

Agenda item 15

**Review of the data management strategy of
the Authority for the period 2023-2028**

**Draft data management strategy of the Authority for the period
2023-2028, draft DeepData user manual for the secretariat data
manager persona, and DeepData reporting template guidance for
the submission of digital data.**

1. In 2015, the Legal and Technical Commission (the Commission) requested that the Secretariat evaluate the methods of managing seabed exploration data and propose actions for a data management strategy to address identified gaps and inadequacies. The Commission also provided recommendations, which included data management: policies and procedures, streamline workflow, generate a database model; select and procure database and mapping software, and finally, to update the hardware and infrastructure upon which the information management system would be hosted¹.
2. The Secretariat responded with a report summarising the results of the review of data management practices and proposed actions to update and improve the system. Based on this report, the Commission in 2016 recommended that the Secretariat implement the elements identified (ISBA/22/LTC/15; ISA 2015).
3. During the 25th session of the Authority in 2019, the Secretariat launched the Deep Seabed and Ocean Database (DeepData) to serve as a spatial, internet-based data management system. Its main function is to host all data related to deep-seabed activities, particularly the data collected by contractors of the Authority during their exploration activities in the Area.
4. Improving the data management of the Authority required the revision of the reporting templates contained in Annex IV of ISBA/21/LTC/15. This was initiated in 2019 and completed in 2021. The revised reporting templates have been made available in ISBA/21/LTC/15/Corr.1. The Secretariat has developed its “DeepData reporting template guidance for the submission of digital data” as contained in Annex I below. This document has been created to assist contractors and other external data providers in effectively utilising the revised reporting templates.
5. To continue improving DeepData and implementing the activities identified, the Secretariat has developed its “draft DeepData user manual for the secretariat data manager persona”, as contained in Annex II below.
6. The Commission resumed its discussion on the Data Management Strategy in March 2023, following the election of its new membership for the period 2023-2028, as contained in Annex III. The Commission welcomed the progress made in improving data management by the Secretariat and considered the draft Data Management Strategy for the period 2023-2028². The Commission recommended that it should prioritise not only short-term strategic directives related to data quality, quantity, and accessibility in DeepData but also include strategic medium-term to long-term objectives for data management in support of the future needs of the Authority.

¹ ISBA/22/LTC/15

² ISBA/28/C/5

7. Pursuant to this recommendation, the Secretariat has further developed the draft Data Management Strategy for the period 2023-2028, as contained in Annex I below.

8. The Strategy advances the results of previous efforts of the Authority in improving its data management, including the development and operationalisation of DeepData. It aims at maximising the value of data managed by the Authority and unlocking its full potential for the implementation of the Authority's Strategic Plan and the Action Plan on Marine Scientific Research in support of the UN Decade of Ocean Science for Sustainable Development³. The data-driven services identified should enhance data quality, availability, accessibility, and use for scientific research and to inform decision-making for the protection and sustainable use of areas beyond national jurisdiction. Specific attention will be given to ensure the participation of developing States and deliver results which enhance their capacities for effective data management and its application to advance marine scientific research.

9. The Commission is invited to endorse the direction and priorities of the draft Data Management Strategy for the period 2023-2028 and requests its continued technical support to the implementation and monitoring of the Strategy. The Commission is also invited to take note of the draft DeepData user manual for the Secretariat data manager persona and endorse the DeepData reporting template guidance for the submission of digital data and recommends its use by ISA contractors.

³ https://www.isa.org.jm/wp-content/uploads/2022/07/ISA_MSR_Action_Plan.pdf
IBSA/26/A/4

Annex I

**International Seabed Authority
Data Management Strategy**

**Draft
DeepData
User Manual
For the
Secretariat Data Manager Persona**

February 15, 2023

Exa Data & Mapping Services, Inc.

Table of Contents

LIST OF TABLES	vi
LIST OF FIGURES	vi
1. INTRODUCTION	1
2. DATA SUBMITTALS	1
2.1 Overall Workflow for Data Ingest.....	1
2.2 Updating File Status	2
2.3 Reviewing Submitted Metadata and Reporting Templates	3
2.4 Reviewing Submitted Unstructured Data Files.....	4
2.5 Policy for Handling Data Submittals that Require Revisions.....	5
2.6 Policy for Handling Revisions to Previously Submitted Data.....	5
3. GEODATABASE.....	7
3.1 Overview	7
3.2 Structure	7
3.3 Geodatabase Dictionary	7
3.4 Policy for Making Structural Changes	8
4. GEODATABASE DATA MANAGEMENT PROCEDURES	10
4.1 Stored Procedures.....	10
4.2 Maintaining Valid Value Look-up Tables	10
4.3 Tracking Changes	10
4.4 Updating Coordinates.....	10
4.5 ISA GIS Geodatabase Maintenance and Automation	12
4.5.1 SQL Server maintenance tasks.....	12
4.5.2 Geodatabase maintenance tasks	13
5. TEMPLATE MAINTENANCE	14
5.1 Updating Valid Value Lists	14
5.2 Editing Templates	15
6. USING FME FOR ETL PROCEDURES.....	16
6.1 Overview of FME Usage.....	16
6.2 Overview of FME Workspace Runners	16
6.3 Procedures for Processing Structured Data	17
6.4 Procedures for Processing Unstructured Data.....	17
7. MANAGING SPATIAL DATA	19
7.1 Connecting ArcMap to the Geodatabase.....	19
7.2 How to Access and add Basemap Layers to a Desktop Project	19
7.3 Handling Submissions for New Contract Areas.....	19
7.3.1 Pre-processing – Create Shapefile	19

7.3.2	Instructions for Calculating Area	22
7.3.3	Insert New Contract Area Records into Geodatabase	22
7.4	Managing Basemap Layers	23
7.4.1	Procedures for adding basemap layers:	23
7.4.2	Procedures for removing basemap layers:	24
7.5	ArcGIS Services	24
8.	WEBSITE ADMINISTRATION PROCEDURES	26
8.1	Receiving Files	26
8.2	Web Server Configuration	27
8.3	Internet Information Server Administration	27
8.4	Create a New User	28
8.5	User-Level Security and Profile Modification	28
8.6	File Upload Procedures	29
8.7	Software Maintenance and Backups	29
9.	SYSTEM REBOOT PROCEDURES	31

LIST OF TABLES

Table 1. Database table types and descriptions.....	T-1
Table 2. ISAGIS database dictionary.....	T-2
Table 3. Unique Indices in the ISAGIS database.....	T-9
Table 4. Summary of database management procedures	T-10

LIST OF FIGURES

Figure 1. Workflow for ingest of structured reporting templates and associated metadata.....	F-1
Figure 2. Workflow for ingest of unstructured data and associated metadata	F-2
Figure 3. Entity-relationship diagram (ERD) for the geodatabase.....	F-3
Figure 4. FME Workbench 32bit Properties.....	F-4
Figure 5. FME Geodatabase Connections Menu	F-5
Figure 6. FME Runners for processing the Geological Reporting Template.....	F-6
Figure 7. FME Runners for processing the Environmental Reporting Template	F-7
Figure 8. Server Manager Dashboard	F-8
Figure 9. Internet Information Services Manager	F-9
Figure 10. Enabling the Python Environment.....	F-10

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, termed DeepData, 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 the data management staff at the Secretariat to handle data submittals and manage the geodatabase and website. Instructions for using the DeepData DMS to submit data and query the geodatabase using the web-based user interface are provided in a separate document.

2. DATA SUBMITTALS

2.1 Overall Workflow for Data Ingest

See Figure 1 and Figure 2 for a graphical representation of the workflow for ingesting structured and unstructured data. The general procedures for processing data submittals for structured data are described below; these generalized procedures are described in more detail in subsequent sections.

1. After the Contractor's file submission is uploaded to the website, move the zip file from the upload folder (U drive) to the processing directory (D:/Processing/Current).
2. Execute the FME procedures for processing the Metadata Template:
D:\FMEWorkspaces\Metadata\Metadata_OrganizeZipContents20180914.fmw. This inserts information from the Metadata Template into the catalogue tables of the SQL Server ISAGIS geodatabase (the geodatabase) and moves the files from the processing directory (D drive) to the archive directory (M drive).
3. Update the File Status on the website.
4. Manually review the Reporting Templates and metadata prior to further processing, to ensure that they consist of the expected data.
5. Update the File Status on the website.
6. If the structured data do not comply with requirements, notify the Contractor to re-submit the data.
7. If the structured data do comply with requirements, save the Reporting Template file as .xlsx and move the file to the processing directory (D:/Processing/Current).
8. Execute the FME procedures for processing the Reporting Template to insert data into the geodatabase: D:\FMEWorkspaces\EnvReporting\ EnvReportingTemplate_RUNNER.fmw or D:\FMEWorkspaces\GeoReporting\ GeoReportingTemplate_RUNNER.fmw.
9. Using the FME Inspector, ensure that points/lines fall within the relevant Contract Areas.
10. Execute stored procedures in the geodatabase and conduct eyes-on review to QA/QC the database insert.
11. If the inserted data passes the QA/QC review, update the File Status on the website; otherwise notify the Contractor to re-submit the data and use the stored procedure to remove the data from the geodatabase.

A table is included in the geodatabase to track the File Status of each submission (tblFileLog). The File Status is updated in the Admin Panel of the website by the Data Manager, and is displayed in the Uploads tab, so the Contractors can log-in to the website and track the progress of data submissions.

2.2 Updating File Status

The Secretariat Data Manager updates the status of each file submitted, using the website Admin Panel (Section 8.1) as follows:

1. Click the LogIn button (right hand side in the dark blue menu bar) and log in as ISA Admin.
2. On the website open the Admin Panel (in the dark blue menu bar).

3. Under Documents, select the file of interest, which opens the Change Document window.
4. Use the Status dropdown menu to select the appropriate status.
5. Enter narrative text in StatusComments, if appropriate.
6. Click the Save button at the bottom right-hand side of the Change Document window.

2.3 Reviewing Submitted Metadata and Reporting Templates

Automated quality assurance/quality control (QA/QC) procedures are embedded in the Contractor Reporting Templates and Metadata Template. The QA/QC function will alert the Contractor of errors, with the intention that these will be resolved prior to inserting into the geodatabase. This automated validation functionality primarily checks format, not content, and was developed under the overarching assumption that the primary mode of validation is an ‘eyes-on’ review by Secretariat Data Managers.

Upon receipt of a Metadata and Reporting Templates, the Secretariat Data Manager should re-run the Error Checker to confirm that all checks pass.

Upon receipt of Metadata and Reporting Templates, the Secretariat Data Manager must execute procedures to ensure that the Contractor did not alter the ‘ValidValues’ and ‘ValidValues_Analysis’ worksheets (the Contractor may be tempted to add or alter records in the ValidValues worksheet to get the automated ‘Error Checker’ QA/QC procedures to pass), and assess any new valid values added by the Contractor. The following are the procedures for conducting this review:

1. In SQL Server, click ‘Open File’, navigate to D:/SP_Queries, and execute the appropriate query, depending on Template type being reviewed: Template_VV_Sheet1_Env, Template_VV_Sheet1_Geo, Template_VV_Sheet1_Meta.
2. Copy/paste the query results into the ‘Sheet 1’ worksheet of the Template being processed.
3. Click the ‘DM Valid Value Checks’ button, on the ‘Admin’ worksheet. The command will report descriptive error messages in the ‘Admin’ worksheet, to assist in identifying the source of the errors.
4. Assess the error messages to determine if changes made by the Contractor are acceptable (e.g., new valid values that are deemed to be acceptable) or errors (i.e., edits to existing entries in the ValidValue worksheet; e.g., for the Analysis TemplateField, changed ‘Calcium (Ca)’ to ‘Ca’).
 - a. If new valid values were added that are deemed to be acceptable, add these to the master dicValidValues table in the geodatabase.
 - b. Reject Template if changes to ValidValue tables are unacceptable.

- c. Make edits to the Template entries (if within ISA policy⁴ for altering data submitted by the Contractor; e.g., in the 'GeoChem_Results' worksheet, edit the 'Ca' entries to 'Calcium (Ca)').

Upon receipt of an Env Reporting Template, the Secretariat Data Manager must execute the 'Prep Biological Data for DB upload' command to translate the data in the 'Biological_Results' worksheet to the format compatible to the geodatabase.

Conduct the eyes-on review to identify errors the automated checks would not catch, for example⁵:

- The file naming convention is acceptable (see Template guidance for the suggested file naming conventions).
- The file names of submitted files match what is included in the Zip File Name fields in the Metadata Template.
- 'Theme' is correct, based on actual content of files.
- If Theme = Photo or Video, validate StationID, SampleID, TowID, OrgNum linkage in the Reporting Templates.
- 'ConfStatus' is appropriate in the PhotoVideoFiles worksheet (assigned by Contractors).
- Assess measurement results for outliers.
- Visually review for anomalies.
- Review content in the Studynotes worksheet and Remarks columns.
- Review for typographic errors.
- Review descriptive fields that lack content (e.g., 'Abstract', 'Description of Dataset').
- Review for incomplete data sets.
- In ENV template, for the following measurements in water, check that the units are mg/L for total, ug/L for dissolved: Calcium (Ca), Iron II (Fe II), Manganese (II).

2.4 Reviewing Submitted Unstructured Data Files

Unstructured data files should be manually inspected prior to entering into the DeepData Library. For example⁶:

- Open the file to ensure it is not corrupted and contains relevant information.

⁴ The Secretariat should develop policy for situations where it is acceptable to edit data submitted by Contractors, as well as a procedure for tracking changes made (e.g., inserting a comment in the 'Remarks' field) and communicating the change to the Contractor.

⁵ The Secretariat should develop a standard checklist for this review.

⁶ The Secretariat should develop a standard checklist for this review.

- Ensure the unstructured file is consistent with the metadata ‘Theme’.

2.5 Policy for Handling Data Submittals that Require Revisions

If data submitted in the Metadata/Reporting Templates do not meet the minimum standards for DeepData (i.e., automated QA/QC checks failed, or issues are found during the Secretariat review), the overarching policy is that the Secretariat is to not edit the templates. If errors or issues are found, the Templates should be rejected, and the Contractor is required to re-submit the file(s)⁷.

For practical reasons, the Secretariat may wish to develop a policy to allow for minor modifications to the Templates. Revisions made by the Secretariat Data Managers should be tracked in Comments/Remarks fields in the Template files.

The Secretariat should develop criteria for rejecting data, for example⁸:

- File submissions with the same name.
- Unexplained failure of automated checks.
- Incomplete metadata.

The procedure for rejecting submitted files is as follows⁹:

1. The Secretary General notifies the Contractor via letter.
2. In the website Admin Panel, the Secretariat Data Manager changes the Status to ‘Rejected’ and add the reason for rejecting the template in StatusComment text box.
3. The rejected file is archived to a sub-directory for rejected files (M:\RejectedData).
4. The Contractor re-submits the file(s). File revisions must have a different name than the original; file submissions with the same file name will be rejected.

2.6 Policy for Handling Revisions to Previously Submitted Data

It is possible that external reviewers may notice errors or issues, or Contractors may submit revisions to previously submitted data. In this situation the Secretariat will need to revise data that has already been entered in DeepData.

The procedure for handling revisions to previously submitted data are as follows:

⁷ Re-submissions must not have the same name as previously submitted files.

⁸ The Secretariat should develop standard criteria for rejecting submitted data.

⁹ This procedure will be amended during future phases of work on the DMS.

1. In the website Admin Panel, the Secretariat Data Manager changes the File Status to ‘Submission Replaced with Update’.
2. Archive the originally submitted file to the sub-directory for rejected files (M:\RejectedData).
3. Notify the Contractor, if necessary.
4. For structured data, execute the ISAWeb.ISA_DeleteTemplate_SubmissionFileName stored procedure to delete the data from ISAGIS and insert it into ISAGIS_RejectedData schema. For structured data, also remove the results from the geodatabase (using Stored Procedure; see Section 4.1).
5. Re-process the revised submission.

For practical reasons, the Secretariat may wish to develop a policy to allow for minor modifications to results in the geodatabase. Revisions made by the Secretariat Data Managers should be tracked in Comments/Remarks fields in the geodatabase tables.

3. GEODATABASE

3.1 Overview

The geodatabase has been designed to store:

- Contract-level information (e.g., Contractors, Contract Areas, mineral type)
- Field information (e.g., coordinates, sampler, depth of sampling, sampling device)
- Geochemical information (e.g., morphology, metal concentrations, geotechnical parameters)
- Biological information (e.g., species, abundance, ecology)
- Water quality information (e.g., water quality measurements from CTD casts)

In addition, the geodatabase has been designed to:

- Enable mapping (includes a spatial component)
- Allow for efficient storage and retrieval of structured data
- Allow for efficient compilation of metadata, facilitating the storage and retrieval of unstructured data files
- Support standard operations
- Associate photos and videos with individual samples and/or results

3.2 Structure

The geodatabase structure has a 4-tier hierarchy:

- Contractor tables
- Sample tables
- Results tables
- Metadata tables

Table 1 includes a listing and description of tables.

