email. The data collected by the PUC from the utility companies are delivered as paper files. These files are then digitized by the PUC. All the agencies in this sector archive their data. FLOW and ANGLEC archive their data after six years. DITES houses all the GoA data in a secured room on file servers. The servers are secure in a reinforced concrete room with a flood-proof metal door. DITES

2 https://www.evident.com/better-care/thrive-EHR-for-hospital-clinic/.
3 https://learningportal.iiep.unesco.org/en/glossary/educational-management-information-system-emis

presently does not have any data in the cloud due to financial and human resource constraints.

All the agencies in the infrastructure sector have their data digitized except the Water Corporation. This agency has data in digital and paper-based form. The meter reading and billing information are digitized; however, the asset inventory information is paper-based.

All agencies that are part of the GoA network have their data stored on a file server that is located at DITES. DITES executes a nightly backup of all GoA digital data on a backup server.

3.2.3 Economic Sector

In the economic sector, none of the agencies are familiar with the data life cycle model processes. All the agencies in this sector have data management security measures except the AHTA. The Ministry of Finance and Economic Unit are GoA agencies and have their data secured on the centralized server at DITES. These two agencies follow the guidelines of the Financial Audit Act, which stipulates that all data should be kept for seven years, after which is the data are archived. All the data shared by the agencies in this sector are sent by email. The AHTA and ACOCI have their own private networks, which are maintained by private companies. These two agencies back up their data weekly or when requested.

The Ministry of Finance, Economic Unit, and Agriculture Unit have all their data in digital format. These agencies are also on the GoA network through fiberoptics and have their data backed up nightly. The Ministry of Finance has recently implemented a Multi-tax Information System. This system helps administer all taxes. The system will also provide intelligence reports and has been integrated with a smart stream financial system. Integration with the Land Information System currently in development is being implemented by the software provider and will be managed by DITES. There are also plans to have the system integrated with the Automated System for Customs Data (ASYCUDA).

The data produced by the Ministry of Finance and Economic Unit are given to the UK Financial Advisor, who is from the Ministry of Finance. This individual ensures the financial data produced are accurate, reliable, timely, and comprehensive. The Financial Advisor attends all EXCO meetings and reports all the financial information to the Government Ministers and Governor when requested. The financial information is also provided to the FCDO upon request. The FCDO verifies the accuracy of the financial information.

3.2.4 Environmental Sector

The agencies in the environmental sector are all on the GoA network. Data management is similar to the GoA agencies in the other sectors where they are on the G-WAN, and the data are located at the centralized GoA DITES. The Department of Lands and Survey is presently implementing a new Land Information System. This system will allow them to digitize all their paper-based cadastral information. This system will also be integrated with the new Tax Information System specifically to generate property tax information.

The Department of Lands and Survey, DPP, and Environmental Unit are among the list of agencies that have spatial data. Each of these agencies has its spatial data disaggregated by cadastral units and updates its datasets regularly as a function of its operations. The Fisheries Unit also has spatial datasets which are disaggregated by points. The spatial datasets used by the Fisheries Unit are updated when needed.

The Fisheries Unit was the only agency in the environmental sector that is familiar with the data life cycle processes. Like some of the other sectors, all the agencies in the environmental sector are GoA agencies. These agencies are on the same GoA network that uses a centralized server which has a central identity management system. All the data shared by these agencies are sent by email. Like the agencies in the other sectors, data collected are not deleted and are archived.

For most of the agencies, the data collected are used to create planning documents and reports. Table 2 shows a list of the agencies and the reports they develop.

Table 2: Developed Reports

Agency Name	Reports and Planning Documents Developed
Ministry of Finance	Budget Report http://www.gov.ai/documents/2020%20Budget.pptx
DOHP	DOHP Annual Report
Unit of Disaster	Anguilla Natural Capital Accounting Report https://data.jncc.gov.uk/data/62e82176-5053-4733-bd05-e33b15bbf88c/ot-nca-sup-sat-19-ang-nca-jun-2019.pdf
HAA	HAA Annual Health Information Report
Economic Unit	Mid Term Economic and Fiscal Plan
PUC	PUC Annual Report https://pucanguilla.com/wp-content/uploads/2020/04/Annual-Report-2018-Dec-13.Final-.pdf
ANGLEC	Annual Report, Financial Statements https://www.anglec.com/documents/annrep2018.pdf
Water Corporation	Financial Reports
DPP	Anguilla Development Manual

The DoE was the only agency that uses a quality assurance process for its data collection. The quality assurance process, however, is not monitored or evaluated due to insufficient human capacity. The infrastructure agencies of FLOW, ANGLEC, and PUC all have quality assurance processes that they follow in collecting their data. These three agencies also monitor and evaluate their quality assurance processes. All the agencies assessed indicated that they do not have a dedicated quality assurance officer to help monitor and review quality assurance procedures.

3.3 Personnel Capacity and Roles

Personnel capacity assessment looked at the individuals in the agencies that perform roles related to data management. For all GoA agencies, job functions and descriptions are defined by the Head of Department with input from the Permanent Secretary. The individuals applying must be approved by the Public Service Commission, a board appointed by the Deputy Governor and responsible for all appointments and transfers of GoA employees.

In the DPP and ASD, there are people who are given the responsibility to manage data, but they are not formally appointed data managers, and many have received limited training in data management. Limited human resources within Small Island Developing States (SIDS) governments preclude significant specialization of job roles, and individuals have to cover multiple responsibilities within a department.

The data management duties in most of the agencies are performed by individuals whose main function is not related to data management. The HAA and the DPP are the only agencies that have personnel who are in official data management positions.

Table 3 provides details of the individuals who perform the data management function in the agency and their substantive roles.

Table 3: Data Management Roles

Agency Name	Name of Role for Data Manager
Ministry of Finance	Budget Officer
PUC	Regulatory Officer
DoE	Education Planner
DDM	Deputy Director
DOHP	Principal Environment Officer
Water Corporation	Water Engineer
DPP	Senior GIS Officer
DITES	IT Manager
Fisheries Unit	Natural Resource Officer
HAA	Health Information Officer

All the individuals listed in Table 3 are in their substantive job role permanently, and data management is one of their official functions. The Education Planner is presently heading the implementation of an EMIS. In this position, the Education Planner will be responsible for collecting and processing all the data collected by the DoE. The AHTA and the ACOCI did not indicate that they employ anyone in a data management role.

3.4 Metadata

The metadata assessment examined how the agencies organize their data and information products to determine if it is done in a logical scheme. An assessment was also done to identify if any data dictionaries or registers exist.

Metadata are available in only a small number of the agencies. The DPP manages metadata for the Fisheries Unit and Environmental Unit and for itself (a legacy of the Coastal Resource Inventory System established in 2005/06). The metadata that is being managed for these agencies uses a built-in software template to manage the data based on ISO 19115 standards. They have a logical naming scheme and a classification scheme for the datasets. The international metadata standard used at the DPP complies broadly with ISO 19115 and uses Federal Geographic Data Committee (FGDC) schema with some local fields. The Department of Lands and Survey manages its own metadata. It has a logical scheme for data organization and the datasets classification and uses the same international metadata standard as the DPP. The ASD indicated it has a metadata model created under a project specifically for the management of health sector data.

3.4.1 Metadata Schema

The design of a metadatabase schema for Anguilla disaster-related statistics incorporated the need to develop a way of practically logging the datasets and statistics seen as important for reporting balanced against the rigor needed for international best practice. A detailed review of several international schema was conducted, consolidating the database design, descriptions, and data models from each individual standard schema. These schema were the following:

- Top-level generic metadata 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 was 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.
- There is a significant overlap of top-level information, in particular between different schema. Adoption of mandatory Dublin Core metadata elements allows the main elements to be covered for all types of data. The other schema was selectively applied in practice, dependent on the nature of the data to be documented.

To complete the schema, appropriate elements from each of the international schema were documented in a master document specifically for Anguilla. A small number of additional non-standard elements were added to support metadata management. Table 4 shows the methods of documentation for metadata elements.

Table 4: Metadata Element Documentation

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 would be machine-readable
Item Name (Terms)	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, 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 (definitions)	A description of the element. Where possible, these descriptions come from the official schema description.

The detailed logic for the schema is included in the schema document (Annex 1 - Metadata Schema Design Document),

which includes the need for adopting international standards, the factors which guided the design of the current schema, and an acknowledgment of the different data types which could be covered.

The schema contains a table describing each element to be documented. It is recognized that some international schema element descriptions are neither immediately clear or ambiguous in their definitions. Where possible, guidance and examples (through links to schema websites) have been given to help clarify the use of each element.

Ancillary tables were identified to support the main metadata table. The primary of these was a contact table (based on ISO 19115 formatting, given it has the most comprehensive documentation of contact details). Both the contact and other tables were added to provide restricted pick-from lists or look-up tables for data entry. Where possible, look-up tables used defined values as in the appropriate international schema and are labelled with standard table names. This should allow ease of transfer if those fields are adopted in any server solution or made interoperable with other metadatabases.

Finally, in the schema document, a comparison is made to map the Anguilla elements against other schemas. As well as the original schema element source, names of the corresponding element label for the other schema are documented here. In some cases, the mapping is not exact, and those fuzzy matches are picked out in italics in the comparison table. This element mapping can serve in the future for any interoperability considerations if the metadata for a particular dataset need importing into an alternative schema format

3.4.2 Metadata Catalogue

For the project, it was agreed to develop the metadata catalogue (Annex 2 - Metadata Catalogue) in an Excel format as that software is widely used throughout the GoA and could allow some control of data validation, data formatting, use of dropdown values, and other tips and supporting methods to ease the data entry process. However, due to its flat file structure, Excel had some limitations in describing more complex metadata relationships. For example, there are several metadata elements where there could be many values related to one value of another element (for example, for one dataset, you may have several classification methods and definitions of variables being used, and for each variable, you may have several subclasses, definitions, and values). It is difficult to produce rigorous representations of those relationships within Excel. The practice of using special characters (commas and semicolons) to separate multiple values stored in one Excel cell provides some future-proofing if the data were to be transferred (that is, parsed) to more appropriate metadata software (for example, pycsw).

All the components needed to manage the metadata were placed in one file for data entry, supporting tools, internal reporting, and mapping to external requirements (for example, the agreed set of statistics).

Elements Considered in the Design of the Database

Within the worksheets, the following design features were considered to help data entry and reporting:

- Color coding of worksheet tabs - grouped to show data entry (orange), supporting look-ups and other management tabs (grey), reporting (blue), and mapping to reporting requirements (green).
- Color coding of field headers to guide data entry to the mandatory elements (red), conditional (pink), advisory (yellow), and optional ones (green).
- Dropdown menus where values had to be limited to a fixed set of values.
- Tool tips on both data entry and element headings to give some information about data to be entered (and guidance on how to enter data related to conditional fields).
- Data validation tools to guide data entry staff—this involved restricting the types of data that could be entered (for example, date or numeric) and lower and upper values for numeric fields. When incorrect data are entered, error or warning messages were designed to guide users. Note that in some cases, the data validation was

relaxed as it is difficult to preempt the range of values that were given (especially numerical values), but the database was designed such that metadata managers could enter extra categories through the look-up table worksheet and retain control over data validation in text fields.

- Some measure of metadata completeness - using simple calculations of the percentage of mandatory fields completed for each section.
- Some measure of catalogue protection - to avoid general users from overwriting or corrupting key parts of the metadata catalogue. Due to some limitations with Excel, this was not entirely successful; protection can cause the protected fields at the extremities of the worksheets to be invisible and inaccessible, and some useful features—sort and filter—become unavailable once table header cells are protected.

Two methods have been used to allow data entry, as a separate entry form which can be used either by any data provider or in an interview with a trained metadata cataloguer. Or else, information can be directly entered into the catalogue. There are pros and cons with each, but both methods are within the data schema, and extensive documentation on the process is provided in the SOP Manual.

3.4.3 Metadata Portal

The rationale for the metadata portal, identifying the parameters to be considered, the recommendations for how to structure metadata in the future (building on the lessons learned within this project), and a suggested architecture are covered sufficiently within the memorandum for the creation of a web-based metadata server (Annex 3 - Memorandum for the Development of a Metadata Portal).

3.4.4 Data Identified

For the population of the metadatabase, the following search strategy was taken, which contained several parallel approaches.

1. Some basic data had been identified as vital at an early stage of the project (from experience and initial engagement with key stakeholders such as the ASD). Other data were identified early on in the Working Group meetings and could be searched for, and finally, once the working group agreed on the final list of statistics and prioritized their importance, a further round of data requests could be made.

   a. In this assessment of data to be identified, the metadata catalogue did not just record the final disaster-related statistical dataset but would look for raw data/depersonalized microdata, which could contribute to a final statistic with simple analysis. This could be datasets that contribute to some measure of vulnerability (for example, administrative data combined with land-use types and exposure to particular hazards) or preparatory data for recording losses in the next disaster (inventory of schools or health facilities, databases to store needs assessment data).

2. Data were provided through several channels:

   a. From previous work in Anguilla, the team was able to draw on existing data and metadata from the 2004–06 Anguilla Coastal Resource Assessment, Monitoring and Management (ACRAMAM) Project, which used ISO 1995 type metadata to catalogue GIS data. It was originally envisaged that this catalogue would form the basis for a National GIS, and the archive has been kept by the DPP since the ACRAMAM project was completed.

   b. Attention was brought to the team of online resources (mainly from the ASD), which had useful minimum requirement metadata.

   c. From the working group and literature searches (including the architecture assessment), requests were sent to several agencies for access to data, including nongovernmental sources.

3. Data were documented using the following strategies:

a. Where metadata existed, it was transferred into the metadata catalogue under the appropriate elements.

b. Access to the original datasets (online or through data requests) allowed several metadata elements to be documented, for example, file size and format.

c. Population of some metadata elements from metadata documentation experience, for example, using the standard scheme guidance.

d. The catalogue has been shared at the training, and feedback on the metadata documentation has been requested.

4. Data from previous projects and from online sources have been documented, but where these datasets have been superseded, only the most recent version of the data has been documented, unless they are part of a recognized temporal sequence.

5. Official statistics and data from government sources were prioritized over other sources for documentation. Primarily data for the DSSAT were prioritized.

6. Several data gaps exist. Where data have been reported to exist through the Working Group sessions or other correspondence, a record has been entered in the metadatabase, but the detailed metadata is left blank. Data gaps have been identified as prioritized to provide a full set of disaster-related statistics. These details are covered by the Statistics Data section of this report (Section 5).

3.4.5 Standard Operating Procedures

A single document was produced (Annex 11 - Metadata Standard Operating Procedures Document), which can serve as both a manual and a broader set of SOPs. The Document covers the metadata management environment (i.e., needs for metadata, roles, and responsibilities), puts the disaster-related statistical data and metadata in context with other international and national data holdings, then describes the metadata catalogue structure, a potential model for management of metadata and provides an overview of the metadata capture process.

For each of the elements of metadata management, individual procedures were written, outlining the aims of each procedure, who should be responsible for executing the procedure, the steps to be taken (where possible, verbose enough that most with some data experience can follow) with selected screenshots, a list of the expected outputs from executing the procedure, and any extra notes to help understand and connect with other procedures.

The range of procedures identified included the following:

- Identification of new reporting needs and which data may be needed
- The initial data capture process
- Quality control of data entry
- Changes in data/updates (for example, from change of responsibilities)
- Annual review checks
- Any modification of the metadatabase schema
- Use of the catalogue to search for, filter, and report where data exist
- Some maintenance tools for adding to the SOP Manual and protecting elements of the catalogue.

3.4.6 Training

Three elements of training have been delivered on metadata throughout the project:

1. At the introductory meeting, a presentation on the planned program was given and some of the parameters for the design of the metadata catalogue were shared.

2. At the general statistics training, some introductory awareness raising was given to the wider stakeholder group about the need for metadata and the prototype catalogue with initial data.

3. A targeted two-hour metadata training session was designed for the ASD and DPP staff which provided more details about the purpose and importance of metadata, tools within the data catalogue, and identification of how it could operate. Several practical sessions were interspersed within the overall training to help students discuss the metadata component (the curtailed set of working group sessions meant this was the only real opportunity to delve deeper into the metadata catalogue elements). At this time, the schema was redistributed along with the SOP Manual and dummy versions of the catalogue to allow students to explore in their own time.

