# An Assessment of the Unified Data-No-Data Single Layer Proof of Concept

October 2022

Prepared in Collaboration with the GEBCO TSCOM

Opportunistic Mapping Resources Working Group

Prepared by:

Earth Analytic, Inc.

Edits and additions by TSCOM OMR WG

**Summary**

This report is presented on behalf of the GEBCO TSCOM Opportunistic Mapping Resources Working Group (OMR WG) in response to the following problem statement:

*“There is no single planning layer that definitively shows seabed mapping coverage and availability. Evaluating areas left to be mapped for proposal presentations, funding requests, planning, and at-sea opportunistic mapping involves finding and navigating multiple data layers, a task that is overly complex for the majority of users and cumbersome in practice for proficient data users.”*.

Drawing on the content presented in the MACHC Web Application, built by the Seabed 2030 Atlantic/Indian Regional Center for MACHC/IOCARIBE, we conducted a comprehensive assessment of available web services for the MACHC region that identified constraints and opportunities related to the project objective. It is the conclusion of this study that at the present time, due to the nature of the available published web services available for the MACHC region (or more specifically, the nature of the subset of web services presented in the MACHC application), it is not possible to generate a single presence/absence data layer without physically downloading all of the underlying data referenced in the application. As such, of the proposed deliverables described in our proof of concept proposal, only a summary report is provided.

This brief report outlines the limitations of the existing web services included in the MACHC application and provides a set of recommendations for strategies and methods that can be used to achieve the key objectives identified in the proposal.

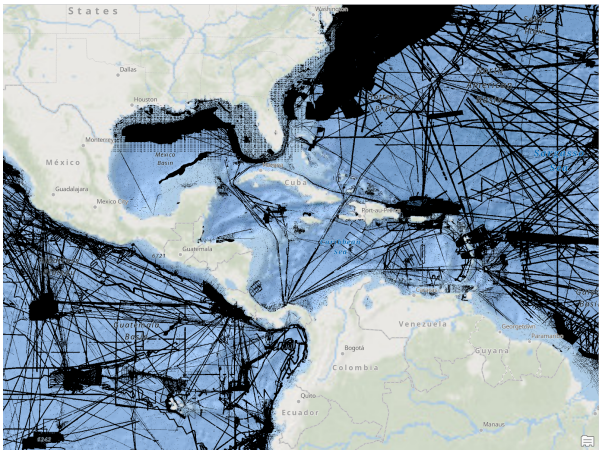
**The Problem**

There is not a single mapping layer that definitively shows seabed mapping coverage and availability. As the GEBCO website states, GEBCO aims to provide the most authoritative,

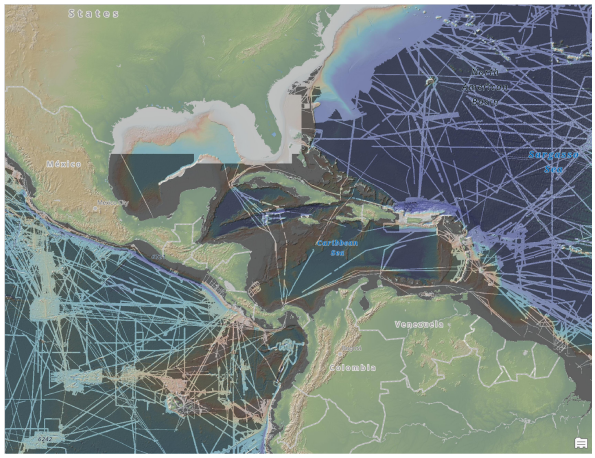
publically available bathymetry data sets for the world's oceans; based on this, one would could presume that the GEBCO grid aims to also be the most complete and as close to *the* definitive seabed mapping coverage layer as possible. But for many reasons, this is often not the case.

- Some historic data sets that are available in other public databases may not have been added to the GEBCO grid yet, or may never be due to restrictions on or desires of the data holders. Those same data sets may be available in other publically available archives or syntheses.
- New data (either newly acquired or newly available) is made publically available at different times. It may be available in a national archive first, then make its way to syntheses such as GEBCO.
- The GEBCO grid is updated yearly, generally in June. Different databases and syntheses are on different release schedules, and bound to have more or different data than those released earlier.