View the crow's foot diagram showing the entity-relationship diagram (ERD) of the ISAGIS geodatabase in SQL Server.

Figure 3 is another version of an ERD which more clearly shows the fields that relationships are based on.

3.3 Geodatabase Dictionary

The geodatabase dictionary is included in Table 2 and includes:

- Listing of fields
- Field descriptions
- Primary Keys
- Foreign Keys
- Valid Value fields
- Required Fields
- Indices (Table 3 also lists the unique indices)

3.4 Policy for Making Structural Changes

An enterprise geodatabase has a framework installed inside the geodatabase that manages the ESRI geographic data and supports the geodatabase API. If structural changes are made in the native SQL Server environment the interaction between the geodatabase API and the geographic data is interrupted. Therefore, there are strict requirements for making structural changes in the geodatabase.

The following policy and procedures are recommended for making structural changes to the geodatabase:

- Removing fields – Not Recommended
- Changing field names – Not Recommended
- Adding tables:
 - Add table in SQL Server Management Studio (SMSS)
 - In ArcCatalog, register with Geodatabase
 - In ArcCatalog, create new relationship class, if necessary
 - Right click on 'Connection to ISA-SQL01.sde' - New - Relationship Class
 - Delete the ArcGIS Service(s) (see Section 7.5)
 - Add table into ArcMap 'ISAMapService' map document
 - After editing is complete, republish ArcGIS Services (see Section 7.5)
- Adding fields:
 - Add field in ArcMap
 - Populate new field in geodatabase or ArcMap
 - After editing is complete:
 - Delete the ArcGIS Service(s) (see Section 7.5)
 - Republish ArcGIS Services (see Section 7.5)
- Deleting fields:

- Delete the ArcGIS Service(s) (see Section 7.5)
- Delete the field in ArcMap
- Republish ArcGIS Services (see Section 7.5)

Note that deleting a field will likely be blocked by a running ArcGIS service as it normally maintains a schema lock. In this case you will need to stop/delete the service before deleting the field.

- Programming is required to display field/table additions on the website (not addressed in this User Manual)
- Track changes in C:/Documents/ISADB_readme.xlsx on the geodatabase server (ISA-SQL01 host)

4. GEODATABASE DATA MANAGEMENT PROCEDURES

4.1 Stored Procedures

See Table 4 for a summary of the stored procedures and queries designed for managing the ISAGIS geodatabase. To execute a stored procedure:

1. In SMSS Object Explorer, expand the Databases folder
2. Expand the ISAGIS geodatabase folder
3. Expand the Programmability folder
4. Expand the Stored Procedure folder
5. Find the Stored procedure of interest, right click and select Execute Stored Procedure
6. For the 'Validate data entry from templates' and 'Delete Data' stored procedures, a specific SubmissionFileName parameter must be entered in the window that opens to Execute the Procedure. SubmissionFileName entries can be looked up in the tblCatalogue table in the geodatabase.

4.2 Maintaining Valid Value Look-up Tables

The master valid value list is stored in the geodatabase in the dicValidValues table. When reviewing submitted Reporting and Metadata Templates, the Secretariat Data Manager will encounter new entries that should be added to the dicValidValues table. This should be done manually in the geodatabase, by editing the dicValidValues table.

Edits to existing entries in the Value field of the dicValidValues table must be made with caution, as this may result in required edits in the main geodatabase tables. Edits to the DBTable, DBField, TemplateWorksheet and TemplateField fields should never be necessary.

4.3 Tracking Changes

The Track Data Changes feature in the geodatabase will log Data Manipulation Language (DML) changes (e.g., insert, update, and delete operations) that are made to tables in the geodatabase. A procedure for rotating change tracking logs ensures that the logs do not consume hard disk space on the ISA-SQL01 host.

4.4 Updating Coordinates

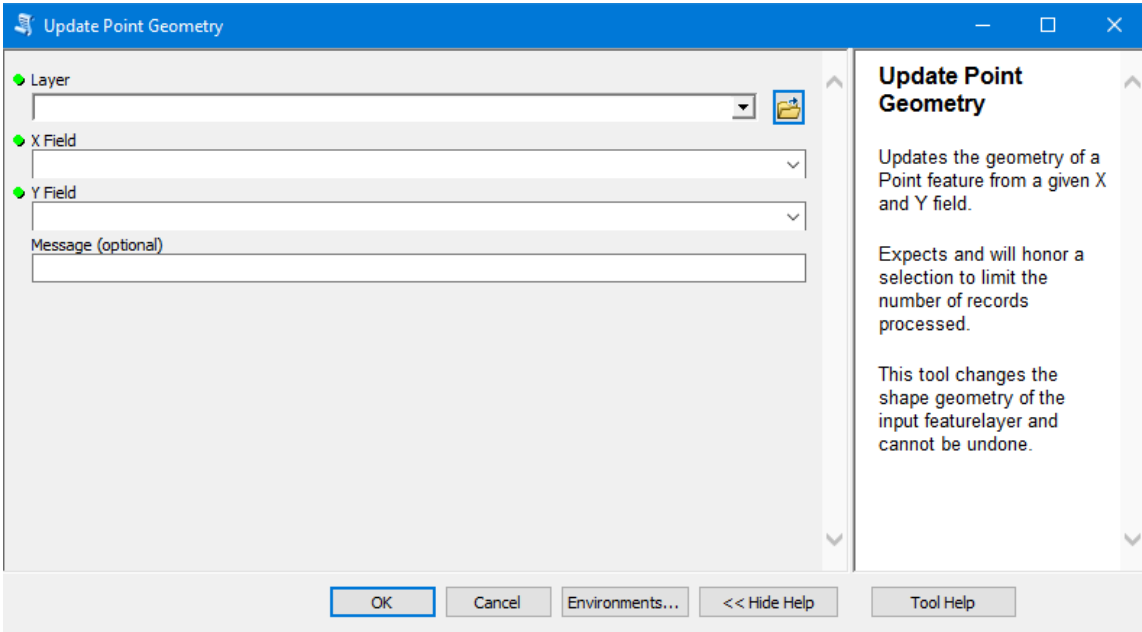
The ArcMap Update Point Geometry tool (in the UpdatePointGeometry toolbox available in the D:/Data/Tools folder on ISA-GIS01 and DEV-GIS01) updates the point geometry of selected features to

values from the Latitude and Longitude fields in the attribute table. Care should be taken to only run the tool against a featurelayer with a selection, as it may take a long time to run against a full dataset.

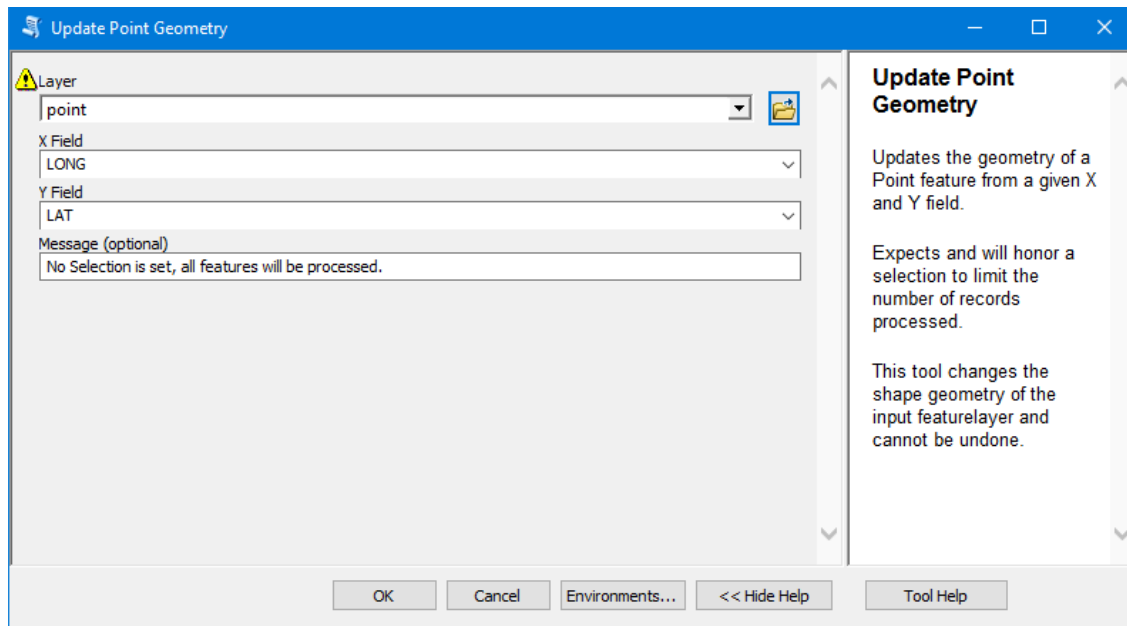
The tool changes the geometry of the input featurelayer.

Using the tool:

1. Open ArcMap and add the point featureclass from the geodatabase.
2. Select the point features that you want to update the geometry for (either in the table or the map).
3. Check that you have valid numeric X and Y field values in the table for the selected features (null values will not update the geometry).
4. Open the tool



5. Select the point layer from the dropdown. The tool will update the message text and set a warning on the layer if there is no selection, but it will still run (see below for example). Close the tool, make a selection and reopen the tool. Choose the X and Y numeric fields from the dropdown options (X is Longitude).



6. After running the tool, you may need to refresh the map view to see the new feature positions. Depending on the changes made you may need to rebuild the spatial index for the featureclass.

4.5 ISA GIS Geodatabase Maintenance and Automation

4.5.1 SQL Server maintenance tasks

General SQL Server maintenance tasks are expected to be performed by the SQL database administrator.

Run periodic SQL Server backups and test backup recovery. Note that if a geodatabase is restored with a different name it is unlikely to function correctly. For testing, the backup should be restored into a separate instance or another SQL Server installation.

Update indexes and statistics not managed by the Esri geodatabase. SQL Server is set to automatically update statistics on the tables and this is generally sufficient for day to day use. If large data loads have occurred, then manually running index updates may be desired.

User management. Additional users can be added by the SQL Server database administrator. The only user that should be used to create new feature classes is the ISAWeb database owner, in order to maintain the existing membership in the ISAWeb schema.

4.5.2 Geodatabase maintenance tasks

Geodatabase maintenance is generally run as either the db_owner (ISAWeb) or a geodatabase admin for system tables.

- Run the usp_Update_V_* stored procedures after data updates to keep the feature class synchronized with the FCLSampleP feature class.
 - The stored procedures are run nightly at XX:XX (TZ) via Task Scheduler on the SQL machine
 - Connect to the SQL Server database, and execute the stored procedure(s) or run the automation script manually

- After major data loads, deletes or edits check the data query speed. If necessary, update GIS statistics on the affected feature classes using the Esri Analyze Datasets tool. An Index refresh/rebuild can also be done via SQL tools and is usually automatically scheduled by the database.
 - Connect to the database and right click to run Analyze
 - OR run the Analyze Datasets tool from ArcToolbox

- Rebuilding spatial indexes may be required when significant numbers of new features are added or deleted. Attribute editing operations do not require updating the spatial index. Note that this can have a significant impact of server resources and should not be done during normal user hours.
 - Open ArcMap and connect to the database
 - Right click and use the Properties dialog of the feature class to access the Index page.
 - Delete and Create the index

- After upgrading ArcGIS client software to a major new release, consider upgrading the geodatabase to take advantage of new functionality and compatibility.
 - Check the prerequisites for the version you are planning to upgrade to
 - Backup the database
 - Remove any customizations that affect the geodatabase system tables. These may cause the upgrade to fail and will need to be reapplied after upgrade
 - Check that the database user has the correct privileges
 - Close all current connections and stop map services
 - Run the Upgrade Geodatabase tool

5. TEMPLATE MAINTENANCE

5.1 Updating Valid Value Lists

Periodically, the new valid values should be incorporated in the Reporting and Metadata Templates and new versions posted. The procedure for this is as follows:

1. Update content in ISAWeb.dicValidValues in geodatabase.
2. Update content in 'ValidValues' worksheet:
 - a. In SMSS, click 'Open File'
 - b. Navigate to D:\SP_Queries\queries (on the ISA-SQL01 host)
 - c. Execute the query relevant to the Template being updated:
 - i. Template_VV_Env_ValidValues
 - ii. Template_VV_Geo_ValidValues
 - iii. Template_VV_Meta_ValidValues
 - d. Copy/paste query results into the relevant 'ValidValues' template worksheet.
 - e. Modify the following TemplateWorksheet entries and then re-sort the table:
 - f. For ENV, edit 'GeoResource_Results/Chem_Results' to 'Chem_Results'
 - g. For GEO, edit 'GeoResource_Results/Chem_Results' to 'GeoResource_Results'
 - h. For ENV, edit 'Towed Gear Sample/Point Sample-E' to 'Towed Gear Sample/Point Sample'
 - i. For GEO, edit 'Towed Gear Sample/Point Sample-G' to 'Towed Gear Sample/Point Sample'
 - j. For META, remove 'MT; '
 - k. For META, edit 'Towed Gear Sample/Point Sample' to 'Multiple'
 - l. For META, edit 'MT; AnnualReport; ReportingTemplates; NonTemplateData' to 'Multiple'
3. Update content in 'ValidValues_Analysis' worksheets:
 - a. In SMSS, click 'Open File'
 - b. Navigate to D:\SP_Queries\queries (on the ISA-SQL01 host)
 - c. Execute 'Template_VV_EnvGeo_ValidValues_Analysis'
 - d. Copy/paste query results into the 'ValidValues_Analysis worksheet' template worksheet in the Env and Geo Reporting Templates.
4. Update version number, date, and file name in 'ReportingTemplates_Guidance_YYYYMMDD.xlsx'.

5. Update date in the file name of each template file and the 'ReportingTemplates_Guidance_YYYYMMDD.xlsx' file.
6. Upload files to the website so they will be available in the 'Downloads' tab (see Section 8.6).
7. Notify Contractors that new versions have been posted.

5.2 Editing Templates

Removing fields or editing field names in the Point Sample, Trawl Sample, GeoChem_Results, Biological_Results or Qualifiers worksheets will necessitate updates to the FME runners and possibly the fields in the geodatabase tables. Adding fields to these worksheets will not affect the FME runners unless the new field must be added to the geodatabase, in which case the FME runners will require updates to facilitate populating the new field.

Procedures for updating the ValidValues and ValidValues_Analysis worksheets are provided in Section 5.1.

It is recommended that only the 'Definition' and 'Default values for Required Fields' columns are edited in the Legend worksheet. Edits in other columns will necessitate related edits in the geodatabase.

The QA/QC checks executed from the ErrorChecker worksheet are programmed in Visual Basic for Applications (VBA). The VBA code can be accessed and edited in each of the Templates by selecting 'Visual Basic' from the Developer menu. The individual checks are documented within the code. When updating the Report Templates, update the version date in the VBA code.

6. USING FME FOR ETL PROCEDURES

6.1 Overview of FME Usage

FME is a specialized extract, transform, and load (ETL) toolset providing anything-to-anything conversion and data manipulation as a software application installed on the ISA-GIS01 host, or the GIS server. FME uses both ESRI and MS SQL Server connections and, due to library constraints, the 32-bit version of FME Workbench is required to be used (Figure 4).

Each user on the server will need to set up their own FME environment, specifically the connections to the geodatabase. In the D:\FMEWorkspaces directory there is an .sde file that can be used to import the connection profile to ensure the naming convention used throughout the FME Workspace runners remains the same. Logical consistency of drive paths, connection names, geodatabase names and so forth is paramount to successful runner maintenance (Figure 5).

Under Tools, click FME Options, right click inside the Geodatabase Connections dialog and click Import, navigate to the .sde file containing the geodatabase connection. Ensure the Connection Name is set to ISA-SQL01.

6.2 Overview of FME Workspace Runners

The three main directories in D:\FMEWorkspaces are Metadata, GeoReporting and EnvReporting, coinciding with the Metadata Template that must accompany any submitted files and the two Reporting Templates.

The other important directory is the D:\Processing\Current directory which is a temporary processing location and must only contain a single file currently queued for processing.

FME Workspace Runners are a way to execute multiple FME Workspaces in FME Workbench with logical processes and execution of quality checks. The user only needs to execute the runner by pressing the 'Play' button, also called 'Run Translation'. Processing output will stream on the bottom of the Workbench for review by the operator.

Although the execution of the Geo and Env runners appears simple, the runners are actually running several ETL operations embedded in the runners. The process by which the Contractor ZIP file submissions are recognized, unpacked, inspected, and processed into normalized data formats and geodatabase structures is depicted in Figure 6 and 7.