4. Development of Disaster-Related Statistics

This chapter presents the results of the review of the operations of the ASD, the proposal for a basic range of disaster-related statistics for Anguilla, an assessment of the gaps in Anguilla's current statistics compared to this basic range and the priorities for filling these gaps, and the conceptual design of a database to house disaster-related statistics in Anguilla. These outputs build on the assessment and outputs of a data inventory and metadata profile presented in the previous chapter.

4.1 Generic Assessment of Statistical Operations

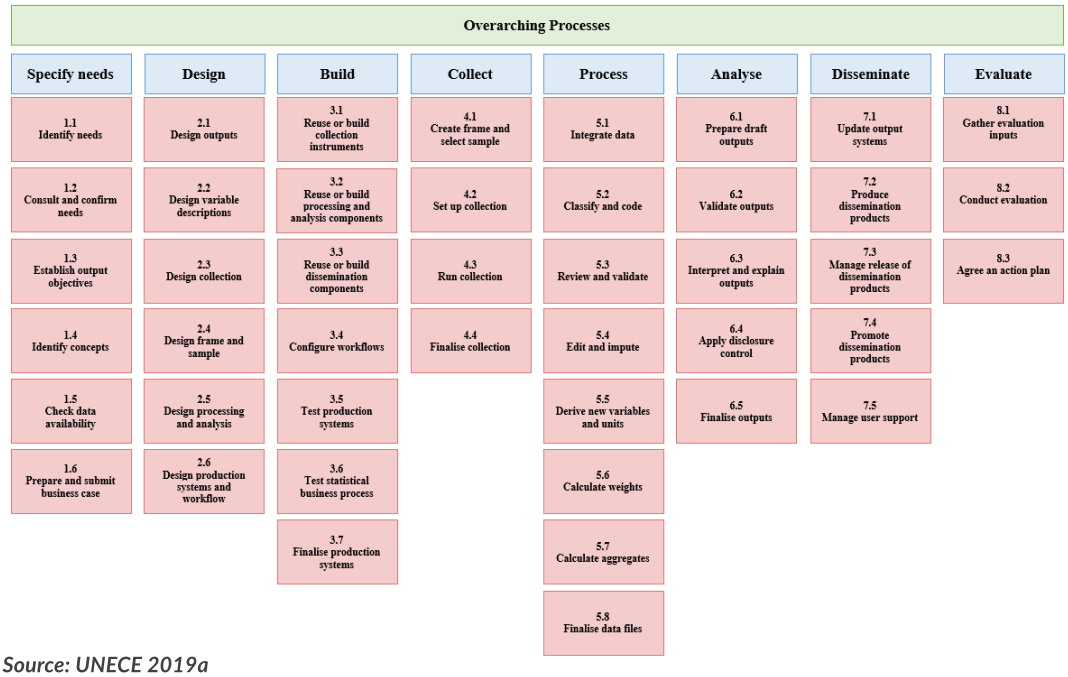
The first component of this review focused on the ASD's overall operations. For this, the Generic Statistical Business Process Model (GSBPM; UNECE 2019a) was used. The GSBPM was developed over the past decade by the United Nations Economic Commission for Europe (UNECE) in collaboration with the broader international statistical community. It describes a set of business processes that are generally accepted to be those involved in the production of official statistics.[4] It provides guidance to statistical offices in refining and modernizing their statistical production processes in line with best international practice.

At its highest level, the GSBPM breaks the production of official statistics into eight phases, each with a number of sub-processes (Figure 2). The extent to which the ASD follows each of the eight processes was assessed through interviews with the ASD. The ASD noted that about half of the staff members are familiar with the GSBPM, which the Caribbean Community (CARICOM) has been encouraging member countries to use since around 2015.[5] Though the model is not yet fully applied in the ASD, the goal is to increase its use. At the moment, the situation could be considered a hybrid, where the GSBPM is explicitly used to guide the ASD's work in some instances but implicitly[6] in others. Since 2017, whenever the opportunity has presented itself and resources have permitted, the GSBPM has been applied. In such cases, formal concept papers,[7] frameworks, and metadata descriptions are put in place.[8] Quality assurance reviews are also undertaken. Before 2017, work was done in a less coordinated fashion and was more reactive and less proactive.

In spite of the goal to increase the use of the GSBPM, the ASD is not always able to 'step back' and plan new or revised statistical activities in this way due to workload pressures and insufficient human and financial resources. Staff turnover is high, and it is time-consuming to train new ones. Because of resource constraints, senior managers are often obliged to do technical work rather than focusing on strategic and managerial issues.

4 Official statistics are generally defined as those published by NSOs and other members of the national statistical system. To be considered official, they should adhere to the UN Fundamental Principles of National Official Statistics.
5 CARICOM is pursuing development of a Regional Strategy for the Development of Statistics that includes a common statistical production architecture based on the GSBPM (CARICOM 2018).
6 'Implicit' use implies that statistical processes are undertaken in a coordinated and logical fashion (as the GSBPM requires) without the GSBPM necessarily being the direct guide.
7 A concept paper is written to outline the importance of statistics in a thematic area. It outlines the need for the development and implementation of a systematic, standardized, and coordinated approach to data collection and dissemination in the context of Anguilla.
8 A statistical domain might be, for example, health statistics, environmental statistics, or national accounts.

Figure 2: The Processes and Sub-processes of the GSBPM



Source: UNECE 2019a

4.1.1 GSBPM Phase 1 - Specify Needs

Basic elements. In Phase 1 of the GSBPM, potential user needs are identified, and users are consulted to confirm the needs. New needs are then compared with existing data to see where gaps exist. Finally, a business case is established to meet the agreed priority needs.

Situation in Anguilla. ASD staff generally know what data are needed in Anguilla and, whenever resources permit, the department is capable of filling these needs. In doing so, the GSBPM is followed to the extent possible. However, as noted earlier, resource constraints do not always permit new statistical activities to be undertaken with GSBPM as an explicit guide.

The major challenge with respect to Phase 1 is that decision-makers and other stakeholders in Anguilla are not accustomed to thinking of statistics as key inputs in their work. The context in Anguilla remains one in which the ASD has to focus on promoting the value of official statistics to potential users rather than users coming regularly to the ASD with demands for new statistics.

Anguilla has a National Statistics Advisory Committee (ANSAC) comprising representatives from the private and public sectors who meet on a quarterly basis. The aim of the ANSAC is to support and coordinate the development of Anguilla's national statistical system to support evidence-based decision-making. The ANSAC promotes the effectiveness of the ASD by advising and working with it to produce quality statistics about Anguilla's society and economy. Among other tasks, this involves fostering collaboration and coordination among data producers and users and identifying the data needs of the business, nonprofit, and public sectors (Ministry of Finance, Economic Development, Investment and Commerce of Anguilla 2014).

The ANSAC provides one mechanism by which users' data needs can be identified in Anguilla, though this is not its primary role. There is no other mechanism by which users' needs are regularly identified. As noted, users' assessments of their needs are less important in shaping the development of the ASD's programs than are the ASD's own assessment of what statistics are needed and what gaps in its existing statistics prevent those needs from being met. This will remain the case until such time users begin to see statistics as 'need to have' rather than 'nice to have'. Disaster-related statistics are a good example. The ASD has long considered this an area of need for Anguilla but it took Hurricane Irma to provide the catalyst for others to recognize this need and argue for it to be filled.

4.1.2 GSBPM Phase 2 - Design

Basic elements. Phase 2 of the GSBPM involves the design of the elements of the statistical system to meet users' needs with high-quality statistics. This includes the design of statistical variables, collection systems, sampling frames, processing systems, and production systems. Metadata protocols and quality assurance procedures are also part of the design phase. The design phase is not undertaken every time a set of statistics is collected but, rather, only when a new or significantly revised set of statistics is planned. International and national standards are applied as much as possible during the design process to enhance the comparability of outputs. For efficiency reasons, reuse or adaptation of design elements from existing processes (for example, existing survey frames) is encouraged.

Situation in Anguilla. In Anguilla, there is heavy reliance on administrative data to meet new data needs, so several elements of the design phase (collection systems and sampling frames) are not relevant in many instances. Instead, effort is devoted to identifying suitable administrative databases and working with agencies responsible for those databases to ensure they are well documented and follow international standards (for example, International Standard Industrial Classification[9]) to the extent possible.

In instances where a new data need is to be met through a survey activity, the ASD applies the steps of GSBPM Phase 2 rigorously. Statistical registers exist for both businesses and households and are kept up to date by tracking business openings and closures, and population changes. However, the quality of these registers is not as high as it could be. In the case of the business register, data on business openings/closures obtained from the responsible ministries are not as well documented as they could be and the ASD does not have the resources to improve them. Though the need to 'work smarter, not harder' is acknowledged, working smarter first requires resources to overcome the existing hurdles.

In the specific case of disaster-related statistics, one of the challenges in using administrative data is inconsistencies in definitions, both across agencies and through time.

4.1.3 GSBPM Phase 3 - Build

Basic elements. Phase 3 of the GSBPM builds and tests production systems to create the complete operational system needed to run the statistical process. New system elements are built only when absolutely necessary, with preference being given to reusing existing elements from within the organization or externally. New elements should be themselves built to be reusable.

Situation in Anguilla. In Anguilla, building system elements for new survey collections does not present a challenge, as the ASD is capable of building and testing the systems it requires. The challenge comes again with respect to the use of administrative data to meet data needs.

4.1.4 GSBPM Phase 4 - Collect

Basic elements. In Phase 4 of the GSBPM, the actual data collection takes place using various collection modes. Data collected are stored for future processing in the next phase. The collection phase includes basic data validation, but the bulk of the data processing occurs in the next phase. Unlike the first three GSBPM phases (which are conducted only for new or substantially modified activities), Phase 4 occurs in each iteration of all activities.

Situation in Anguilla. In Anguilla, the first two elements of the collection phase (selecting the survey sample and setting up the collection system) are generally implemented without trouble. The difficulties come in the last elements of the phase, where the actual data collection occurs. Recipients of ASD surveys are often reluctant to respond, partly because a culture of willingly providing data is not yet fully developed in Anguilla and also because respondents' time is limited. This obliges the ASD to commit additional resources from its already limited budget to ensure reasonable response rates. Response rates to business surveys are particularly problematic. Response rates to household surveys, including the population census, are better. The ASD reminds respondents of their legal requirement to provide data, but this only goes so far in encouraging response.

9 https://unstats.un.org/unsd/classifications/Family/Detail/27.

Obtaining administrative data from agencies and ministries can also be problematic. Even though the data have already been collected for whatever administrative purpose they serve, the additional burden of providing them to the ASD is not always welcomed, as the agencies and ministries have their own resource constraints. In the specific instance of disaster-related statistics, this means that data related to households (for example, household emergency preparedness) will be easier to obtain than data related to businesses (for example, business spending on risk reduction or business losses due to disasters).

4.1.5 GSBPM Phase 5 - Process

Basic elements. In Phase 5 of the GSBPM, data collected in Phase 4 are processed and prepared for analysis through integration, classification, and validation. Like Phase 4, Phase 5 occurs in each iteration of all statistical activities. The elements in the processing phase apply to both survey and administrative data (with the possible exception of weighting, which generally applies only to survey data). Activities in this phase can begin before those in the preceding phase (collection) are complete. This permits provisional results to be compiled more quickly in instances where timeliness is an important concern for users.

Situation in Anguilla. In Anguilla, the ability to execute this phase is challenged by the quality of the sample frames (business and/or household registers), and their completeness and weighting of sample data to produce aggregate estimates, for example, is hampered by the incompleteness of the registers. However, all efforts are made to analyze data in an integrated manner, using cross-references wherever possible. The ASD uses international classifications and definitions in the compilation of data series (for example, International Standard Industry Classifications, International Standard Classification of Occupations, and International Standard Classification of Education). In 2017, the ASD led the preparation of a protocol calling for the adoption of the International Standard Industrial Classification in classifying industrial data across government. This protocol was endorsed by EXCO of Anguilla.[10]

4.1.6 GSBPM Phase 6 - Analyze

Basic elements. In Phase 6 of the GSBPM, which occurs in every iteration of all statistical activities, statistical outputs are produced. Outputs include commentary on the data or technical notes describing them. The steps involved include ensuring outputs are 'fit for purpose' before dissemination by, for example, subjecting data to confidentiality analysis (to ensure that no single respondent's personal information can be revealed) and comparing data with other sources to confirm their reasonableness. The analysis phase applies to all statistical activities, whether based on survey or administrative data. Analysis can be undertaken in parallel with the preceding phase (processing), and it is often helpful to do so, as analysis can reveal instances in which additional data processing may be needed. Like the processing phase, analysis can also begin before the conclusion of the collection phase.

Situation in Anguilla. In Anguilla, initial outputs in the form of raw tables are run for review and analysis. When and where necessary, the tables are discussed with subject matter specialists and/or research is undertaken to validate or corroborate the data. When dissemination tables are built, they are reviewed to ensure confidentiality of respondents' data is maintained. Where necessary, the information is aggregated to ensure confidentiality or presented in a manner that does not disclose confidential data.

4.1.7 GSBPM Phase 7 - Disseminate

Basic elements. In Phase 7 of the GSBPM, which occurs in every iteration of all statistical activities, the statistical products produced during the preceding phases are released to users in various formats. This phase is key to providing users with access to the data they require to carry out evidence-based decision-making. Statistical agencies have a role to play not just in releasing data products but also in promoting them to potential users. They also are obligated to provide users with support when they have queries regarding the data products.

Situation in Anguilla. In Anguilla, the ASD uses a number of mediums to disseminate data. For a number of the data series, statistical summaries are disseminated to the general press and those requesting those summaries simultaneously by email. Once the summaries are released, all other tables are updated and released on the internet: ASD website[11] and the

10 http://www.gov.ai/documents/exco/Mn159-18.pdf.
11 http://statistics.gov.ai/StatisticsDept/MetaData1_4.

Government of Anguilla Facebook page.[12] The data are disseminated in summary format and downloadable tables where possible. The ASD also produces some infographics targeted at persons who may better understand data in that format. These are posted to the government's Facebook page for easy access (see, for example, this infographic on GDP[13]).

When queried, the ASD provides users with explanations of how its data should be interpreted. Additionally, if data are requested, they are provided, and, if possible, individuals are directed to the website for future updates. If required, the ASD may provide clarifications on misinterpretation by persons of its data.

4.1.8 GSBPM Phase 8 - Evaluate

Basic elements. In Phase 8 of the GSBPM, the overall process of production for a given set of statistics is evaluated. The evaluation can take place at the end of the process but can also be done throughout. Evaluation relies on inputs gathered throughout the preceding phases. It relies on a range of quantitative and qualitative inputs and results in the identification and prioritization of improvements. Evaluation should, in theory, occur for each iteration of all statistical processes. However, in the case of long-standing and well-established processes (for example, for production of the labor force survey), evaluation might be carried out less frequently.

Situation in Anguilla. The ASD is aware that there is a need to monitor and evaluate the processes undertaken in the production of statistics. However, limited human resources constrain this activity. Wherever possible, the ASD does step back to review processes with a view to learning what went well and what requires improvement.

4.2 Assessment of Statistical Operations in Relation to Disaster-related Statistics

Part I of the DSSAT focuses on the overall institutional and organizational structure of the statistical system and on specific information in terms of policy frameworks, mandates, institutional setup, organization, collaboration, resources, international cooperation, and uses of disaster-related statistics. Part II assesses the national relevance, importance, availability, and sources of the individual disaster-related statistics and helps identify data gaps and a plan for filling.

The team used Part I of the DSSAT for assessment of the ASD's operations with respect to the production of disaster-related statistics. The results of this assessment are summarized below.[14] The team used Part II for the assessment of the gaps in existing disaster-related statistics in Anguilla, which is discussed in Section 4.4.

