International Seabed Authority
Data Management Strategy

User Manual

Version 1.1

November 19, 2018

Exa Data & Mapping Services, Inc.

Table of Contents

ACRONYMS

AJAX	Asynchronous JavaScript and XML
API	Application program interface
CCFZ	Clarion-Clipperton Fracture Zone
CRC	Cobalt-rich ferromanganese crusts
CTD	Conductivity, temperature, and depth
DMS	Data Management System
EEZ	Exclusive economic zone
ERD	Entity-relationship diagram
ESRI	Environmental Systems Research Institute
ETL	Extract, transform, load
FME	Feature Manipulation Engine
GB	Gigabyte
GEBCO	General Bathymetric Chart of the Oceans
GIS	Geographic Information System
GPS	Global Positioning System
IIS	Internet Information Server
ISA	International Seabed Authority
ISO	International Standards Organization
IT	Information technology
LTC	Legal and Technical Commission
MS	Microsoft
OS	Operating System
PDF	Portable document format
PK	Primary key
PMN	Polymetallic nodules
PMS	Polymetallic sulphides
MVC	Model View Controller
QA/QC	Quality assurance/quality control
SAN	Storage area network
SQL	Structured query language
SSL	Secure sockets layer
SU	Standard unit
UI	User interface
WGS	World Geodetic System

GLOSSARY OF TERMS

Contractors
The groups granted a contract by the International Seabed Authority for the exploration of polymetallic nodules, polymetallic sulphides and cobalt-rich ferromanganese crusts.

CTD
Pertaining to a system for measuring Conductivity (a measurement of salinity), Temperature and Depth (determined from pressure measurements). Additional parameters, such as pH and dissolved oxygen concentration, can be measured if optional sensors are installed. Often has Niskin bottles attached for taking water samples at specific depths.

FME Runner
A script that automates specific procedures using the FME software.

Geodatabase
A database structure including a spatial component designed to enable mapping.

Legal and Technical Commission
The Legal and Technical Commission comprises 22 members elected by the Council of the International Seabed Authority on the basis of their qualifications in the fields of mineral resources, oceanography, protection of the marine environment, or economic or legal matters relating to ocean mining. Members of the Commission are elected for a period of 5 years.

Metadata
A set of data that describes and gives information about other data.

Metadata Template
Template for recording information associated with all uploaded files (unstructured data and raw template files).

Reporting Template
Templates for submitting structured data.

Structured data
Contractor data submittal of required environmental, biological, geological, and chemical data entered in the Reporting Templates.

Template
A template is a file that serves as a starting point for a new document.

Unstructured data
Contractor data submittal of files and information not accommodated by the Reporting Templates; e.g., photos, videos, maps, calibration or QA/QC data, sidescan/seismic files.

1. INTRODUCTION

The International Seabed Authority (ISA) is an intergovernmental organization established under the Law of the Sea Convention to administer the resources of the seabed outside the boundaries of national jurisdictions. Central to the ISA's mission is the management of data collected from the seabed and surrounding environment. The Authority's Secretariat is responsible for maintaining a data repository containing data from Contractors and other sources relating to seabed mineral resources. There are three types of mineral resources currently under exploration: polymetallic nodules (PMN); polymetallic sulphides (PMS) formed at hydrothermal vents; and cobalt-rich ferromanganese crusts (CRC) formed on top of seamounts.

In order to more effectively manage oceanographic, environmental, and mineralogical data collected in support of seabed mineral exploration, the Secretariat has implemented a Data Management Strategy (the Strategy) that comprehensively addresses information management policies and procedures. As part of the Strategy, a data management model has been implemented through the development of a spatial, internet-based Data Management System (DMS) to warehouse data collected by Contractors and other sources. The DMS streamlines receipt, storage, and access of data and enables key users to easily search for and retrieve data over a web-based user interface (UI). In addition, the DMS includes tools and procedures to standardize, review, and incorporate data collected by the Contractors into the geodatabase.

The purpose of this document is to provide instructions for utilizing the Data Management System to submit data and query the geodatabase using the web-based user interface. Instructions for the data management staff at the Authority to manage the DMS are provided in a separate document.

2. DATA WORKFLOW AND PERSONAS

Data workflow is the conceptual process that delineates each step of data management. Data workflow describes the processes for incorporating new data, as well as sharing the data. In the following sections, we provide the workflow for data ingest, data sharing, and data management. To facilitate the process of describing workflow, explicitly defined user types, or personas[1], have been developed as target audience members for different components of the system.

The ISA Data Management Strategy personas are:

- Contractor: A company under contract with the Authority that submits data and an annual report as part of their contractual requirements.
- LTC Member: A member of the Legal and Technical Commission who has interest in searching for and retrieving data classified as both public and confidential as part of their advisory role in the Authority.
- Public Individual: A member of the public who is interested in viewing or downloading information relating to the deep-sea environment.
- Secretariat Data Manager: Secretariat staff who are tasked with managing incoming data from the Contractors and other sources, and overall maintenance of the DMS.

A generalized workflow showing how the different personas interact with the DMS is depicted in Figure 1. Contractors, members of the LTC, and public users will access the DMS via the website, although the LTC will have user permissions enabling them to download confidential data, and Contractors will have user permissions that allow them to upload data and information. Contractors can only access confidential data that they submitted for specific contracts. Access to information over the website depends on log-in credentials:

- Contractor persona: confidential results accessible only for the logged-in Contract ID.
- LTC persona: confidential results for all Contractors are accessible.
- Public persona: confidential results are not accessible.
- Secretariat Data Manager persona: confidential results for all Contractors are accessible.

---

[1] In a user-centered design, a *persona* is a character created to represent a user type that might use a website in a similar way.

Members of the Secretariat data management team interact with the DMS over the website (to manage data submittals) and also directly with the in Microsoft (MS)™ SQL Server geodatabase and GIS mapping data.

Subsequent sections will provide greater detail on how each persona interacts with the DMS:

- Section 3: Contractor persona – templates and data submittals.
- Section 4: LTC persona – website navigation.
- Section 5: Public persona – website navigation.
- Section 6: Secretariat Data Manager persona – managing data submissions; managing the geodatabase and spatial data; website administration; roles and responsibilities.
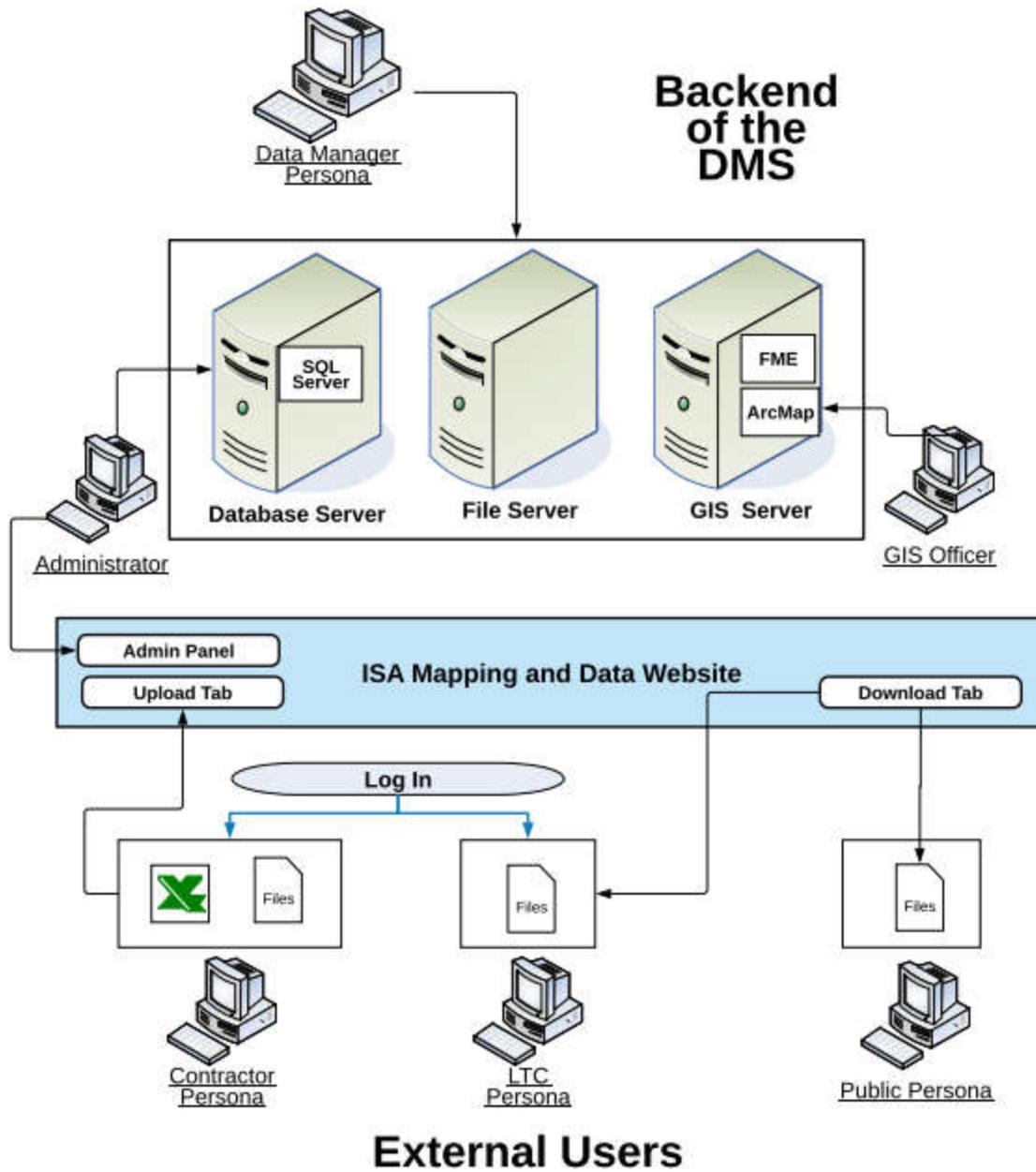
Figure 1. Generalized workflow diagram showing interaction between personas and the DMS.