6.3 Procedures for Processing Structured Data

The procedures for processing Reporting Templates to insert data into the geodatabase are as follows:

1. If a previous version of the data exists in the geodatabase, delete using the 'ISA_DeleteTemplate_SubmissionFileName' and 'ISA_DeleteMetaData' stored procedures and remove the data files from the M:\Archived\- 2. Delete all files in the D:\Processing\Current directory.
- 3. Copy a zipped Geo or Env Reporting Template file from one of the U:\<ContractorID> directories and paste it into the D:\Processing\Current directory.
- 4. Open FME, open D:\FMEWorkspaces\Metadata_OrganizeZipContents.fmw and click "Play".
- 5. Review warnings in the Translation Log (blue text) and determine if any warnings need action.
- 6. When processing the FIRST file in a batch (group of files associated with a single metadata template), execute the Metadata runner a second time.
- 7. In the ISAGIS geodatabase, confirm that the correct number of records have been inserted into tblCatalogue and tblFileIndex.
- 8. Confirm that the file has been saved into the M:\Archived\- 9. Save the reviewed (see Section 2.3) Template file as .XLSX and copy it into the D:\Processing\Current directory (i.e., not the directory the zip file was extracted to).
- 10. Open D:\FMEWorkspaces\GeoReportingTemplate_RUNNER.fmw or EnvReportingTemplate_RUNNER.fmw and click 'Play'.
- 11. The FME Data Inspector will open – visually inspect the sample points/lines to ensure they fall within the appropriate Contract Area.
- 12. In the ISAGIS geodatabase, confirm that the correct number of records have been inserted into each of the tables by executing the 'ISA_ListTemplateRecordCountsForSubmissionFileName' stored procedure and cross-referencing the record counts in the geodatabase against the record counts in the Reporting Template file.

6.4 Procedures for Processing Unstructured Data

The procedures for processing unstructured data files are as follows:

1. If a previous version of the data exists in the geodatabase, delete using the 'ISA_DeleteTemplate_SubmissionFileName' and 'ISA_DeleteMetaData' stored procedures and remove the data files from the M:\Archived\- 2. Delete all files in the D:\Processing\Current directory.

3. Copy the zipped unstructured data file (e.g., reports, photos/videos, bathymetric data) from the U:\<ContractorID> directories and paste it into the D:\Processing\Current directory.
4. Open FME, open D:\FMEWorkspaces\Metadata_OrganizeZipContents.fmw and click 'Play'.
5. When processing the FIRST file in a batch (group of files associated with a single metadata template), execute the Metadata runner a second time.
6. Review warnings in the Translation Log (blue text) and determine if any warnings need action.
7. In the ISAGIS geodatabase, confirm that the correct number of records have been inserted into tblCatalogue and tblFileIndex.
8. Confirm that the file has been saved into the M:\Archived\<ContractorID><ContractualYear> directory (e.g., IOMPMN12014).

7. MANAGING SPATIAL DATA

7.1 Connecting ArcMap to the Geodatabase

The procedures for establishing a connection between ArcMap and the geodatabase are as follows:

1. Ensure SQL Server Native Client 2012 (32 bit) is installed on the computer.
2. In ArcCatalog click Geodatabase Connection, click Add Geodatabase Connection
 - a. Choose SQL Server
 - b. Instance: xxx.xxx.xx.xx (IP address for the ISA-SQL01 host)
 - c. Authentication: Database Authentication (users are set up with SQL Server logins and users; request credentials from Secretariat Data Manager).
3. Choose the geodatabase to connect to.

7.2 How to Access and add Basemap Layers to a Desktop Project

To access the basemap layers that are used by the website, copy D:\Data\Additional_layers.gdb and share this folder on the network. It is not recommended that the D:\Data\Additional_layers.gdb file is used directly, as any alterations will impact the map display on the website.

7.3 Handling Submissions for New Contract Areas

The following sections describe the procedures for processing and conducting QA/QC checks for proposed Contract Areas, as well as the procedures for inserting new Contract Areas into the geodatabase, and thus displaying them on the website map.

In the contract application, a Contractor submits information delineating the new Contract Areas. The Secretariat is developing a policy for guiding the information Contractors provide regarding the proposed Contract Areas, which is included in the Data Management Strategy Report.

7.3.1 Pre-processing – Create Shapefile

In ArcMap, generate a shapefile delineating the Contract Areas using the coordinates provided by the Contractor:

1. Ensure that the geographic coordinate system of the new shapefile is GCS_WGS_1984.
2. The shapefile¹⁰ attribute table MUST be in a standard format. The following fields must be included in the shapefile:
 - AreaKey - Primary Key for the area, sub-area, cluster, or block (lowest level area)

¹⁰ During the next phase of development it is recommended that management of the fclContractAreas table is conducted in a geodatabase, instead of shapefiles.

- ContrID – ID assigned to the contract
 - SubArea - ID for sub-areas of non-contiguous areas, applicable only to PMN
 - BlkCID - ID for the block, applicable only to PMS and CRC
 - BlkCell - Indicates if the entry in the 'Exploration Block ID/Cell ID' column is a Block ID or Cell ID. Cell ID applies to relinquished PMS areas.
 - ClustrID - ID for the cluster, applicable only to PMS and CRC
 - AreaType - The regulated classification of a seabed area specific to the contract/metal type. Entries constrained to valid values.
 - AreaKM2 - Area of each polygon in square kilometers, as provided by the Contractor
 - AreaKISA - Area of each polygon in square kilometers, as calculated by the ISA (see Section 7.3.3)
 - ActDate - Date of activity (changes in status, additions, etc.) for the geographic area (AreaKey)
 - Status - A description of the current activity taking place at the contracted or reference seabed area. Entries constrained to valid values.
 - Remarks - Any other relevant information
 - Blk_Orig - Block ID original designated by the Contractor
 - Cls_Orig - Cluster ID original designated by the Contractor
3. **QA/QC** - Determine the mineral type of the import source file. This could be obtained from the ContractID, or the ContractType (PMN, PMS, CRC). The type of metal will determine the workflow for import, generation, and QC of the new Contract Areas.
4. **QA/QC** - Use ArcMap tools to calculate the total area in square kilometers for the polygons in the source file using Cylindrical Equal Area World projection, WGS84 datum. See Section 7.3.3 for specific instructions on calculating the areas. If any Contract Area (i.e., sum area of all blocks) is larger than the regulated allowable maximum size for the mineral type, it will be rejected:
- a. PMS - 20,000 km²
 - b. CRC - 6000 km²
 - c. PMN - 150,000 km²

5. **QA/QC** – Compare the new Contract Areas in the shapefile with the existing active Contract Areas to identify overlap. The file will be rejected based on overlap with an existing Contract Area, EEZ, or APEI. Run the ‘Select by Location’ tool (in the Selection menu) using the shapefile with new areas as the ‘target layer’ and fclContractAreas as the ‘source layer’. This will select/highlight any polygons that intersect with pre-existing Contract Areas. There may be other checks to add later (e.g., overlap with areas of critical environmental concern [ACECs] or inside territorial boundaries).
6. **QA/QC** – Review and accept/reject the blocks and clusters (PMS/CRC only) proposed by the Contractor. For PMS contracts, each Block will be 10 km², and each contract will consist of a maximum of 100 Blocks. Clusters will be defined to delineate groups of 20 Blocks. For CRC contracts, each Block will be 20 km² and each contract will consist of a maximum of 150 Blocks. Clusters will be defined to delineate groups of 5 Blocks. Then assign BlockIDs and ClusterIDs for each. Blocks are assigned a value between 1-n, Clusters are assigned a value between A-Z. Then populate the AreaKM2_ISA field as described in Section 7.3.3.
7. **Populate AreaType field** – The Contractor will have identified two halves of the Contract Area, indicating areas of equal value. Arbitrarily, designate half the area as AreaType = Exploration and half the area as AreaType = Reserve.
 - a. PMS - 10,000 km² as Exploration 10,000 km² as Reserve
 - b. CRC – 3000 km² As Exploration, 3000 km² as Reserve
 - c. PMN - 75,000 km² as Exploration 75,000 km² as Reserve
8. **Populate AreaKey field** – Using ArcMap tools, populate the AreaKey field, concatenated from the source information ContractID, BlockID, and Area Type code . This can be done using the “Calculate Field” tool with the following string: [ContractID] & AreaTypeCode & [BlockID or SubArea]¹¹. AreaType codes are as follows:
 - i. X – Exploration
 - ii. XA – Exploration (Allocated)
 - iii. R – Reserved
 - iv. RC – Reserved (Contributed)
 - v. RQ – Relinquished
 - vi. OA – OutsideArea

¹¹ During the next phase of development it is recommended to develop a script for generating AreaKey codes to ensure internal consistency.

9. **Populate Status field** – Populate the Status field with the appropriate entries (Active, Inactive, OutsideArea)).
10. **QA/QC** – Ensure that required fields are included and populated in the shapefile attribute table.

7.3.2 Instructions for Calculating Area

Use the following procedures for populating the AreaKM2_ISA field in the fclContractAreas table in the geodatabase:

1. Create a new map document.
2. Add the fclContractAreas feature.
3. Set the Data Frame properties to Cylindrical Equal Area World projection.
4. Calculate the areas (right-click on the AreaKM2_ISA field and select Calculate Geometry tool).
Be sure to select the result to be displayed in square kilometers, since this will not be the default option.

7.3.3 Insert New Contract Area Records into Geodatabase

Using the automated FME procedure, insert the final set of records into the master Contract Area table (fclContractAreas) in the geodatabase:

1. Ensure there are records in tblContracts and luContractors for the ContractID associated with the Contract Areas that are to be inserted.
2. Delete all files in the D:\Processing\Current directory.
3. Copy all files associated with the shapefile into the D:\Processing\Current directory.
4. Open FME, open D:\FMEWorkspaces\ContractAreas\ContractArea_append.fmw and click 'Play'.
5. Review warnings in the Translation Log (blue text) and determine if any warnings need action.
6. The FME Data Inspector will open - visually inspect the sample points/lines to ensure they fall within the appropriate Contract Area.
7. In the ISAGIS geodatabase, confirm that the correct number of records have been inserted into the fclContractAreas featureclass.

Finally, create a shapefile of the revised Contract Areas and send to the Contractor to notify them of the assigned ClusterID and BlockIDs.

7.4 Managing Basemap Layers

Basemap Map Document name: **Reference.mdx**

7.4.1 Procedures for adding basemap layers:

Apply these procedures in the development environment, test, then implement the same procedure in the production environment.

First you add the new layers to the ArcMap map document:

1. Stop the ArcGIS service (see Section 7.5)
2. Save the shapefiles of interest onto the server under D:/Map Documents/Shp from ISA (or some other location)
3. On the GIS01 host, open ArcMap and then open the basemap map document.
4. Click on the Catalog button, navigate to the shapefile you want to add, right click and choose Export, then 'Geodatabase (single)'
5. In the Feature Class to Feature Class window, the Input Feature field is already completed. In the Output Feature Class' field, navigate to D:\Data\Additional_Layers.gdb. In the 'Output Feature Class' field, enter a name for the layer (this will not be the name that appears on the website but chose a descriptive name).
6. Click OK. This will add the data to the Additional_Layers.gdb geodatabase and the data will automatically appear on the map.
 - a. If the new layer is at the bottom then layer IDs will be maintained. If the new layer is not at the bottom of the table of contents then service ID references will need to be updated in the website code.
7. Check that the layers are referencing the correct data location.
8. Save the basemap map document.
9. Re-publish the ArcGIS service (see Section 7.5)

The new layers need to be added to the website Basemap Layers legend (requires edits to the web UI code):

- Open the REST endpoint and make a note of the layer names and ID values.
- In the file isa/templates/isa/map.html beginning on line 843 there is a "sublayers" list of basemap layers. New IDs need to be added to this list. Any changes to ID values will require other code updates where referenced.

7.4.2 Procedures for removing basemap layers:

Apply these procedures in the development environment, test, then implement the same procedure in the production environment.

1. Open ArcGIS Server Manager and stop the Reference service.
2. Open the basemap ArcMap map document.
3. In the Table of Contents pane, click the checkbox to the left of the layers to be removed (so the box is NOT checked). For example:
 - a. EEZ_land_v2_201410
 - b. ECS_Polygons_Submissions
 - c. ECS_Polygons_Recommendations
4. Save the ArcMap document.
5. Re-publish the ArcGIS service (see Section 7.5)
6. Edit the web UI code to remove these layers from the Basemap Layers legend. In the file isa/templates/isa/map.html beginning on line 843 there is a "sublayers" list of basemap layers. Remove id: 0, id: 1 and id: 2 dictionary objects from the list. The changes will be reflected immediately after a browser page refresh.

7.5 ArcGIS Services

Deleting the ArcGIS services will cause a disconnect from the website (ArcGIS services cannot be reached). To minimize the downtime duration, do not delete or stop a service until you are ready to publish the new one.

To delete ArcGIS Services:

1. Click on ArcGIS Server Manager (icon on the Desktop).
2. Select 'Continue to this website' when the Certificate Error pops up.
3. Enter username and password (request credentials from Secretariat Data Manager).
4. Select the service of interest¹², click the X on the far-right hand side.
5. Sign out.

To republish ArcGIS Services:

1. In ArcMap, click File - Share As – Service.

¹² 'ISAMapService' includes the spatial information from the geodatabase (i.e., Contract Areas, sample points and trawl lines); 'Layers_for_mapservice' includes the spatial layers for the basemap layers (e.g., seamounts, seafloor bathymetry)

2. In the Share as Service window, select 'Publish a Service', click Next.
3. In the Publish a Service window: in 'Choose a connection', select 'arcgis on localhost_6080 (publisher)'; in 'Service name' enter ISAMapService; click Next.
4. Select 'Use existing folder' and select [root] from the dropdown list; click Next.
5. Click Continue.
6. Set max record count to 5000 each time you re-publish:
 - a. Click on ISAMapService
 - b. Click Parameters
 - c. Under Properties, enter 5000 next to 'Maximum Number of Records Returned by Server'.

8. WEBSITE ADMINISTRATION PROCEDURES






8.1 Receiving Files

The following are the procedures the Secretariat Data Manager will use to receive and review files uploaded by Contractors:

1. Click the LogIn button (right hand side in the dark blue menu bar) and log in as the ISA administrator.
2. Click AdminPanel and then select Documents. The Secretariat Data Manager can view the list of files that have been uploaded, along with their File Status, date of upload, and current location on the U drive.
3. After identifying the file(s) of interest, the Data Manager navigates to where this file is saved on the U drive of the ISA-GIS01 host and copies it to their desktop computer to review.

For uploaded Reporting Templates, the first step in the review process is to re-run the ‘Error Checker’ tool in the ‘ErrorChecker’ worksheet to ensure all embedded QA/QC checks have Passed (all error messages in the Checker Report will be green text). The Secretariat Data Manager then conducts additional review steps (see Section 2.3). For uploaded unstructured data files, the Secretariat Data Manager conducts review steps (see Section 2.4). Based on the results of the review, the File Status is updated (Section 2.2).

In the list of File Uploads, the three icons to the left of the File Name indicate the status of the file. 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. The icons indicate the following data processing steps:

-  Upload status
-  Review status - based on review of Secretariat Data Managers for data quality and format
-  Archived status - indicates the data has been processed into DeepData and is now searchable through the Search tab (for structured data) or the Library tab (for unstructured data)
-  Rejected – indicates the data has not passed the review requirements
-  Updated – indicates an earlier submission has been updated

The Secretariat Data Manager has the following choices for File Status:

Uploaded Successfully	File was successfully uploaded; this status is automatically assigned at upload (files that did not upload will not appear in the list of uploads)
Metadata Processed Successfully	Information in the Metadata Template was successfully inserted into the geodatabase and passed ISA Review
Passed ISA Review	Reporting Templates have passed automated and eyes-on review
ISA Review Pending	Indicates that the data submission is pending review
Archived	Data submission has been filed in the Library (unstructured data) or inserted into the geodatabase (structured data)
Rejected	Indicates the data submission was rejected because it did not pass minimum review requirements and will need to be re-submitted by the data provider
Submission Replaced with Update	Indicates a previous data submission has been updated with a revised version.

8.2 Web Server Configuration

The configuration of the web server located at ISA-WEB01 was performed as part of the installation and can be found in the IT Installation Report. No special modifications were performed on the base OS as part of this task.

The Server Manager Dashboard (Figure 8) is a convenient gateway to the server administration tools, including Internet Information Servers (IIS), located under the Tools menu.

8.3 Internet Information Server Administration

The Internet Information Servers (IIS) service running on the web server ISA-WEB01 is a standard installation running a single website bound to the same host as the server. The 'Default Web Site' is always disabled in the current configuration. The only administration tasks directly related to the ISA website are stopping and starting the 'Production' website using the 'Stop' and 'Play' buttons. Select the 'Production' website from the 'Sites' listing under Connections on the left side of the IIS application (Figure 9). On the right side of the page a 'Manage Website' menu will appear with the Stop and Start buttons.