4.2.1 Policies, Mandate, and Organization of Disaster-related Statistics

Anguilla has several policies relevant to disaster management, including a Comprehensive Disaster Management Policy [15](Government of Anguilla 2013) and a National Disaster Plan[16] (Department of Disaster Management 2012). Both of these are the responsibility of the DDM. It is worth noting that while both policies speak implicitly and explicitly of the need for disaster management to be evidence-based, neither devotes significant space to the discussion of disaster-related statistics. They focus mainly on matters related to the provision of information to the public during an active emergency. This fact is consistent with the Working Group's assessment that the need for better statistics to support disaster preparedness and post-disaster assessment is of high importance in Anguilla.

Anguilla has a national statistical law[17] (Government of Anguilla 2000) that mandates the ASD to produce statistical information relative to the commercial, industrial, social, economic, and general activities and conditions of Anguilla and to do so in collaboration with other departments and agencies. There is a national statistical system[18] in place in principle in Anguilla, with all government departments and agencies as part of it and the ASD as the lead. In practice, however, the ASD is the main provider of official statistics. Work toward a national strategy for the development of statistics[19] under the auspices of PARIS21 has been underway for several years but has not yet come to fruition due to a lack of resources.

12 https://www.facebook.com/anguillagovernment.
13 https://www.facebook.com/anguillagovernment/photos/a.10158154088478725/154340358724/?type=3&theater.
14 Full details of the Part I assessment are found in Annex 4.
15 http://www.gov.ai/documents/Anguilla Draft National CDM Policy.pdf.
16 http://ddmaxa.org/documents/Anguilla National Disaster Plan.pdf.
17 http://www.gov.ai/statistics/images/StatisticsAct151200.pdf.
18 A national statistical system is a formal system for producing and disseminating official national statistics that includes all agencies with a mandate to produce statistics and is coordinated centrally, usually by the NSO.
19 https://paris21.org/national-strategy-development-statistics-nsds.

Despite this, no interinstitutional working is in place to coordinate the production of disaster-related statistics. The National Disaster Management Committee (NDMC) and its subcommittees play an informal role in this regard and do request data ad hoc. The main barriers to more formal coordination are lack of resources and time, low priority accorded by institutions, and competing interests among institutions.

No direct mention is made of disaster-related statistics in the Statistics Act, but the Act does provide for statistics on 'vital occurrences and morbidity' and 'other matters', both of which could be taken to provide a mandate for the collection of disaster-related statistics. The ASD does, in fact, have such a mandate, though resource constraints impede its work in the area.

4.2.2 Production and Use of Disaster-related Statistics

The Working Group found that most of the existing disaster-related statistics in the country are produced by the ASD. Many existing statistics rely on data that are collected by other departments and agencies, however, so there is a degree of interagency cooperation and coordination involved.[20] The production of these statistics is, at the moment, somewhat uncoordinated, as the Working Group reported that no national plan for disaster-related statistics is currently in place. At the same time, the ASD is aware of and makes use of several of the major international frameworks and guidelines for disaster-related statistics, including the Sendai Framework for Disaster Risk Reduction (United Nations 2015) and the UNECE recommendations for statistical offices on disaster-related statistics.[21] The ASD does not currently make use of the DRSF prepared by the Expert Group on Disaster-related Statistics in Asia and the Pacific (UNESCAP, 2018).

Existing disaster-related statistics in Anguilla mainly cover the impacts of past disasters on citizens and physical infrastructure. Statistics on vulnerability to future disasters are a gap, as are statistics on the impact of disasters on the environment (see the results of Part II of the assessment below for further details on gaps). Existing disaster-related statistics are made available through the ASD's website, though there is no section of the website currently devoted to these statistics; rather, they are simply categorized under their primary heading (for example, population statistics or GDP statistics). Disaster-related statistics are shared with several regional and international organizations, including the Caribbean Disaster Emergency Management Agency (CDEMA), the Caribbean Electric Utility Services Corporation, PAHO, the Caribbean Regional Fisheries Mechanism, and the UNECLAC. Such sharing is more ad hoc than routine.

The Working Group reported that existing statistics are used during non-crisis periods to inform disaster management policies by, among others, the Strategic Planning Subcommittee of the NDMC. During crisis periods, existing statistics are used in a limited fashion to help estimate the number of emergency responders that are required. During the post-crisis period, existing statistics serve to estimate aid requirements and plan for relocation of affected citizens.

4.2.3 Resources Available for Disaster-related Statistics

Overall, Anguilla has few resources devoted to disaster-related statistics, and this situation has not improved over time. The ASD has just seven full-time staff members, including the Chief Statistician, to cover its entire statistical mandate, including disaster-related statistics. The total resources available within the ASD for disaster-related statistics amount to much less than one full-time employee (FTE), which is not surprising given the wide range of statistics under the ASD's responsibility. The Anguilla DNR has one person devoted to disaster risk management, including the gathering of disaster-related statistics. The DNR also occasionally benefits from the assistance of consultants through project-related work. The DPP has staff in three units working on disaster risk mitigation, including gathering disaster-related statistics.[22] Within the DDM, the Mitigation Working Group of the NDMC focuses on means to mitigate the impacts of disasters. It has six members from the DPP.

20 The agencies and departments most involved in providing data for disaster-related statistics are the PUC, the Ministry of Finance, the Anguilla Air and Seaport Authority, the DNR, the DDM, and the DPP.
21 https://unece.org/DAM/stats/publications/2019/ECECESSTAT20193.pdf.
22 The three units are Development Control Unit (ensures that planned developments meet certain disaster risk management criteria), Development Planning/GIS Unit (prepares policies and maps), and Building/Electrical Unit: (assesses structural/plumbing/electrical plans to ensure integrity).

4.2.4 Engagement with Regional and International Organizations

Engagement with regional and international disaster management agencies is limited, according to the Working Group's assessment. The work of CARICOM and UNECLAC is followed regionally by the ASD. The work of the UNECE Task Force on Measuring Hazardous Events and Disasters (Geneva) and the UN Statistics Division (New York) is also followed by the ASD. Other departments and agencies of the GoA do not engage in regional or international work on the topic.

Anguilla has taken limited advantage of regional and international training opportunities related to disaster risk management and statistics, participating in training to enhance skills in conducting damage and loss assessment using the UNECLAC model. One Anguillan expert has assisted in providing training to other countries in the region.

4.2.5 Plans for Improving Disaster-related Statistics

The Working Group reported that Anguilla currently has no formal plan to strengthen and develop disaster-related statistics, though it is acknowledged that such work is required in numerous areas. The legal framework, institutional setup, and resource allocation for the collection of disaster-related statistics all deserve attention. All kinds of regional and international technical assistance could be of value to Anguilla to this end.

4.3 Determining a Basic Range of Disaster-related Statistics for Anguilla

The international community has devoted considerable attention in recent years to the development of frameworks for disaster risk management. Among these efforts, the Sendai Framework for Disaster Risk Reduction (United Nations 2015) stands out as particularly significant. Developed in the aftermath of the 2011 Japanese earthquake and tsunami, the Sendai Framework's goal is to prevent new and reduce existing disaster risk through the implementation of a broad range of measures to prevent and reduce hazard exposure and vulnerability to disaster, increase preparedness for response and recovery, and strengthen resilience. The framework outlines seven targets and four priority action areas related to this goal (Table 5).

Table 5: Targets and Priorities of the Sendai Framework for Disaster Risk Reduction

Targets	Priority Areas for Action
1. Substantially reduce global disaster mortality by 2030 2. Substantially reduce the number of affected people globally by 2030 3. Reduce direct disaster economic loss by 2030 4. Substantially reduce disaster damage to critical infrastructure and disruption of basic services by 2030 5. Substantially increase the number of countries with national and local disaster risk reduction strategies by 2020 6. Substantially enhance international cooperation to developing countries by 2030 7. Substantially increase the availability of and access to multi-hazard early warning systems and disaster risk information and assessments by 2030	1. Understanding of disaster risk 2. Strengthening of disaster risk governance to manage disaster risk 3. Investing in disaster risk reduction for resilience 4. Enhancing disaster preparedness for effective response and to 'Build Back Better' in recovery, rehabilitation, and reconstruction

Recognizing that frameworks and indicators for disaster risk management are of no value without the basic statistics necessary to compile them, the global community has been active in recent years as well to improve the statistical basis for measuring disaster risks, impacts, and reduction efforts. One initiative of particular note is the DRSF[23] prepared by the Expert Group on disaster-related statistics in Asia and the Pacific. Though the core concepts and indicators for disaster risk

23 https://communities.unescap.org/asia-pacific-expert-group-disaster-related-statistics/content/drsf.

reduction are defined in the Sendai Framework (and the related Sustainable Development Goal [SDG] indicators), there was a need to translate the agreed concepts and definitions into specific instructions and technical recommendations for production and dissemination of statistics. To this end, the DRSF provides a basic range of disaster-related statistics that all countries are recommended to consider compiling to support reporting against the Sendai Framework. The basic range is broken into a set of seven topics (or tables) and an account for measuring expenditures on disaster risk reduction. Each of the tables is divided into sub-tables that present the actual statistics making up the DRSF basic range. This is summarized in Table 6 and presented in full detail in Annex 7 -UN-ESCAP Basic Range of Disaster-related statistics.

Because the DRSF was developed expressly to support the compilation of the basic statistics required to compile the Sendai Framework indicators, it was used as the starting point for developing an 'ideal' basic range of disaster-related statistics for Anguilla. The same working group convened to implement the DSSAT was tasked with reviewing the DRSF basic range of statistics and determining its relevance to Anguilla. The Working Group concluded that essentially the entire DRSF basic range was relevant to Anguilla, with three minor exceptions: statistics related to slum dwellings, statistics involving regional breakdowns of economic variables (such as GDP by region), and statistics related to transfers from central to local governments.

Table 6: DRSF Basic Range of Disaster-related Statistics

Table A. Disaster Occurrences
A.1. Extreme event/disaster (geophysical; hydrological; meteorological and climatological; biological; other)
A.2. Location
A.3. Magnitude (Large, medium, small)
A.4. Date of occurrence
A.5. Duration
Table B. Background Statistics, Exposure to Hazards and Coping Capacity
B1a.1. Population by region
B1a.2. Households by region
B1a.3. Median household disposable income by region
B1a.4. GDP by sector by region
B1a.5. Population in hazard area by region
B1b.1. Population by age class, gender, income level, etc.
B1b.2. Population in hazard areas by age class, gender, income level, etc.
B2.1. Critical infrastructure in hazard areas
B2.2. Land area
B3.1. GDP by region
B3.2. GDP per capita by region
B3.3. Median household disposal income by region

B3.4 – Number of households with slum designation by region

B3.5 – Population living in areas with slum designation by region

B3.6. Population covered by early warning systems by region

B3.7. Household disaster emergency preparedness by region

B3.8. Environmental resilience by region

B3.9. Spending on risk reduction activities by region

B3.10 – Disaster-related transfers from central to local Government

Table C. Human Impacts

C.1. Deaths or missing

C.2. Injured or ill

C.3. Displaced individuals

C.4. Dwellings damaged

C.5. Loss of jobs/occupations

C.6. Number of people evacuated or receiving aid

Table D. Direct Material Impacts (Physical)

D1.1. Direct physical impacts on fixed assets or consumer durables

D1.2. Direct physical impacts on valuables

D1.3. Direct physical impacts on natural resources

D1.4. Direct physical impacts on critical infrastructure

D1.5. Direct physical impacts on cultural heritage

D2.1. Disruption to basic services

Table E. Direct material impacts (monetary)

E.1. Direct monetary impacts on fixed assets or consumer durables

E.2. Direct monetary impacts on valuables

E.3. Direct monetary impacts on natural resources

E.4. Direct monetary impacts on critical infrastructure

E.5. Direct monetary impacts on cultural heritage

Table F. Material Impacts to Agriculture

F.1. Impacts on crops

F.2. Impacts on livestock

F.3. Impacts on fisheries

Table G. Material Impacts to the Environment

G.1. Impacts on land and coastlines

G.2. Loss of critical ecosystems

G.3. Loss of natural water resources

G.4. Impacts on atmosphere and greenhouse gases

Source: UNESCAP 2018.

Note: Text in red indicates a variable from the DRSF basic range of statistics considered not relevant to the Basic Range of Disaster-related Statistics for Anguilla.

4.4 Determination of Gaps in Disaster-related Statistics and Priorities for Filling these Gaps

Once the Basic Range of disaster-related statistics for Anguilla (the Basic Range from here on) was agreed upon, the Working Group was then tasked with identifying the gaps in Anguilla's disaster-related statistics by comparing existing statistics with the Basic Range. The Working Group identified existing disaster-related statistics in Anguilla using Part II of the DSSAT (Annex 4- Disaster-related Statistics Self-Assessment Tool). Comparing the results of Part II with the Basic Range, a number of gaps were identified. Recognizing that the concept of a data gap is not absolute, as data can be partially available, the Working Group categorized gaps as either 'full' (that is, no statistics exist at all in Anguilla matching those called for in the Basic Range) or 'partial' (statistics exist in Anguilla, though not exactly as called for in the Basic Range).

Once the data gaps were identified, the Working Group's final task was to prioritize the gaps for filling. To do this, the Working Group used a scale of 1 to 3, with 3 being the highest priority for filling. Table 7 shows the statistics that the Working Group identified as full or partial data gaps, along with the group's assessment of the priorities for filling. As can be seen, the gaps in Anguilla's existing statistics were found to be split roughly equally between full and partial gaps. Most of the full gaps are related to statistics that are relatively difficult to measure, such as direct material impacts of disasters. Except for statistics in DRSF Table F (Material Impacts to Agriculture), all the statistics identified as full gaps were considered by the Working Group to be of only medium priority for filling.

All the statistics identified as gaps in DRSF Table C (human impacts of disasters) were identified as partial gaps, and the Working Group considered them all to be a high priority for filling. This is obviously appropriate, as one would expect that filling gaps in statistics related to the human impacts of disasters would be a higher priority than filling gaps related to, for example, material impacts. Similarly, the Working Group gave high priority to filling gaps related to impacts on agricultural production capacity (DRSF Table F). The only other high priority gap identified was for statistics on the exposure of critical infrastructure to disaster-related hazards, which the Working Group found to be a partial gap. It is worth noting that the Working Group did not give low priority to filling any data gap, indicating that its members generally accorded considerable importance to improving Anguilla's evidence base for disaster risk management. Complete details of the Working Group's prioritization of data gaps are presented in Annex 8.

Table 7: Gaps in Anguilla's Basic Range of Disaster-related Statistics with Priority for Filling

Statistics*	Full Gap	Partial Gap	Priority for Filling**
Table B. Background Statistics and Exposure to Hazards			
B1c.5. Household disaster emergency preparedness		X	2
B2.1. Critical infrastructure in hazard areas		X	3
B2.2. Land area		X	2
B3.6. Environmental resilience	X		2
B3.7. Spending on risk reduction activities	X		2
Table C. Human Impacts			
C.1. Deaths or missing		X	3
C.2. Injured or ill		X	3
C.3. Displaced individuals		X	3
C.4. Dwellings damaged		X	3
C.5. Loss of jobs/occupations		X	2
C.6. Number of people evacuated or receiving aid		X	3
Table D. Direct Material Impacts (Physical)			
D1.1. Direct physical impacts on fixed assets or consumer durables	X		2
D1.2. Direct physical impacts on valuables	X		2
D1.3. Direct physical impacts on natural resources	X		2
D1.4. Direct physical impacts on critical infrastructure	X		2
D1.5. Direct physical impacts on cultural heritage	X		2
D2.1. Disruption to basic services	X		2
Table E. Direct material impacts (monetary)			
E.1. Direct monetary impacts on fixed assets or consumer durables	X		2
E.2. Direct monetary impacts on valuables	X		2
E.3. Direct monetary impacts on natural resources	X		2
E.4. Direct monetary impacts on critical infrastructure	X		2
E.5. Direct monetary impacts on cultural heritage	X		2
Table F. Material Impacts to Agriculture			
F.1. Impacts on crops	X		3

F.2. Impacts on livestock	X		3
F.3. Impacts on fisheries	X		3
Table G. Material Impacts to the Environment			
G.1. Impacts on land and coastlines		X	2
G.2. Loss of critical ecosystems		X	2
G.3. Loss of natural water resources		X	2
G.4. Impacts on atmosphere and greenhouse gases		X	2

Source: Anguilla Disaster-statistics Data Assessment Working Group.