3. DATA SUBMITTALS – CONTRACTOR PERSONA

This section of the report describes the Reporting Templates to be used by the Contractors to compile data (Section 3.3) and an additional data submission template developed to contain documentation and tracking information about the data submittals, referred to as the Metadata Template (Section 3.2). The procedures the Contractors will use to submit data to the Secretariat over the website interface is described in Section 3.4.

A critical concept is that the DMS is designed to accommodate two data streams:

- Structured data (required environmental, biological, geological, and chemical data [ISA 2015] entered in the Reporting Templates).
- Unstructured data (all files and information not accommodated by the Reporting Templates; e.g., photos, videos, maps, calibration or QA/QC data, sidescan/seismic files).

Each data stream is handled differently by the DMS (Figure 2), including how the data are submitted by the Contractors, warehoused in the DMS, and queried over the website. The remainder of this report will reference either structured or unstructured data and it is important for the reader to understand the fundamental differences between these types of information.
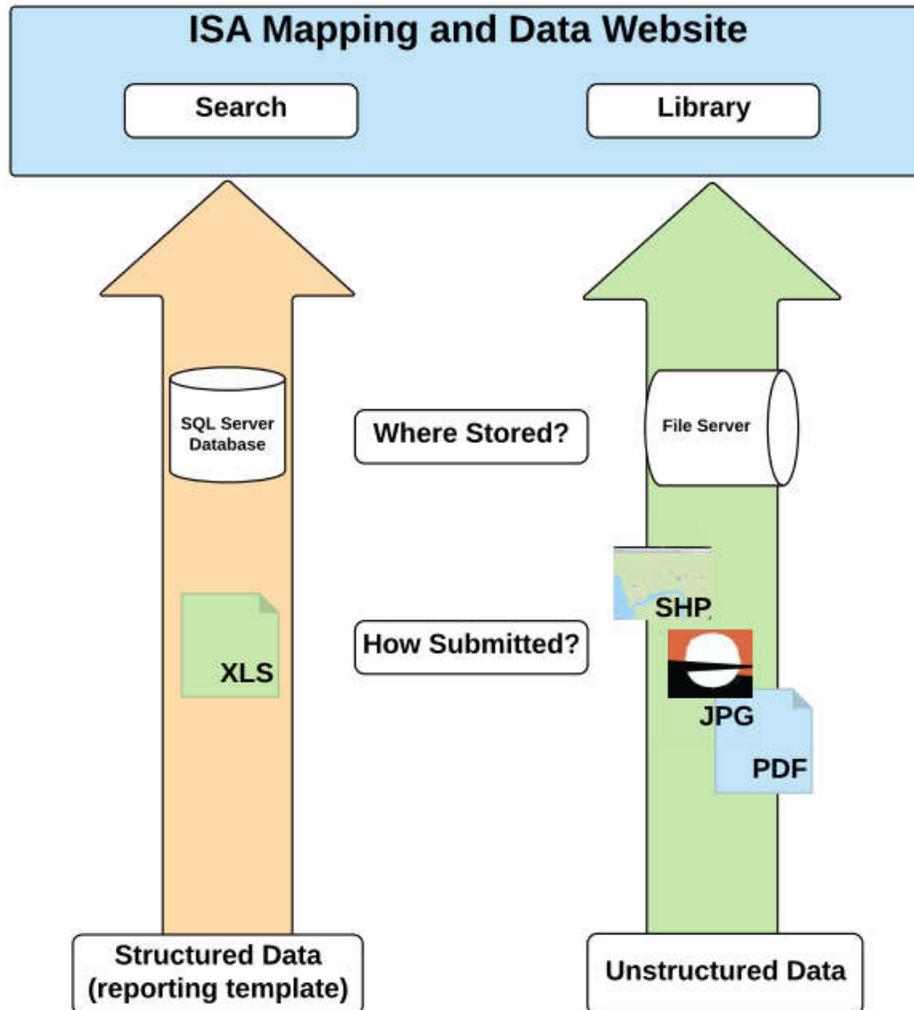
Figure 2. Data streams accommodated by the Data Management System.

3.1 User Authentication

Each Contractor will be issued credentials for each contract, giving them access to all public data and the confidential data associated with the contract. On the website, the Contractor clicks the LogIn button on the right-hand side in the dark blue menu bar and enters the secure credentials.

To change a password:

1. Log in, as usual.

2. Change URL to http://data.isa.org.jm/isa/password/

3. Follow instructions to update the password.

3.2 Metadata Template

3.2.1   Overview

The Metadata Template is in MS Excel format and its purpose is to categorize and document information about the files submitted over the website, and to streamline the receipt, import, and classification (confidential, non-confidential) of both structured and unstructured data provided by the Contractors.

The information provided in the Metadata Template is captured in the geodatabase to allow website users to identify and download specific files, and to download metadata about those files (e.g., contact information and descriptive information).

Information on different file types are entered in different worksheets in the Metadata Template. The following data entry worksheets are included:

Annual Report	Complete this worksheet when submitting Annual Report document(s) in a compressed (zip) file.
ReportingTemplates	Complete this worksheet when submitting Reporting Template files in a compressed (zip) file.
NonTemplateData	Complete this worksheet when submitting all other data files (i.e., data that is not accommodated by the Reporting Templates), including compressed files containing photos and videos.
PhotoVideoFiles	If the NonTemplateData worksheet includes files with Theme = Photo_Video, then complete this worksheet with information on the individual photo and video files (see Section 3.2.4 for more information for filling out this worksheet).

Guidance for Contractors populating the Metadata Template is included in Appendix 1[2] of this report, and the 'Instructions' worksheet of the Metadata Template file.

---

[2] The instructions included in the Template package, available on the website, will be periodically updated and should be considered the primary source of information on how to populate the Templates.

3.2.2 Preparing the Data Submission and Metadata Template

The Metadata Template must be completed following specific instructions, to ensure the submission is accepted and the metadata information is processed accurately, as follows:

1. Compile the files for submission.

2. Group files by Theme, as each data submission must contain files associated with ONLY a single Theme (see Section 3.2.3 for more information on Themes). For example, photos and videos must be included in one submission and documents must be included in a separate submission.

3. Fill out rows in the Metadata Template for the files that will be included in the submission. Each file submitted must have one row filled out in the Metadata Template.
   Importantly, a <u>single</u> Metadata Template is used for <u>all data associated with the Contractual Year</u> (i.e., all data associated with an Annual Report). This means that the same Metadata Template is filled out cumulatively for a Contractual Year, and this Metadata Template is included with multiple submittals.[3]

4. Run the QAQC checks in the ErrorChecker worksheet of the Metadata Template.

5. Name the Metadata Template file according to the following naming convention: master file name + Contractor ID + Contract Type + Contractual Year (e.g., MetadataTemplate_20181005_DORDPMN2015).

6. Compress the files to be submitted in a zip file, including the current version of the Metadata Template. <u>EACH</u> zip file submitted <u>must</u> include a Metadata Template file.

7. Reporting Templates must be submitted with only one Reporting Template per zip file.

8. Name the zip file data submission according to the following naming convention: Contractor ID + Contract Type + Contractual Year + Theme + descriptive information (e.g., DORDPMN2015_AnnualReport_appendices; JOGMECCRFC2016_Bathymetry_CruiseA).

Additional instructions for submitting the files over the website are included in Section 3.4.

3.2.3 Theme

For the NonTemplateData worksheet, Contractors are required to identify the 'Theme' of the submitted files to categorize it based on data type (Table 1). The entries in the Theme field of the Metadata Template are constrained by a specific set of restricted entries, termed valid values. Each

---

[3] Within a single Metadata Template file, there cannot be two 'Zip File Name' entries that are the same.

Theme is associated with a confidentiality status, except for Photos/Videos; each submitted photo or video is classified as confidential or non-confidential in the Metadata Template by the Contractor.

Information provided in the AnnualReport worksheet will be automatically identified as the 'Annual Reports' Theme; information provided in the ReportingTemplates worksheet will be automatically identified as the 'Unprocessed Templates' Theme.

Table 1. Theme metadata tags and associated confidentiality status.

Theme	Description	Confidentiality Status
Navigation	Electronic files with the coordinates of each of the following items: (i) Station locations; (ii) Multibeam, sonar and seismic track lines; (iii) Ship track. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public
Bathymetry/Confidential	Confidential, processed bathymetric data as digital xyz files in the ASCII format or a common GIS format. The processing sequence must be fully described. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public
Sidescan_Seismic	Digital files (SEG-Y or XTF) and/or as high-resolution images (JPG, PDF, TIFF, etc.). Enter in 'NonTemplateData' worksheet of Metadata Template.	Confidential
Electromagnetic	Digital grids in a common GIS format. Enter in 'NonTemplateData' worksheet of Metadata Template.	Confidential
Electric self-potential	Digital grids in a common GIS format. Enter in 'NonTemplateData' worksheet of Metadata Template.	Confidential
CTD_Water Quality	Tables (Excel, txt, etc.) and graphs in digital format. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public
Photo_Video	High-resolution representative images (JPG, PDF, TIFF, etc.). Enter in 'PhotoVideoFiles' worksheet of Metadata Template.	Ind files
Current Meter	Flow measurements by mechanical (rotor current meter), tilt (Tilt Current Meter), acoustical (ADCP) or electrical means. Not required. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public
Unprocessed Templates	Submitted raw Reporting Templates. Enter in 'Templates' worksheet of Metadata Template.	Confidential
Document/Confidential	Confidential reports and documents (PDF, DOC, etc) that are not Annual Reports. Enter in 'NonTemplateData' worksheet of Metadata Template.	Confidential
Annual Reports	Annual reports (PDF, DOC, etc). Enter in 'AnnualReport' worksheet of Metadata Template.	Confidential
Document/Non-Confidential	Non-confidential reports and documents (PDF, DOC, etc) that are not Annual Reports. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public
Bathymetry/Public	Publicly-available, processed bathymetric data as digital xyz files in the ASCII format or a common GIS format. The processing sequence must be fully described. Enter in 'NonTemplateData' worksheet of Metadata Template.	Public

3.2.4   Metadata for Photo/Videos

There are specialized procedures for logging photos and videos in the 'PhotoVideoFiles' worksheet of the Metadata template. The individual[45] photo and video files are logged, and information must be

---

[4] Each photo or video file submitted must have one row filled out in the PhotoVideoFiles worksheet.
[5] There cannot be two 'VP File Name' entries that are the same in the PhotoVideoFiles worksheet.

provided to relate these files to specific samples or results. This will enable users to identify (and export) photos and videos that are associated with results from the geodatabase.