8.4 Create a New User

To add a new local Django user to DeepData , use the following procedures:

1. In the Admin Panel, select Users. This opens the Authentication and Authorization panel.
2. Select the ADD USER button on the right-hand side. This opens the Add User panel.
3. Enter Username and Password, then click the SAVE button on the bottom right-hand side.
4. Open the profile and modify the properties. For Contractors, enter the appropriate code in the ContractorID field.

To add a new LDAP user, use the following procedures:

1. Create a LDAP user and share credentials with the new user
2. New user logs into DeepData
3. ISA DM updates the profile/properties of the new user.

8.5 User-Level Security and Profile Modification

All user profile and security maintenance are performed from the Django administration page after logging in to the website as the Administrator. Click on the Admin Panel menu located in the top right of the navigation bar on the home page to access the administration page.

There are different methods for assigning security to a user account:

- Populate the ContractorID in the user's profile
- Assign roles to the user profile
- Combine both methods.

From the default Django administration page click on the Users under Authentication and Authorization and select the Administrator user (e.g., isaadmin). This page include the typical username, email address, password and profile information that can be modified. The Administrator is a superuser in the system, thus has 'Staff status' and 'Superuser status' checked under the Permissions header. Groups and individual user permissions are not currently used for any user account. The last header, 'PROFILE', has Bio, Location and Contractor ID fields. The Contractor ID field is special to this website where the ID has meaning in the underlying permissions on the main website pages and in the geodatabase.

Current user profiles include:

- ISA – Views all pages, files and query results and CAN upload files. Only to be used for internal ISA users.
- READONLY – Views all pages, files and query results but CANNOT upload files. This profile is typically reserved for LTC members.
- The Contractors are assigned the associated geodatabase ContractorID to ensure each Contractor can only see the proprietary files and query results associated with a specific Contract (each ContractorID is associated with a unique ContractID). This also allows each Contractor to upload files under specific ContractIDs which is critical for accurate ETL operations and subsequent correlation during website query functionality.

8.6 File Upload Procedures

To post files on the Download tab of the website, the Secretariat Data Manager logs in as the Administrator and uses the same upload process as used by the Contractors (the Upload tab on the website). Then, from the Django administration page select 'Documents', select the uploaded file, set the 'DownloadPageVisible' checkbox, and click Save at the bottom right-hand side of the panel.

8.7 Software Maintenance and Backups

The IIS server is backed up at the server level and restoring OS, IIS or other non-Django software-related functionality should be done under consultation with an IT professional. In the case of simply updating the Django code that comprises the ISA website there are a few methods available to the Secretariat. The website code is in the C:\Apps\Production path. Modification of the website code generally requires Django management commands to be issued to "compile" the website and ensure all supporting files are copied properly.

The python virtual environment is in the C:\ProgramData\miniconda3\envs\isaweb39\ path (Figure 10). A combination of conda and pip is used to populate the isaweb39 virtual environment using the environment.yaml and requirements.txt file found in the C:\Apps\Production folder.

```
`conda env create -f environment.yaml`
```

This will create the isaweb39 Python 3.9 conda virtual environment and install all the prerequisites to run the DeepData Django application.

To enable the python environment (Figure 9) in a new console, open an Administrator console and type

```
`conda activate isaweb39`
```

then change directory to C:\Apps\Production and issue the command ``python manage.py``. There are a number of management commands available for Django. It is highly recommended to read the documentation prior to issuing any Django management commands. Some useful commands are listed below:

1. ``python manage.py createsuperuser``
 - a. ONLY to be used if a new superuser account is needed or if the current isaadmin superuser account password is lost and needs to be reset.
2. ``python manage.py makemigrations``
 - a. Step 1 after updating any of the models.py files.
3. ``python manage.py migrate --run-syncdb``
 - a. Step 2 after creating the migrations to apply them to the default database (MSSQL).
4. ``python manage.py collectstatic``
 - a. Copies any static files in the app directories to C:\Apps\Production\static for use by IIS to serve static content.
5. ``python manage.py flush``
 - a. WARNING: ONLY to be used if clearing out the entire Django database is required, which is not typical.

C:\Windows\System32\inetsrv\config\applicationHost.config is an important configuration file for the IIS system and defines the Python/Django app environment.

The Django ``settings.py`` file located in C:\Apps\Production\Production contains common and security settings, such as the database connection credentials, and should be properly secured or redacted shared among developers.

The files located in C:\Apps\Production can be overwritten directly using the Windows copy and paste method between a local development machine and the server over Remote Desktop Protocol (RDP). As good practice, always back up your server, database, and Django files prior to making any changes.

9. SYSTEM REBOOT PROCEDURES

Each server must be rebooted in a specific order:

1. ISA-AD01
2. ISA-AD02
3. ISA-SQL01
4. ISA-GISLIC01
5. ISA-GIS01
6. ISA-WEB01

Each server needs to come up before the next server is restarted. All other servers can be rebooted after AD01 and AD02 come back online.

Annex II

**International Seabed Authority
Data Management Strategy**

**DEEPDATA
REPORTING TEMPLATE GUIDANCE
FOR THE
SUBMISSION OF DIGITAL DATA
VERSION 1.8**

Table of Contents

GLOSSARY OF TERMS.....	iii
REPORTING TEMPLATES –GENERAL INSTRUCTIONS	1
1. INTRODUCTION.....	1
1.1 Organization of this Guidance Document.....	1
1.2 Which Template to Use?.....	1
1.3 File Name Convention	2
1.4 Modifying the Template Structure	2
1.5 Reporting Template Worksheets.....	2
1.6 Legend worksheet	4
2. EXPLANATION OF KEY IDENTIFIER FIELDS	5
2.1 Expedition Identifiers.....	5
2.2 Station and Sample identifiers	5
2.3 Relating Information in Different Worksheets.....	6
3. INSTRUCTIONS FOR ASSIGNING STATION IDS.....	6
4. INSTRUCTIONS FOR ASSIGNING SAMPLE IDS	7
5. CODING SAMPLES WITH BOTH CONFIDENTIAL AND PUBLIC DATA	7
6. INSTRUCTIONS FOR ENTERING CTD CAST DATA.....	8
7. QUALIFIER CODES	8
8. VALID VALUE LISTS.....	9
8.1 ValidValue_Analysis Worksheet.....	9
8.2 ValidValue Worksheet.....	10
9. STUDYNOTES WORKSHEET.....	10
10. INSTRUCTIONS ON LINKING RESULTS TO IMAGE FILES.....	10
11. ERRORCHECKER.....	11
REPORTING TEMPLATES –SPECIFIC INSTRUCTIONS	12
12. SPECIFIC INSTRUCTIONS FOR THE GEO REPORTING TEMPLATE.....	12
12.1 Data Entry Worksheets	12
12.2 Data and Information to Submit for Geological and Resource Assessment	12
12.3 Provisional Results.....	14
12.4 Data Requirements.....	14
12.5 Data Reporting Specifications.....	15

13.	SPECIFIC INSTRUCTIONS FOR THE ENV REPORTING TEMPLATE.....	16
13.1	General.....	16
13.2	Details on Identification Methods and Taxonomy Information.....	17
13.2.1	Morphological ID.....	17
13.2.2	Genetic ID.....	18
13.2.3	Remote ID.....	19
13.3	Details on Vouchers.....	19
13.4	Submitting Ancillary Information.....	20
13.5	Provisional Results.....	20
13.6	Ecological and Organism details.....	21
13.7	Chem_Results in the ENV Template.....	21
	METADATA TEMPLATE & DATA SUBMISSION.....	23
14.	METADATA TEMPLATE.....	23
14.1	File Naming Convention for Metadata Template.....	23
14.2	Modifying the Template Structure.....	23
14.3	Metadata Template Management.....	23
14.4	Instructions for Populating.....	24
14.5	Photo_Video Theme.....	24
14.6	File Naming Convention for Data Submissions.....	25
14.7	ErrorChecker.....	26
15.	SUBMITTING DATA.....	26
15.1	Resubmittals.....	27
16.	REFERENCES.....	27
	Annex A.....	28
	Annex B.....	29
	Annex C.....	30

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.

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.

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.

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.

10. REPORTING TEMPLATES – GENERAL INSTRUCTIONS

1. INTRODUCTION

1.1 Organization of this Guidance Document

This document includes guidance and instructions for populating the Reporting Templates (GEO and ENV) and the Metadata Template. The document has three main parts:

1. **Reporting Templates – General Instructions.** These instructions apply to both the GEO and ENV Reporting Templates (Section 1 – 11).
2. **Reporting Templates – Specific Instructions.** These are specific instructions for entering results into the GEO Reporting Template (Section 12) and the ENV Reporting Template (Section 13)
3. **Metadata Template & Data Submission.** These instructions apply to preparing the Metadata Template and data submission packages to upload to the DeepData website (Section 14 - 15).

1.2 Which Template to Use?

The Reporting Templates are to be used by Contractors and other stakeholders (termed data providers) to submit Structured data (required biological, geological, and environmental chemistry data collected during expeditions) for PMN, CFC, and PMS environments.

Use the GEO_Template for geological results from samples that are considered to represent mineral resources (e.g., geotechnical and mineralogy properties of PMN, PMS or CFC resources). These results will be flagged as confidential in DeepData (ISA 2015)¹. See Section 12 for specific instructions on populating the GEO template.

Use the ENV_Template for biological and environmental chemistry results from samples that are not considered to represent mineral resources (e.g., pelagic sediment chemistry, water chemistry, CTD cast sampling). These results will be flagged as publicly available in DeepData (ISA 2015)¹. See Section 13 for specific instructions on populating the ENV template.

The Reporting Templates can be accessed from the ‘Downloads’ tab of the DeepData website (<https://data.isa.org.jm/isa/map/>). A Template Example Package is also available for download and will be an important reference for populating the Templates.

¹ Regulations pertaining to confidentiality of data and information included in ISA (2015): PMN: ISBA/19/C/17: Part II, Regulation 7; Part VI Regulation 36 and Regulation 37
PMS: ISBA/16/A/12/Rev.1: Part II, Regulation 7; Part VI Regulation 38 and Regulation 39
CFC: ISBA/18/A/11: Part II, Regulation 7; Part VI Regulation 38 and Regulation 39

For questions about how to fill out the Reporting Templates, contact the ISA Data Management Team at: data@isa.org.jm.

1.3 *File Name Convention*

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

Use the following naming convention for the Reporting Template file:

ContractID² + Contractual Year_ + master file name_ + descriptive details

For example: BGRPMN12015_Geo_Template_2014rock.xlsm

The file name may not exceed 100 characters.

1.4 *Modifying the Template Structure*

The Template structure is designed to be compatible with ETL procedures to append results into the DeepData database. As such, the structure is not flexible:

- 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.
- Do not use non-standard characters in the template; a Unicode character set must be used.
- Do not make any updates to the legend or valid values sheets.

1.5 *Reporting Template Worksheets*

Each Reporting Template includes several worksheets which are described in Table 1.

An important concept regarding the organization of the Reporting Templates is that Station- and Sample-level information is captured in ***different worksheets*** than the Results-level information obtained from those samples.

- Example of Station-level information: coordinates
- Example of Sample-level information: sampling date, sampling depth
- Example of Result-level information: metal concentrations

² A list of ContractIDs can be found in Annex A.

Table 1. Listing of the worksheets in the Reporting Templates.

<u>Worksheet</u>	<u>Purpose</u>
GEO Reporting Template	
Point Sample	Enter Station- and Sample-level information for samples collected at a point location (e.g., cores, grabs).
Towed Gear Sample	Enter Tow- and Sample-level information for samples collected a line location (e.g., trawl or dredge samples).
GeoResource_Results	Enter geological results for field collected samples.
Resource Information	Enter information about sampling locations that are specific to mineral-type.
Resource Assessment	Enter site-level information regarding resource assessment.
ENV Reporting Template	
Point Sample	Enter Station- and Sample-level information for samples collected at a point location (e.g., cores, grabs).
Towed Gear Sample	Enter Tow- and Sample-level information for samples collected a line location (e.g., trawl or dredge samples).
Chem_Results	Enter environmental chemistry results for field collected samples.
Biological_Results	Enter biological results.
Worksheets in both GEO and ENV Reporting Templates	
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.
Qualifiers	Stores definitions for qualifier codes included in GeoChem_Results.
ValidValues	Listing of allowable entries for fields constrained by valid values.
ValidValues_Analysis	Listing of allowable entries for the Category, Analysis, and Units fields in the GeoResource_Results and Chem_Results worksheets.
refQualifiers	Listing of example qualifier codes commonly used with analytical data. Examples are provided as a reference only.
StudyNotes	Enter any narrative information about the study in this worksheet.
ErrorChecker	The 'Error Checking' command executes a series of QAQC checks to ensure that the data set, as a whole, is reported correctly.

1.6 Legend worksheet

Strictly refer to the **Legend** worksheet of the Reporting Templates. Each of the fields in the Template are described in this worksheet and provides an important resource for understanding how to populate the Template. Table 2 provides descriptions of each of the columns included in the **Legend** worksheet.

Table 2. Description of information provided in the Legend worksheets.

<u>Column</u>	<u>Description</u>
Worksheet	Indicates which Template worksheet the field is included in
Sample Worksheet	Indicates if the field is included in the 'Point Sample' worksheet, 'Towed Gear Sample' worksheet, or both. The purpose of the Sample Worksheet column is to avoid repetition of the fields that are common between the 'Point Sample' and 'Towed Gear Sample' worksheets.
Category	Indicates the Category of information; corresponds to the entries in Row 1 of the data entry worksheets.
Field	Name of the field included in the data entry worksheets
Data Type	Indicates the data type format in the DeepData database.
Definition	Description of the field. Includes important information for understanding the intended content of the field.
Valid Value Field	Indicates if the field is constrained to the list of valid values (included in the ValidValue and ValidValue_Analysis worksheets). See Section 8 for more information.
Required Field	Indicates if the field must be populated. 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.
Maximum Length of Entry	Indicates the maximum number of characters allowed for the field.
Relevant Mineral Types	Indicates if the field is only relevant to certain mineral types.
Default values for Required Fields	For Required fields, indicates the default value to use if true results were not obtained.

2. EXPLANATION OF KEY IDENTIFIER FIELDS

Fields that create the unique key for records in the **Point Sample** and **Towed Gear Sample** worksheet are:

ContractID, Cruise Name, Leg Number, Station ID (or Tow ID), Sample ID

2.1 Expedition Identifiers

In the **Point Sample** and **Towed Gear Sample** worksheets, the key fields that identify data collected during an expedition are: **ContractID**, **Cruise Name** and **Leg Number**.

All data collected during an expedition (regardless of data type, sampling equipment or sampling method) are associated by assigning the same codes in these three fields.

2.2 Station and Sample identifiers

The codes entered in the Station ID and Tow ID fields are identifiers for *positions*, as indicated by different geographic coordinates.

Station ID is used in the **Point Sample** worksheet and is the identifier for samples collected from a point location (e.g., sediment core sample). Tow ID is used in the **Towed Gear Sample** worksheet and is the identifier for samples collected from line locations (e.g., tows or trawls).

The codes entered in the Sample ID field are identifiers for samples collected (e.g., depth intervals from a sediment drill sample).

In this context, a Station ID does NOT identify a feature (e.g., a specific chimney edifice). If a group of Station IDs are associated with a feature, this should be entered in the Station Description field. In Annex B, Example 1, the same feature (Edifice 123) was sampled on 3 different days; the coordinates for each Station ID are slightly different, and each different ROV dive is given a different Station ID; an entry is provided in the Station Description field to allow grouping by the sampled feature.

Conversely, Example 2 is a drill sample that was sub-sampled into intervals. The Station ID, Sample Date and coordinates are the same, and the 4 different depth intervals are indicated in the Sample ID and Upper/Lower Depth fields.

For a single expedition, the combination of Station ID+Sample ID (or Tow ID+Sample ID) fields must be coded to identify unique field-collected samples.

Using the same Station ID (or Tow ID) associates all data collected at that location; using different Sample IDs differentiates samples collected that are different matrices (sediment or water or rock), collected on different dates and/or different depths.

See Section 6 for special instructions for coding Station and Sample identifiers for CTD casts.

2.3 Relating Information in Different Worksheets

For GEO Template:

- Station ID+Sample ID entries in the **Point Sample** worksheet must match to entries in the **GeoResource_Results** worksheet.
- Tow ID+Sample ID entered in the **Towed Gear Sample** worksheet must match to entries in the **GeoResource_Results** worksheet.
- Station ID or Tow ID entries in the **Resource Information** worksheet must match to entries in the **Point Sample** or **Towed Gear Sample** worksheet.
- Exploration Block ID/Cell ID or Exploration Subarea entries³ in the **Resource Assessment** worksheet must match to entries in the **Point Sample** or **Towed Gear Sample** worksheet.