Note: *Statistics are labelled following the categorization of the DRSF. The first three (or fewer) characters refer to the DRSF table, sub-table, and sub-sub-table, and the number following the period refers to the specific statistic within the tables. **1 = low priority; 2 = medium priority; 3 = high priority.

The prioritization of statistics for the DSSAT clearly lays out the end products required. These have been formulated into the database design in Section 4.5. Creating these outputs requires assimilation of multiple datasets and steps to aggregate and process to create the final statistics. There are several factors to consider when creating the final set.

The need to divide the territory into region-related statistics relies on the need to use spatial data and have recognized administrative sub-divisions in the country (in Anguilla's case, the district level, although enumeration district and electoral districts also exist).

The DSSAT identifies the need to create regional statistics that document the vulnerability due to a range of different hazards. Therefore, individual hazard datasets are required for each of the hazard types (for example, storm surge, earthquake, flood, and high wind zones) before intersecting that with the administrative areas and summarizing (usually counting) exposed features (that is, population sector of the infrastructural unit). This means that for one broad statistic, there is the need for the combination of a number of input datasets covering each dimension of the final statistics.

Given these requirements, there are some major basic datasets that are required as inputs to describe the potential exposure:

- Demography/census
- Basic infrastructure (buildings and roads)
- Critical infrastructure
- Natural capital (both land and marine-based)
- Cultural, heritage, and touristic value datasets.

Beyond these data, there are still other statistics for the national level only and are captured/aggregated with different methodologies.

In some instances, the perfect dataset that can service one of the identified disaster-related statistics does not currently exist, but there are surrogates that can partially assist. The datasets have been categorized in the metadatabase as a full match, a fair match (they can be used without modification at present but could be improved), fuzzy match (the dataset can be used to some extent but would not completely fulfil the requirement), and needs adaptation (where some data around the theme exist but they do not have the required structure and content to satisfy the requirement) The reasons why this

data are a fuzzy match or needs adaptation for the perfect statistics include the following:

- The data needs regular updating—for example, for new infrastructure such as electricity poles/transformers. No regular updates are scheduled or have ever been instigated.
- The data are missing some aspects useful for the full statistics.
- Datasets only partially cover the need. For example, there is no comprehensive layer showing water sources, but both a generalized aquifer map and some data on the location of desalination plants are available. No data are available on localized rainwater capture. All three would be needed to identify the exposure of water sources to disaster.
- Data do not match the requirement but provide some useful indication (for example, fish catch baseline as an indicator of the level of fishery production).

4.5 Conceptual Design of a Database for Disaster-related Statistics in Anguilla

To support the ability of the GoA to produce statistics for disaster risk assessment, response, impact assessment, and recovery, the project included the design of a database to provide access to data for disaster response, PDNA, recovery planning, and subsequent disaster risk reduction. DADS is intended as a general-purpose database that would be feasible for implementation and management by the ASD, contributed to by multiple national data holders, and used by the NDMC to manage disasters. Implementation of this database (which was not considered in this project) would enhance the capacity of the ASD to contribute to the needs of Anguilla in terms of disaster-related activities.

No 'readymade' database design satisfying all the requirements of the project objectives was found. The European Union (EU) concluded in 2018 (Rios Diaz and Marin Ferrer 2018) that no country had a comprehensive, multi-disaster database platform. Research is also ongoing to develop risk assessment databases from public data (for example, Choi and Suh 2020). International databases focus on country-reported losses (Desinvantar and Sendaimonitor), events (Centre for Research on the Epidemiology of Disasters Emergency Events Database[24]), or country-contributed data (United Nations Office for the Coordination of Humanitarian Affairs Humanitarian Data Exchange[25]) or do not include smaller countries, such as Anguilla (INFORM Risk[26]). Therefore, the design proposed here should be considered experimental, even if based on international frameworks, examples, and statistical frameworks to the extent possible, especially the following:

- The Sendai Framework (2015) for monitoring progress toward agreed targets of increasing preparedness and reducing impacts
- The DRSF
- The PDNA (GFDRR 2013), which provides guidelines on what data to collect after a disaster to assess impacts and plan for recovery.

DADS builds upon and links to other components of this project, notably, to the assessment of Anguilla's information architecture (Sections 3.1 to 3.3), to the metadata catalogue of disaster-related data (Section 3.4), and to the assessment of disaster-related data presented earlier in this section. It is designed to facilitate the integration of currently available data and provides recommendations for proxies where detailed data are not available.

DADS is conceived as a set of tables at the statistical district or enumeration district geographic levels[27] with extensions to the ocean to encompass marine ecosystems. Implementing the design and operationalizing the resulting database will provide the following:

- An annual risk table including

24 https://www.emdat.be/
25 https://data.humdata.org/
26 https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Risk
27 Districts and enumeration districts are statistical units employed by the ASD for the collection and dissemination of a variety of statistics; most notably, population census statistics.

- A compilation of the current risks by hazard type;
- The exposure of people, the economy, infrastructure, and environment to multi-disaster risks, singling out vulnerabilities and coping capacities where feasible; and
    - A summary risk assessment suitable for prioritizing risk-reduction activities

- An input to disaster response by indicating exposed, vulnerable, and low-coping capacity districts
- A framework for data collected post-disaster as well as initial estimates of impact, including
    - Standard reporting of the disaster event and estimates of severity;
    - Population, economic and environmental impacts, and service disruptions; and
    - A summary impact assessment suitable for contributing to the PDNA.

An MS Excel worksheet designed to facilitate testing of DADS is found in Annex 9.

4.5.1 Institutional and Technical Database Requirements

4.5.1.1 Institutional Requirements

In designing DADS, we accounted for Anguilla's small size and limited human, institutional, and technical capital. Data collection and analysis necessarily focuses only on priorities in such a context. This provides an argument for standardization, data reuse, and avoidance of duplication of efforts.

It is proposed that the ASD assumes the lead in managing the maintenance of DADS. Other stakeholders will still collect operational data, but the ASD will remain responsible for the integration of data into DADS. Maintaining the database will require a regular flow of data among GoA stakeholders. It is recommended that one ASD staff be designated as the 'DADS Coordinator', who will regularly interact with designated stakeholder data providers.

An example of the kind of collaboration we envisage for DADS is the relationship between the ASD; the Attorney General's Office; and the Registry of Births, Deaths, and Marriages. The ASD conducts a detailed population census approximately every 10 years. Intercensal population estimates, required for managing the country, are based on annual data on migration and citizenship from the Attorney General's Office and data on births and deaths from the registry.

While intercensal population estimates by the statistical district will be included in DADS, the database will require dozens of similar activities of data preparation and regular communication to the ASD. This implies not only agreement on what is to be provided but also a regular process of data preparation, data transmission, validation, and if required, revision.

Many of the data preparation activities will require spatial analytical work to aggregate detailed data to the statistical district level (for example, number of residences), allocate data from other geographic boundaries to the district level (for example, hazard risk and length of road), or downscale more general data to the district level (for example, protected areas). With this in mind, it is recommended that the DPP take a lead role in supporting the data preparation.

Given that the ultimate user of DADS will be the NDMC, it would benefit the process to set a regular annual deadline for when the annual database is expected to be available. Working back from that date will establish the updated schedule. Roles and responsibilities are discussed in further detail in Section 4.5.6.

An integral aspect of the compilation of DADS will be the regular testing, review, and assessment of the database. Quality assessment/quality control checks, described in more detail in the next section on technical requirements, will provide guidance on ensuring the data are as 'useable' as possible. Given that many of the data elements will be estimated, accuracy and precision are of less concern than fitness for use.

Further, an essential aspect of implementing and managing DADS is the human and financial resources available. This is

discussed further in the database implementation workplan (Section 4.5.7).

4.5.1.2 Technical Requirements

The current conception of DADS is as a set of spreadsheets, implementable on desktop (for example, Microsoft Excel) or in the cloud (for example, Google Sheets). Contributions to the spreadsheet by the ASD and other data providers will use technology already in place in the office of the data provider. Most database updating operations will consist of copying data into appropriate cells in the spreadsheet.

As with regular updates to the metadata catalogue (Section 3.4), it is expected that most database updates will be done by data providers sending a filled spreadsheet template to the DADS coordinator. This will not be straightforward during times of active disasters and shortly afterward. The following sections presume an accessible copy of DADS on the cloud to which key users (NMDC, ASD, and others) have access during disasters. This would make the database accessible by cell phone, tablet, or laptop. As an ultimate backup, a printed copy of the relevant tables would be kept.

Since DADS consists almost entirely of aggregated data, there should be little concern about confidentiality. However, the ASD will need to ensure that any concern about confidentiality and privacy be addressed in additional provisions.

Since the design is not based on any specific technology, it can be enhanced as capacities improve. Future versions may include more of the data preparation within a spatial and relational database operation rather than depending on data providers to take on the bulk of the preparation work.

4.5.1.2.1 QA/QC Rules

DADS consists not only of some observed data but also much estimated data. As with intercensal population estimates, the estimated data are 'fit for purpose' but not necessarily highly accurate. DADS will be based on the best available estimates, though some will inevitably be out of date. This emphasizes the need to

- Understand the quality of the source data and record this quality in the metadata catalogue;
- Check the allocations to the statistical district, because they will have no basis for comparison other than allocations done previously; and
- Certify the results as 'fit for use' by the Chief Statistician.

In terms of the source data, the metadata catalogue includes several quality entries, including the following:

- Main data
  - Period of validity (the years to which the data refer)
  - Specific usage (for what the data are intended)
  - Use limitations
  - Other constraints
- Statistical
  - Response rate
  - Estimates of sampling error
  - Weighting.

In terms of the period of validity, if data are too far out of date, they will be less valid. Data collected for one period can be

used to estimate a future period, but the quality will vary depending on the phenomenon being measured. For example, land cover does not change substantially from one year to another unless there is an observed reason such as a natural disaster or extensive construction. The population is more dynamic and requires a process of intercensal estimates, which may contain high, medium, and low estimates. The economy is dynamic and assumptions from one period may not be valid in the next. The general recommendation is to apply 'fit-for-use' rules that consider the age of the data and dynamics of the phenomenon being measured.

The metadata catalogue variables for specific usage, use limitations, and other constraints should indicate if the data are of low or high quality. Some data included in DADS may be of relatively low quality but not identified as such in the metadata catalogue. Response rate, estimates of sampling error, and weighting are terms applied to statistical data collected from 'sample surveys', such as those commonly used to gather business and household data. The terms are interlinked in that low response rates lead to high sampling errors and the need arises to increase weighting to represent the population. If response rates are low or sampling error or weighting are high, then the data may be of lower quality.

The simplest ways to check the validity of data include the following:

- Comparison with control totals. If data have been allocated at the district level, the totals should add to the national total. This will apply to most population, infrastructure, economic, and environmental data.
- Calculation of change over period. When making an estimate for a year not measured (either interpolation or extrapolation), the change over the period should be 'feasible', that is, the population usually grows at 1–2 percent per year. In the past, the economy of some countries has grown by 5–6 percent per year, but this is now more in the order of −1 percent to +1 percent. Individual sectors, unless there is a known reason, usually do not grow or decline by more than 10 percent per year. Using this method requires the analyst to understand the local dynamics of the phenomenon being measured.
- Proportionality. For estimates where there is neither a control total nor a previous period for comparison, the proportions among statistical districts should be 'reasonable'. For example, most estimates of population vulnerability would be proportional to the population in each district. Deviations from this proportion may be explained in terms of known variations in age, income, or health but usually these would be small.

Timeliness is an important element of data quality. As noted earlier, data that refer to the situation too far in the past are less fit for use. Scope for improving the timeliness of data provided to DADS should be constantly investigated and be part of the annual DADS quality statement.

Each data element should have the assurance of the DADS Coordinator that source data and estimates have been checked and are 'reasonable' and as timely as possible. Such an annual quality statement would indicate priority source data and estimates that require improvement over the next cycle. This will also facilitate certification by the Chief Statistician.

4.5.1.2.2 User Access, Security, and Logging

DADS is designed as a simple spreadsheet database that can exist in multiple locations. The main concerns in terms of access are availability during times of disaster and confidentiality of the data. As a failsafe backup, it is recommended that the database be printed and stored for safekeeping. Security of the cloud version of DADS can be maintained by ensuring password protection. The spreadsheet files should also be locked with passwords available only to the designated users.

Although DADS consists of mainly estimated data aggregated to the district or enumeration district level, there may still be confidentiality concerns. For example, knowing that an enumeration district has one manufacturing establishment and its revenues are US$1 million would reveal confidential information about that establishment. To avoid such confidentiality concerns, it is recommended that all persons accessing the database be designated employees of the ASD. This would necessarily include members of the NDMC. They would then be obliged to adhere to the same rules as the ASD in terms of maintaining data confidentiality.

The only alternatives would be to (a) anonymize the data or (b) allow access only to ASD staff. Either of these alternatives would defeat the purpose of DADS. Anonymizing data would, for example, move or remove outliers. These outliers may be the most vulnerable people and exposed businesses. Restricting access to ASD employees would require that only a few people would have access at times of most need.

Logging access to DADS on the cloud can be done automatically. Google Sheets has a facility for viewing who has viewed the worksheet on 'work accounts' (that is, not personal accounts). Logging access to the spreadsheet file version would require the DADS Coordinator to log the name of any recipients of the file after obtaining assurance they would not share the file under oath to the Statistics Act.

Logging should also be performed as DADS is updated and prepared for the annual distribution to key users (all of whom must be under oath to the Statistics Act to maintain confidentiality). This will be relatively simple because only the DADS Coordinator will need to make the final updates. This could be implemented by adding three rows to the spreadsheet that record the date of acquisition of each data element from the data provider, the date the data element was updated in the spreadsheet, and the date the data were quality checked.

Each annual cycle of updates creates a new version of the spreadsheet file. Individual elements are overwritten as they are updated. In a disaster, it is recommended that the most recent cloud version of DADS be used. The working version may have some more updated data, but it will not have been quality assured and certified.

4.5.1.2.3 Output Formats and Reporting

DADS, in spreadsheet form, should be ready to set up to print for reporting and archival purposes. It is suggested that a print version be created, shared with key staff, and stored in safe places that would be accessible during a disaster. Figure 3 contains an example of one printed page.

Figure 3: Sample DADS Print Page

Table: Hazard Risk

Source:										
Date received:										
Date updated:										
Date certified:										

	Units	km^2	High / Medium / Low	High / Medium / Low	High / Medium / Low	High / Medium / Low	High / Medium / Low	High / Medium / Low	High / Medium / Low	Yes / No
District_Name	**District ID**	**Area**	**Hurricane Risk**	**Earthquake Risk**	**Flood Risk**	**Storm Surge Risk**	**Island Flooding Risk**	**Beach Change Risk**	**Suspended Solids Risk**	**Any Risk High**
West End	A01									
South Hill	A02									
Blowing Point	A03									
Sandy Ground	A04									
North Hill	A05									
George Hill	A06									
The Valley	A07									
North Side	A08									
The Quarter	A09									
Stoney Ground	A10									
The Farrington	A11									
Sandy Hill	A12									
East End	A13									
Island Harbour	A14									
Total Land Area										

Marine Area 1	AM1									
Marine Area 2	AM2									
Marine Area 3	AM3									
Marine Area 4	AM4									
Total Marine Area										

The Data Dictionary, with description of the variables, control totals, sources, and links to the metadata catalogue should also be printed in the same document. The DADS report would serve as an annex to the annual risk assessment provided by the NDMC.

4.5.1.2.4 Other Requirements for Establishing within the National Statistical System and GoA ICT

DADS, as designed, requires no additional equipment or personnel to operationalize. It can be implemented with existing computers, software, and personnel. However, it is recommended that the staff member assigned as the 'DADS Coordinator' and data providers undergo training on the concepts of disaster-related statistics and specifically on the operation of DADS. The general training and training on the metadata catalogue provided during the project should serve as the basis for training on DADS operation.