If the NonTemplateData worksheet includes files with Theme = Photo_Video, then complete the 'PhotoVideoFiles' worksheet with information on the individual photo and video files. Enter information in the StationID, SampleID, TrawlID, and/or TaxaID fields to link individual photos and videos to specific samples, trawls, or results. If photos/videos are not associated with specific samples, trawls, or results, then enter 'NA' in the 'Associated Result' field and leave the StationID, SampleID, TrawlID, TaxaID fields null.

Six different scenarios have been identified where photos or videos may be procured and accommodated in the Metadata Template, Reporting Templates, and the geodatabase structure. The six photo/video scenarios identified in Figure 3 provide examples of appropriate data entry in the Metadata Template for photos and videos.

3.2.5   Valid Values

The purpose of valid value lists are to standardize the entries in a particular field so that the data can be more consistently filtered, sorted, and analyzed. Columns constrained by valid values are identified in the Legend worksheet, and in row 3 of the data entry worksheets.

Entries in columns that have constrained content must exactly match the entries in the ValidValues worksheet, which contains all acceptable entries. If values are entered that are not included on this list the Error Checker (see Section 3.2.6) will create an error message so the error can be easily identified and corrected. New items may not be added to the ValidValues worksheet, *and existing entries must not be modified.*

3.2.6   Validation Procedures

The Metadata Template includes an Error Checker that will be executed by the Contractor preparing the template. The Error Checking command executes a series of automated quality assurance/quality control (QA/QC) checks to ensure that there are no inconsistencies or errors in the file set. A Checker Report is generated, listing descriptive error messages so the user can easily identify problems.

Table 2 includes the list of validation checks performed in the Metadata Template.

Table 2. Listing of QAQC checks embedded in the Metadata Template.

Check type	Check #	Check
Structural	A	Template file name is prefixed with 'MetaData_Template_yyyymmdd' and does not exceed 120 characters.
Structural	B	Checks that worksheets are present and correctly named.
Structural	C	All field names in Template are checked to ensure they are correct.
Check required fields	D	Required fields are populated.
Check required fields	E	The number of characters are checked for length restrictions.
Structural	F	Clear out content in columns to the right, and rows to the bottom.
Check Unique records	1	Check for duplicate records based on key fields (Template/Worksheet - key fields):
		Meta/AnnualReport - Annual Report Title + Zip File Name
		Meta/ReportingTemplates - Contractual Year + Template Type + Zip File Name
		Meta/NonTemplateData - Contractual Year + Theme + Zip File Name
		Meta/PhotoVideoFiles - StationID + SampleID + TrawlID + VPFileName
Valid Value checks	2	Values in selected fields are checked against the allowable values in the ValidValues worksheet.

Scenario	Scenario 1	Scenario 2	Scenario 3*	Scenario 4	Scenario 5	Scenario 6
Description of Scenario	Photo of sample taken at a point location	Photo of organism taken from a point sample	Photo of the seafloor to record mineral abundance	Photos of organisms collected from trawl sample	Video track	Photo frames from video track
Field in PhotoVideoFiles worksheet						
VP File Name	Prt001.jpg	ytt004.jpg	DSCN1005.jpg	DSCN2095.jpg	IMG_5854.mp4	VidTrack1_211.mp4
Short Title for Photo or Video File	Stn 1922-1 retrieved sample	Organism from Stn 1922-1	seafloor nodules	Specimen A from 3546t trawl	Video for 3546t trawl	Organism B from 3546t traw
Associated Result	Sample - point	Individual - point	Sample - point	Individual - trawl	Sample - trawl	Individual - trawl
Station ID	1922-1	1922-1	ABC	--	3546t	3546t
Sample ID	1	1	1	57	57	57
Trawl ID	--	--	--	3546t	--	--
Taxa ID	--	3	--	--	--	5
Frame Code	--	--	--	--	--	456
Area of Image	0.2	0.2	0.5	0.2	--	--
Camera Specs	Nikon F mount 1.5 x lens focal width	Nikon F mount 1.5 x lens focal width	Nikon F mount 1.5 x lens focal width	Marshall CV502-WPMB Full HD 3.7 mm lens	Marshall CV502-WPMB Full HD 3.7 mm lens	Marshall CV502-WPMB Full HD 3.7 mm lens
Video/Photo Date/Time Stamp	11/22/2014 14:15	11/22/2014 15:15	2/15/2014 17:05	11/27/2014 15:15	6/11/2014 15:20	6/11/2014 15:20
Confidential Status	Public	Public	Confidential	Public	Public	Public
Remarks					Date/time is the start time of video track	Date/time is the start time of video track

*Photos that are not associated with samples or results may be submitted [e.g., a geological feature (black smoker), the seafloor, an organism swimming in the water column].

Figure 3. Example entries in the PhotoVideoFiles worksheet of the Metadata Template for various scenarios for collecting photos/videos.

3.3 Reporting Templates

The Contractor persona can find and download the most recent version of required template files from the website. The Download tab includes resources relevant to the Contractors, including the most recent versions of the Reporting Template and Metadata Template files. Clicking on a file in the list will export the file. The valid value lists in the Reporting Templates will be updated regularly, so Contractors must always seek the latest version of the Template files prior to preparing a data submission.

3.3.1   Overview

The purpose of the Reporting Templates is to streamline the receipt, validation, import, and classification (confidential, non-confidential) of sampling data provided by the Contractors. There are two Reporting Templates:

- The Env_Template - for biological, geological and chemical results from samples that are not considered to be mineral resources. These results will be flagged as publicly available in the ISA Data Management System.

- The Geo_Template - for geological and geochemical results from samples that are considered to be mineral resources. These results will be flagged as confidential in the ISA Data Management System.

The Templates are provided in MS Excel format, and are designed for Contractors to manipulate and enter original data sets into a geodatabase structure that will allow efficient conversion to the project geodatabase structure, termed the 'ISAGIS' geodatabase. The Reporting Templates can store data from samples collected from multiple cruises.

Both Reporting Templates contain a series of worksheets that are structured to be more flexible than that required by the ISAGIS geodatabase structure. They are intended to:

- Be flexible for data population.
- Structure the data in a format directly compatible with the ISAGIS geodatabase.
- Allow for automated validation prior to inserting into the geodatabase.

Refer to detailed guidance for populating the Reporting Templates in Appendix 2 and Appendix 3[6], which includes instructions on the following topics:

- General information

- Instructions for assigning Station IDs

- Instructions for assigning Sample IDs

- Instructions for entering data from CTD casts[7]

- Qualifier codes

- Valid Value lists

- Instructions for entering results in the GeoChem_Results worksheet

- Instructions for entering results in the Biological_Results worksheet

- Instructions on Linking results to photo/video files

- Error Checker

- Submitting Reporting Template files to the Authority

- Guidance for entering Biological results into Reporting Templates

---

[6] The instructions included in the Template package, available on the website, will be periodically updated and should be considered the primary source of information on how to populate the Templates.
[7] Data from CTD casts may be handled differently in next version of the Reporting Templates.

3.3.2 Reporting Template Worksheets

Each Reporting Template includes several worksheets which are described in Table 3.

Table 3. Listing of the worksheets in the Reporting Templates.

Worksheet	Purpose	Env or Geo?
Legend	Field descriptions for each worksheet in the Template; identifies valid value fields, required fields, default values, maximum field lengths and default values for required fields.	Both
Point Sample	Enter Station- and Sample-level information for samples collected at a point location (e.g., cores, grabs).	Both
Trawl Sample	Enter Trawl- and Sample-level information for samples collected a line location (e.g., trawl samples).	Env only
GeoChem_Results	Enter geological, geochemical, and chemical results.	Both
Biological_Results	Enter biological results.	Env only
Qualifiers	Stores definitions for qualifier codes included in GeoChem_Results.	Both
ValidValues	Listing of allowable entries for fields constrained by valid values.	Both
ValidValues_Analysis	Listing of allowable entries for the Category, Analysis, and Units fields in the GeoChem_Results worksheet.	Both
StudyNotes	Enter any narrative information about the study in this worksheet.	Both
ErrorChecker	The 'Error Checking' command executes a series of QAQC checks to ensure that the data set, as a whole, is reported correctly.	Both

3.3.3 Valid Values

The purpose of valid value lists are to standardize the entries in a particular field so that the data can be more consistently filtered, sorted, and analyzed. Columns constrained by valid values are identified in the Legend worksheet, and in row 2 or 3 of the Point Sample, Trawl Sample, and Results worksheets.

Entries in the 'Category', 'Analysis' and 'Units' columns of the GeoChem_Results worksheet must exactly match the entries in the ValidValues_Analysis worksheet. Entries in other columns that have constrained content must exactly match the entries in the ValidValues worksheet.

The ValidValues and ValidValues_Analysis worksheets contain all acceptable entries. If values are entered that are not included on these lists the Error Checker (see Section 3.3.4) will create an error message so the error can be easily identified and corrected.

Contractors may collect information and measurements on parameters that are not currently included in the ValidValues_Analysis list. These new parameters should be added to the ValidValues_Analysis

worksheet, *but existing entries must not be modified*. New items must include information in the Description field. Before adding an item to ValidValues_Analysis, be sure to do a thorough search to ensure the item is not already on the list (i.e., use the Find command; search for the item using different key phrases). New items in ValidValues_Analysis will be reviewed by the Secretariat Data Manager and incorporated into the master list, if approved. These new items will appear in the ValidValues_Analysis list in a revised version of the Templates.

New items may only be added for the 'Sampling Device' template field in the ValidValues worksheet.

3.3.4   Validation Procedures

The Reporting Template includes an Error Checker that will be executed by the Contractor preparing the template. The Error Checking command executes a series of automated quality assurance/quality control (QA/QC) checks to ensure that there are no inconsistencies or errors in the file set. A Checker Report is generated, listing descriptive error messages so the user can easily identify problems.