For ENV Template:

- Station ID+Sample ID entries in the **Point Sample** worksheet must match entries in either the **Chem_Results** worksheet or the **Biological_Results** worksheet.
- Tow ID+Sample ID entries in the **Towed Gear Sample** worksheet must match entries in either the **Chem_Results** worksheet or the **Biological_Results** worksheet.

The Error Checker will ensure that Station ID/Sample ID/Tow ID in the sample-level worksheets are represented in the results-level worksheets [an exception to this is tow samples that only have video(s) associated with them].

3. INSTRUCTIONS FOR ASSIGNING STATIONIDS

Different Station IDs should be assigned when the Latitude/Longitude are different. The exception is CTD casts (see Section 6 below).

Preferably, the Latitude/Longitude should be measured with the instrument at the seafloor (bottom samples), or the sample collection point (for mid-water sampling). The location of the instrument recording coordinates must be specified in the Position of GPS/Transducer field.

³ A shapefile with the Exploration Block ID/Cell ID and Exploration Subarea entries can be accessed from the 'Downloads' tab of the DeepData website (<https://data.isa.org.jm/isa/map/>).

The formats for coordinates must be provided in Decimal Degrees (DD); all other coordinate formats are not acceptable [e.g., Degrees/Minutes/Seconds (DMS)].

Information on station locations must be made to the most accurate geographical and depth measurements, whenever possible.

4. INSTRUCTIONS FOR ASSIGNING SAMPLE IDs

When samples are collected at different depths, each sample interval should be assigned a different Sample ID. See Annex B, Example 2.

Different matrices sampled must have unique Station IDs+Sample IDs (e.g., a sediment and porewater sample collected at the same station should be considered two separate samples, and Sample IDs such as 01S and 01W should be assigned). See Annex B, Example 3.

Similarly, if total and dissolved fractions are measured in the same field-collected water sample, each fraction must be assigned a different Sample ID (e.g., 01T for the total fraction and 01D for the dissolved fraction). See Annex B, Example 4.

Sample IDs for geochemistry samples must be different from the Sample IDs given to biological samples. See Annex B, Example 5.

Since **Point Samples** and **Towed Gear Samples** are captured in different worksheets, it is important that the Station IDs and Sample IDs do not overlap between the two worksheets. So, if both Point Samples and Towed Gear Samples are collected during the same Cruise, the same Sample ID must not be used in both the Point Sample worksheet and the Towed Gear Sample worksheet.

5. CODING SAMPLES WITH BOTH CONFIDENTIAL AND PUBLIC DATA

Most cruises will collect samples that will be designated as EITHER confidential OR public data. These **separate** samples will be differentiated by being included in EITHER the GEO Template OR the ENV Template, respectively.

There is the special example of **individual** samples that include both confidential AND public data. An example is a box core sample that is sub-sampled for nodules and sediment and both are analyzed. In

this case, the nodule results must be entered in the GEO Template; the sediment results must be entered in the ENV Template. To identify the samples as sub-samples:

- In both templates, enter the word "Paired" in the Sample Collection Method field.
- For these 'Paired' samples, use the same ContractID/Cruise Name/Leg Number/Station ID in both templates.
- Use the same root Sample ID in both templates and add a "P" suffix (for paired) to the Sample ID in the ENV Template.

6. INSTRUCTIONS FOR ENTERING CTD CAST DATA

CTD data is publicly accessible and therefore must be entered in the ENV Template.

Samples and sensor readings from the same cast should have the same Station ID, *even if there are unique coordinates for each Sample ID* (this is an exception to the rule that each set of unique coordinates should have a unique Station ID; see Section 3).

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

In the **Chem_Results** worksheet, results from CTD sensor readings must have Category = 'Water Properties'. The depth and sample date for CTD sensor readings are entered in the CTD_Depth and CTD_SampleDate fields.

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

See the Template Example Package for specifics on entering CTD cast data in the ENV Reporting Template.

7. QUALIFIER CODES

Analytical data is accompanied by qualifier codes that flag results with certain problems and scenarios (e.g., U = Analyte was not detected at or above the reporting limit; J = estimated result). The laboratory will assign qualifier codes to individual results to assist users with interpreting the data. There may also be an independent 3rd party validation (data quality assessment) in which the laboratory data are scrutinized, and the validators add validation qualifier codes.

Examples of qualifier codes commonly used with analytical data are included in the **refQualifiers** worksheet. These example codes are provided for reference only.

The qualifier codes accompanying the results from the analytical laboratory should be captured in the 'Qualifier' column in the **GeoResource_Results** (GEO Template) or **Chem_Results** (ENV Template) worksheet.

Qualifier codes in the **GeoResource_Results** or **Chem_Results** worksheets must be included in the **Qualifiers** worksheet and defined.

8. *VALID VALUE LISTS*

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

Fields constrained by valid values are identified in the **Legend** worksheet, and in row 2 or 3 of the **Point Sample**, **Towed Gear Sample**, **GeoResource_Results/Chem_Results** and **Biological_Results** worksheets.

8.1 *ValidValue_Analysis Worksheet*

- Entries in the 'Category' and 'Analysis' columns of the **GeoResource_Results/Chem_Results** worksheet must exactly match the entries in the **ValidValues_Analysis** worksheet.
- Entries that are not included in this list will be identified by the Error Checker (see Section 11) and an error message will be given so that the invalid entry can be easily identified and corrected.
- The UnitsSolid and UnitsWater fields in the **ValidValues_Analysis** worksheet provides an indication of the acceptable unit of measurement to use for each for each parameter, depending on whether measured in a solid or aqueous matrix.
- The **ValidValues_Analysis** worksheet includes the current list of analytes for the ISA DeepData database, however, not all analytes must be reported for each template and matrix type.
- 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 Management Team and incorporated into the master list, if approved.

8.2 *ValidValue Worksheet*

- The allowable entries for all other valid value fields (i.e., not including Category and Analysis fields from the **GeoResource_Results/Chem_Results** worksheet; see Section 8.1) are included in the **ValidValues** worksheet.
- Entries in these valid value fields must exactly match the entries in the **ValidValues** worksheet.
- Entries that are not included in this list will be identified by the Error Checker (see Section 11) and an error message will be given so that the invalid entry can be easily identified and corrected.
- The ValidValue list includes special codes if information for valid value fields is not available (e.g., 'Not Applicable' or 'Not Reported'). The Error Checker will only accept these values as correct entries.
- It is anticipated that the current valid value list for 'Sampling Device' is not comprehensive, therefore it is acceptable to enter a Sampling Device that is not included in the valid value list – in this case ignore the Error Checker error. New items will be reviewed by the ISA Data Management Team and incorporated into the master list, if approved.

9. *STUDYNOTES WORKSHEET*

Contractors can include any narrative information about the study in the StudyNotes worksheet.

Examples of information to include:

- For boxcore samples, it is important to mention where the boxcorer was deployed, i.e. from the back or the side of the ship, as this has important consequences for sample quality (macrofauna data).
- Parameters used in modelling calculations
- For calculated ecological indices (e.g., Relative abundance, Relative Dominance, Taxon Density) - Contractors are encouraged to provide details on the calculation procedures used

10. *INSTRUCTIONS ON LINKING RESULTS TO IMAGE FILES*

Image files are an important aspect of documenting and characterizing the deep seabed environments.

Examples of image files: video tracks taken from an ROV; photographs of organisms retrieved from collected samples, a frame from a video displaying a feature or organism observed in the water column.

Information on individual photos and videos is required to allow linkages to specific samples, tows, or results (i.e., organisms), which can then be accessed via the DeepData website Photo/Video Gallery tab.

Linkages between samples/results and image files must be made in BOTH the Reporting Template and the associated Metadata Template.

In the Reporting Template, information on photo/video files is entered in the following worksheets and fields:

Point Sample - Image File Name

Towed Gear Sample - Image File Name

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

If multiple image files are submitted for a single sample or result, include multiple entries in these fields, separated by commas.

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

The Metadata Template:

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

See Section 14.5 for specific instructions on populating the Metadata Template with information on image files.

11. *ERRORCHECKER*

The 'Error Checking' command, in the **ErrorChecker** worksheet executes a series of QAQC checks to ensure that the data set, as a whole, is reported correctly. A Checker Report is generated with messages describing the errors found, 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 (this is to ensure that the Checker Report is updated to reflect the most recent changes in the Templates).

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

11. REPORTING TEMPLATES – SPECIFIC INSTRUCTIONS

12. *SPECIFIC INSTRUCTIONS FOR THE GEO REPORTING TEMPLATE*

This section provides specific instructions for entering confidential data into the GEO Reporting Template. The general instructions for entering sampling information (i.e., Point Sample and Towed Gear Sample) apply to the GEO Reporting Template, and this section focusses specifically on entering resource assessment results that are specific to the GEO Reporting Template.

12.1 *Data Entry Worksheets*

When entering results from geological prospecting and sampling into the GEO Reporting Template, Contractors provide information in five different worksheets:

1. **Point Sample** - Cruise, Station and Sample information for samples collected at point locations (e.g., drill samples)
2. **Towed Gear Sample** - Cruise, Towed Gear location, and Sample information for samples collected at tow locations (e.g., dredge samples)
3. **GeoResources_Results**: Results from geochemical analysis are captured in this worksheet - Station ID (or Tow ID), Sample ID, Category, Analysis, Result, Units, Qualifier, Limits of Detection (LOD), Analytical Technique, Lab Matrix, Total or Dissolved, Measurement Basis, Laboratory, Remarks.
4. **Resource Information** – information associated with specific mineral types to characterize resources at sampling locations (e.g., Hydrothermal Activity, Nodule abundance, Crust Thickness)
5. **Resource Assessment** – information associated with resource assessment for specific mineral types, at a site-level (e.g., Extensions of PMS site, Inferred Tonnages, Area of Crust Coverage)

12.2 **Data and Information to Submit for Geological and Resource Assessment**

Mineral exploration aims at the identification and assessment of mineral resources. Identification of mineralized areas and subsequent definition of deposits with potentially economic ore grades, and earliest spatial reduction of the exploration (i.e. license) area for cost reduction are of key interest of land-based exploration, and the same approach is required for economic seafloor exploration. All three mineral resources in the Area have specific characteristics for their formation and enrichment to deposit-sizes that largely depend on the regional to local geological and oceanographic conditions. Geological models were successfully developed which reliably describe the conditions of formation for all three types. However, ore deposit formation requires specific processes and favorable conditions, and their detailed documentation is critical for deposit characterization.

Specific local characteristics and mineral endowment as well as overall variation of geological and oceanographic conditions in un- or less-mineralized regions of the Area can ideally be described by the following (categorized as Unstructured or Structured data):

Unstructured

1. Detailed swath bathymetric survey(s) of the license area with resolutions better than 20 meters for areal coverage of manganese nodules (PMN) and better than 5 meters for local mineral enrichments (polymetallic sulphides, PMS, cobalt-rich crusts, CFC). Detailed and high-precision swath bathymetry allows for spatially high-resolution sample localization and characterization.
2. Detailed geophysical surveys for subsurface identification and modeling: Gravimetry, magnetics, (backscatter) sonar (PMN, PMS, CFC), electromagnetics (PMS, CFC), seismics (3D, PMS).
3. Detailed geological mapping and photo documentation of the identified deposit (towed video sled, ROV, AUV, HOV).

These unstructured data types are submitted as raw files and cataloged in the Metadata Template (see Section 14).

Structured

1. Resource assessment and classification by determination of coverage density (PMN) and by drilling (PMS, CFC).

These structured data types are submitted in the GEO Reporting Template format.

Analysis and assessment of these data ultimately will result in:

- Scientific assessment of geological exploration data and deposit modeling, and,
- Detailed geological characterization and assessment, and definition of prospective areas according to geological evolution of the license area: Structural geology, volcanism, sedimentation processes, mineral precipitation, remobilization and redistribution.

Implementing these data collection efforts during exploration facilitates the assessment of resources, deposit characterization and classification. Survey data are submitted in raw format and logged into the Metadata Template (e.g., bathymetric surveys); analytical data are re-formatted to DeepData's confidential GEO Template to be incorporated to the DeepData database as structured data. The documentation assists in the compilation, understanding, interpretation and refinery of exploration data to official mineral title application forms, market analysis of mineral resources and required technical resource classification reports (e.g., NI 43-101, CRIRSCO, ISA Reporting Standard template).

12.3 *Provisional Results*

Submissions to DeepData should include provisional results during exploration. It is recognized that the assessment, characterization, and deposit definition are usually a long and iterative process. Provisional results should be submitted to DeepData, with the understanding that the different identified sites and areas will be revisited continuously, geological and analytical results may differ over short distances and the assessment will be refined and re-submitted as exploration and sampling continues, and analyses are refined to a greater level of detail. Information on the status of characterization and the deposit identification process are indicated by the Cruise and Station Information, the Sample Date; specific indications of the status of analytical results can be indicated in the Qualifier field of the **GeoResource_Results** worksheet. It is understood that entries in certain fields will be incomplete when the status of knowledge is interim.

When results from a previously submitted Reporting Template are revised (e.g., Qualifiers are added; additional parameters are added; Results are changed), the Reporting Template must be re-submitted. **Re-submissions must include the entire dataset, not just the records that have been updated.** See Section 15.1 for the detailed procedures for submitted revised Reporting Templates.

Note that previous submissions will be removed from DeepData and archived; the re-submitted dataset will replace the original version. A record of the progression of deposit definition and resource assessment over time will be stored in the DeepData Archive.

12.4 *Data Requirements*

The **ValidValues_Analysis** worksheet lists information on a vast variety of different chemical, physical and compositional analytical categories. These include element analyses, geotechnical parameters, mineralogical compositions, physical properties, and nutrient compositions. The type of samples collected during seafloor exploration is quite variable and includes organic and inorganic materials, solid and diverse volcanic rock samples, soft and indurated sediment samples of different origin (e.g., volcanoclastic, terrigenous, biogenic), highly diverse mineral precipitates; pore water, bottom-interface water and open ocean seawater, and also mixtures and intergrowths of all the former. It is evident that not all the analytical categories are applicable to every single sample, that a number of samples need severe sample processing for a proper analysis of single categories, and that the sampling process as well as the sample treatment can severely alter the sample composition.

This version of the GEO Template does not specify the required measurements over an exploration period, nor does the Error Checker verify that the required measurements have been submitted. This verification will be conducted on the data compiled for the entire exploration period (i.e., the checks will be conducted by the ISA Data Management Team on the data compiled in the database for the contract).

12.5 Data Reporting Specifications

An increasing number of chemical and physical parameters can be analyzed in the field, and both on-line measurements and lab measurements (e.g., pH values and element concentrations, respectively) need to be differentiated by providing information in the Sample Collection Method and Sampling Device fields. It is necessary to supply information on the sampling methods and analytical techniques applied and to report critical detection limits as well as analytical uncertainties, in the LOD (Limit of Detection) and Qualifier fields, respectively. Analytical caveats, however, should not prevent from reporting as continuing exploration helps to review, refine and further state the analytical information more precisely. Also, the potential general variability in compositions (at least for some geological samples) will provide important information on natural variability within the license area, with consequences for the exploration as well as for environmental baseline surveys.

The detection of physical and geotechnical parameters will largely influence any potential development of mining technologies and future mining strategies. Measurements are critical for successful technical solutions and improve the development of low-impact collector, mining and recovery systems. Proper selection of the type of analysis, its execution, critical operation analysis and documentation of the analysis and results are necessary and therefore suggested.

For each Category/Analysis reported in the **GeoResources_Results** worksheet, it is required to report in the following fields: Category, Analysis, Result, Units. It is advantageous to also provide information in the following fields: Analytical Technique, Qualifier and LOD (Limit of Detection). In addition, it is important to provide additional information for the following Categories, to aid in data interpretation:

- For the Elements and Rare Earth Elements Category - Loss on Ignition (wt%) should be included
- For the Nodule Properties Category
 - Analyzed Segment should be reported, to indicate whether the nucleus, interior or rim was the analyzed segment.
 - Layer Sampled (in the **Point Sample** worksheet) should be reported, to indicate the layer sampled, relative to the surface of the nodule, in mm).
- For the Mineralogy Category
 - It is recognized the quantitative mineralogy is a complicated process and is not necessary to be determined for every sample. Percentages of specific minerals (e.g., ‘Minerals – Feldspar’ need only be reported when available and should be accompanied with notes on analytical conditions [enter in the Remarks field; for example: timing, x-ray conditions (shaking or light), method preparation].

At this time, laboratory QA/QC results (e.g., method blanks, lab replicates) should not be incorporated in the Reporting Template. This information can be submitted as raw data via the Metadata Template (see Section 14).