4.5.2 Conceptual and Logical Data Models

4.5.2.1 Conceptual Data Model

The accompanying DADS test spreadsheet (Annex 9 –Excel spreadsheet for testing DADS) describes 180 variables in 13 tables. There are two main sets of tables: risk assessment and impact assessment. The risk assessment table group is intended to be reviewed annually as new data become available. It is not expected that all data elements will be updated annually, but the risk assessment should reflect ongoing efforts to reduce and manage risks. Having this risk assessment data available during an event will provide some indication of which districts may require priority assistance. It also provides initial input to the impact assessment.

The impact assessment group is completed shortly after each event. Initial estimates of impacts may be made by applying an estimate of event severity to the population, infrastructure, economy, and environment exposed. More refined data (for example, deaths and service disruption) would need to be collected by data providers during and shortly after the event.

Individual variables, control totals, sources, links to metadata, and the Basic Range are described in the 'Information' sheet. Each table is prototyped in a separate sheet. Each table includes a row for each district. This may be revised for all tables if data are required at the enumeration district level. Placeholders are provided for marine areas, to be used if appropriate marine areas can be designated.

Following is a brief discussion of each table:

- Risk Assessment Group
    - R-Hazard risk. This includes the land area of each district (and marine area) as well as estimates of the levels of risk (high/medium/low) of each expected hazard type in each district. An overall hazard risk estimate for the district indicates whether any hazard has a high risk in the district. This is relevant when calculating exposure and vulnerability.
    - R-Population. This includes the main population indicators, estimated by district. Age groupings, employment status, and disability status contribute to the assessment of vulnerability. Household readiness and access to early warning contribute to the population risk assessment (see the Risk Assessment table).
    - R-Economy. This includes an estimate of the economic production by sector in each district. This may be allocated from national GDP figures using labor statistics. For example, using the proportion of national employment in manufacturing in the sector to allocate national contribution of manufacturing to GDP to the district.
    - R-Infrastructure. This includes estimates of infrastructure in each district. This covers dwelling and business buildings and their value, roads, health facilities, food stocks, electrical connections, generators, fuel, water, cell towers, hotels, and early warning systems. Populating this table will require spatial analysis to allocate to the district level. It is proposed that the DPP takes a lead role in this. There is no entry for food and water stores in the metadata catalogue or in the Basic Range of statistics. These would need to be developed.

- R-Environment. This includes estimates of environmental areas of concern (beach, farm, mangrove, coral, forest, biological reserve, freshwater, and marine parks) and fish stocks in each district. Note that coral reefs, mangroves, and marine parks would exist mostly in marine areas.

- R-Risk assessment. This summarizes the exposure, vulnerability, and coping capacity of each district. For exposure estimates, DADS applies an overall risk assessment for each district (Variable: Any Risk High = 'Y') to the district variables described in the population, economy, infrastructure, and environment table. For example, the entire population in a district designated as having 'Any Risk High' is considered exposed. The only vulnerabilities calculated in the DADS spreadsheet are for vulnerable populations ('Pop under 18' + 'Pop Seniors' + 'Pop disabled') and low food and water supply. Coping capacity is calculated for population based on numbers of households ([not educated in disaster management] + [not prepared for disasters] + [no IT access] + [no access to early warning]). There are recognized overlaps between these variables but no simple means of eliminating double counting. One approach may be to select the highest number of the four indicators.

- Impact Assessment Group

  - I-Event. This describes the event according to CRED EMDAT naming and classification standards. Initial impact assessments in the following tables are based on a severity estimate in terms of percent damage, that is, if an initial estimate of 20 percent damage is assigned to a district, then this 20 percent is applied to the infrastructure, economy, and environment tables to provide initial damage estimates. Population impacts are not based on this general severity estimate.

  - I-Population. This records the number of deaths, injured, displaced, job losses, and evacuated in the district.

  - I-Economy. This records the loss in economic production (contribution to GDP) based on the severity of the event. This initial estimate would be replaced by more accurate data when available.

  - I-Infrastructure. This records the damage to infrastructure. Initially, this may be estimated by applying the severity estimate in the I-Event table to the R-Infrastructure table. Airports and ports are not included in the R-Infrastructure table but are included in the I-Infrastructure table. This is because airports and ports are few in number and their risks are national in scope. They should, however, be included in the infrastructure impact estimates.

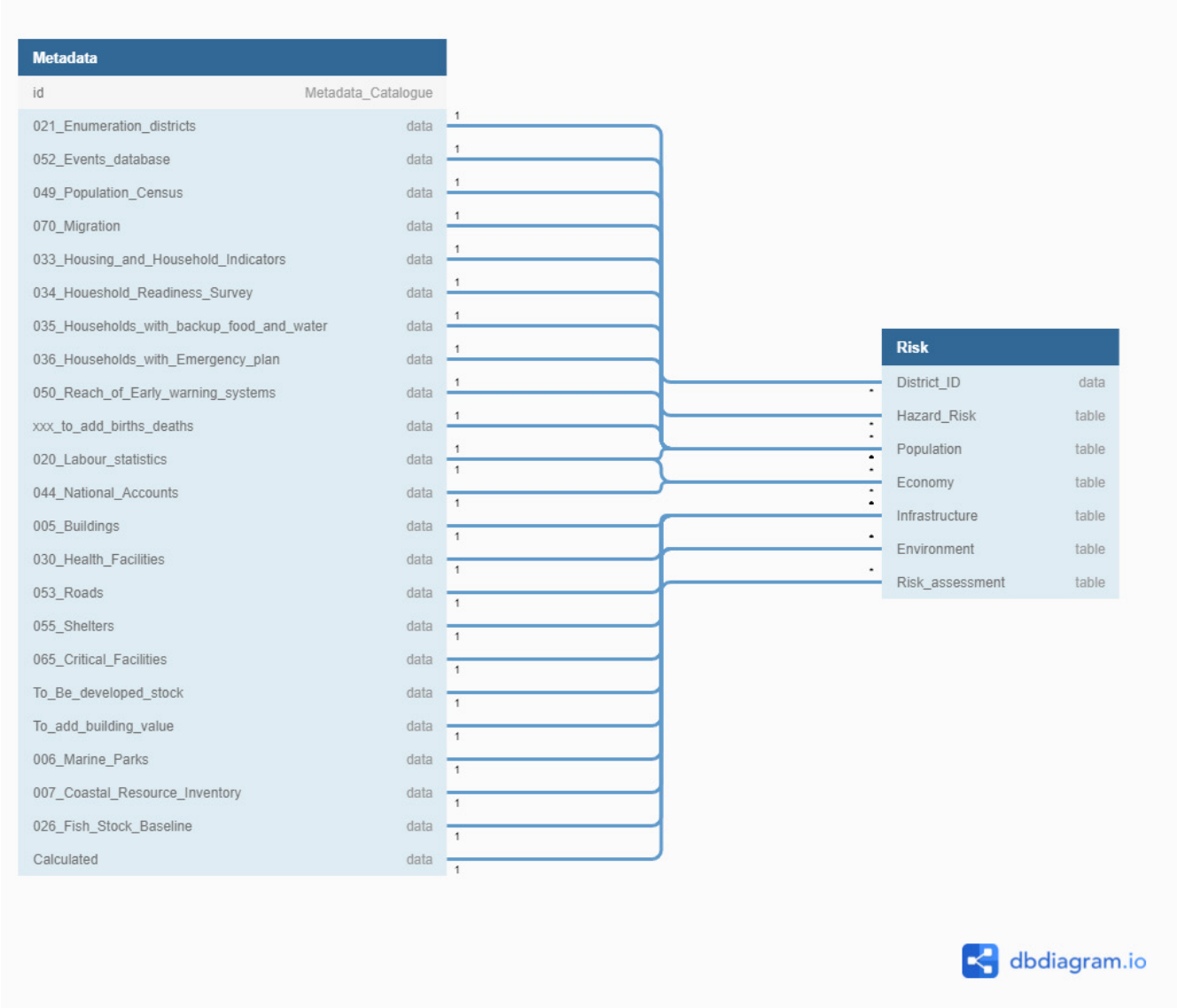
  - I-Environment. This records the damages of environmental areas of concern (beach, farm, mangrove, coral, forest, biological reserve, freshwater, and marine park) and fish stock. It also provides a place to include estimates of the economic costs of these damages. Initial area damages may be estimated by applying the event severity estimate to the areas to the R-Environment table. Note that the value of environmental areas goes well beyond their economic contribution. The areas lost are a better indicator of damage. However, to provide an overall economic estimate, it is suggested that more detailed estimates may be made based on replacement costs or ecosystem services lost for environmental areas excluding biological reserves and marine parks.

  - I-Service Disruption. This is not linked to any tables in the risk group. It records an estimate of the health services, education, public administration, transportation, electricity, water, and ICT services loss in each district due to the event. It is suggested to calculate a pro-rated national disruption in days for each element. An experimental method for calculating this is provided in the spreadsheet.

  - I-Damage. This summarizes the damages from the event in terms of population affected, infrastructure damage cost, environment damage cost, economy damage cost, and service disruption. The detailed Impact Assessment Group tables are much more informative, but the I-Damage table may be a simple means of identifying districts with the most damage. Not all impacts are included in this table because further detail would simply repeat the component impact tables.

Figure 4 provides a simple 'conceptual model' overview of the data sources for the Risk Assessment Group. These are linked to the metadata catalogue rather than individual data sources and departments. No conceptual model is provided for the Impact Assessment Group tables because data are either collected upon the event or estimated.

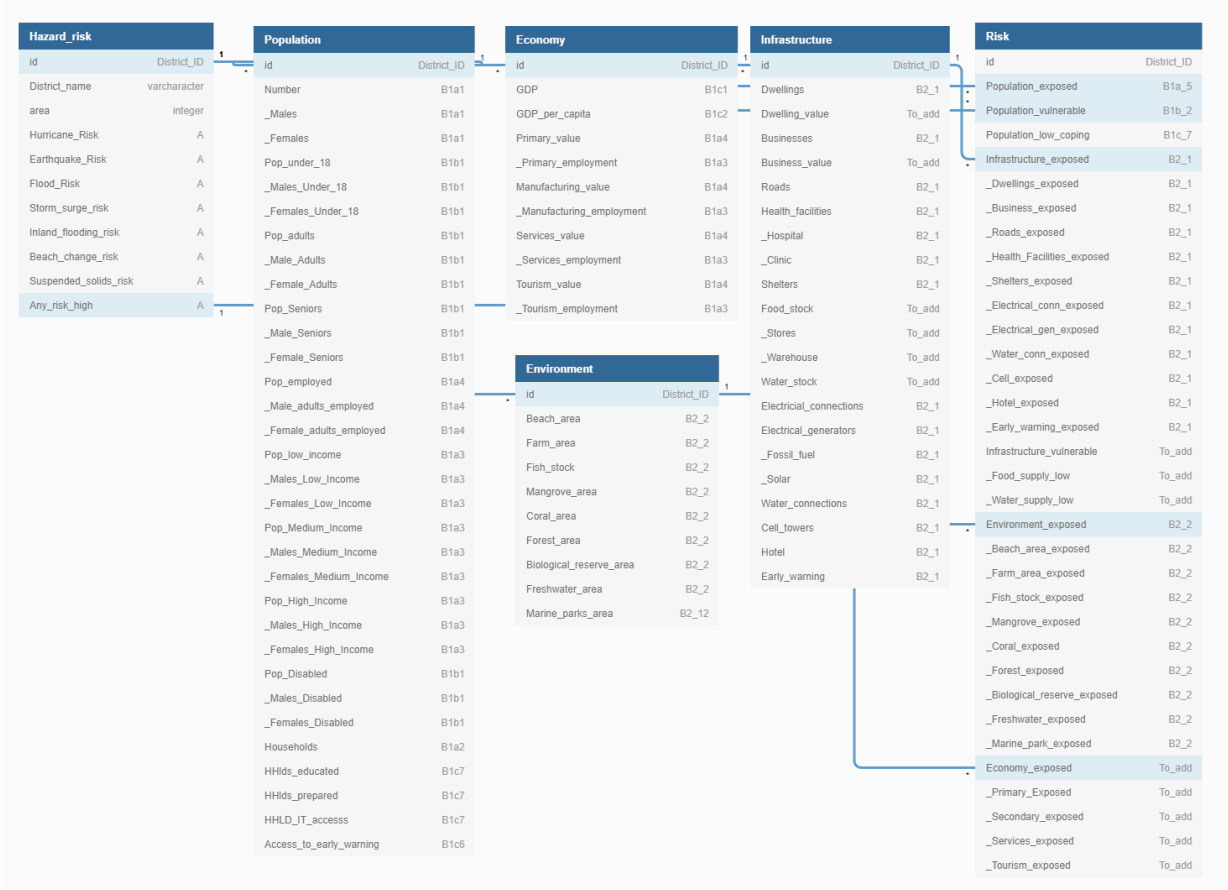
Figure 4: Conceptual Model Entity Relationship Diagram for Risk Assessment Group Tables



4.5.2.2 Logical Data Models

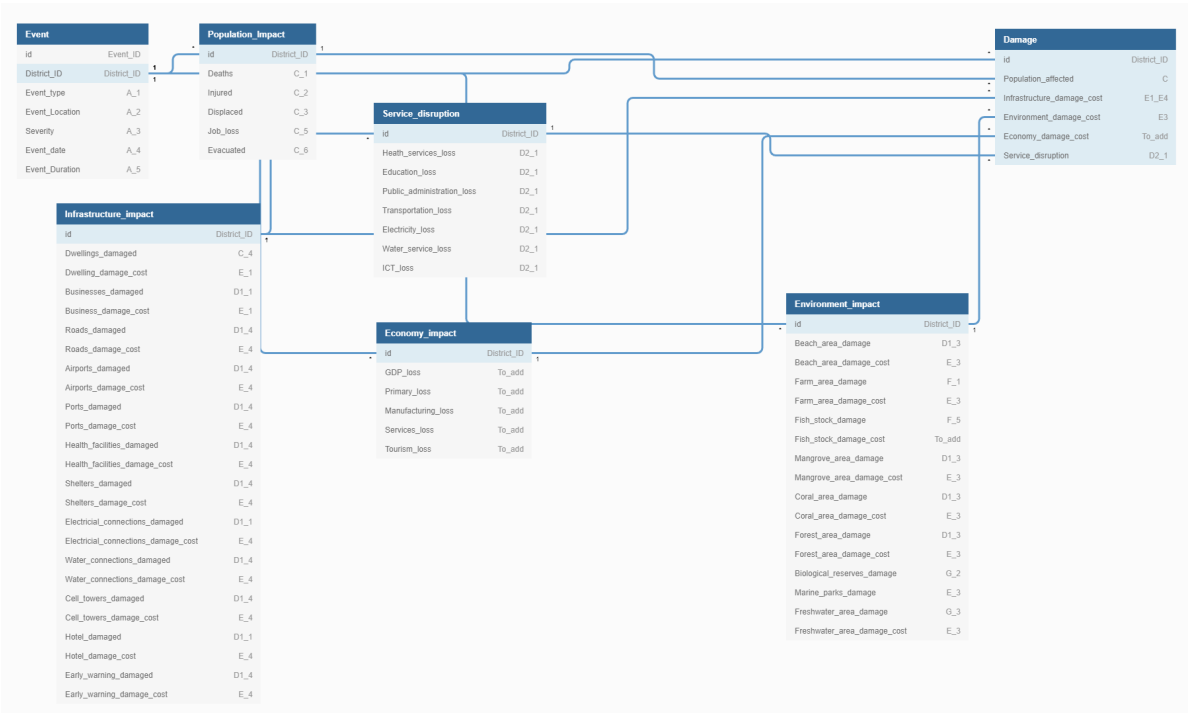
The logical model of both the Risk Assessment Group (Figure 5) and Impact Assessment Group tables (Figure 6) are relatively simple. The only element linking them is the district identifier. DADS does not repeat the descriptions of the individual data sources already provided in the metadata catalogue.

Figure 5: Logical Model Entity Relationship Diagram for Risk Assessment Group Tables



Note: Variable specifiers refer to element of the Basic Range of Disaster-related Statistics.