Table 4 includes the list of validation checks performed in the Reporting Templates; the entry in the 'Env or Geo?' column indicates to which template the QAQC check is relevant.

Table 4. Listing of QAQC checks embedded in the ENV/GEO Reporting Templates.

Check type	Check #	Check	Env or Geo
Check structure	Prelim Check - A	Template file name includes 'Env_Template_' or 'Geo_Template_' and does not exceed 60 characters.	both
Check structure	Prelim Check - B	Checks that worksheets are present and correctly named.	both
Check structure	Prelim Check - C	All field names in Template are checked to ensure they are correct.	both
Check required fields	Prelim Check - D	Required fields are populated	both
Check required fields	Prelim Check - E	The number of characters are checked for length restrictions.	
Check structure	Prelim Check - F	Clear out content in columns to the right, and rows to the bottom.	both
Check relationships	1A	For each Station ID and Sample ID combination in **Point Sample** worksheet, there must be a match in **GeoChem_Results** _or_ **Biological_Results** worksheet.	both
Check relationships	1B	For each Station ID and Sample ID combination in GeoChem_Results and Biological_Results worksheet, there must be a match in Point Sample worksheet.	both
Check relationships	1C	For each Trawl ID in Trawl Sample worksheet, there must be a match in Biological_Results worksheet; For each Trawl ID in Biological_Results worksheet, there must be a match in Trawl Sample worksheet.	env only
Check relationships	2	Check that qualifier codes entered in the 'Qualifier' field in the GeoChem_Results worksheet have an entry in the Qualifiers worksheet.	both
Check Unique records	4	Check for duplicate records based on key fields (Template/Worksheet - key fields):	both
		Env/Point Sample - ContractID + Cruise Name + Leg Number + Station ID + Sample ID	
		Env/Trawl Sample - ContractID + Cruise Name + Leg Number + Trawl ID	
		Env/GeoChem_Results - Station ID + Sample ID + Category + Analysis	
		Env/Biological_Results - Station ID + Sample ID + Trawl ID + TaxaID	
		Env/Qualifiers - Qualifier	
		Geo/Point Sample - ContractID + Cruise Name + Leg Number + Station ID + Sample ID	
		Geo/GeoChem_Results - Station ID + Sample ID + Category + Analysis	
		Geo/Qualifiers - Qualifier	
Check Unique records	5	Biological_Results worksheet - if Trawl ID is populated, then Station ID and Sample ID may NOT be populated; check that if Station ID is populated, Sample ID is populated; and Trawl ID is NOT populated	env only
Additional Checks	6	Check that Time Datum = UTC in Point Sample and Trawl Sample	both
Valid Value checks	7	Values in selected fields are checked against the allowable values in the ValidValues worksheet.	both
Valid Value checks	8	Values in the Category, Analysis and Units fields of the GeoChem_Results worksheet are checked against the allowable values in the ValidValues_Analysis worksheet.	both
Valid Value checks	9	Check that Point Sample Depth Units = cm and Water Depth Units = m	both
Valid Value checks	10	Check that Trawl Sample Trawl Depth Units = m	both

There are six primary categories of checking routines included in the Error Checker that are applied to both Reporting Templates:

- Check structure – checks that the worksheets and fields within the Template have not been altered
- Check required fields - checks that required fields are populated; required fields are specified in the Legend worksheet of the Templates
- Check relationships - checks that relationships between the worksheets are maintained:
  - For each Station ID and Sample ID combination in Point Sample worksheet, there must be a match in GeoChem_Results _or_ Biological_Results worksheet
  - For each Station ID and Sample ID combination in GeoChem_Results and Biological_Results worksheet, there must be a match in Point Sample worksheet

- For each Trawl ID and Sample ID combination in Trawl Sample worksheet, there must be a match in Biological_Results worksheet[8]
- For each Trawl ID and Sample ID combination in Biological_Results worksheet, there must be a match in Trawl Sample worksheet
- Check that qualifier codes entered in the 'Qualifier' field in the GeoChem_Results worksheet have an entry in the Qualifiers worksheet
- Check Unique records - tests to ensure that each record is unique based on the primary key fields in each worksheet
- Additional Checks - conducts a variety of additional checks to ensure that content is valid and geodatabase conventions are adhered to
- Valid Value checks (checks that fields constrained to valid values have appropriate entries; valid value fields are indicated in the Legend worksheet of the Templates)

3.4 Website Map Options: Uploads (Data Submittal)

The DMS includes a module for the Contractor persona to upload their data, files, and reports[9] as part of their required submissions to the Secretariat. The following are the procedures a Contractor will use to upload a file to the website:

1. Prepare the file for submission (see Section 3.2.2).

2. On the website, click the LogIn button on the right-hand side in the dark blue menu bar, and enter credentials.

3. Click Map Options and then select the Upload tab. This area is where Contractors will submit data and documents and includes a list of files that have been uploaded, along with their status and date of upload.

4. To upload a file, click the Choose File button, navigate to the location of the file to be uploaded, select the file, then click Upload. The file will now appear in the list of uploaded files, with the status of 'Successfully Uploaded'.

In the list of File Uploads, the three icons to the left of the File Name indicate the status of the file (Figure 4). If the icon is green, the data processing step has been completed, if the icon is grey the data processing step has not been completed.

---

[8] Entries in Trawl Sample worksheet that have no match in Biological_Results worksheet are allowed if the 'Video File Name' field is populated (this is to accommodate video tracks that are not associated with any results).

[9] Policies for Contractors to follow when submitting proposed Contract Areas are included in the Data Management Strategy report.

The icons indicate the following data processing steps:

- Upload status

- Review status - based on the Secretariat Data Manager review for data quality and format

- Archived status - indicates the data has been processed into the DMS and is now searchable through the Search tab (for structured data) or the Library tab (for unstructured data)
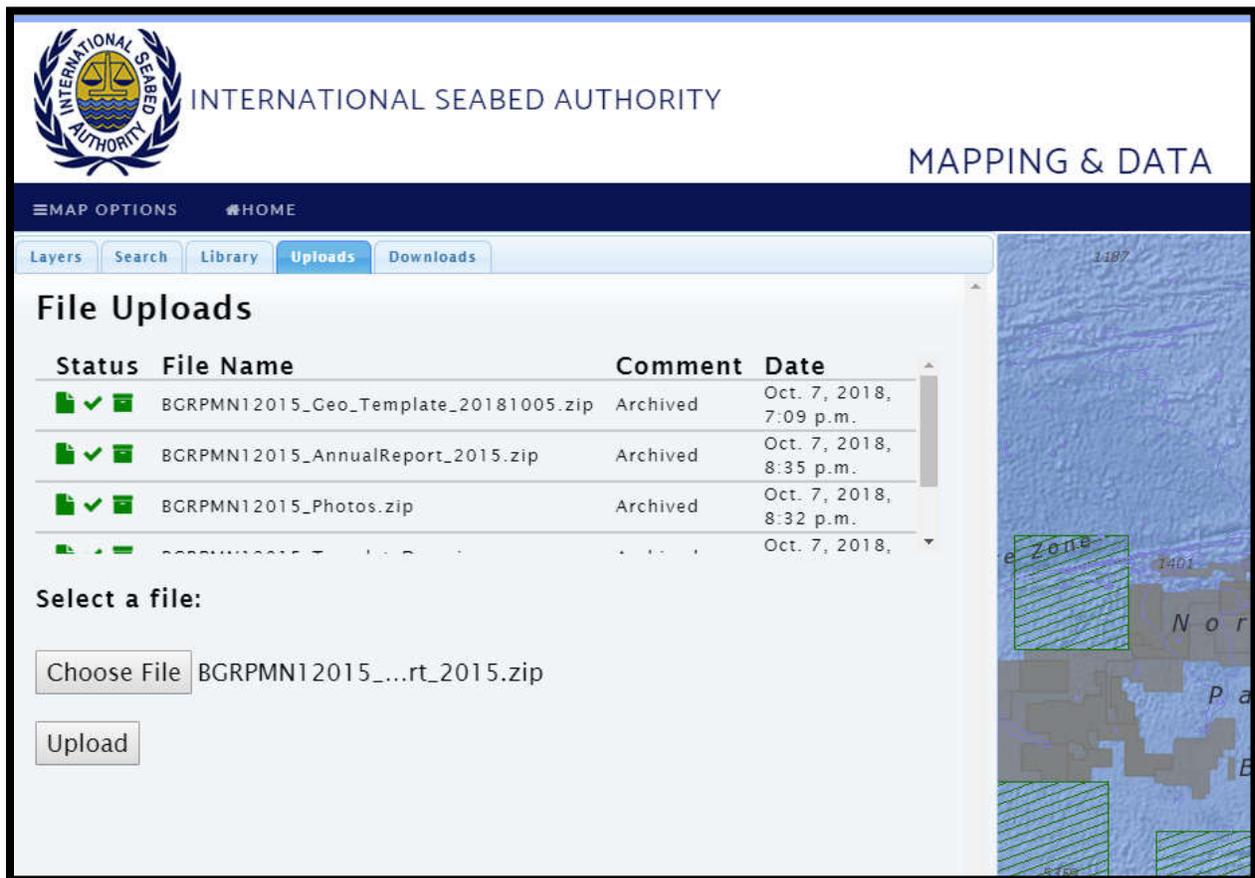

Figure 4. Representation of website portal for Contractor data submittal.

The Contractor can refer to the Uploads tab to a) add more files or update files; b) check on the status of previously submitted files.

Contractors should only upload files that are less that 2 GB in size. For larger files, contact the Secretariat Data Manager to make other arrangements. Files that are not submitted over the website must still be accompanied by a Metadata Template and will undergo the same data processing procedures as files submitted over the website.

Contractors that re-submit data must rename the file (i.e., multiple files with the same name cannot be accommodated by the DMS).

4. WEBSITE USER INTERFACE – LTC PERSONA

This section of the report describes user interaction with the website for the LTC persona.[10]

4.1 User Authentication

Each LTC member will be issued credentials, giving them access to all data (including confidential data) from the geodatabase. On the website, the LTC member clicks the LogIn button on the right-hand side in the dark blue menu bar and enters the secure credentials.