Analytical results stored in the **GeoResources_Results** worksheet must be made to the highest precision and qualified confidence level, whenever possible. Analytical results must be reported in standardized units, which are listed in the **ValidValues_Analysis** and **ValidValues** worksheets.

Other specifics to note regarding the Matrix field:

- Sediment porewater is interstitial water, which is water extracted from the sediment column, NOT water collected at the sediment/water interface.
- For sediment porewater samples, Upper Depth and Lower Depth entries will reflect the sediment sampling depth (although Matrix Type = porewater).
- All analytical results for solid matrices (i.e., rock, sediment) must be reported on a dry weight measurement basis.

13. *SPECIFIC INSTRUCTIONS FOR THE ENV REPORTING TEMPLATE*

This section provides specific instructions for entering publicly available data into the ENV Reporting Template. The general instructions for entering sampling information (i.e., Point Sample and Towed Gear Sample) apply to the ENV Reporting Template, and this section focusses specifically on entering biological results that are specific to the ENV Reporting Template.

For Station ID + Sample ID that have associated results in the **Biological_Results** worksheet, the Matrix Type must be assigned as 'Biological sample'.

13.1 *General*

For entering results from biological sampling into the **Biological_Results** worksheet of the ENV Reporting Template, the following six categories of information are available:

1. **Taxa ID** (i.e., the taxonomic identification of individual organisms; Class, Family, Species, etc.; identification method and status)
 - a. The entries in the Species field should be the species name without the author citation (e.g. "hoffmeisteri" not "hoffmeisteri Claparède, 1862"). The author citation should be entered in the 'Taxonomic author citation' field.
 - b. Note that the 'Scientific Name' field is a Required field and is a key field for verifying taxonomic identifications against the World Register of Marine Species (WoRMS) taxonomic database (see below).
2. **Taxonomy Information** (i.e., information regarding traceable data or evidence from which the taxonomic identification was obtained, e.g., voucher specimens, animal tissues, and DNA sequence information)
3. **Identifier Information** (name, email, institution, whether the identifier is a taxonomist)
4. **Ecology** (e.g., number of individuals collected, biomass, etc.)
5. **Organism Details** (e.g., sex, life stage of individual organisms, etc.)

6. **Photo/Video Information** (if applicable, this information is captured to allow a linkage to the file collection Library, enabling database users to quickly identify and access photos/videos of a specific sample or organism using the DeepData website).

13.2 *Details on Identification Methods and Taxonomy Information*

Following the guidance provided by the Legal and Technical Commission (ISBA/25/LTC/6/Rev.1, Annex I, para. 47), more accurate and consistent taxa identification is provided by two sets of information: morphological descriptions (e.g., images, drawings) and genetic sequences. For new species descriptions, the two sets of information are then published in scientific journals as well as added to international databanks and curated collection facilities (voucher specimens, molecular samples, including DNA extracts).

There are three different types of Identification Methods that can be entered in the Reporting Templates: **Morphological ID**, **Genetic ID** and **Remote ID**. Specific information is required to be reported for each Identification Method (see details below). In addition, there is an option to indicate ‘combination’ Identification Methods (see the **ValidValue** worksheet); this allows data providers to indicate instances where multiple identification methods were used (i.e., genetic + morphological; Remote ID + genetic/morphological).

Under the Taxonomy Information category, the data provider is encouraged to report on taxonomic databases [e.g., World Register of Marine Species (WoRMS), GenBank, etc.] accessed to record the information that would allow an end user to verify the Scientific Name reported; this information is entered in the following fields: *Taxonomic Database*, *Database Taxa ID*, *LSID*, *INSD accession number*, as well as details on vouchers (see specific details below). **The purpose of tracking the ancillary data associated with a taxonomic identification is to allow regulators (and other data users) to:**

- 1) locate specimens;
- 2) track information on how identifications were determined; and
- 3) validate the identification.

It is critical to correctly populate the taxonomic categories (i.e., Kingdom, Phylum, etc), with respective scientific name under each taxonomic category, up to the lowest taxonomic rank that can be determined.

13.2.1 Morphological ID

If identification was made by an expert using morphology and phenotypic characteristics (Identification Method = Morphological ID), the following Template fields should be populated:

- a. **Taxonomic Database** – Reference database used to verify the reported taxonomic name (e.g., WoRMS). For a listing of available databases, refer to Table 1 in: Rabone *et al.* 2019.

- b. **Database Taxa ID** - The associated taxon identifier from the taxonomic database (e.g., if using the WoRMS website, the AphiaID code would be entered; note that a Database Taxa ID is required for species level as well as higher-level taxa identification),
- c. **Taxonomic Status** – Status of taxon details, as indicated in the taxonomic database; in case of pending/interim identifications, “NA” should be entered).
- d. **Taxonomic Identification Qualifier** - A controlled value to express the determiner's doubts about the Identification (see **Legend** worksheet, Templates Example Package, and Horton *et al.* 2021).
- e. **Notes on Taxonomic Identifications** – Additional remarks on identification status, explaining taxonomic identification qualifier (see Templates Example Package for example entries).
- f. **LSID** – the Life Science Identifier

The Life Science Identifier (LSID) is critical, as it provides a ‘naming standard underpinning for wide-area science and interoperability’ (LSID, n.d.). Data providers are encouraged to use the WoRMS database to verify taxonomic nomenclature because WoRMS tracks changes to taxonomic names, and the ISA will use WoRMS to verify the taxonomic information submitted by Contractors. The WoRMS website provides a Taxon Match tool, to automatically match a species or taxon list with the WoRMS database. Guidance for using this tool can be found at:

- 1) <http://www.marinespecies.org/aphia.php?p=match>
- 2) <https://www.youtube.com/watch?v=Mk1IiEwvsF8> (from 5min48sec to 13min55sec)

13.2.2 Genetic ID

If identification was made using genetic methods, such as DNA barcoding, (Identification Method = Genetic ID), the following Template fields should be recorded, for a Final identification:

- a. **Taxonomic Database** - Database used for data submission (e.g., DDBJ, ENA, NCBI)
- b. **INSD accession number** - INSD accession number(s) assigned by the data submission process of the DDBJ, ENA or NCBI databases.
- c. **Description of DNA Marker** - describe the gene sequence (e.g., COI, 16S genes, 18S genes, etc.)
- d. **Tissue Descriptor** – information on the tissue type submitted for DNA analysis.
- e. **LSID** – the Life Science Identifier (see above)

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 databases and a single accession number is propagated to all three sources. Further information about DNA accessions is available on the DDBJ (DNA Data Bank of Japan) website:

http://www.ddbj.nig.ac.jp/sub/acc_def-e.html.

In cases where the identification is made solely by DNA sequencing, information on phylogeny should be provided to the lowest possible taxonomic rank (i.e., entries in the high-level taxonomic categories; Kingdom, Phylum, etc.), as well as “Morphotype”, if the Identification Status is Interim or Final at the time of Template submittal. The WoRMS LSID should be entered for the lowest taxonomic level determined.

13.2.3 Remote ID

If identification is made remotely, i.e. using video or image observations, without physical specimens (Identification Method = Remote ID), the following Template fields should be recorded:

- a. **PhotoVideo Information** – Images that are associated with the organism identified by Remote ID should be submitted to DeepData and associated with the specific organism using the following fields: Photo file name, Video frame code, Video photo frame file name.⁴⁵ Note that multiple images can be submitted for a single organism identification using comma separated entries (see Section 10 on instructions for linking results to image files).
- b. **Taxonomic Database** – If the images used in the remote identification were submitted to an image database, include the database in the Taxonomic Database field (e.g., Biigle).
- f. **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 Taxonomic Database referenced.
- g. **LSID** – the Life Science Identifier (see above)

13.3 Details on Vouchers

Following the guidance provided by the Legal and Technical Commission (ISBA/25/LTC/6/Rev.1, IV.A.19; IV.B.22), “*Specimens **must** be archived for comparison with taxonomic identifications from other sites and to understand the details of changes in the composition of species over time. If species composition does change, it might be subtle, and reference back to the original animals (where there might only have been a putative identification) is essential. It is recommended that samples be archived as part of national or international collections and that appropriate funding be identified to facilitate this.*”

⁴ See Section 10 and 14.5 of this guidance, Section 3.2.4 of the User Manual, and the Template Example Package for detailed instructions on submitting photo and video files and recording information in the Metadata Template to document these files.

A specimen voucher is the verifiable reference that can be used (and revisited) to confirm an identification. This means that the identifications can be compared between surveys, by checking the voucher specimens. There are two types of voucher specimens, as follows:

- (1) Type specimens – upon which formally accepted names of taxonomic units are based;
- (2) Taxonomic support specimens – which document identifications in taxon-based studies other than nomenclatural studies.

DeepData requires a taxonomic support specimen to be retained for each identification made. One or two representative organisms must be vouchered, given a code, and stored. **It is suggested that the samples are submitted to a collection registered on this site:** <https://www.gbif.org/grscicoll>

The following information about the voucher must be captured in the Template⁶:

- a. **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);
- b. **Voucher code** - Code assigned by the submission process the voucher is submitted to; identifies the stored voucher specimen;
- c. **Voucher institution code** – Code for the museum/institution where voucher specimen is stored.

13.4 *Submitting Ancillary Information*

Data providers are strongly encouraged to submit the following types of information that are associated with taxonomic identifications:

- 1) Images of organisms obtained during field sampling, and/or laboratory studies (e.g., microscope photos, drawings);
- 2) Publications;
- 3) Protocols used during sampling and identification, including adjustments to the Polymerase Chain Reaction (PCR) protocols.

Files containing this information are submitted as unstructured data, using the Metadata Template to log and categorize files. See Section 14 of this guidance, Section 3.2 of the User Manual, and the Templates Example Package for detailed instructions on submitting ancillary information.

13.5 *Provisional Results*

It is recognized that identifying species can be a long and iterative process. Submissions to DeepData accommodate provisional taxonomic results, with the understanding that the results may be refined and re-submitted at a later date. Information on the current status of the identification process should be included in the following Template Fields: *Taxonomic Identification Qualifier*, *Identification Date*,

⁶ If multiple vouchers are submitted, include multiple entries in these fields separated by commas.

Identification Status (i.e., Final or Interim) and *Notes on Taxonomic Identifications*; for information on the use of *Taxonomic Identification Qualifiers* (e.g., *stet.*, *indet.*, *inc.*, *aff.* and *cf.*) see examples and explanations provided in the **Legend** worksheet, Templates Example Package, and Horton *et al.* 2021. It is understood that entries in certain fields will be incomplete when the Identification Status is Interim (e.g., entry in the ‘Database Taxa ID’ should be set to ‘Not Available’).

When results from a previously submitted Template are revised (i.e., Identification Status and Identification Date is changed; Taxa ID information is updated; **Chem_Results** are revised), the Template must be re-submitted. **Re-submissions must include the entire dataset, not just the records that have been updated.** See Section 15.1 for the detailed procedures for submitting revised Reporting Templates. Note that previous submissions will be removed from DeepData and archived; the re-submitted dataset will replace the original version. A record of the progression of an identification over time will be stored in the DeepData Archive.

13.6 *Ecological and Organism details*

Information on taxa abundance (i.e., number of individuals) and biomass is entered under the Ecology section of the Biological results. Other calculated ecological indices (e.g., Relative abundance, Relative Dominance, Taxon Density) may also be provided, but are not required to be submitted by Contractors in the Template. If this information is reported in the Template, Contractors are encouraged to provide details on the calculation procedures used, either in the ‘Additional notes about ecology’ field, the StudyNote tab, or in associated documentation submitted as unstructured data. Other biological indices (e.g., spatial extent, density of particular taxa, diversity) that have been determined should also be submitted as unstructured data. When available, organism details (e.g., life stage, sex, etc.) should be entered in the template for each taxa.

13.7 *Chem_Results in the ENV Template*

If there is ancillary data collected along with biological sampling (e.g., grain size information from a sediment sample, presence of coral assemblages), this information should be entered in the **Chem_Results** worksheet. Specifics to note about entering results in the **Chem_Results** worksheet:

- For the CFC mineral type, there is a requirement to measure parameters in water samples collected as close as possible to the rock/water interface (epibenthic boundary layer). For these samples, the Habitat Type must be assigned as ‘Epibenthic boundary’.
- For the PMN mineral type, there is a requirement to measure parameters in the pore water matrix. For these samples, the Matrix Type should be assigned as ‘Sediment Porewater filtered’ or ‘Sediment Porewater unfiltered’.
 - For these samples, Upper Depth and Lower Depth entries will reflect the sediment sampling depth;

- 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.

METADATA TEMPLATE & DATA SUBMISSION

14. *METADATA TEMPLATE*

The Metadata Template is used when submitting all data (structured and unstructured) to the International Seabed Authority. The purpose of this template 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.

In addition to this guidance, the User Manual (Section 3.2) and the Template Example Package should be referenced for instructions on populating the Metadata Template.

14.1 *File Naming Convention for Metadata Template*

Use the following naming convention for the Metadata Template file:

MetadataTemplate_ + ContractID + Contractual Year (e.g., MetadataTemplate_DORDPMN12015).

The file name may not exceed 100 characters.

14.2 *Modifying the Template Structure*

The Template structure is designed to be compatible with ETL procedures to append results into the DeepData database. As such, the structure is not flexible:

- All fields in these worksheets are text and should remain as text.
- Do not change the content of the first three to 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.
- Do not make any updates to the legend or valid values sheets.

14.3 *Metadata Template Management*

The 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 logged in the Metadata Template 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 individual file names that are the same.

14.4 *Instructions for Populating*

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

A description of the worksheets included in the Metadata Template file is shown in Table 3.

Table 3. Description of worksheets in Metadata Template.

Worksheet Name	Descriptions
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.
ValidValues	Listing of allowable entries for fields constrained by valid values.
ErrorChecker	The 'Error Checking' command executes a series of QAQC checks to ensure that the data set, as a whole, is reported correctly.

The individual files included in the compressed (zip) files do not need to be cataloged in the Metadata Template, **except for files submitted under the Photo_Video Theme.**

14.5 *Photo_Video Theme*

Information on individual photos and videos is required to allow linkages to specific samples, tows, or results (i.e., organisms), which can then be accessed via the DeepData website Photo/Video Gallery tab.

Linkages between samples/results and image files must be made in BOTH the Reporting Template and the associated Metadata Template. See Section 10 for specific instructions for populating the Reporting Template with information on image files.

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 Annex C provide examples of appropriate data entry in the Metadata Template for photos and videos.

Multiple image files are included in a compressed file (.zip), and this file is logged in the **NonTemplateData** worksheet with an appropriately named 'Zip File Name' (see Section 14.6), and Theme = Photo_Video. The *individual* image files included in the zip file are cataloged in the **PhotoVideoFiles** worksheet, including the name of the individual file in the 'VP File Name' field (e.g., DSCN2095.jpg) and an entry in the 'Photo or Video field' (i.e., to indicate if the file is a photo or a video). Also enter information in the Station ID, Tow ID Sample ID, OrgNum and Image Number fields to link individual photos and videos to specific point samples, tow samples, or results. Each 'VP File Name' is associated with a 'Zip File Name', that was logged into the **NonTemplateData** worksheet.

If multiple image files are submitted for a single sample or result, differentiate these by entering an arbitrary identifier (e.g., 1, 2, 3) in the Image Number field.

If photos/videos are not associated with specific samples or results, then enter 'NA' in the Associated Results field and leave the Station ID, Tow ID, Sample ID, OrgNum and Image Number fields null.

Records in this **PhotoVideoFiles** worksheet must be unique based on the following fields:
Station ID + Sample ID + Tow ID + OrgNum + VP File Name + Image Number

14.6 *File Naming Convention for Data Submissions*

A data submission is a single zip file. Each zip file includes a Metadata Template file and one or more data files. The name of the data submission file is entered in the Zip File Name field in the appropriate worksheet in the Metadata Template. The zip files must adhere to a naming convention⁷.

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

The file name may not exceed 100 characters.

Use the following naming convention for Reporting Template submissions:

⁷The file names are important as they are referenced in automated data processing procedures, which are used to accurately log and categorize the various types of data types submitted.

ContractID + Contractual Year_ + master file name_ + descriptive details
For example: BGRPMN12015_Geo_Template_2014rock.zip

Use the following naming convention for other data submissions:

ContractID + Contractual Year_ + Theme_ + descriptive details

For example:

- DORDPMN12015_AnnualReport_appendices.zip
- JOGMECCRFC12016_Bathymetry_CruiseA.zip

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

14.7 *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 may not be accepted by the ISA if the Checker Report includes Fail error messages.