Figure 6: Logical Model Entity Relationship Diagram for Impact Assessment Group Tables



Note: Variable specifiers refer to element of the Basic Range of Disaster-related Statistics.

4.5.3 Linking DADS, the Metadata Catalogue, and Basic Range of Disaster-related Statistics

Data Obtained and Catalogued

Data were identified from the working group meetings. After being suggested by the team, the Anguilla stakeholders debated the relevance of those indicators and statistics and identified which agencies (if any) held data that contributed to the statistics. These could be the following:

- Statistics, usually held in Excel format
- Baseline mapped data (for example, population census and land cover mapping) which could be held in databases or geographical information systems files
- In exceptional cases, reports containing tabular data, maps, or other statistical information.

For the data obtained (Table 8),

- A total of 71 datasets were identified as either potentially providing the key disaster-related statistics or being source data contributing to one or more disaster statistics.
- Of these, 33 datasets were identified that could be used to provide key disaster-related statistics or contribute to them;
- Of these, 11 datasets can be found online and downloaded freely.
- The other 22 are available offline on written request to the responsible agency.
- Most of these datasets are sourced from the DPP (9) and ASD (5). Few other agencies responded to data requests, although several only had one dataset that was identified as relevant; and
- Of the datasets obtained, the majority were geographically related (vector shapefile or coverage or raster grid). All statistical datasets were in XLSX format. A small number of reports have been included as they contain a reference to key datasets.

Table 8: Total Number of Datasets Provided by File Format for Statistical Data, Geographical Information, and other Formats (Reports or Image Formats).

Format	Count
Statistical Datasets	9
XLSX	9
Other Database	1
Geographical Datasets	22
ArcInfo Coverage	1
GRID	1
SHP	19
MDB	1
Other Reports	5
PDF	5
JPG	2

- Some key datasets, although identified as existing, were not shared with the consultants. Two key ones were the following:
    - The 2011 population census detailed data, although metadata are shared on the ASD website.
    - The 'Districts' GIS layer which would be the aggregating level for regional data in any disaster-related statistics database going forward.

Data Completeness. A measure of 'Data Completeness' has been estimated based on various factors. While this is a subjective assessment, it gives some indication of data readiness.

- Level 1 (Green in the catalogue). This dataset is complete, available, up to date, well documented, and appropriate to disaster-related statistics.
- Level 2 (Yellow). This dataset is complete and appropriate to disaster-related statistics but may be out of date or could benefit from some alterations. It also includes data known to be appropriate but for some reason (broken web links) has not been made available. This also includes some surrogate data sources where a replacement could improve the quality of information being provided.
- Level 3 (Orange). This dataset is known to exist but has not been made available. This can include datasets related to emergency event reporting (for example, affected populations) but while a report exists, the data have not been archived in a systematic format.
- Level 4 (Brown). Some data are known to exist but are not in an appropriate format to be useful for disaster-related statistics (micro or raw data, not aggregated, and not stored in a convenient systematic manner for analysis and integration).
- Level 5 (Red). Dataset needed for disaster-related statistics but is not known to exist for Anguilla.

Table 9: Tally of Evaluation of Dataset Completeness (based on descriptions of level above).

Level	Number of Datasets
1	19
2	24
3	15
4	7
5	13

Table 9 shows that a significant proportion of the datasets needed are either available and in good condition to be used, although many would benefit from some simple redesign or updating. It also suggests that only a smaller number have data which are not in good condition for use in disaster-related statistics. However, there is a significant portion of datasets which are needed for Anguilla's disaster-related statistics and identified as required, which have not been created or identified (17).

The discrepancy between the number of available datasets (33) and those at levels 1 and 2 of data readiness (43) is that the exact location of some datasets and the responsible agencies have not been determined, but metadata has been obtained in the catalogue for completeness.

Table 10: Data Themes and Sources

Theme - Agency	Boundaries	Economy	Environment	Society	Utilities Communication	Structure	Planning Cadaster	Disaster	Transportation	Location	Health	Elevation	Oceans	Total
Agriculture Unit		2												2
ANGLEC					1									1
AHTA		1												1
Anguilla National Trust				1		1								2
ASD2		1		4										5
CDEMA														0
DDM								2						2
DITES					1									1
DPP	3		2	1			1		1	2				10
EM-DAT								1						1
Environment Systems Ltd.			1									2		3
Environment Unit, DNR			1									1		2
Fisheries Unit	1		1										1	3
Joint Nature Conservancy Council			1											1
Princess Alexandra Hospital											1			1
Water Corporation of Anguilla														0
Total	**4**	**4**	**6**	**6**	**2**	**1**	**1**	**3**	**1**	**2**	**1**	**3**	**1**	**35**

Table 10 shows a high proportion of datasets cover environmental themes, boundaries, and economics (these being the official statistics) but there are significant gaps in the utilities sector, transportation, and ocean-based datasets.

4.5.4 Comparison against the Key Datasets List

The ideal suite of statistics for Anguilla (DSSAT) has seven agreed topics. Each topic has been divided into several sub-topics within which there are several individual statistics; 54 individual statistics were identified as needed for the ideal set.

Some of these statistics can be identified from single data sources (for example, population), and others require several datasets analyzed to create the final statistical set. Several of the statistics required further disaggregation before being able to identify the key datasets needed, for example, 'loss of critical biological reserves and other designated ecosystems, ecosystems hosting endangered species, and other critical ecosystems (by hazard type, land cover type, and region)'. Multiple datasets will be required to cover these factors which include the following:

- Different types of reserved land/marine area are held in multiple datasets.
- Hazard mapping is needed in several layers to describe risk (for example, from storm surge, wind damage, and earthquake).
- Regions can be potentially categorized in different ways (for example, by country or by district).

Working through these different parameters, the number of individual statistics/dataset mixes required was 87.

A number of these are a fuzzy match to the required data. Recommendations on how to improve the match are summarized below and contained within the catalogue under the metadata log tab.

In total, 22 of the individual statistics required could not be matched with any existing data as catalogued in the workbook. This accounts for 40 percent of all the statistics required. This does not mean the data do not exist, only that the data for those statistics were not made visible to the consultants for cataloguing.

4.5.5 Recommendations for data improvement

Stakeholders made it clear during the working groups when they knew that the required data for disaster-related statistics needed improvement. The review of data that was shared with the consultants also gave an opportunity to examine how datasets could be improved to be 'fit for purpose'.

The full list of recommendations is found within the metadata management tab of the catalogue and are not repeated here. But the types of recommendation made are the following:

1. Dataset templates can be designed to be useful for collecting data post the event (for example, assessments of percentage infrastructure damaged). Many templates are being developed for standard assessment methods (for example, CDEMA developing Damage and Needs Assessment standard forms/datasets).
2. Systematic methods are needed to take survey data and archive in a single database for a theme (for example, for fish stock assessment data).
3. It is advised to establish a complete inventory stock taking of some data holdings (for example, GIS data) to ensure it is up to date. Triangulation of data holdings from different agencies may be necessary (for example, combine the Hotels Association property listings with the GIS layer).
4. Transfer of data storage techniques from simple reporting (for example, for the After Action Hurricane Irma report) into standard data spreadsheets—the raw data are more easily accessible and useful for other data managers and save effort in data transfer.
5. Some redesign of datasets should be considered to make them fit a wider set of output needs. For example, a detailed marine area dataset was created for the ACRAMAM Project but the multiple zones could be aggregated to 4–5 areas for reporting on the impact of disasters on the nearshore waters of Anguilla.
6. Increased attribution (both new columns of information and detail within each data element) will increase the richness and usefulness of single datasets. This would fit the mantra of 'create once, use many times'.
7. Ensure that any consultants working on new datasets that have multiple uses leave an archive with the host agency. All GIS data created in externally contracted projects should be deposited with standard metadata with the DPP. The data derived from the Natural Capital report, for example, was only partly available from the GoA.
8. International datasets are often useful where no nationally created datasets exist but localized data repositories will be more appropriate as they can be defined, held, and updated with more control. Relying on others to create datasets means data definitions and content may change without consultation, or the initiative may end at a point leaving Anguilla with no access to an important dataset for its use.
9. Some datasets would benefit from a more comprehensive quality control procedure. Reduction of spelling mistakes in the data and more standardized categorization (which helps in both visualization and in aggregation/analysis) will result in a better dataset for use by a wider set of stakeholders.

4.5.5.1 Metadata Completeness

Using a simple measure to look at completeness of mandatory metadata in each of the four sections of the catalogue, on average, those datasets which were identified and catalogued, the status of metadata completeness is 36 percent.

Table 11: Completeness of Mandatory Metadata in Each of the Four Sections of the Catalogue

Metadata Theme	Percentage of Datasets where Mandatory Fields Are Complete
Main or Dublin Core metadata	33.0
Geographical	38.5
Statistical	30.0
Metadata Management	43.0

Most of the missing Dublin Core elements are related to multiple date fields where the information was not forthcoming or could not be documented as they were not relevant to the particular dataset.

The statistical datasets have very limited metadata that is relevant to geographical mandatory fields, which suppresses the reported percentage of completeness. Similarly, the detail of geographical datasets precludes the need to document statistical mandatory fields, so this explains some of the low percentage of completeness. But while this measure is crude, it does show that it is possible to add more richness to each dataset's metadata.

A review of the relevant datasets by agencies could substantially improve the metadata completeness figures. More detailed metadata would also enrich the usefulness of the catalogue for searching for relevant datasets and identifying gaps and management actions (for example, to update frequencies).

4.5.5.2 Training of Staff

For the attendees of the awareness raising seminar and introduction to metadata segment in the general training session, see the section on Working Groups and General Statistics Training.

For the metadata training, six people were trained on Thursday, April 15, 2021: four staff from the ASD (including the Chief Statistician) and two staff/consultants from the DPP. The feedback from the session was that, for the junior staff, it was a useful introduction to the importance and usefulness of metadata as well as a chance to experiment with the tool for documentation. For the head of the ASD, the session was useful to provide her validation of arguments for the need for good documentation.

There was not enough time allocated to go through detailed examples of metadata entry or experiment between the two methods of data entry. This has meant that extensive testing of the catalogue and embedding it within staff systems was not possible. It is hoped that the design considerations and SOP Manual will allow its take up beyond the project lifetime.

The DADS spreadsheet (Annex 9) provides detailed links between the variable and the metadata catalogue. The metadata catalogue describes the institutional context and contents of key databases. As source databases change (for example, a question in a survey or categories in a classification), the metadata catalogue and the DADS design will require updating.

The Basic Range of disaster-related statistics identifies key data gaps, which when addressed, will also contribute to the continued development of DADS.

4.5.6 Database Maintenance

It is recommended that the ASD assume responsibility for populating DADS including coordination of inputs from data holders. This could be formalized by assigning one ASD staff member as 'DADS Coordinator'. The DADS Coordinator would be responsible for establishing a regular schedule of updates. Updates would be in order in the following circumstances.

- New data are collected. New surveys or administrative data may be initiated either specifically for disaster management or other purposes. This should be seen as an opportunity to improve the relevance, quality, and timeliness of DADS. This could result in the addition of variables or detail to DADS.
- Data are revised. Revisions to existing data sources may result in additional, more accurate, or more timely data being available to DADS.
- New data sources are discovered. For example, as part of the ongoing maintenance of the metadata catalogue, existing data sources may become available. International agencies are also improving their curation of national data and use of 'big data'. Annex 10 contains a list of datasets in HUMDAT referring to Anguilla. While some are empty, others, such as the Humanitarian Open Street Map data may be more up to date than national data.

Besides providing input to an annual disaster risk assessment for the NDMC, DADS will also provide an overview of the quality of key disaster management data. By highlighting the vintage and quality of input data, an annual DADS review would help set priorities for improving source data and metadata.

It is recommended that all departments collecting disaster-related data contribute district-level summaries to DADS; the key roles and responsibilities are the following:

- Anguilla Statistics Department
  - Coordinate integration of data into DADS, assure fitness for use, report annually to the NDMC, ensure copies of DADS are securely maintained and accessible to key disaster response individuals, conduct an annual assessment of DADS, and coordinate the revision of DADS when required.
  - Establish a network of DADS data providers in departments and a regular schedule of updates.
  - Provide district-level estimates of population, social, infrastructure, and economic variables under its domain.
  - Contribute to the computation of reporting of key risk assessment and impact assessment variables in DADS.
- Department of Physical Planning
  - Provide district-level (and marine area-level) data under its domain.
  - Support other departments with less spatial analytical capacity in allocating their data to the district level.
- Department of Disaster Management
  - Provide data under its domain at the district level. If necessary, collaborate with the DPP to allocate data to the district level.

The project provided two main training sessions: one as an introduction to disaster-related statistics and the other specifically on the metadata catalogue. These could serve as the basis for ongoing training of the DADS Coordinator and staff designated as DADS data providers.

The main components of the overview of disaster-related statistics were the following:

- Introduction. What is a disaster? Why do we need statistics?
- Disaster-related frameworks. Overviews of the Sendai Framework, DRSF, SDGs, UNECE Recommendations on the role of official statistics, CRED EMDAT, PDNA, and others
- Data requirements for risk assessment, disaster response and impact assessment
- Key statistical concepts for providing disaster-related statistics
- Introduction to data sources and how they are collected.

The main components of the training on metadata were the following:

- Metadata concepts and standards - why do we need this?
- Data types
- The proposed metadata schema - what is to be recorded about each dataset?
- The proposed metadata catalogue - structure and operation
- Proposed roles and responsibilities.

In addition to these introductory trainings, the DADS Coordinator and key DADS data providers, especially those in the DPP and DDM, would include basic training in the subject matters covered by DADS. These would include concepts, classifications, data collection, and analytical methods in the following:

- Social and demographic statistics (census, surveys, registers, and administrative data)
- Economic statistics and national accounts (business surveys, registers, and administrative data)
- Environment statistics, including geographic and spatial analysis (ecosystems, natural resources, biodiversity, and environmental quality).

Some examples of existing course programs that could be taken online or used as models include the following:

- PDNA online training (https://olc.worldbank.org/content/post-disaster-needs-assessment-pdna-online-training)
- Environment statistics online at ESCAP: https://communities.unescap.org/environment-statistics
- General online official statistics courses at the Statistical Institute for Asia and the Pacific (SIAP): https://siap-elearning.org/course/index.php. In previous years, there has been a face-to-face comprehensive course on official statistics conducted over about 12 weeks in Chiba, Japan.

4.5.7 DADS Implementation Work Plan (3 years)

This work plan describes a three-year project, which after completion is absorbed into the GoA regular budget. After three years, maintenance and upkeep of DADS would be required in the order of one-quarter person year (PY) from the Anguilla Department of Statistics and perhaps another quarter PY distributed across the government. Other direct costs will be minimal; however, this does not include improving and documenting the source data.

As with most statistical development activities, this three-year project will require intensive dedication of human resources. One option would be to engage a local consultant for the first 18 months to support the development of DADS and improvement of the metadata catalogue. This would include reviewing the design with stakeholders, making any necessary revisions, populating the initial database, and revising the metadata catalogue. This would limit the time required by the

DADS coordinator for the ongoing allocation of one-quarter PY. The local consultant could also work with data providers to set up a regular data provision process, including spatial analysis.

The three-year project is planned in three phases: (a) prototyping, (b) scaling up, and (c) regular maintenance. Funds would be required to (a) engage a local consultant for 18 months and (b) engage junior staff at the ASD to replace regular duties of the DADS Coordinator and other staff engaging in training and working more intensively on DADS over the first three years. Data providing departments may also require additional staff to support the project. This totals about 18 months of local consultant time and approximately 6 PYs of junior staff time over the three years. The following describes the work plan in more detail.