To change a password:

1. Log in, as usual.
2. Change URL to http://data.isa.org.jm/isa/password/
3. Follow instructions to update the password.

4.2 Website Design and Navigation

The website is optimized to be used with the Chrome browser and designed to be map-centric while providing access to both structured and unstructured datasets in various download and visualization formats. The map interface has basic navigation functionality including zoom and pan. Pre-selected map data and imagery forms the background setting that provides the detail necessary to orient the location of the Contract Areas and associated data. Basemap layers available for viewing from the main interface include:

- ISA Surface Sediment Points
- ISA Seamount Biology Points
- ISA Hydrothermal Vent Biology Points
- GEBCO Bathymetry
- APEIs (Areas of Particular Environmental Interest)
- ISA Abyssal Biology Points
- Exclusive Economic Zones (EEZs)

---

[10] In addition to uploading data submittals, the Contractor persona also has the ability to use the website in the same way as the LTC persona. The Layers, Search, and Library tabs behave the same for all users; for the Contractor persona, querying the data through the Search and Library tabs will include confidential results for the contract that the Contractor is logged into.

Bathymetry, EEZs, and APEIs are displayed on the Home page by default, as well as the Contract Areas.

Users can also choose to view the location of collected samples. The DMS is organized to display samples collected from point locations, such as grab samples and CTD casts ('Sample – Points' layer), as well as data collected from trawls or along track lines ('Sample – Lines' layer). Additionally, the user can choose to turn on layers that show only features that are selected through the Layers tab (Section 4.3) and Search tab (Section 4.4); these layers are termed 'Selected Contract Areas', 'Selected Trawl Lines', and 'Selected Sample Points'.

To view the attributes associated with Contract Areas and Samples (points or lines) select a map feature of interest. This produces a pop-up window including the attributes in tabular format. If the map selection resulted in more than one feature, the features are included in different pages of the pop-up window, which can be scrolled through by using the arrows on the top right-hand side of the pop-up window. Alternatively, click on the menu button on the top right-hand side of the pop-up window, which will result in a drop-down list which can be used to identify and select the feature of interest.

The user can also choose from a selection of world map backgrounds and view the legend and symbology for each of the layers that are turned on, using the icons on the top right-hand side of the map display.

The Map Options button, available from the Home page, displays five tabs when the LTC persona is logged in (Figure 5):

- Layers - This tab allows interactive filtering and selection of categories related to contracts and is available to all users.
- Search - This tab allows interactive querying of structured data from the geodatabase as imported from the Reporting Templates and is available to all users.
- Library - This tab allows interactive querying of unstructured data and is available to all users.
- Upload – This tab is only visible to personas with log-in permissions (Contractor, LTC member, Secretariat Data Manager). It is the area where Contractors will submit data and documents. This tab is viewable by LTC members, but uploading documents is not permitted.

- Download – This tab is only visible to personas with log-in permissions. It contains the most recent templates, and any other guidance or reference documentation for the Contractors. The LTC persona can download these reference documents as well.
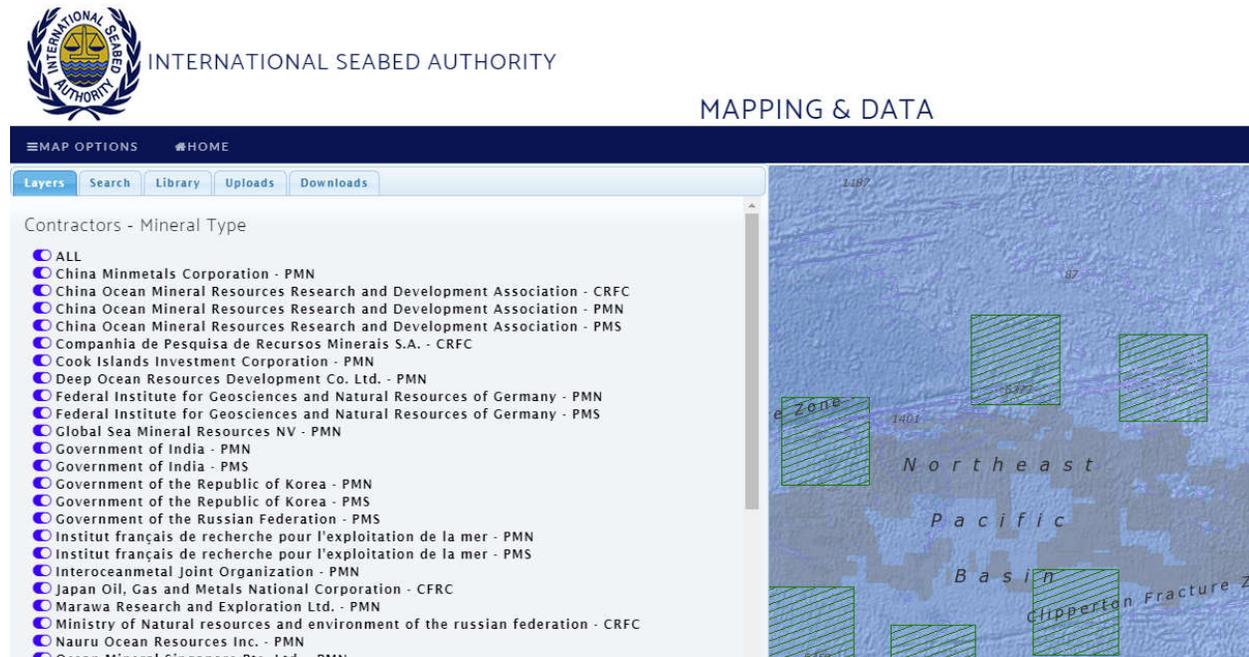

Figure 5. Representation of website tabs available for the LTC persona.

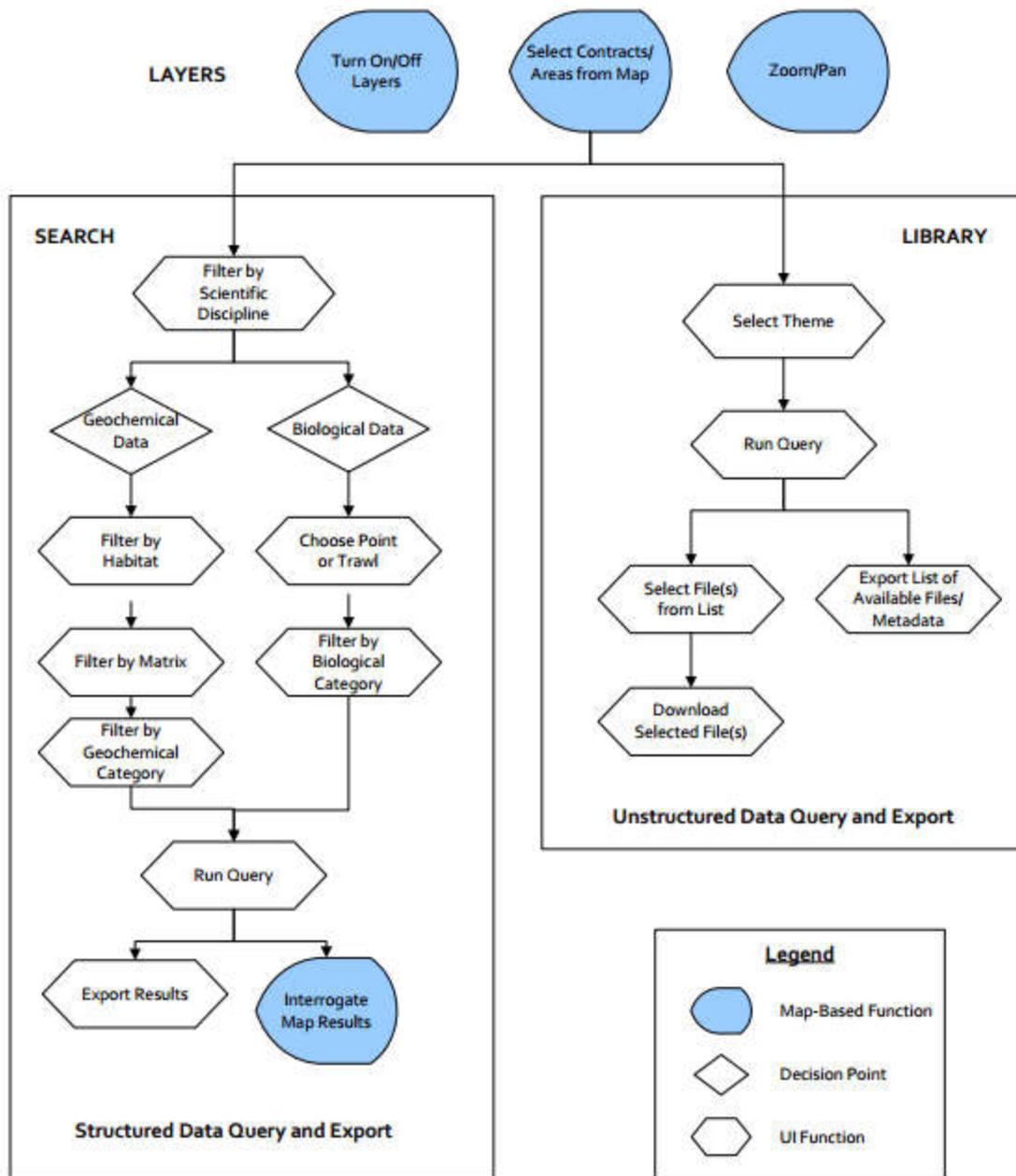An overview of the workflow for navigating the Layers, Search, and Library tabs is shown in Figure 6.

Figure 6. Website navigation workflow.

The following sections provide more detail on how the LTC persona (as well as other users) interact with the functions available in the Layers, Search, and Library tabs.