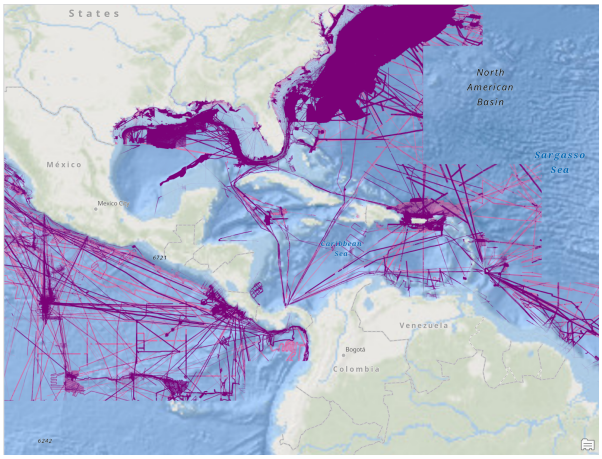
The series of images below shows four commonly used web services for evaluating what has been mapped – the *GEBCO TID measured data only* Web Map Services layer, the *Global Multi-Resolution Topography synthesis Masked* Web Map Services layer, the *NOAA Bathymetry Gap Analysis for the U.S. and MACHC region* Map Image layer, and the *NOAA NCEI Multibeam Bathymetry Mosaic: Shaded Relief Visualization* Image Service Layer.



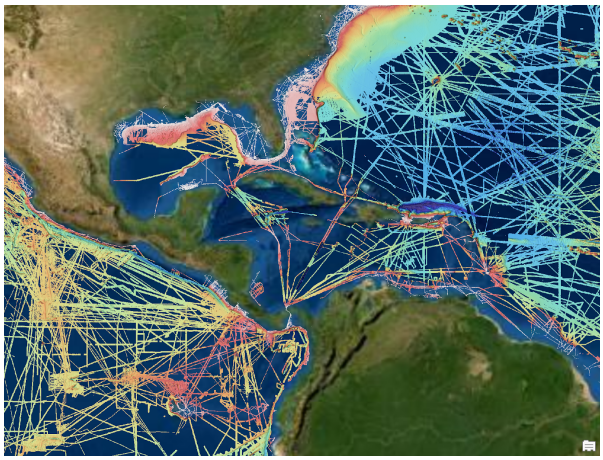

*GEBCO 2022 WMS TID layer, measured depths only (includes sources other than multibeam), updated June/July 2022 (https://www.gebco.net/data_and_products/gebco_web_services/).*




*GMRT v4 WMS Masked Bathymetry layer (highlights area of high resolution data), updated January 2022 (https://www.gmrt.org/services/index.php).*




*NOAA Bathymetry Gap Analysis Map Image Layers for U.S. (https://noaa.maps.arcgis.com/home/item.html?id=4d7d925fc96d47d9ace970dd5040df0a) and MACHC region (generated specifically for MACHC efforts, https://gis.ngdc.noaa.gov/arcgis/rest/services/bathy_gap_analysis_machc/MapServer).*



*NOAA NCEI Multibeam Bathymetry Mosaic: Shaded Relief Visualization Image Service Layer, updated May 2022 (https://noaa.maps.arcgis.com/home/item.html?id=7c7dd79ba4dc4385be199e923d1b7c94).*

*Figure: Four commonly used web services.*

When visualized together, you can clearly see that there are differences in mapping coverage:

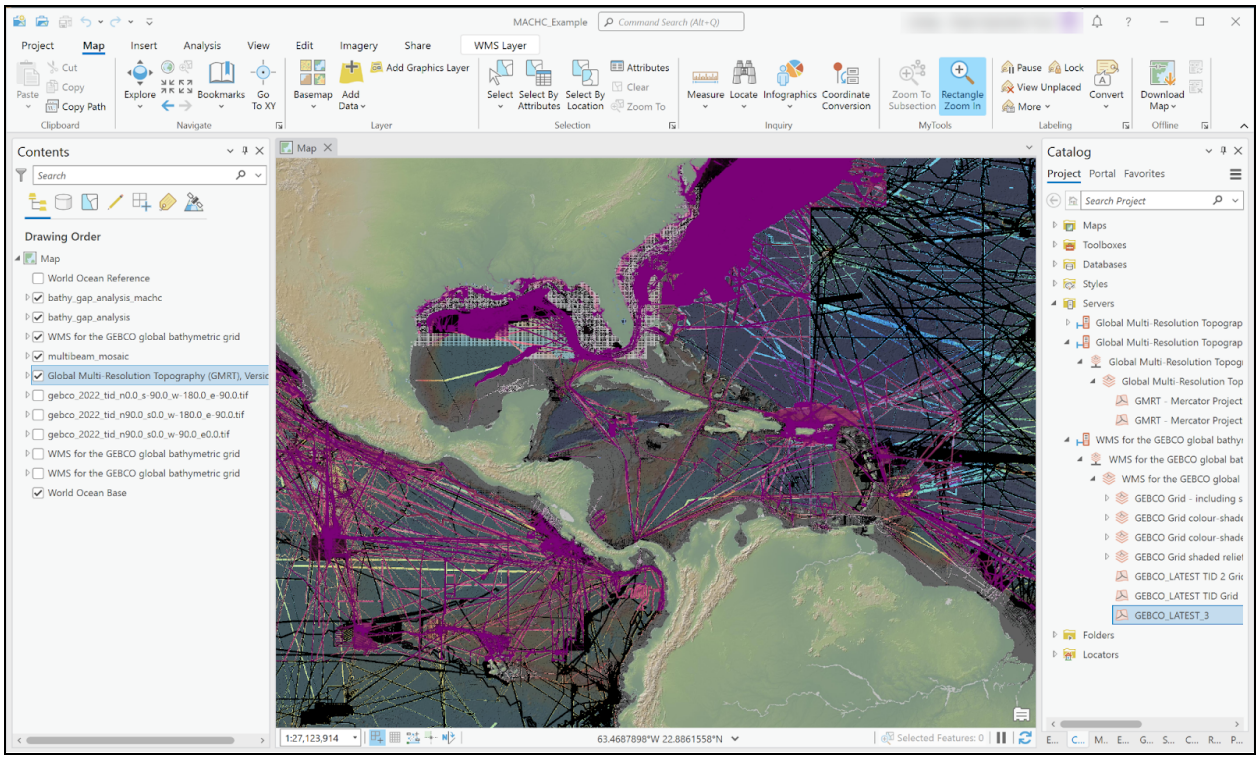



*Figure: The four image services shown together in ArcGIS Pro, as might be done when planning a mapping campaign.*

Some of this difference is due to differences in the product resolution – for example, the GEBCO web service is based on the 15 arc-second (~400 m) resolution GEBCO grid and TID, whereas the NOAA Bathymetry Gap Analysis is based on 100 m grid resolution. Some is due to product purpose – the NOAA Bathymetry Gap Analysis is focused on U.S. waters (and in this specific case, a MACHC-specific gap analysis done by request), whereas the other web services are global. But some of it is also due to the reasons suggested previously – different release dates and different data holdings that may or may not ever be 'resolved' with all data ending up in the GEBCO grid.

When one takes into account all (or even most) of the data layers for a specific region that somehow describe what has been mapped, the results are even more visually overwhelming, overly complex, and often disparate. The example below is from the MACHC Web Application, built by the Seabed 2030 Atlantic/Indian Regional Center for MACHC/IOCARIBE (https://iocaribe.ioc-unesco.org/en/SBE2030):

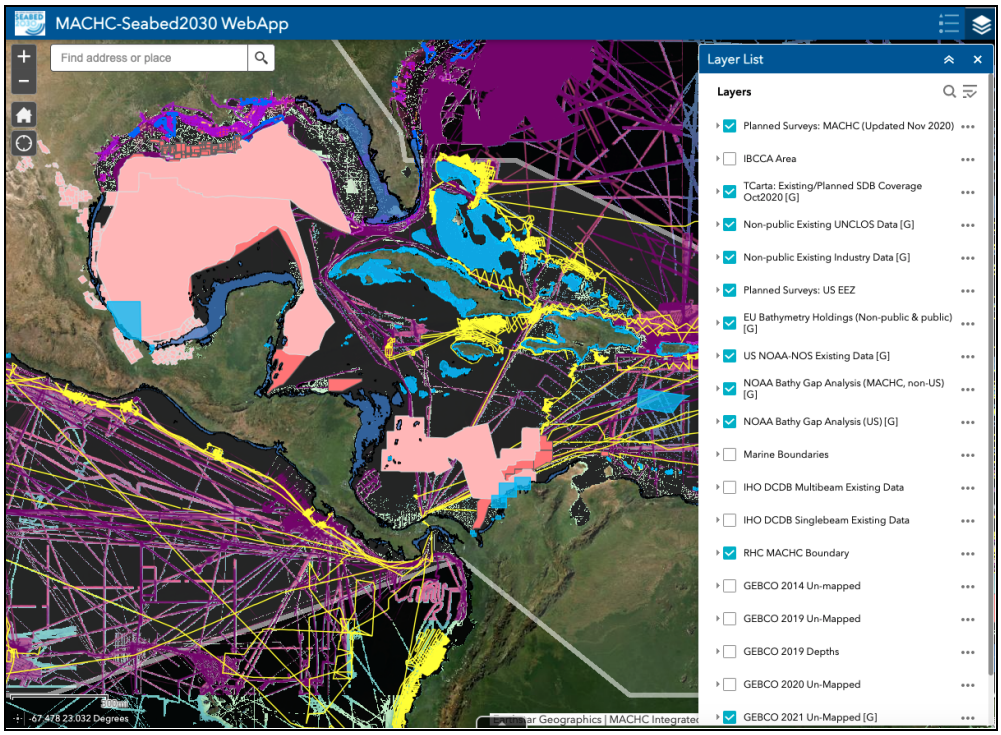



*Figure: An example of pulling in and navigating multiple data layers to estimate true coverage in the MACHC region.*