- Prototyping (Months 0 to 6)
    - Establish the DADS Coordinator and network of data providers and review database design (and conduct training).
        - Identify priority missing data and establish a two-year plan for incrementally adding new data elements.
    - Assemble most recent data in the DADS spreadsheet, link to metadata, review metadata catalogue, and identify priority improvements.
        - This initial incorporation of data into DADS will necessarily focus on the most readily available data and identify data elements which will require more work in the scale-up phase.
        - Test the online version of DADS and establish access protocols and individuals with access. It is recommended that access to DADS be limited to persons certified by the Statistics Act, including key personnel in data-providing departments and local consultants.
    - Report on challenges, output tables, and recommendations for revision to steering committee.
    - Produce proposal for the following:
        - One-year project to scale up to regular production (costs, personnel, institutional requirements, design changes, and major milestones)
    - Institutional requirements may include improved data sharing arrangements and additional capacity building.
    - Design changes may include additional data elements (from improved access or new data collections) and substitution of sources (for example, international proxies may have local sources).
    - Additional data development may be required to complete DADS. This would include, for example, creating consistent digital files at departments, reclassifying data, georeferencing data, conducting spatial analysis to overlay, and allocating data to the district level.
        - 18 months ongoing maintenance as funded activity
        - Budget for stakeholder departments' support personnel, IT infrastructure, and external consultants.
- Scaling up (Months 7 to 18)
    - Based on experience in initially populating DADS in the prototyping phase and identification of priority data development required, support the ASD and data providers in additional data development.
        - This would also include improvements in data quality, timeliness, and metadata of source data.

- By Month 15,
  - The historical event and damage data should be of sufficient quality and consistency to Desinventar (UN DRR https://www.desinventar.net/);
  - A prototype risk assessment report, certified by the Chief Statistician, should be available for discussion with the NDMC; and
  - A prototype disaster data assessment should be available to identify priorities for further data and metadata improvement.
- By Month 18,
  - Establish a schedule of regular updates to DADS and work plan for improving data and metadata.

- Ongoing maintenance (Months 19 to 36)
  - Conduct regular updates to DADS and the metadata catalogue.
  - Conduct ongoing data and metadata improvement as established in Month 18.
  - By Month 27,
    - Provide the first comprehensive DADS risk assessment and data assessment to the NDMC and key stakeholders and
    - Provide an updated disaster data assessment with revised priorities for data and metadata improvement.
  - By Month 30,
    - Submit a project evaluation and proposal for ongoing budget.

5. Conclusions and Recommendations

This document has outlined a possible approach which the GoA could pursue toward an enhanced capability for capturing the impacts of disasters. The authors of this report took care to propose solutions which are within the GoA's implementation capability. However, implementation of these solutions will require a long-term commitment from the ASD and its partner agencies as well as donor support to supplement the technical capabilities of the GoA. The following sections outline the specific recommendations for the ASD with regard to disaster-related statistics and for the wider GoA concerning the data governance arrangements necessary to enable the sustainable creation of disaster-related statistics.

5.1 Conclusions and Recommendations for the ASD

The authors offer the following recommendations to the ASD for improving Anguilla's disaster-related statistics:

The ASD should continue with its existing good practices. The ASD is to be commended for a number of existing practices and for managing a wide-ranging program of statistical outputs with a small team of professionals and budget. Good practices we noted that should be continued include the following:

1. **Effective statistical business practices.** Our review of the ASD's generic business practices against the international standard GSBPM demonstrated that despite its limited financial and human resources, the department does a remarkably good job of employing good practices. The existence of a statistical law mandating the ASD to collect official statistics on behalf of the government and guaranteeing its independence from the political side of government demonstrates the seriousness with which the GoA views evidence-based decision-making. The ASD's use of population and business registers, international standard classifications, and modern dissemination tools (Facebook and infographics) all reflect best practices. Similarly, the review of the ASD's operations specifically with reference to disaster-related statistics showed a strong, professional foundation upon which to build even if the department has very few resources (considerably less than a single FTE) to devote to the topic. The policy environment for disaster-related statistics is favorable, with a national disaster management policy in place (Government of Anguilla 2013) that calls expressly for evidence-based decision-making to "ensure that the comprehensive disaster management objectives at the community and national levels are achieved…based on effective disaster monitoring information." At the same time, the ASD noted that it did not have a specific plan in place for the improvement of disaster-related statistics and that work on the National Strategy for the Development of Statistics (NSDS), which had been started some years ago with the aid of PARIS21, had stalled. Given this, we recommend that the ASD seeks to reinitiate work on the NSDS with disaster-related statistics given high priority in the plan.

2. **Effective multi-stakeholder engagement.** The ASD is clearly capable of creating and managing multi-stakeholder engagement processes focused on improving official statistics. When asked to create a multi-stakeholder working group for this project, the ASD was both willing and able to so and managed the group effectively to achieve its ends. Having invested the time and effort to create this group and facilitate its work, we recommend that the ASD keep the group in place for the foreseeable future with a renewed mandate to serve as an interagency working group focused on improving Anguilla's official disaster-related statistics. The ASD should consider co-chairing this group with another agency (ideally the DDM) so that its ownership is shared with an important user of disaster-related statistics. This will help ensure that users' needs are fully reflected in the group's work.

3. **Openness to user consultation.** The ASD is clearly aware of the need to engage users to understand their statistical needs and at least one formal mechanism—the ANSAC—exists for this purpose. At the same time, the ASD reported that it operates in a context in which users (or potential users) of statistics do not reflexively think of statistics as a key input into delivering on mandates. Rather, Anguilla remains a society where the use of statistics in decision-making and program delivery remains something users need to be reminded of. We also heard that identifying users' needs is just one of many roles of the ANSAC and definitely not its main role. Given

this, we recommend that the ASD strengthens its capacity to identify user needs by creating an annual forum for user consultation. This forum should be organized and run by the ANSAC, with a formal report provided to the Chief Statistician. Such a forum would have two benefits. First, it would provide the ASD with regular feedback on the needs of its users. Second, it would provide an opportunity for the GoA, though the ANSAC, to emphasize the importance of evidence-based decision-making. Over time, such a forum could go a long way toward building a culture of statistics users within Anguilla.

4. **Effective dissemination practices.** We noted several effective statistical dissemination practices in Anguilla. The ASD has recently launched a modern website that facilitates users' access to its statistics. The department also makes effective use of other dissemination channels, including the GoA's Facebook page. At the same time, there is no current statistical portal dedicated to disaster-related statistics nor is it obvious how disaster-related statistics would be integrated into the ASD's new website. The statistical holdings on the website are structured following the UN Classification of International Statistical Activities (CISA)[28], which does not include an explicit category for disaster-related statistics. Rather, disaster-related statistics would be potentially found under any of a number of headings of the CISA. We see two options for addressing this issue. Preferably, the ASD would consider restructuring its website so that it does not explicitly follow the structure of the CISA. This would permit it to develop a dissemination portal devoted specifically to disaster-related statistics. It would also avoid one of the problems we noted with the website, which is that the ASD has no date offerings under a number of the CISA domains (for example, social protection and income and consumption). In such cases, clicking the link on the website leads nowhere, leaving the visitor to wonder whether the link is broken or whether data do not exist. Restructuring the website to eliminate domains for which the ASD has no data and focus on those where it does would not only improve users' experience with the website but also permit the ASD to dedicate parts of the website to important cross-cutting issues such as disaster-related statistics. If restructuring of the website is not possible, the next best option would be to add disaster-related statistics as an explicit heading under either domain 3.2 (regional and small-area statistics) or domain 3.3 (environment and multi-domain statistics) of the CISA. While neither of these explicitly includes disaster-related statistics as a sub-domain, there is no other domain of the CISA that is better suited to housing them. In either case (restructuring of the website or not), we would recommend that the ASD remove from the website links to any domain for which it has no statistics available.[29]

In addition to continuing its good practices, the ASD should build upon what has been accomplished during the course of this project to continue down the path of improving disaster-related statistics.

1. Several of the UNECE's recommendations to NSOs engaged in improving disaster-related statistics have been implemented as part of this project: user consultation has been undertaken to identify needs for disaster-related statistics, reveal the gaps in existing statistics, and prioritize those gaps for filling. The next step is to establish a clear plan for compiling the statistics necessary to fill those gaps, starting with those considered to be highest priority. As noted above, such a plan could be a centerpiece in a renewed effort to create a national strategy for the development of statistics.

2. **The ASD is encouraged to make full use of the metadata catalogue that has been developed as part of this project** (Section 3.4). The catalogue has been designed specifically with use of use and maintenance in mind and the ASD is encouraged to keep it up to date and to ensure that any new disaster-related statistics developed over time are recorded in the catalogue.

28 https://unstats.un.org/unsd/iiss/Classification-of-International-Statistical-Activities.ashx.
29 We also noted a number of domains for which the website implies the ASD has data holdings where the data tables apparently stored on the website cannot be opened (giving a 'File or directory not found' error when clicked). This is the case, for example, for domain 2.5.2 (public sector debt) and domain 1.10 (justice and crime statistics). We recommend that these domains also be removed from the website—or that the broken links to the data tables be repaired—to improve the user experience.

3. Recognizing that metadata are necessary but not sufficient for good statistics, the ASD is encouraged to create and implement a plan to turn the conceptual design for the DADS into a reality. Like the metadata catalogue, DADS has been designed to be simple to implement and maintain. The ASD should begin to populate DADS with the existing disaster-related statistics that were identified during this project (Section 4.3). As new statistics are developed, they should be added to both DADS and the metadata catalogue. Further recommendations to this end are given in the next section.

5.1.1 Recommendations for the Design of a Database for Anguilla Disaster-related Statistics

The design for the proposed DADS outlined in Section 4.5 was based on the following:

- Ongoing consultation with the GoA though a series of four working group meetings
- Input from the project components on assessment of the GoA Information Architecture and metadata catalogue
- Training sessions on introduction to disaster-related statistics and metadata
- Feedback from the Anguilla Disaster-statistics Data Assessment Working Group that identified the proposed Basic Range of disaster-related statistics and priority data gaps (see sections 4.2 and 4.3).

The design distributes responsibilities for data provision to data holders, who are responsible for the data and have expertise in using it. Given human resources and time, a more sophisticated design would be possible by allocating more responsibility to the ASD. A dedicated team of 2–3 staff with expertise in the subjects covered and technical capacity to conduct spatial analysis could support such a sophisticated design and absolve data providers of much of the responsibility for preparing data. Such a design, rather than allocate all data to the statistical district level, would maintain data in their original spatial formats and allow for more spatially detailed risk and impact assessments. This would therefore allow for tracking of individual people, households, and assets as proposed by Rios Diaz and Marin Ferrer (2018).

Following are the main recommendations related to DADS resulting from this project:

- **Structure**
  - DADS should be conceived as a set of tables at the statistical district or enumeration district geographic level, with extensions to the ocean to encompass marine ecosystems.
  - Implementing the design and operationalizing the resulting database would provide
    - An annual risk table including
  - A compilation of the current risks by hazard type
  - The exposure of people, the economy, infrastructure, and environment to multi-disaster risks, singling out vulnerabilities and coping capacities where feasible, and
  - A summary risk assessment suitable for prioritizing risk-reduction activities.
    - An input to disaster response by indicating exposed, vulnerable, and low-coping capacity districts.
    - A framework for data collected post disaster as well as initial estimates of impact including
  - Standard reporting of the disaster event and estimates of severity;
  - Population, economic and environmental impacts, service disruptions; and
  - A summary impact assessment suitable for contributing to the PDNA.

- **Institutional requirements**
  - The ASD should assume responsibility for populating DADS, including coordination of inputs from data holders.
  - All departments collecting disaster-related data should contribute statistical district-level summaries to DADS.
  - One ASD staff should be designated as the 'DADS Coordinator', who would regularly interact with designated stakeholder data providers.
  - The DPP should take a lead role in supporting the data preparation.
  - A regular, annual deadline database update and release should be established.
- **Technical requirements**
  - An accessible copy of DADS should be placed in the cloud to which key users (NMDC, ASD, and others) would have access during disasters.
- **Quality Assessment/Quality Control**
  - 'Fit for use' rules that consider the age of the data and dynamics of the phenomenon being measured should be applied.
  - Some data sources mentioned in this DADS design may be of uncertain quality but not identified as such in the metadata catalogue
    - The metadata catalogue entries for these should be updated to include quality information.
  - An annual quality statement should indicate priority source data and estimates that require improvement over the next cycle.
- **User access, security, and logging**
  - All persons accessing the database should be designated employees of the ASD for DADS
    - This would necessarily include members of the NDMC.
- **Output formats and reporting**
  - DADS, in spreadsheet form, should be ready to set up to print for reporting and archival purposes.
  - The DADS report would serve as an annex to the annual risk assessment provided by the NDMC.
- **Database maintenance**
  - In addition to introductory trainings, the DADS Coordinator, and key DADS data providers, especially those in the DPP and DDM, would include basic training in the subject matters covered by DADS.
- **Database implementation**
  - A local consultant should be engaged for the first 18 months to support the development of DADS and the improvement of the metadata catalogue.
  - This would include reviewing the design with stakeholders, making any necessary revisions, populating the initial database, and revising the metadata catalogue.

Design and Implementation Scheme for Using Data and Updating the Metadata

The cataloguing of metadata should not be seen as an archiving process to understand the spread of current data but should be the first step in implementing a program to improve the reporting and use of disaster-related statistics. To that end, steerage is needed to ensure that the recognized outputs or those statistics are identified using the GSBPM process, the datasets are identified or created, the documentation is maintained, and association between the metadata and the output is made. Currently this matching of data to output is done through the DSSAT process, but there should be no issue of adding more reporting mechanisms (both national and international) and identifying those data for any new processes.

It is important to ensure that the metadata are consistently updated regularly (can be infrequent but has to be regular). For Anguilla, two particularly critical points would be when reporting is coming up (for example, for Sendai or SDG targets) and in advance of key seasons, for example, hurricane season.

Use this Report as a Lobby and Roadmap for Better Data Governance

The metadata stream of this project has worked through a strategy for implementing metadata for one sector of government business, namely being able to report disaster statistics. It has identified a process that involves describing internationally recognized data elements in a coherent and comprehensive schema, implementation of that for data documentation, training and manuals to support the procedures and protocols of managing metadata, and identification of a way of translating that into a shared online catalogue. Given Anguilla's small size and human capital constraints, it would be inefficient to continue metadata management just for this small section of government data management. Instead, it should be used as a case study to advocate for better data governance throughout Anguilla's administration. Identifying schema which suit various types of dataset and comprehensively documenting these will raise awareness of data more widely and provide much richer detail to support archiving, end uses, and strategically managing updates.

The tools and instruments identified within this project support the Chief Statistician in his/her drive to improve how statistics can be obtained for disaster management, encourage interdepartmental cooperation and liaison, and smooth data exchange.

5.2 Conclusions and Recommendations for Data Governance and Integration

A sustainable system for disaster-related statistics is dependent on the wider data governance context and infrastructure. These were investigated in detail in Section 3 of this report. Strengths and weaknesses were identified. In totality, the network infrastructure of Anguilla was found to be a strength which could support greater integration of systems and wider sharing of information. However, policies around government-wide information management and sharing of nonexistent and ad hoc arrangements are the norm for data sharing. This is a strength when persons are well acquainted, have positive working relationships, and the time to respond to requests. This is not the case in a disaster response scenario as persons are not available or capable of responding to requests on time and unfamiliar persons are often making these requests. In these situations, data need to be available ahead of time and a formal process put in place to allow those from outside of Anguilla to request data. This project encountered difficulties in persuading agencies to share data to examine the status and document information about the data.

Formal data sharing relationships between agencies and a formal process for obtaining data access for outside persons and agencies would greatly assist in regularizing the movement of information such that critical information may be available before a disaster and not need to be requested during the response phase. While MoUs are being used currently to set up these formal relationships, fundamental enabling legislation around access to information rights, cybersecurity, data protection, and data privacy are nonexistent. These legislations will improve digital protection and will regulate the collection, organization, storage, and retrieval of data. The creation of a data governance framework can help organize the implementation of various legislations mentioned. Changes to legislation will require the leadership of decision-makers and the EXCO to implement.

Private sector agencies were noted in the assessment as having a greater capacity than the public sector for data management. Lessons could be learned from these agencies and applied within the GoA. Formal data exchange relationships between the GoA and the telecommunications sector were also noted as a strength. The authors note that a similar relationship could be formalized with tourism operators for the sharing of information required for disaster-related statistics.