4.3 Website Map Options: Layers

When Map Options is selected, the Layers tab becomes viewable. The purpose of the Layers tab is to allow the user to make selections about the seabed areas contracted from the Authority. It displays a list of the current contracts, the mineral type associated with those contracts, and other attributes assigned to each contract. The user can select and/or filter the map display based on these attributes:

- Contractors - Mineral Type – Each contract is specific to a Contractor and a mineral type
- Contract Status – Options are active, expired, or extended
- Sponsoring State – The country that sponsors each contract
- Mineral Type – Options are Cobalt-rich Ferromanganese Crust, Polymetallic Nodules, and Polymetallic Sulphides
- Location – General geographic area of the contract

As the above options are selected or de-selected, the following actions occur:

- All other categories in the Layer tab adjust to show the valid options; for example, if a Mineral Type is selected, only those contracts and Sponsoring States with that Mineral Type is selected.
- The map displays the selected Contract Areas and highlights those areas in blue.

The user should use the Layers tab first to find the specific Contract Areas of interest. If they do not make any selections, the default is that all contracts are selected. Selections from the Layers tab are active when the user chooses to query structured data (Search tab) or unstructured data (Library tab). Therefore, the selections made in the Layers tab are the highest-level of filtering for subsequent queries of structured data (Search) or unstructured data (Library).

4.4 Website Map Options: Search

The purpose of the Search tab is to allow all users to query and download **structured data** (Figure 2) from the geodatabase as imported from the Reporting Templates. If there is no active selection from the Layers tab, then all data in the geodatabase are queried. The user first chooses the Scientific Discipline they would like to query: Geochemical Data or Biological Data. Geochemical data include results obtained from an analytical laboratory (e.g., water quality parameter concentrations, seabed metal concentrations) and other parameters that are geological in nature (e.g., mineralogy and

morphological parameters). Biological data include taxonomy and ecological results and information about specific organisms collected from the deep-sea environment. Follow-on filters are determined by this selection.

If the user selects Geochemical data, follow-on filters include Habitat Type, Matrix Type, and Geochemical Category. Habitat Type describes the general environment of the sample; examples include Sediment, Crust, and Water Column. Matrix Type describes the material that was analyzed; examples include Sediment, Water unfiltered, Mineable substrate – rock and Mineable substrate - nodules.

If the user selects Biological data, follow-on filters include Sample Type, and Biological Category. The choices for Sample Type are 'Point' or 'Trawl Line'. One or several options can be selected in the list displayed in the Geochemical Category and Biological Category drop-down lists (Table 5).

Table 5. Categories used in data Search functionality.

Discipline	Category	Category Definition
Biological	Ecology	Community structure metrics for the benthic and pelagic communities.
Biological	Organism Details	Detail about the organism collected (e.g., sex, life stage).
Biological	PhotoVideo Information	Information to associate specific Samples/Taxa to photo and/or video files.
Biological	Taxonomist information	Contact information for the taxonomist responsible for the identification.
Biological	Taxonomy ID	Taxonomic identification information (e.g, order, genus, species) and identification method used.
Biological	Taxonomy information	Information to link the identified organism with existing identifier codes in taxonomic databases.
Geochemical	CTD Sensor Reading	Measurements obtained from CTD sensors (e.g, temperature, salinity).
Geochemical	Elements (Metals)	Concentrations of metals not classified as rare earth metals (e.g., mercury).
Geochemical	Elements (Rare Earth)	Concentrations of rare earth metals (e.g., holmium).
Geochemical	Geochemical Compounds	Concentrations of geochemical compounds (e.g., magnesium oxide).
Geochemical	Geotechnical Parameters	Physical properties of seabed substrates (e.g., shear strength).
Geochemical	Mineralogy	Mineralogical composition; type of mineralization.
Geochemical	Morphological Information	Measurement of the seabed geological formations (e.g., nodule abundance, thickness, texture).
Geochemical	Nutrients	Concentrations of nutrients (e.g., nitrate).
Geochemical	Radionuclide	Pb-210 activity
Geochemical	Sulphide Properties	Concentrations of sulphide compounds (e.g., hydrogen gas).
Geochemical	Water Quality	Measurements of water quality in water column samples (e.g., pH, turbidity).

Choosing to 'Run Geochemical Query' or 'Run Biological Query' results in a summary table of results based on the selections made, allowing the user to review the results and speedily determine if these

are the results of interest. The LTC persona will have access to confidential and public results, while other users will be restricted to view only results that meet their security authentication (Section 2). When the desired results have been queried, the results can be downloaded by choosing to 'Export Geochemical Query' or 'Export Biological Query' which generates a .csv file, viewable in a spreadsheet program, containing all results associated with the search selections. The export file includes many additional columns as compared to the summary table and includes records for all results.

4.5 Website Map Options: Library

The purpose of the Library tab is to allow searching for **unstructured data** (Figure 2) by Theme. If there is no active selection from the Layers tab, then all files in the library are queried. Choosing a Theme from the drop-down list results in a table of metadata for the available files based on the selections in the Layer tab and Theme. The LTC persona will have access to all files, while other users will be restricted to view only files that meet their security authentication (Section 2). The user can then choose to download the metadata results by choosing to 'Export Library Query.' In the display of the available Library files, there is a column called 'Action' that allows the user to download file(s) directly from the website by clicking on the icon.

5. WEBSITE USER INTERFACE – PUBLIC PERSONA

The experience for a member of the public will be identical to that described above for the LTC persona except they have no access to confidential data as they do not have login credentials. In addition, the Public persona does not view either the Uploads or Downloads tabs on the website.

6. SECRETARIAT DATA MANAGER PERSONA

The overall data workflow for receiving new data, incorporating it into the DMS, and sharing this data is complex and involves many data processing steps carried out by either the Contractor or the Secretariat Data Manager, as well as the automated steps carried out by the different software components of the DMS.   Structured and unstructured data are received from the Contractors, extracted from the submitted compressed files, which then go through a series of validation, transformation and loading operations. After the data are loaded into the DMS, there are QA/QC and maintenance procedures to be implemented in the geodatabase and file collections. The Secretariat Data Manager is also responsible for managing spatial data and administering the website.

This section of the report generally describes the interaction the Secretariat Data Manager persona has with the DMS. This persona is tasked with managing incoming data from the Contractors and other sources, and overall maintenance of the DMS. More detailed procedures are included in the User Manual for the Secretariat Data Manager Persona.

6.1 Managing Data Ingest

The Secretariat Data Manager is responsible for receiving data submitted by Contractors (and other sources) over the website, subsequently tracking the data processing steps of each file received and updating the File Status on the website. In addition, the submitted files are reviewed to ensure they meet the minimum standards for the DMS. If errors or inconsistencies are found, the Secretariat Data Manager is responsible for notifying the Contractor to ensure the data are corrected and re-submitted.

Approved data submissions then undergo the extract, transform and load (ETL) procedures which are responsible for the import, QA/QC, and archiving of all incoming data from the website upload tool or manual uploads direct to the server by an administrator. The software toolset procured to accomplish these ETL processes is FME Workbench. In brief, files will be extracted from the submitted compressed files, which then go through a series of validation, transformation and loading operations using the FME software.

The FME software is also used to append new Contract Areas into the geodatabase after the Secretariat Data Managers conduct QA/QC checks and conduct further data processing on proposed new Contract Areas.

6.2 Managing the SQL Server Geodatabase

Data submitted by Contractors and other sources are stored in a SQL Server geodatabase. The geodatabase structure was designed to enable mapping, including a spatial component compatible with the ESRI ArcGIS spatial software. The intent of the geodatabase is to allow efficient storage and retrieval of structured and unstructured data from the Reporting Templates and Metadata Template, respectively, and to support the Secretariat's Data Manager persona by supporting standardized operations.

The objectives for the geodatabase structure include combining field sample collection information and results of the chemical and biological analyses into a relational database designed to minimize data redundancy and effectively organize data in a structure useful for a variety of needs such as data archiving, data analyses, and use with GIS.

Specific data standards have been delineated in the geodatabase structure, as follows:

- Formatting standards (specifying the format for date, time, and coordinates).
- Scientific standards (units).
- Valid values.

These standards are documented in the Reporting Templates and Metadata Template, and the geodatabase structure is designed to be consistent with the templates. Furthermore, these data standards will be enforced by the embedded QA/QC checks in the templates, as well as stored geodatabase procedures that were developed to be executed by the Secretariat Data Manager to identify errors and inconsistencies.

6.3 Managing Spatial Data

The Secretariat Data Manager persona is responsible for managing the spatial data displayed in the website user interface. This includes ensuring that the basemap layers are relevant and up-to-date, handling submissions for new proposed Contract Areas, and managing the ArcGIS services.

6.4 Website Administration

The web application includes system administration tools, accessible using a Django admin account. This toolset facilitates the management of users, permissions, contractor-uploaded files, and

Secretariat-uploaded files. The user and permission interface has been augmented from the base Django interface to include ISA-specific rule sets regarding the visibility and accessibility of contractor datasets, reports, images, and science data.

The Secretariat Data Manager also can view, audit, and manage each compressed file uploaded by the Contractors as part of the overall workflow process. This includes file acceptance, rejection, and setting the status of each file as described in Section 3.4.

The Secretariat Data Manager can upload files to the system using the same upload form used by the Contractors, and then adjusting the settings in the admin interface so that these files are displayed in the 'Downloads' tab of the website. This area will contain the most recent templates and any other guidance or reference documentation for the Contractors.

6.5 Development and Production Environments

The DMS was designed with a 2-tier architecture: development and production. Changes to the system are generated in the development environment and tested, prior to being rolled out to the production environment. The production environment serves the end-user (personas). The DMS was designed with a 2-tier architecture to allow development to continue without disrupting the experience of the end-user.

REFERENCES

International Seabed Authority (ISA).2015. *Recommendations for the guidance of contractors on the content, format and structure of annual reports*, ISBA/21/LTC/15.

International Seabed Authority

Data Management Strategy

User Manual Appendices

Version 1.1

November 19, 2018

Exa Data & Mapping Services, Inc.

Table of Contents

1. APPENDIX 1 – INSTRUCTIONS FOR POPULATING THE METADATA TEMPLATE

The following instructions are included in the Instructions worksheet of the Metadata Template:

International Seabed Authority (ISA)
Metadata Template for Annual Data Submission
Version 1.3
Updated October 5, 2018

1.1 General

This is the spreadsheet template to be used when submitting all data to the International Seabed Authority.