It should be noted that the MACHC Web Application is no more than a viewer, corralling a slightly customized subset of publicly existing, and in most cases authoritative, web services for a specific region.

These examples are meant to show that evaluating areas left to be mapped for proposal presentations, funding requests, planning, and at-sea opportunistic mapping involves finding and navigating multiple data layers, a task that is overly complex for the majority of users and cumbersome in practice for proficient data users.

**Proposed Solution**

When the TSCOM OMR WG first met in May of 2021 to discuss community needs for opportunistic mapping tools and how best to utilize the funding the GEBCO Guiding Committee had allocated for the purpose of generating these tools, it became apparent that this problem – no definitive single source for seabed coverage, and overly complex data mining and navigation to discover what has been mapped – needed to be solved first, and indeed we felt it to be the

number one priority for furthering the goals of GEBCO and Seabed2030. Our conclusion was that a single data/no data layer summarizing all of the available modern mapping data, to best public knowledge and with regular iterative update and improvement, was required. Discussions with the Nippon Foundation - GEBCO Seabed2030 Global Data Assembly and Coordination Centre (GDACC) in late 2022 made it clear that Seabed2030 had also identified this as a priority, with plans to allocate funding for the GDACC to tackle the problem in 2022-2023. Working with the GDACC with hopes of furthering the goal of a single definitive source showing where seabed coverage existed (or in our words, data-no data), without replicating work, we proposed utilizing some of the allocated funding to work through a unified data-no data single layer proof of concept. We suggested evaluating a select few of the many layers showing bathymetric coverage that are available as various forms of web services in order to generate a methodology and supporting tools and templates for the semi-automated identification and extraction of best-available bathymetry data. We were aiming first for a single metadata-rich web service layer, with the ultimate goal of enabling the automated extraction of authoritative data with metadata for a given area of interest, based on user-specified rules (e.g. which datasets to include). For this initial proof of concept we decided to focus on layers intersecting a specific region, in order to narrow the number of available services and make this case study/proof of concept feasible within a limited time frame. We chose the Meso American and Caribbean Sea Hydrographic Commission (MACHC) region; this region was selected as it had the required overlap of multiple data sources, we had in-house expertise for many of these data sources who could provide input to the project, and because the region would likely benefit from improved access. TSCOM OMR WG members (and the Seabed2030 GDACC as available) were to provide technical input, testing, and feedback, with the proposed work will be completed by a contracted service provider (Earth Analytic, Inc). Our intended proposal deliverables included (1) a repeatable methodology for interrogating and integrating multiple web services, (2) a single data/no data web service layer integrating the layers of interest, and (3)

recommendations for metadata to include in these services that will allow for this type of integration and can enable more advanced analysis.

**Data Assessment**

In the following section, we identify a variety of web services types with varying functionality for use in a web-based tool for generating a single data/nodata status dataset on the fly (or at least at regular intervals). These are summarized below.

- Map Service Layers: Reference only; not viable for direct use in proposed Data/No Data layer.
- WMS (OGC Web Mapping Service): Do not lend themselves to web-based data extraction. Not viable for direct use in the proposed Data/No Data layer.
- WFS and WCS services: Useable; some custom scripting required; feature limits (max no. of features) and pixel limits must be sufficient to cover max extent of analysis (user-defined area of interest).
- Feature Services (include one or more Feature Layers): Useable; polygonal features from one or more feature layers can be extracted, converted to binary raster (data/nodata) and combined with other inputs in the final raster overlay function of the proposed web-based tool.
- Image Service: Recommended; easily leveraged in the proposed web-based tool.

Given the focus on generating a data/no data layer based on direct measurements only, several layers included in the app are not included in this assessment. These include: planned surveys (US EEZ and MACHC), the boundaries of the IBCCA area, and TCarta existing/planned coverage. The remaining layers are summarized below.