The GoA is implementing a series of information systems for health, education, and land information Management. These systems were investigated during this report and the digitization of information was noted as a strength. However, these systems are not being integrated in any way and data exchange between systems will require manual intervention, which will put further demands on the limited number of trained GoA staff. There is no ICT framework or strategy in place guiding the implementation of these systems and they are being put in place on a per-agency/project basis with no plan for integration. Metadata are a critical component of systems and data integration and the work presented in this report could inform the creation of a GoA-wide metadata profile which could form part of an eventual government-wide systems and data integration exercise.

Major weaknesses were noted in the application of data life cycle processes. Most of the agencies are not familiar with these processes and this reduces their ability to effectively manage the data. Training in these processes was noted as a need. The lack of data quality assurance officers was noted as a challenge. There is little monitoring of the quality and consistency of data and Section 4.5.5 of this report makes recommendations for improvements which could be considered not just for statistical data.

While the network infrastructure of the GoA was noted as a strength, the current data storage arrangements are a potential vulnerability. Presently, most of the data servers are located at a centralized location, which increases the risk of damage to the infrastructure and data loss. Utilizing secure off-island backup infrastructure for data and systems should be considered.

Regional collaboration and support for improving the enabling environment for data management, sharing, and the development of disaster-related statistics will be key. Examples of international best practices for regional data sharing and management initiatives include the Infrastructure for Spatial Information in the European Community (INSPIRE) Knowledge Base.[30] This is a collaboration between EU member states to systematically catalogue and share data at different scales and over different temporal revolutions, so as to be easily found for environmental purposes. There is a diverse community of interest that understands the need to document and make ready data that can be shared, the tools and data are available, and it is backed by legislation across the European Union. The position INSPIRE is in is only due to 14 years of difficult negotiation, learning, and building, based on the foundations of previous initiatives and significant investment and support. It also has a large user base and technical support available.

Anguilla has benefited from regional initiatives implemented by agencies such as CDEMA and the Organisation of Eastern Caribbean States (OECS). The OECS implemented Electronic Government Regional Integration Project (E-GRIP) assisted Anguilla in the setup of its financial management system and provided a possible model for an enterprise integration platform for OECS countries. This model was never adapted however and would need updating as it is over a decade old, and the technology and standards presented have become outdated. The OECS is in the process of implementing the Caribbean Digital Transformation Project (CARDTP)[31] which is assisting in the development of policies and standards for data sharing and development of key systems (such as national identifications) for Eastern Caribbean States. Anguilla could benefit from some of the outcomes of this project to inform the development of its own policies and standards. CDEMA offers trainings in damage and loss assessment as well as disaster information management. Officers from Anguilla attended trainings offered in January 2020 for the Caribbean Risk Information System and the topics covered in this training were applicable to the development of DADS and disaster-related statistics.

30 https://inspire.ec.europa.eu/.
31 https://www.oecs.org/en/our-work/knowledge/library/projects/caribbean-digital-transformation-project-cardtp

Within the Caribbean region, the most holistic and active data sharing has perhaps been within Jamaica (https://data.gov.jm/). Their open data portal (funded through the World Bank and UK aid) primarily for budget readiness (cf. the parallels with Anguilla's need to be ready with disaster-related statistics). This initiative, ministerial championing, backed by steady lobbying by permanent secretaries and department heads for money and space, works out how best to manage data sharing. In particular, it has been created for geographic data because there has been a series of champions and development of strategic partnerships both within government and outside (for example, with software providers and technology/training support). Even with this good design and initial uptake, the Jamaica open data portal was only updated up to 2015. Effort is clearly needed not just to create such a system but also to sustain the effort. If that maintenance angle is not sustained, the up-front costs of building new systems can deter future expenditure to resuscitate activity.

Several of the metadata schema used within this project have supporting documentation and user groups that could aid any larger implementation of metadata systems across government. For example, the DCMI includes a community which helps the development and implementation of good practice and standards for that core top level metadata.[32]

For geographical information, the United Nations is promoting better information management arrangements though the 'Integrated Geospatial Information Framework'.[33] This is a practical framework which countries can adopt and has been recently expanded to a more detailed implementation guide for member states. It is expected that funding will continue to be made available (partly through World Bank programs) to support member state implementation, while this is not applicable to Anguilla at this time, a plethora of supporting materials and experiences is openly available through various forums and websites.

32 https://dublincore.org/themes/community/.
33 https://ggim.un.org/IGIF/.

REFERENCES

CARICOM. 2018. The CARICOM Regional Strategy for the Development of Statistics. Georgetown, Guyana: Caribbean Community (CARICOM) Secretariat. https://statistics.caricom.org/Files/RSDS2020.pdf .

Chai, S. S., and D. Suh. 2020. "Design and Implementation of a Multi-Disaster Risk Assessment Database System in South Korea." Information 11 (1): 8. https://doi.org/10.3390/info11010008.

Department of Disaster Management. 2012. Anguilla Draft Comprehensive Disaster Management Policy, The Valley, Anguilla: Government of Anguilla. http://www.gov.ai/documents/Anguilla Draft National CDM Policy.pdf.

GFDRR (Global Facility for Disaster Reduction and Recovery). 2013. Post-disaster Needs Assessments: Volume A Guidelines. https://www.gfdrr.org/en/publication/post-disaster-needs-assessments-guidelines-volume-2013.

Government of Anguilla. 2000. Statistics Act, R.S.A c. S60, The Valley, Anguilla: Government of Anguilla. http://www.gov.ai/statistics/images/StatisticsAct151200.pdf.

Government of Anguilla. 2013. Comprehensive Disaster Management Policy, The Valley, Anguilla; Government of Anguilla. http://www.gov.ai/documents/Anguilla Draft National CDM Policy.pdf.

Ministry of Finance, Economic Development, Investment and Commerce of Anguilla. 2014. Anguilla National Statistics Advisory Committee - Terms of Reference. Anguilla: Government of Anguilla.

Rios, Diaz F., and M. Marin Ferrer. 2018. Loss Database Architecture for Disaster Risk Management. EUR 29063 EN, Publication Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-77752-3, doi:10.2760/647488, JRC110489. https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/loss-database-architecture-disaster-risk-management.

UNECE. 2019a. Generic Statistical Business Process Model, v5.1. Geneva: United Nations Economic Commission for Europe (UNECE). https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1.

UNECE. 2019b. Recommendations on the Role of Official Statistics in Measuring Hazardous Events and Disasters. ECE/CES/STAT/2019/3, Geneva: United Nations Economic Commission for Europe (UNECE). https://unece.org/statistics/publications/recommendations-role-official-statistics-measuring-hazardous-events-and-disasters.

UNESCAP. 2018. Disaster-related Statistics Framework. Bangkok: UN Economic and Social Commission for Asia and the Pacific (UNESCAP). https://communities.unescap.org/asia-pacific-expert-group-disaster-related-statistics/content/drsf.

United Nations. 2015. Sendai Framework for Disaster Risk Reduction. New York: United Nations. https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030.

United Nations. 2016. Report of the Open-ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction. General Assembly report A/71/644, New York: United Nations. https://www.undrr.org/publication/report-open-ended-intergovernmental-expert-working-group-indicators-and-terminology.

UNSD. 2017. Framework for the Development of Environment Statistics 2013. New York: United Nations Statistics Division (UNSD). https://unstats.un.org/unsd/envstats/fdes.cshtml.

ANNEXES

Annex 1 - Metadata Schema Design Document

To access Annex 1 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 2 - Metadata Catalogue

To access Annex 2 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 3 - Memorandum for the Development of a Metadata Portal

To access Annex 3 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 4- Disaster-related Statistics Self-Assessment Tool

To access Annex 4 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 5- Terms of Reference - Anguilla Disaster-related Statistics Data Assessment Working Group

Introduction

The Anguilla Statistics Department is the lead agency for a World Bank-funded initiative to improve the ability of the Government of Anguilla to produce statistics vital for disaster response, impact assessment and recovery. The objectives of the project are to assess Anguilla's disaster-related information architecture, to assess Anguilla's existing disaster-related statistics and the gaps therein, to develop a metadata catalogue and to design a database for disaster-related statistics.

Developing a comprehensive national system for disaster-related statistics requires engagement of departments that share a mandate for economic, social and environmental data necessary for effective disaster response, impact assessment and recovery.

To ensure engagement across departments is effective, timely and beneficial for all participants, a working group comprising representatives of key Government of Anguilla departments and agencies will be formed to assist the World Bank project team to:

- assess the current state of disaster-related statistics in Anguilla, identify a complete set of disaster-related statistics consistent with international standards and assess gaps in existing statistics vis à vis the complete set
- support training on disaster-related statistics
- provide advice and feedback on the design of a database which reflects the priorities and needs of the Government of Anguilla.

Requirements/expectations

The working group members will:

- be knowledgeable about the disaster-related statistics collected and used by the Government of Anguilla
- manage and participate in an assessment of Anguilla's existing disaster-related statistics and a comparison of those statistics with international standards
- promote and attend (or nominate attendees to) an online training session on disaster-related statistics of about four hours (to be scheduled for mid-February 2021)
- respond to occasional requests from the World Bank project team for written advice on (a) a proposed set of disaster-related statistics for Anguilla; (b) a draft metadata catalogue design (February 2021); and (c) a draft database design (March 2021).
- participate in three meetings (approximately two hours each) to carry out the working group's core mandate:
    - Meeting 1– (January 25, 2021; 90 minutes): Review DSSAT (to be in advance prepared by World Bank team) and agree on approach to its application for information gathering for Anguilla; review international standards for disaster-related statistics and their relevance to Anguilla
        - Deliverable – Finalized self-assessment tool and implementation plan; finalized list of international standards for use in study

- Meeting 2 (February 10, 2021; 2.5 hours):
  - Review and confirm information on disaster-related statistics gathered through self-assessment tool
  - Review and confirm proposed "ideal" list of disaster-related statistics for Anguilla based on international standards (to be prepared in advance by World Bank team)
  - Deliverables – Finalized self-assessment tool results and "ideal" set of disaster-related statistics for Anguilla
- Meeting 3 (February 24, 2021; 2 hours):
  - Review and confirm list of gaps in Anguilla's disaster-related statistics based on comparison of self-assessment tool findings with "ideal" set (to be prepared in advance by World Bank team)
  - Review recommendations for improving Anguilla's disaster-related statistics (to be prepared in advance by World Bank team)
  - Deliverable – Finalized list of gaps in Anguilla's disaster-related statistics and finalized set of recommendations

Timing

The working group will be formed as soon as possible and will remain in place until the end of April 2021. All working group meetings will take place in January and February 2021. Any engagement with the working group members after the end of February will be by phone or e-mail only.

Governance

The ASD will chair the working group and the World Bank project team will provide secretariat services. Direct contact of working group members with the World Bank project team by phone or e-mail will be required from time to time.

Annex 6 - Membership - Anguilla Disaster-statistics Data Assessment Working Group

Name	Title	Department
Lori-Rae Alleyne-Franklin	Chief Statistician	Anguilla Statistics Department
Rhina Meade	Director of Economic Planning	Ministry of Economic Development, Commerce, Information Technology & Natural Resources
Ambrell Richardson	Director	Department of Health Protection
Randall Richardson	Fisheries Officer	Fisheries & Marine Resources
Linette Carty	Health Information Coordinator	Health Information Unit, Health Authority of Anguilla
Vaughn Hazell	Director	Department of Information Technology and E-Government Services
Carencia Rouse	Director of Environment	Department of Natural Resources
Damian Barker	Deputy Director	Disaster Management
Julian Hughes	Senior GIS Officer	Department of Physical Planning
Silvia Erni	Senior Planner (Development)	Department of Physical Planning
Stafford John	Planner	Department of Physical Planning
Isabel Rosario	Plant Protection Officer	Department of Natural Resources
Kafi Gumbs	Director of Fisheries	Fisheries & Marine Resources
Karen Pina	Administrative Assistant	Anguilla Chamber of Commerce & Industry
Desron Bynoe	Country Manager	Cable & Wireless Communications (FLOW Anguilla)

Annex 7 - UN-ESCAP Basic Range of Disaster-related statistics

To access Annex 7 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 8- Full details of the Anguilla Disaster-statistics Data Assessment Working Group's response to DSSAT Part II (Statistical-level Assessment)

To access Annex 8 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 9 –Excel spreadsheet for testing DADS

To access Annex 9 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications

Annex 10 -Anguilla data in HUMDAT

Name and link (sort by Last Modified; Feb. 2, 2021)	Note
Survey on Gender Equality At Home	Global, no data on Anguilla
Anguilla - Prices	FAOSTAT: Several datasets: CPI, deflators, exchange rates
Current IATI aid activities in Anguilla	International Aid Transparency Initiative
Anguilla-healthsites	Global Healthsites Mapping Project; Available in GIS formats (GISJSON, SHP) and CSV
Anguilla - Data on forcibly displaced populations and stateless persons	UNHCR; there are 17 Venezuelans
Anguilla - Education Indicators	UNESCO
Anguilla - Population Density	WorldPop; GEOTIFF format; 1km grid by lat/long, annual
Anguilla - Population Counts	WorldPop; GEOTIFF format; 1km grid by lat/long, annual
Anguilla (United Kingdom) - Requirements and Funding Data	OCHA FTS ($27M required; $585K received)
HOTOSM Anguilla Points of Interest (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
;HOTOSM Anguilla Waterways (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Roads (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Buildings (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Populated Places (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
Anguilla - Age and sex structures	WorldPop (one file per single year of age x sex)
Anguilla - Pregnancies	WorldPop, GEOTIF
Anguilla - Births	WorldPop, GEOTIF
HOTOSM Anguilla Airports (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Sea Ports (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)

HOTOSM Anguilla Education Facilities (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Financial Services (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
HOTOSM Anguilla Health Facilities (OpenStreetMap Export)	Humanitarian OpenStreetMap Team (HOT)
Anguilla - Internally displaced persons - IDPs	Internal Displacement Monitoring Centre (IDMC)
HOTOSM Anguilla Schools	OCHA ROLAC
Excel Mapping tool for Caribbean Islands (admin0)	OCHA FISS. National level. Choose 2 data items from list.
Proportion of births attended by skilled health personnel	Africa only. Why in Anguilla?
Anguilla - Subnational Administrative Boundaries	OCHA ROLAC
Airports in Anguilla	OurAirports
Anguilla: High Resolution Population Density Maps + Demographic Estimates	Facebook (??) total and 5-year and functional age groups.
Anguilla - Subnational Population Statistics	OCHA ROLAC. 2011 only.
InterAction member activities in Anguilla	InterAction. (link not working)
Caribbean Hurricanes: Regional Who is Doing What Where (3W)	OCHA ROLAC. Only 3 entries for Anguilla in Sept. 2017.
Hurricane Maria - gust footprints	British Red Cross Maps Team
Hurricane Irma - Windspeed and Gusts	British Red Cross Maps Team
Damages assessment for Hurricane Irma.	International Federation of Red Cross and Red Crescent Societies (IFRC). Links to online Google Docs worksheet
Damage Assessment in Anguilla Island as of 8 September 2017	UN Operational Satellite Applications Programme (United Nations Satellite Centre [UNOSAT]). Very initial physical damage assessment (automated?)
Tropical Cyclone IRMA-17 : Population Exposure estimates	UN Operational Satellite Applications Programme (UNOSAT). Looks like one number for the country.
Tropical Cyclone IRMA-17 : Path and Wind Speed Zones	Tropical Cyclone IRMA-17 : Path and Wind Speed Zones UN Operational Satellite Applications Programme (UNOSAT)
Global - Epidemiological update on Zika Virus	International Federation of Red Cross and Red Crescent Societies (IFRC). 2016 data. 20 suspected cases in Aguilla
Epidemiological update on Zika Virus	International Federation of Red Cross and Red Crescent Societies (IFRC). Sept. 2016.
GAR15 Global Exposure Dataset for Anguilla	United Nations Office for Disaster Risk Reduction (UNDRR)

Annex 11 - Metadata Standard Operating Procedures Document

To access Annex 11 please go to the Anguilla Statistics Department (ASD) website: http://statistics.gov.ai/StatisticsDept/Disaster_Publications