For questions about how to fill out this spreadsheet, contact Sheldon Carter at scarter@isa.org.jm or +876 922-9105 ext 260.

Use the following naming convention for the Metadata Template file: master file name + Contractor ID + Contract Type + Contractual Year (e.g., MetadataTemplate_20181005_DORDPMN2015). The file name may not exceed 120 characters.

Contractor fills out a single Metadata Template for all data associated with the Contractual Year (i.e., all data associated with an Annual Report).

The same Metadata Template file is submitted with multiple submittals (i.e., update a single Metadata Template file for all data associated with an Annual Report).

Include the current version of the Metadata Template with each compressed (zip) file submitted (i.e., EACH zip file must include a Metadata Template file).

Each compressed (zip) file must contain files associated with ONLY a single theme (see ValidValues worksheet for a listing of themes).

Reporting Templates must be submitted with only one Reporting Template per zip file.

Within a single metadata template file, there cannot be two 'Zip File Name' entries OR 'VP File Name' entries that are the same.

All fields in these worksheets are text and should remain as text.

Do not change the content of the three four rows of each worksheet (do not rename or remove the columns).

Do not rename or move the worksheets in the template.

Do not merge cells in the template worksheets.

Do not use non-standard characters in the template; a unicode character set must be used.

Required fields must be populated. Other fields can have null entries.

Do not use non-standard characters in the template.

1.2 Description of Worksheets in this Metadata Template file:

Legend
Field descriptions for each worksheet in the Template
Annual Report
Complete this worksheet when submitting Annual Report document(s) in a compressed (zip) file.
ReportingTemplates
Complete this worksheet when submitting Reporting Template files in a compressed (zip) file.
NonTemplateData
Complete this worksheet when submitting all other data files (i.e., data that is not accommodated by the Reporting Templates), including compressed files containing photos and videos.
PhotoVideoFiles
- If the NonTemplateData worksheet includes files with Theme = Photo_Video, then complete this worksheet with information on the individual photo and video files.
- Information on individual photos and videos is required to allow linkages to specific samples, trawls, or results.
- If photos/videos are not associated with specific samples, trawls, or results, then enter 'NA' in the Associated Results field and leave the StationID, SampleID, TrawlID, TaxaID fields null.
- Records in this worksheet must be unique based on the following fields: StationID + SampleID + TrawlID + VPFileName

1.3 Instructions for naming convention for 'Zip File Name'

Each 'Zip File Name' within a Metadata Template for a Contractual Year must be unique.

Use the following naming convention for the Zip File Name: Contractor ID + Contract Type + Contractual Year + Theme + descriptive information (e.g., DORDPMN2015_AnnualReport_appendices; JOGMECCRFC2016_Bathymetry_CruiseA). The file name may not exceed 120 characters.

Submitted files must not have the same file name as previously submitted files.

1.4 ErrorChecker

The 'Error Checking' command executes a series of overall checks to ensure that the data set, as a whole, is reported correctly.

A Checker Report is generated so the user can easily identify problems. Note that three different types of error messages are generated: Pass, Fail, Alert.

After resolving errors or making any modifications, it is necessary to re-run the 'Error Checking' command in the ErrorChecker worksheet (this is necessary to ensure that the Checker Report is updated to reflect the most recent changes in the template).

The template will not be approved by the ISA if the Checker Report includes Fail error messages.

2. APPENDIX 2 – INSTRUCTIONS FOR POPULATING THE REPORTING TEMPLATES

The following instructions are included in the guidance for populating Reporting Templates ('ReportingTemplates_Guidance_20181005.xlsx'):

International Seabed Authority (ISA)
Guidance for Populating Reporting Templates (Env_Template_20181005.xlsm;
Geo_Template_20181005.xlsm)
Version 1.3
Updated October 5, 2018

2.1 General

These are the templates to be used when Contractors submit sampling results for PMN, CRC, and PMS environments.

Use the Env_Template for biological and geo/chem results from samples that are not considered to be mineral resources. These results will be flagged as publicly available in the ISA Data Management System.

Use the Geo_Template for geo/chem results from samples that are considered to be mineral resources. These results will be flagged as confidential in the ISA Data Management System.

You may not submit multiple template files with the same file name.

For questions about how to fill out this spreadsheet, contact Sheldon Carter at scarter@isa.org.jm or +876 922-9105 ext 260.

Use the following naming convention for the Reporting Template file: ContractID + Contractual Year + master file name (e.g., BGRPMN12015_Geo_Template_20181005.xlsm).

Optionally, other descriptive details can be added as a suffix (e.g., BGRPMN12015_Geo_Template_20180622_Chloropyll_a.xlsm). The file name may not exceed 100 characters.

All fields in these worksheets are text and should remain as text.

Do not change the content of the first four rows of each worksheet (do not rename or remove the columns).

Do not rename or move the worksheets in the template.

Do not merge cells in the template worksheets.

Required fields must be populated - the 'Legend' worksheet indicates default values to use if true results were not obtained.

Fields that are not required should be left null if true results were not obtained - do NOT populate with placeholder values such as -8 or NR.

All date and time entries must be reported in the UTC (Coordinated Universal Time) datum.

Do not use non-standard characters in the template; a Unicode character set must be used.

The 'RequiredMeasurements' worksheet lists the measurements required for each Reporting Template and Matrix Type, over the exploration period.

This is provided for information purposes but is not used in any embedded QAQC checks, as it is not expected that any single Reporting Template file will include all of the required measurements.

For Geo Template:

Station and Sample information is entered in the 'Point Sample' worksheet; results associated with each StationID/SampleID are recorded in the 'GeoChem_Results' worksheet.

For Environmental Template:

Station and Sample information is entered in the 'Point Sample' worksheet for samples collected at a point location (e.g., grab samples, core samples).

Trawl and Sample information is entered in the 'Trawl Sample' worksheet for samples collected from a trawl.

Results associated with each StationID/SampleID for Point Samples are recorded in either the 'GeoChem_Results' or the 'Biological_Results' worksheet.

Results associated with each TrawlID/SampleID for Trawl Samples are recorded in the 'Biological_Results' worksheet.

The combination of StationID/SampleID fields must be coded to identify unique field-collected samples.

The Error Checker will ensure that Station/SampleID/TrawlID in the 'Sample' worksheets are represented in the 'Results' worksheets (an exception to this is trawl samples that only have video(s) associated with them).

2.2 Instructions for assigning Station IDs

Different StationIDs should be assigned for locations with different Actual Coordinates. The exception is CTD casts (see below).

If the same location (same Actual Coordinates) is sampled during one cruise, then the same StationID should be assigned to all samples from that location.

Preferably, the Actual Latitude/Longitude should be measured with the instrument at seafloor (bottom samples), or a the sample collection point (for mid-water sampling).

The location of the instrument recording coordinates must be specified in the GPS Position field.

Coordinates must be provided in Decimal Degrees (DD) format; Degrees/Minutes/Seconds [DMS] coordinate format is not acceptable.

2.3 Instructions for assigning Sample IDs

When samples are collected at different depths, each sample should be assigned a different SampleID. For example:

Station ID	Sample ID	Upper Depth	Lower Depth	Depth Units
1923	01s	0	0.5	cm
1923	02s	10	15	cm
1923	03s	25	30	cm

Different matrices sampled must have unique StationIDs/SampleIDs (e.g., a sediment and water sample collected at the same station should be considered two separate samples, and SampleIDs such as 01S and 01W should be assigned).

Similarily, if total and dissolved fractions are measured in the same field-collected water sample, each fraction must be assigned a different SampleID (e.g., 01T for the total fraction and 01D for the dissolved fraction).

Sample IDs for geochemistry samples must be different from the Sample IDs given to biological samples.

For Env Template: A sample is either a Point sample or a Trawl sample, and sample information must be entered in the appropriate worksheet (Point Sample or Trawl Sample).

If both types of samples are collected during the same cruise, the sample IDs must not overlap (point and trawl Sample IDs must be unique).

2.4 Instructions for entering data from CTD casts

CTD data is publically accessible and therefore must be entered in the Env_Template.

Samples and sensor readings from the same cast should have the same StationID, even if there are unique coordinates for each SampleID (this is an exception to the rule that each set of unique coordinates should have a unique StationID).

Each sensor reading must be treated as a separate sample. Enter the depth of the sensor reading in the SampleID field.

Each sensor reading must have Actual Latitude, Actual Longitude, Sample Date, and Water Depth recorded.

For water samples collected from Niskin bottles, collected at the same depth as a sensor reading, use the same root SampleID, and add a "S" suffix (for sample).

In the GeoChem_Results worksheet, results from CTD sensor readings must have Category = CTD sensor reading.

In the Point Sample worksheet, samples associated with CTD casts must have Sampling Device = CTD sensor and Profile = Cast.

2.5 Qualifier codes

Codes can be added to the 'Qualifier' column in the GeoChem_Results worksheet, to provide additional information about the results.

Qualifier codes in the GeoChem_Results worksheet must be included in the Qualifiers worksheet, and defined.

Qualifier codes used in any particular cruise must be unique (i.e., do not use one code with multiple definitions).

2.6 Valid Value lists

The purpose of ValidValues lists is to standardize the entries in a particular field so that the data can be more consistently filtered, sorted, and analyzed.

Columns constrained by valid values are identified in the Legend worksheet, and in row 2 or 3 of the Point Sample, Trawl Sample, and Results worksheets.

Entries in the 'Category', 'Analysis' and 'Units' columns of the Results worksheet must exactly match the entries in the ValidValues_Analysis worksheet. Entries in other columns that have constrained content must exactly match the entries in the ValidValues worksheet.

The ValidValues_Analysis worksheet includes the complete list of analytes for the ISA Data Management System, however, not all analytes must be reported for each template and matrix type.

Refer to the 'RequiredMeasurements' worksheet to identify the analytes and measurements required for each template and Matrix Type, over an exploration period.

The ValidValues worksheets contain all acceptable entries. Should you enter values that are not included on these lists, the Error Checker will create an error message so that you can easily identify and correct the invalid entry.