15. *SUBMITTING DATA*

Data submissions are uploaded through the 'Upload' user interface on the ISA website:
<https://data.isa.org.jm/isa/map/>

Large data submissions (greater than 20 MB) should not be submitted over the website user interface; contact the ISA Data Management Team (data@isa.org.jm) to make other arrangements.

Subsequent to upload, the file status can be viewed in the Upload user interface on the website (e.g., 'Uploaded Successfully' indicates Template has been received, 'Archived' indicates the data has been added to the database or file library).

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

15.1 Resubmittals

When data from a previous submission are revised (i.e., errors were found and corrected, or the dataset was expanded) the data must be re-submitted. **Re-submissions must include the entire dataset, not just the portion of the data that were changed/revised.**

Also, data may have been submitted but not processed and included in DeepData (this can occur if the ISA Data Management Team find errors or issues that require the Contractor to update the file).

Re-submissions must not have the same name as previously submitted files. The file name should include a suffix indicating it is a revision. For example:

- original file name: BGRPMN12015_Geo_Template_2014rock.zip
- revised file name: BGRPMN12015_Geo_Template_2014rock_20210410rev.zip

The re-submitted file must be flagged as such in the Metadata Template by including a comment in the Remarks field.

Note that previous submissions will be removed from DeepData and archived; the re-submitted dataset will replace the original version. The archived data is not available over the DeepData website, but can be requested from the ISA.

16. REFERENCES

International Seabed Authority (ISA). 2015. Consolidated regulations and recommendations on prospecting and exploration. Revised Edition.

International Seabed Authority (ISA). 2020. Recommendations for the guidance of contractors for the assessment of the possible environmental impacts arising from exploration for marine minerals in the Area. Issued by the Legal and Technical Commission, ISBA/25/LTC/6/Rev.1

Life Science Identifier (LSID) resolution project. (n.d.). Retrieved from <http://www.lsid.info/>

Horton T, Marsh L, Bett BJ, Gates AR, Jones DOB, Benoist NMA, Pfeifer S, Simon-Lledó E, Durden JM, Vandepitte L and Appeltans W (2021) Recommendations for the Standardisation of Open Taxonomic Nomenclature for Image-Based Identifications. *Front. Mar. Sci.* 8:620702. doi: 10.3389/fmars.2021.620702

Rabone M, Harden-Davies H, Collins JE, Zajderman S, Appeltans W, Droege G, Brandt A, Pardo-Lopez L, Dahlgren TG, Glover AG and Horton T (2019) Access to Marine Genetic Resources (MGR): Raising Awareness of Best-Practice Through a New Agreement for Biodiversity Beyond National Jurisdiction (BBNJ). *Front. Mar. Sci.* 6:520. doi: 10.3389/fmars.2019.00

Annex A

11.1.1 Listing of current ContractIDs to use when submitting data to DeepData.

ContractID	Description
BGRPMN1	Federal Institute for Geosciences and Natural Resources of Germany - PMN
BGRPMS1	Federal Institute for Geosciences and Natural Resources of Germany - PMS
BPHDCPMN1	Beijing Pioneer Hi-Tech Development Corporation - PMN
BRAZILCRFC1	Companhia de Pesquisa de Recursos Minerais S.A. - CRFC
CIICPMN1	Cook Islands Investment Corporation - PMN
CMMPMN1	China Minmetals Corporation - PMN
COMRACRFC1	China Ocean Mineral Resources Research and Development Association - CRFC
COMRAPMN1	China Ocean Mineral Resources Research and Development Association - PMN
COMRAPMS1	China Ocean Mineral Resources Research and Development Association - PMS
DORDPMN1	Deep Ocean Resources Development Co. Ltd. - PMN
GSRPMN1	Global Sea Mineral Resources NV - PMN
IFREMERPMN1	Institut français de recherche pour l'exploitation de la mer - PMN
IFREMERPMS1	Institut français de recherche pour l'exploitation de la mer - PMS
INDIAPMN1	Government of India - PMN
INDIAPMS1	Government of India - PMS
IOMPMN1	Interoceanmetal Joint Organization - PMN
JOGMECCRFC1	Japan Oil, Gas and Metals National Corporation - CRFC
KOREACRFC1	Government of the Republic of Korea - CRFC
KOREAPMN1	Government of the Republic of Korea - PMN
KOREAPMS1	Government of the Republic of Korea - PMS
MARAWAPMN1	Marawa Research and Exploration Ltd. - PMN
NORIPMN1	Nauru Ocean Resources Inc. - PMN
OMSPMN1	Ocean Mineral Singapore Pte. Ltd. - PMN
POLPMS1	Government of the Republic of Poland - PMS
RUSFEDPMS1	Government of the Russian Federation - PMS
RUSMNRRCRFC1	Ministry of Natural resources and environment of the russian federation - CRFC
TOMLPMN1	Tonga Offshore Mining Limited - PMN
UKSRLPMN1	UK Seabed Resources Ltd. - I - PMN
UKSRLPMN2	UK Seabed Resources Ltd. - II - PMN
YUZHPMN1	Yuzhmorgeologiya - PMN







11.2 Annex B

11.2.1 Example of coding unique Station IDs and Sample IDs for different types of sample collections.

ContractID	Cruise Name	Leg Number	Station ID	Station Description	Sample ID	Sample Date	Latitude	Longitude	Upper Depth	Lower Depth	Depth Units	Sampling Device	Matrix Type	Vertical Datum	Remarks
Example 1: Sampling a single sulphide chimney structure during 3 different ROV dives															
IOMPMN1	Yuzhmorgeologiya	1	ROV1	Edifice 123	Sample A	4/21/2014	10.876133	-119.450667	0	5	cm	ROV	Sediment	SWI	sediment sampling
IOMPMN1	Yuzhmorgeologiya	1	ROV2	Edifice 123	Sample A	4/22/2014	10.865767	-119.63215	-9	-9	cm	ROV	Water filtered	NA	fluid sampling
IOMPMN1	Yuzhmorgeologiya	2	ROV3	Edifice 123	Sample A	9/2/2014	10.844667	-119.83275	0	0	cm	ROV	Rock	SWI	rock sampling
Example 2: Drill sample															
IOMPMN1	Yuzhmorgeologiya	2	Core A		Interval 1	5/15/2014	11.073117	-119.661817	10	20	cm	Drill	Sediment	SWI	
IOMPMN1	Yuzhmorgeologiya	2	Core A		Interval 2	5/15/2014	11.073117	-119.661817	40	70	cm	Drill	Sediment	SWI	
IOMPMN1	Yuzhmorgeologiya	2	Core A		Interval 3	5/15/2014	11.073117	-119.661817	100	110	cm	Drill	Sediment	SWI	
IOMPMN1	Yuzhmorgeologiya	2	Core A		Interval 4	5/15/2014	11.073117	-119.661817	110	130	cm	Drill	Sediment	SWI	
Example 3: Different matrixes collected at the same Station															
IOMPMN1	Yuzhmorgeologiya	2	1923		o1S	5/24/2014	11.067583	-119.65775	0	20	cm	Gravity corer	Sediment	SWI	
IOMPMN1	Yuzhmorgeologiya	2	1923		o1PW	5/24/2014	11.067583	-119.65775	0	20	cm	Gravity corer	Sediment Porewater filtered	SWI	
Example 4: Total and dissolved fractions sampled from the same water sample.															
IOMPMN1	Yuzhmorgeologiya	2	1942		o1T	5/26/2014	11.0743	-119.65455	3210	3211	cm	Niskin bottle	Water unfiltered	AW	
IOMPMN1	Yuzhmorgeologiya	2	1942		o1D	5/26/2014	11.0743	-119.65455	3210	3211	cm	Niskin bottle	Water filtered	AW	
Example 5: Sediment collected from the same Station is sub-sampled and used for biological and chemistry testing															
IOMPMN1	Yuzhmorgeologiya	1	3502		Sample A_bio	5/30/2014	10.844667	-119.83275	0	15	cm	Box corer	Biological sample	SWI	
IOMPMN1	Yuzhmorgeologiya	1	3502		Sample A_chem	5/30/2014	10.844667	-119.83275	0	15	cm	Box corer	Sediment	SWI	

Annex C

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

						
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	Prtoo1.jpg	yttoo4.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 trawl
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	--	--
OrgNum	--	3	--	--	--	5
Frame Code	--	--	--	--	--	456
Area of Image	0.2	0.2	0.5	0.2	--	--
Height above seafloor	--	--	4.2	--	23	13
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
Confidentiality Duration	--	--	5	--	--	--
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].						

Annex III

DATA MANAGEMENT STRATEGY OF THE INTERNATIONAL SEABED AUTHORITY FOR THE PERIOD 2023-2028

I. Introduction

1. As part of its specific and exclusive mandate assigned to it by the United Nations Convention on the Law of the Sea (the Convention) and in the Agreement relating to the implementation of Part XI of the Convention (the 1994 Agreement), the International Seabed Authority is mandated to facilitate and support deep ocean science, particularly in relation to the Area.

2. Under Article 143, paragraph 2, of the Convention, the Authority shall promote and encourage the conduct of marine scientific research in the Area and shall coordinate and disseminate the results of such research and analysis when available. As part of its responsibilities, the Authority also has the duty to encourage appropriate programmes to be designed and implemented for the benefit of developing States and technologically less developed States, with a view to (a) strengthening their research capabilities, (b) training their personnel in the techniques and applications of research and (c) fostering the employment of their qualified personnel in research in the Area (*ibid.*, art. 143 (3)).

3. The Authority is undertaking several initiatives to discharge its responsibilities effectively. These initiatives are detailed in the Strategic Plan of the Authority for the period 2019–2023 adopted by its Assembly in 2018.¹³ Accordingly, the members of the Authority have established the vision of the Authority in relation to the protection of the marine environment (strategic direction 3) and the promotion and encouragement of marine scientific research in the Area (strategic direction 4). They identified the need to ensure the collaborative and transparent collection and sharing of environmental data and improve the Authority's capacity to facilitate public access to environmental information (performance indicator 3.4).¹⁴ They also highlighted the need to establish strategic alliances and partnerships to contribute effectively to enhancing and expanding the sharing of data and information (performance indicator 4.3)¹⁵ and the ability of the Authority to disseminate research results and analyses through its database (performance indicator 4.2).¹⁶ This vision is being implemented according to the High-Level Action Plan for 2019–2023 adopted by the Assembly in July 2019¹⁷ and embedded in the Action Plan for Marine Scientific

¹³ [ISBA/24/A/10](#)

¹⁴ ISBA/25/A/15, p.5.

¹⁵ ISBA/25/A/15, p.6.

¹⁶ ISBA/25/A/15, p.6.

¹⁷ ISBA/25/A/15

Research in support of the UN Decade of Ocean Science for Sustainable Development (ISA MSR Action Plan) adopted by the Assembly in December 2020¹⁸ (see Annex 1).

4. This Data Management Strategy (the Strategy) is being developed in line with the strategic research priority 5 of the ISA MSR Action Plan that aims at “promoting dissemination, exchange and sharing of scientific data and deep-sea research outputs and increasing deep-sea literacy”. The Strategy also has strong interlinkages with strategic research priority 1 (Advancing scientific knowledge and understanding of deep-sea ecosystems, including biodiversity and ecosystems functions, in the Area); strategic research priority 4 (Enhancing scientific knowledge and understanding of potential impacts of activities in the Area); and strategic research priority 6 (Strengthening deep-sea scientific capacity of Authority members, in particular developing States).

5. The Legal and Technical Commission (the Commission) resumed discussion of the Strategy in March 2023, following election of the new membership of the Commission for the period 2023-2027. The Commission acknowledged the progress made in improving the management of data by the Secretariat. It also considered possible objectives of a draft Data Management Strategy for 2023-2028.¹⁹ The Commission advised that it should prioritise not only short-term strategic directives related to data quality, quantity, and accessibility in DeepData but also include strategic medium-term to long-term objectives for data management in support of the future needs of the Authority.

6. This document presents the Data Management Strategy for the period 2023-2028, building on the deliverables associated with the development of DeepData as requested by the Commission in 2016.²⁰ The document also outlines five strategic components of data management within the Authority, as well as the objectives of the Strategy for 2023-2028.

II. Background

7. The collection of data and information about the exploration of the deep seabed was initiated before the Authority was even established. From the 1980s to the 2000s, the six registered pioneer investors collected sampling data which was submitted to the Preparatory Commission for the Authority and the International Tribunal for the Law of the Sea.²¹ With the establishment of the Authority, the data provided by the pioneer investors was then placed in the custody of the Authority in an early database called POLYDAT.²²

8. The Secretariat began to assemble and digitize these data into the first Central Data Repository (CDR) from 2000, as a successor to POLYDAT. This involved processing valuable historical data into a common format in a digital database, which marked an important milestone

¹⁸ ISBA/26/A/17

¹⁹ ISBA/28/C/5

²⁰ ISBA/22/LTC/15

²¹ LOS/PCN/153 (Vol 1)

²² ISBA/5/C/6

in improving the accessibility and quality of deep-sea data at the global level. Discussions on collecting and centralising all data and information on marine mineral resources began soon after the signature of the first exploration contractors for polymetallic nodules in 2001. In 2002, the Commission formally identified this work as important. According to the Commission's request, the CDR was expanded to include information on all mineral resources under exploration.²³

9. In 2012, the Secretariat convened an informal meeting with representatives of all current contractors to update the Authority's data management requirements.²⁴ The Commission took note of a programme of work that was agreed upon at this meeting, which included redesigning the database to accommodate new environmental data formats and other activities to improve the data quality. The Commission also noted the lack of budgetary provision for these activities and highlighted the need for additional resources to undertake the work. The following year, the Commission emphasised that the management of data is fundamental to the credibility of the Authority and that there was a need for a clear data management strategy to ensure adequate quality and displays of data for appropriate assessments.²⁵

10. In 2015, the Commission took note of the preliminary work initiated by the Secretariat to improve the management of data by the Authority and emphasised the need for adequate principles, objectives and processes to describe the use of the data by the Authority. The Commission requested the Secretariat to develop a data management plan and make an assessment of the financial implications of its implementation.²⁶ The Commission stressed that the resources of the Secretariat should be allocated giving priority to the implementation of the plan, which has also been identified as a high priority action in the context of the development of the exploitation regulations. The Commission also issued recommendations for the guidance of contractors on the content, format and structure of annual reports.²⁷ This included reporting templates²⁸ required for contractors' use in their data submission from 2016, which significantly contributed to the data standardization and quality.

11. Subsequently, in February 2016, the Commission formed a working group, which conducted a preliminary review of the data and information management arrangement and the existing hardware and infrastructure supporting that arrangement. The Commission continued to work on this topic during the 22nd session and considered the data management plan proposed by the Secretariat.²⁹ With strong support for this project, the Commission provided a recommendation to ensure the availability of resources, including two new staff positions on this work. This two-year plan aimed to deliver a fit-for-purpose data management system for the Authority.

12. Starting in August 2016, the Secretariat began to migrate all historical data provided by contractors into a digital format and geo-referenced data. By August 2017, approximately 1.8

²³ ISBA/9/LTC/3

²⁴ ISBA/18/C/20

²⁵ ISBA/19/C/14

²⁶ ISBA/21/C/16, para.36

²⁷ ISBA/21/LTC/15

²⁸ <https://www.isa.org.jm/exploration-contracts/reporting-templates/>

²⁹ ISBA/22/LTC/15; ISBA/22/C/17

million data points from areas in which plans of work have been approved have been migrated, in accordance with international standards, to ensure their compatibility with other databases worldwide.³⁰

13. In parallel, work has been progressed by the Secretariat to address all recommendations of the Commission, including to: i) updating policies and procedures; ii) streamlining workflows; iii) generating a database model; iv) selecting and procuring database and mapping software; and v) updating the hardware and infrastructure upon which the information management system would be based. These were successfully delivered during the development of DeepData and encompassed the entire data lifecycle within the Authority. This work also described data governance initiatives undertaken to ensure data standardisation, storage, and accessibility for users, while maintaining appropriate levels of confidentiality and security.

14. In October 2018 and March 2019, two beta versions of the database were launched for the Commission before the official public launch in July 2019³¹ under the name “DeepData”. In July 2019, the Authority launched its DeepData database during its 25th anniversary celebration of the Authority.³² In September 2020, the Secretariat hosted a workshop on the data management strategy with a focus on DeepData,³³ which aimed at facilitating effective submission and referencing of data, creating a mechanism to review DeepData, and informing the Commission, contractors, and the public of the use of DeepData.

15. Launching the DeepData website³⁴ and database in 2019 marked a significant achievement of the Authority’s efforts to improve data management, according to the initial recommendations of the Commission³⁵ as described above. DeepData contains environmental data, which are publicly available, including biological, physical, and geochemical parameters of the marine ecosystems from the seafloor to the ocean surface. It also contains maps, photographs, videos, graphics, and relevant publications in peer-reviewed journals from ISA contractors. It has since served as an effective tool to facilitate the sharing of environmental data in an open and transparent manner, thereby advancing the scientific understanding of the Area.