The ValidValue lists include special codes if information for required fields is not available (e.g., 'none' or 'not reported'). The Error Checker will only accept these values as correct entries.

New items may be added to the ValidValues_Analysis worksheet, but existing entries must not be modified. New items must include information in the Description field.

Before adding an item to ValidValues_Analysis, be sure to do a thorough search to see if the item is already on the list (i.e., use the Find command; search for the item using different key phrases).

New items in ValidValues_Analysis will be reviewed by the ISA Data Manager and incorporated into the master list, if approved.

New items may only be added for the 'Sampling Device' Template Field in the ValidValues worksheet.

2.7 Instructions for entering results in the GeoChem_Results worksheet

For CRC mineral type entered in the Env template, some dissolved metals have liquid units (mg/L) indicated for sediment matrix. These measurements are to be measured in a water sample collected as close as possible to the the rock/water interface (epibenthic boundary layer).

For PMN mineral type entered in the Env template: Some dissolved metals have liquid units (mg/L) indicated for sediment matrix. These measurements are to be measured in sediment porewater samples.

For sediment porewater samples, Upper Depth and Lower Depth entries will reflect the sediment sampling depth (although Matrix Type = porewater).

Sediment porewater is interstitial water, which is water extracted from the sediment column, NOT water collected at the sediment/water interface.

All analytical results for solid matrices (i.e., rock, sediment) must be reported on a dry weight measurement basis.

2.8 Instructions for entering results in the Biological_Results worksheet

Refer to Appendix 3, or 'Guidance_Biological_Results_Draft_20180418.doc'

2.9 Instructions on Linking results to photo/video files

Information on photo/video files can be entered into the following worksheet fields:

> Point Sample - Photo File Name

> Trawl Sample - Video File Name

> Biological_Results - Photo File Name, Video Frame Code, Video photo frame file name

When information on photos/videos is entered into the Reporting Template, they must also be logged into the 'PhotoVideoFiles' worksheet of the associated Metadata Template.

The Metadata Template:

1. allows storage of additional information on the photo/video files (e.g., Camera Specs, DateTime Stamp, description) and,
2. provides the mechanism for linking and accessing photo/video files through the website Search interface.

2.10   ErrorChecker

The 'Error Checking' command executes a series of QAQC checks to ensure that the data set, as a whole, is reported correctly.

A Checker Report is generated so the user can easily identify problems. Note that three different types of error messages are generated: Pass, Fail, Alert.

After resolving errors or making any modifications, it is necessary to re-run the 'Error Checking' command in the ErrorChecker worksheet (this is necessary to ensure that the Checker Report is updated to reflect the most recent changes in the Templates).

The Template may not be approved by the ISA if the Checker Report includes any Fail error messages.

2.11   Study Notes

Narrative information about the study can be entered in the StudyNotes worksheet. This can be any information that would assist the Secretariat in interpreting the results.

2.12   Submitting Reporting Template files to the ISA

Each Reporting Template file must be logged into a Metadata Template. The Reporting and associated Metadata templates are then compressed (zipped) and submitted as one zip file.

(One Reporting Template file and one Metadata Template file per zip file).

The Reporting Template, along with the associated Metadata Template, is uploaded through the 'Upload' user interface on the ISA website. Note that the Metadata Template can log additional files and the upload can consist of multiple files.

Subsequent to upload, the file status can be viewed in the Upload user interface on the website (e.g., 'DM Approved and Final' indicates Template has been reviewed and approved; 'Appended to DB' indicates the

Template data has been inserted into the database and can be queried over the website).

Submitted files must not have the same file name as previously submitted files.

3. APPENDIX 3 - GUIDANCE FOR ENTERING BIOLOGICAL RESULTS INTO REPORTING TEMPLATES

The following instructions are included in the guidance for populating Reporting Templates ('Guidance_Biological_Results_Draft_20180418.docx'):

3.1 General instructions

- Taxonomic identifications must be made to the species level, whenever possible, and as per regulations.
- A representative specimen of each species or morphotype collected must be submitted to a museum/institution for storage; all samples collected must be saved unless destroyed (i.e., through the DNA barcoding process).
- These instructions assume that only one of the identification methods are used for a given organism. If multiple identification methods were used, only the final and/or most reliable method should be entered in the Reporting Template.
- For information about DNA accessions, the DDBJ (DNA Data Bank of Japan) is informative http://www.ddbj.nig.ac.jp/sub/acc_def-e.html. There are three members of the INSD (the International Nucleotide Sequence Databases): DDBJ, ENA, NCBI (=Genbank). Only one accession needs to be made at any of the three and then a single accession number is propagated to all three sources.
- A WoRMS account is not necessary to access taxonomic information in this database, however, if a Contractor wishes to add or edit information in the WoRMS website an account must be created.
- The Result entry for Analysis = Species should be the species name without the author citation (e.g. "hoffmeisteri" not "hoffmeisteri Claparède, 1862"). The author citation should be entered as Analysis = Taxonomic author citation.
- To confirm if voucher specimens currently exist for a given taxon (and are logged into the WoRMS database), follow these instructions:
  a. Go to the WORMs website (http://www.marinespecies.org/index.php); you do not need to be logged in to access taxonomic information.
  b. On the top left of the Home page, enter Scientific or Common name into the Quick Search box, which will direct to 'WoRMS taxon details' (note that the AphiaID code is the first item in the list of taxon details).
  c. At the bottom of the 'WoRMS taxon details' page, there is a menu bar with tabs for related information. Select the 'Specimens' tab. NOTE that some species may not have associated voucher specimens.

3.2 Guidance for entering information on Taxonomic identifications

When entering results from biological sampling into the Reporting Templates, Contractors provide information on six different Categories of information. The Categories, and the Values to be entered into the Analysis field of the Results worksheet, are listed in the ValidValues_Analysis worksheet of the Template file. The Categories of information are:

1. Taxonomy ID (i.e., the taxonomic identification of individual organisms; Class, Family, Species, etc.)
2. Taxonomy information (i.e., information regarding how the taxonomic identification was made)
3. Taxonomist information (name, email, institution)
4. Ecology (e.g., number of individuals collected, species density)
5. Organism Details (e.g., sex, weight, life stage of individual organisms)
6. Photo/Video information (if applicable). This information is captured to allow a linkage to the file collection Library, to allow database users to quickly identify and access photos/videos of a specific organism.

For the Taxonomy Information category, the Contractor is required to access a taxonomic database [e.g., World Register of Marine Species (WoRMS) website, National Center for Biotechnology Information (NCBI)] to record information that would allow an end user to cross-reference the taxonomic identification with a reliable taxonomic source; this information is entered as Analysis = Identification Method.

The purpose of tracking the ancillary data associated with a taxonomic identification is to allow regulators (and others) to: (1) locate specimens; (2) track information on how identifications were made; (3) understand the reliability of the identification; and, (4) validate the identification.

In the Clipper-Clarion Fracture Zone (CCFZ), it is expected that species will be observed that will not be included in taxonomic databases, because much of the biodiversity in this large region remains undescribed. The ISA has created an Atlas of Abyssal Megafauna Morphotypes of the CCFZ (http://ccfzatlas.com/wiki/index.php?title=Main_Page) that is to be used to assist taxonomic identification of megafaunal species.

Four different types of Identification Methods can be entered in the Reporting Templates, and slightly different information is required for each Identification Method. Note that the Reporting Template QAQC checks will review that the appropriate Analysis entries have been entered, depending on the Identification Method. If Identification Method = Unknown, then treat the entry of taxonomic information as a Morphological type identification.

3.3 Details on Identification Methods

1. If identification was made by an expert using morphology and phenotypic characteristics (Identification Method = Morphological), the following Analysis entries must be recorded:
    a. **Taxonomic Database** - Database used to confirm taxonomic identification (e.g., WoRMS)
    b. **Database Taxa ID** - the associated taxon identifier from the confirmatory database (e.g., if using the WoRMS website, the AphiaID code would be entered; note that a Database Taxa ID is required for higher order taxa identification),
    c. **Taxonomic Status** – Status of taxon details, as indicated in the taxonomic database.

2. If identification was made using DNA barcoding (Identification Method = DNA), the following Analysis entries must be recorded:
   a. **Taxonomic Database** - Database used for data submission (e.g., DDBJ, ENA, NCBI)
   b. **INSD accession number** - INSD accession number assigned by the data submission process of the DDBJ, ENA or NCBI databases.
   c. **Description of DNA sequence** - describe the gene sequence (e.g., COI, 16S genes, 18S genes, etc)
   d. **Putative species name or number -** an informal species name that represents species that have not yet been described in the literature, or specimens that have not been identified to a particular species

3. If identification is made from a photo or video (Identification Method = Remote ID), the following Analysis entries must be recorded:
   a. **Taxonomic Database** - Database used for data submission (e.g., CCFZ Atlas)
   b. **Morphotype** - An informal group of taxa with similar or identical morphology; often preferred to describe morphological variation within a species, especially if a formal description as a variety appears not strongly supported or premature. _For cross-referencing purposes, the identifier entered in the Template must match exactly the morphotype title/identifier in the CCFZ Atlas._
   c. **The higher order taxon groups (Kingdom, Order, Family) – these taxon groups,** as shown in the CCFZ Atlas, must also be recorded for each Morphotype record in the Template

   NOTE: Contractors are strongly encouraged to enter information to the Atlas on the species observed in the Clipper-Clarion Fracture Zone (CCFZ) region.

4. If the organism identified has been submitted as a voucher specimen (Identification Method = Voucher), the following Analysis entries must be recorded to facilitate tracking the progression of the submitted organism:
   a. **Taxonomic Database** - Database used for data submission (e.g., WoRMS),
   b. **Putative species name or number –** an informal species name that represents species that have not yet been described in the literature, or specimens that have not been identified to a particular species,
   c. **Voucher status –** a description of the stage of submission and approval/acceptance of the voucher specimen (e.g., submitted, frozen, rejected, destroyed, accepted, pending approval),
   d. **Voucher code -** Code assigned by the submission process of the database the voucher is submitted to; identifies the stored voucher specimen,
   e. **Voucher institution code -** Museum/institution where voucher specimen is stored.