16. In 2019, the Commission also initiated a further review of the reporting templates for contractors,³⁶ and the revised reporting templates were completed in 2021.³⁷ They support the extraction, transformation and loading procedures associated with inserting new data into the primary data repository of the Authority. The revised templates also saw the implementation of several quality assurance and control mechanisms to improve the overall quality of data submitted to the Authority.

³⁰ ISBA/23/C/13, para.29

³¹ ISBA/26/C/12, para.35

³² ISBA/25/C/19/Add.1

³³ <https://www.isa.org.jm/events/workshop-on-deepdata-focusing-on-data-management-strategy>

³⁴ <https://data.isa.org.jm/isa/map/>

³⁵ ISBA/22/LTC/15

³⁶ ISBA/26/C/12/Add.1

³⁷ ISBA/LTC/15/Corr.1

17. The Authority's efforts towards advancing data and information sharing through DeepData were further promoted through strategic partnerships with other UN agencies and scientific organisations. For instance, the Secretariat launched the AREA2030 Initiative in partnership with the International Hydrographic Organization (IHO) to facilitate the high-resolution mapping of the Area by 2030. The Authority is also the first UN organisation to serve as a node for the Ocean Biodiversity Information System (OBIS), sharing its biodiversity data with the OBIS network through DeepData.

18. Today, the Authority is the custodian of over ten terabytes of data through DeepData, representing the largest repository of seabed data in the world. Its data inventory has contributed significantly to the management of the Area. Specifically, these data provided a scientific basis for developing the regional environmental management plan (REMP) for the Area in the Clarion-Clipperton Zone (CCZ), protecting over 1.9 million square kilometres of the ocean from future exploitation activities.³⁸ Various data management activities continue to be undertaken to provide a solid foundation for data management within the Authority, including metadata management, data governance, migration, integration, and quality control.

III. Rationale

19. Since the first scientific exploration of the world's deep sea by the Challenger Expedition in 1872, many activities have been carried out to increase our understanding of the deep sea. In comparison to other marine scientific research, conducting deep-sea research requires more technological, financial, and human resources, due to the nature of the deep-sea environment. However, substantial investments to undertake exploration in the Area were made and led by industries since the 1970s, following the identification of seabed minerals as potential resources. Despite the apparent value of these efforts, the absence of a regulatory framework in the Area prior to the establishment of the Authority made it difficult to enable global coordination of deep-seabed research and to ensure the collected data are appropriately managed, used and shared for the benefit of mankind as a whole.

20. The specific mandates assigned to the Authority by the Convention place the Authority in the unique position to encourage, facilitate, and ensure sharing of data collected from activities in the Area, while promoting coordination of deep-sea research efforts by the industry, the scientific community, and governments at the international level. Under the current regulatory framework, contractors are required to submit the data collected from their exploration activities to the Authority. Since 2001, contractors have invested nearly 1.6 billion USD in exploration activities in the Area, of which over 260 million USD was spent to conduct environmental studies. This represents the world's largest research efforts in the seabed, both at temporal and spatial scales, consistently generating data across various disciplines that have been increasing our understanding of the seabed at an unprecedented rate.

³⁸ https://www.isa.org.jm/wp-content/uploads/2022/06/ISBA_26_C_43-2110787E.pdf

21. In accordance with the current regulatory framework of the Authority, environmental data collected from exploration activities are made publicly available. These data have significant value for marine scientific research, given the extent of resources and expertise required to conduct research in the deep sea. For instance, DeepData provides curated oceanographic data on deep-sea currents, which rarely exist, especially as public data, yet are highly demanded to understand the global marine system better. With over 60,000 data records on species occurrence alone, the Authority also contributes to advancing the scientific understanding of deep-sea biodiversity, including its Sustainable Seabed Knowledge Initiative (SSKI).³⁹

22. Ensuring data quality is a prerequisite for robust scientific research. Identifying, tracing, and standardising data, while avoiding duplications, have been long-standing challenges in the field of data management. These are arguably the most apparent in managing deep-sea biodiversity data where most species are undescribed. Therefore, applying adequate quality control and assurance measures from the submission of data to uploading into the database is essential in minimising the risks of data inaccuracy. Sustainability of the Authority's efforts in collecting, processing, and sharing data will be crucial to ensuring continuous advancement in deep-sea science and effective governance in the Area informed by the best available science.

23. While only the environmental data are openly accessible, all data are used to inform the work of the Authority for effective management of activities in the Area. This includes decisions related to ensuring effective monitoring and compliance of contractors' activities and protection of the marine environment. Different tools to support data analyses can facilitate interpretation and application of data, while customised data products for priority needs of the Authority can ensure that key discussions are readily informed by the best available data and information. This will also facilitate the use of the data to support other discussions on the protection and sustainable use of areas beyond national jurisdiction, maximising the value of the data for the benefit of all.

24. Data and information provide the ability to adequately identify problems, adapt to changes, and develop solutions. Increased access to credible data and enabling stakeholders, especially developing States, to manage and use the available data are particularly important in ensuring multilateral approach to managing the global commons. Data management capacity of developing States will be more important when future potential exploitation activities begin to generate a large volume of new data, in order to ensure their effective participation in decision-making. Increase in the global capacity on managing, sharing, and using data and information will support equal participation in decision-making and development of innovative solutions based on scientific evidence.

25. Since the establishment of the Authority, there have been growing interests in mineral resources development, scientific and technological advancements, and progress towards developing a regulatory framework for effective management of the Area under the Authority. This has led to an increase in the amount of data being collected and submitted to the Authority and wider use of such data, which calls for additional efforts to ensure effective data management. It is vital for the Authority to continue expanding its capacity as the custodian of a global deep-

³⁹ <https://www.isa.org.jm/sski/>

sea database and leverage the best practices. Successful implementation of this Strategy will enable the Authority to effectively manage data in its full life cycle from collection, curation, and standardization of data to sharing and use of data for establishing evidence-based governance in the Area, and further, the global oceans.

26. The importance of ocean data has also been increasingly highlighted at the international level. In 2017, the United Nations General Assembly⁴⁰ proclaimed the Decade of Ocean Science 2021-2030 to stimulate interdisciplinary ocean research and promote the generation and dissemination of data, information, and knowledge to support the 2030 Agenda for Sustainable Development.⁴¹ Through the Lisbon Declaration of the 2022 UN Ocean Conference, governments committed to improving the timely sharing and dissemination of data and knowledge, including by making data widely accessible through open-access databases, standardising data, ensuring interoperability between databases, and synthesising data into information for policy- and decision-makers, and support capacity building in developing countries to improve data collection and analysis.⁴²

27. In this context, this Strategy advances the results of previous ISA efforts in data management referred to in para. 10 above primarily focused on creating and operationalising a database successfully achieved with the launching of DeepData. By scaling up the scope and objectives of the initial work done, the present Strategy will further enable the Authority to use, manage, and share its data for delivering its mandates to ensure the protection and sustainable use of the Area.

IV. Objectives

28. After nearly thirty years of existence and activities, the Authority has proven that it is uniquely positioned to regulate and organise activities in the Area, including through the promotion and encouragement of marine scientific research, which directly contributes to promoting the orderly, safe and responsible management and development of the resources of the Area for the benefit of mankind as a whole. This involves ensuring the effective protection of the marine environment in accordance with sound principles of conservation and contributing to agreed international objectives and principles, including the Sustainable Development Goals.

To this end, the Data Management Strategy 2023-2028 aims at ensuring that:

- i) data is stored most effectively and efficiently leveraging new technologies;
- ii) data is Findable, Accessible, Interoperable, and Reusable (FAIR) for both humans and machines;
- iii) data remains accurate, current, and relevant for the protection and sustainable use of the Area; and

⁴⁰ A/RES/72/73 UNGA resolution

⁴¹ https://decada.ciencianomar.mctic.gov.br/wp-content/uploads/2021/01/IP_v2_eng.pdf

⁴² https://sdgs.un.org/sites/default/files/2022-06/UNOC_political_declaration_final.pdf

iv) effective data governance procedures are in place for safe and sustainable use of the Authority's data.

The Strategy aims at maximising the value of data management for the Authority and unlocking its full potential for the implementation of its Strategic Plan. The data-driven services identified should enhance data availability, accessibility, and use for scientific research and inform decision-making for the protection and sustainable use of areas beyond national jurisdiction. Further promoting marine scientific research through the use of the best available deep-sea data, will enable the Authority and evidence-based governance of the Area for the benefit of humankind. Specific attention will be given to ensuring the participation of developing countries and delivering results which enhance their capacities for data management and its application to marine research.

VI. Strategic Components

The initial efforts on data management by the Authority⁴³ primarily focused on data storage. They resulted in the development of the DeepData database and website, as well as associated data tools and products. To further support the Authority's comprehensive data management, this Strategy advances the work to encompass all data flow steps, from data submission and uploading to sharing and use.

The Strategy comprises five strategic components that underpin its implementation:

1. Data identification and standardization.

The Strategy will develop tools and procedures to ensure proper identification and representation of all data, including defining and standardising file names, formats, and vocabularies. It will also ensure that all datasets are rich in metadata and uniquely identifiable with a persistent identifier. This will support the identification of data gaps and improvement in data quality for strengthening the scientific basis to sustainable ocean governance. Successful delivery of this component will effectively increase the utility and usability of the data by the organs of the Authority for decision-making, as well as the scientific community for advanced research activities.

2. Data storage

The Strategy will further enhance DeepData's data storage capacity, building on the achievements made during the development of DeepData. Data storage incorporates the 'Findable, Accessible, Interoperable and Reusable' (FAIR) data principles, which serve as the guiding principles for the open sharing of non-confidential data. The Strategy aims to increase the FAIR data maturity⁴⁴ ratings of DeepData. Another key result of this component will be the

⁴³ ISBA/22/LTC/15

⁴⁴ <https://storage.googleapis.com/jnl-up-j-dsj-files/journals/1/articles/1241/submission/proof/1241-1-8492-2-10-20201027.pdf>

expansion of DeepData from a data repository to a database with modules specific to various scientific disciplines related to the Area.

3. Data interoperability

One of the key priorities of the Strategy is to ensure data interoperability to enable their effective sharing through different databases. This will allow the data to be used for various purposes by different users beyond the data's original purpose. To this end, the Strategy aims to enhance the interoperability and reusability of data for both humans and machines. This includes using broadly applicable languages to ensure machine readability without needing specialised or ad-hoc algorithms, translators, or mappings. The Strategy will continue facilitating interconnection between DeepData and other databases for broader data sharing and use by further developing the application programming interface.

4. Data processing.

This component involves integrating data from diverse sources into DeepData to ensure the data are relatable in the database and easily identified by users. Data processing involves several steps, from preparing and transforming raw data to making the data ready to use. The Authority's reporting templates play a key role in facilitating the processing of data, which also includes automated quality assurance and control mechanisms to ensure that the data submitted by contractors can be adequately added to DeepData ready for various uses.

5. Data governance.

Establishing effective data governance through developing and communicating guidelines on the use of the data is critical to the success of the Strategy. Such guidelines should address various aspects, such as data accuracy, use, and manipulation. The Strategy will develop mandatory licensing and referencing guidelines for using and sharing data with the Authority.

VII. Expected outcomes and outputs

29. The successful implementation of the Strategy will contribute to the following outcomes:

- Decision-making processes of the Authority are informed by best available data and information.
- Marine scientific knowledge of the Area and its resources are advanced through increased quality and quantity of deep-sea data.
- Capacity of developing States to use the best available deep-sea data is strengthened.
- Capacity of the Authority for data management is further strengthened.

The expected outputs that contribute to achieving the outcomes above are outlined below.

Outcome 1: The best available data and information effectively inform the decision-making of the Authority.

DeepData hosts a vast range of data that inform the decision-making of the Authority on its management of activities in the Area. For instance, the data are used to develop, implement, and review REMPs, which are critical to ensuring the protection of the marine environment in the Area. Contractors' data submissions are used to monitor and review their activities to ensure their accountability and compliance with the obligations set out by the regulatory framework of the Authority. The data are also used to support contractors in preparing for environmental impact assessment reports, especially to provide evidence of the temporal monitoring of environmental variables in their contract areas. To this end, various data analysis and visualisation tools further enable decision-makers and other stakeholders to adequately understand and interpret the data, facilitating the integration of scientific evidence into the decision-making process.

Output 1.1: DeepData effectively supports the scientific approach to developing, implementing, and reviewing REMPs.

Output 1.2: DeepData effectively supports compliance, monitoring, and review of the ISA contractors' activities.

Output 1.3: New data visualisation and analysis tools are created to facilitate using and understanding of the data shared through DeepData.

Outcome 2: Marine scientific research is advanced by increased quality and quantity of data in DeepData.

Environmental data stored in DeepData are publicly available for use by anyone around the world. Given various challenges associated with collecting data and information in the deep ocean, especially in areas beyond national jurisdiction, DeepData is crucial in facilitating deep-sea research, which contributes to advancing our understanding of the ocean, especially the Area. By increasing the quality and quantity of the data, the Strategy will further advance marine scientific research, which will also support establishing a robust scientific basis for successfully implementing the Authority's mandates. In so doing, the Strategy responds to the growing global demands and calls for increased scientific understanding of the ocean, especially to ensure the protection and sustainable use of marine resources, as well as to address other global environmental issues, including climate change and marine pollution. The Strategy will also foster new and strengthen existing partnerships to enhance data sharing, increase access to data, and catalyse novel research.

Output 2.1: The quality and standardisation of data are enhanced by the effective use of the revised reporting templates⁴⁵ by the ISA contractors.

Output 2.2: Metadata is collated and integrated into DeepData.

Output 2.3: Historical data is integrated into DeepData.

⁴⁵ ISBA/21/LTC/15/Corr.1

Output 2.4: The interoperability of DeepData is improved and sharing of data enhanced.

Outcome 3: Capacity of developing States to use the best available deep-sea data is strengthened.

One of the key objectives of the Strategy is to promote access to data and information related to the deep sea, in support of the Authority's duties set out by UNCLOS. This aligns with the Authority's commitment to increasing global deep-sea literacy and building capacity related to activities in the Area, particularly for developing States that may not have the resources or expertise to collect and analyse the data. Acknowledging the importance of capacity-building to ensure that all stakeholders can contribute to the sustainable management of the Area, the Strategy will undertake activities designed to develop capacities related to data management. Relevant priority needs of developing States will be identified to provide a basis for implementing the activities. Compilation and dissemination of purpose-driven data for developing States will also be facilitated to enhance their access and understanding of data that address specific needs within those states.

Output 3.1: Priority needs of developing States related to data management are identified and inform the development of relevant capacity-building activities.

Output 3.2: Capacity-building activities dedicated to enhancing data management expertise in developing States are developed.

Output 3.3: Training materials and knowledge products on managing deep-sea data are created and disseminated.

Outcome 4: Capacity of the Authority in managing a global deep-sea database is further strengthened.

The Strategy will contribute to further advancing the Authority's capacity to manage the global database of the Area by developing tools and guidelines for improved data infrastructure. For instance, sound data governance guidelines are paramount to ensure the data are managed consistently and protected from potential misuse and misrepresentation. While environmental data are shared as open data, appropriate guidelines and identifiers must conform to the FAIR data principles. Successful delivery of the Strategy will see the development of rules, principles and guidelines that provide the framework to manage data within the Authority.

Output 4.1: Further developing and implementing persistent identifiers to uniquely identify datasets within the Authority.

Output 4.2: A dedicated archive of publications and knowledge products of the Authority is created.

Output 4.3: Guidelines on data governance and use are developed and disseminated.

VIII. Implementation

30. The Strategy includes short-term and long-term outputs, which will be operationalised through annual work plans. The Secretariat will elaborate the plans in collaboration with the Commission. Details on the timeline for each output can be found in Annex 1. The strategy's implementation and monitoring will be the Secretariat's responsibility, which will work in close collaboration with the Commission. It should be noted that some activities are contingent on the adoption of the exploitation regulations.

31. The Secretariat is committed to collaborating with all stakeholders to establish a global deep-sea database that is comprehensive, reliable, and accessible to all. The Authority remains committed to continuous improvement and aims to implement client-oriented tools.

32. Implementing the Strategy and addressing its priorities will continue to require sustained budgetary support through the regular budget of the Authority. The data-driven services produced by the Strategy should provide the opportunity to design a business model based on a cost-recovery basis for the services benefiting exploration and, eventually, exploitation activities. In parallel, dedicated efforts will be engaged by the Secretariat to secure extra-budgetary resources and in-kind contributions from interested stakeholders and partners.

33. A technical expert group will be formed and engaged on an ad-hoc basis to provide technical advice on specific issues to support the Secretariat in implementing the Strategy.