1. Non-public Existing UNCLOS Data: This Feature Service layer is made available by the IHO DCDB and contains polygons representing zones that have been surveyed with

direct measurement techniques, but the bathymetry is not available to the public. Despite the lack of availability to the public, these zones have been mapped and as such could be used as input to a script or desktop GIS workflow to delineate the Data/ NoData Coverage Planning Layer for authorized users. That said, given the stated objectives of the project, data for these zones might never be accessible to members of the public at large. As such, we recommend that the proposed No Data Coverage Planning Layer generation tool exposes a parameter allowing end-users to exclude or include these areas.

2. Non-public Existing Industry Data: This data layer is made available by the IHO DCDB and is a Feature Service comprised of polygons delineating areas that have been mapped with direct measurement techniques. While it is possible that these industry-related datasets might become available for public use at some time in the future by the licensee, we again suggest that these polygons are exposed as an inclusion/exclusion parameter in the proposed No Data Coverage Planning Layer tool. At runtime, the end-user can decide whether or not to check the “include as no data” option, which in turn will affect the output.
3. EU Bathymetry Holdings (Non-public & public): This layer group, made available by EMODnet and also integrated into the IHO DCDB viewer, contains a suite of polygonal and linear features delivered via a WMS (OGC Web Mapping Service). Although the title of the layer indicates that both public and non-public datasets are included, the information popups associated with the features in the service do not include whether or not they are publicly accessible. One must click on each polygon or track to get a link to the underlying report and download page. The WMS get capabilities metadata page indicates that these services are “WMS services for marine bathymetric datasets used in EMODnet HRSM and provided by SeaDataNet”, suggesting that they have been “cooked” into the EMODnet data product. Furthermore, because WMS services do not

lend themselves to web-based data extraction, this service is not viable for direct use in the proposed No Data Coverage Planning Layer Tool. Importantly, there are separate WFS and WCS services available from the EMODnet, which are potentially useful for the proposed tool development effort. All of these OGC services are described in more detail here: https://portal.emodnet-bathymetry.eu/services/. Upon deeper investigation, we determined that the WFS service does not include any data in the MACHC study area. In other words, the WMS and WFS services have different content. As such, we cannot extract polygons showing coverage from the WFS for use in the proposed tool. If the same layers presented in the WMS service were available in the WFS service, and if those features included an attribute indicating whether or not the data was publicly available (instead of a link to a report), one could programmatically leverage this resource in the proposed No Data Planning tool.

4. US NOAA-NOS Existing Data – This map service group layer, made available by NOAA, and integrated in the IHO DCDB viewer, contains four layers: surveys with BAGS, Surveys with Digital Data, Surveys without Digital Data and BAG footprints. Each of these layers provides a URL to metadata for the various surveys. The BAG footprints layer actually shows the outlines of areas covered with survey data. If these layers were delivered in a feature service, they could be used in the proposed No Data Planning tool. As it is, these layers require manual investigation to determine the source info and their representation as map images do not lend themselves to web-based analysis to delineate what is covered and what is not. It would be helpful if the layer containing polygons for surveys with digital data included an attribute indicating the sensor type. Further investigation of the REST folder containing this map service reveals the availability of a useful Image Service with the actual bathymetry data, which is not included in the MACHC application. Given that this Image Service represents source data obtained via multibeam echosounder, it could be used in the proposed No Data

Planning Tool. While there are limits on the number of pixels (total area) one can request at one time via a web request (20,000 by 20,000), these limits are relatively high, which is quite helpful. More discussion of how image services can be used in the proposed tool are discussed in the following section, but it should be noted that a secondary service or image band comprised of integer values representing the sensor type used to collect the data would provide a more direct, less resource-intensive avenue for programmatically identifying areas of no data / data.

5. NOAA Bathy Gap Analysis (MACHC, non-US) and NOAA Bathy Gap Analysis (MACHC, US) – These two Map Service layers, made available by NOAA, contain raster layers indicating the density of digital soundings present. As described on this website, “These reclassed grids show (…) what is “mapped” in and around U.S. waters are viewable as a geospatial web service (below). While “mapped” is any cell with at least one sounding, a second sub-category of “mapped” is a grid cell supported by 3 or more soundings. This second category is often referred to as “better mapped” (purple), while the first category is “minimally mapped” (pink)…This map and corresponding table represent bathymetry that was compiled in January 2022.” These are very useful web services that could be made even more useful by enabling image service capabilities, a relatively simple process that would enable programmatic delineation of no data /data in the proposed tool. Given that multiple direct-measurement data sources went into the workflow required to generate this map service, it would be helpful if each individual source type was published as a separate image service, enabling one to exclude single beam data and/or crowd-sourced bathymetry for a more refined output from the proposed data/no data tool. As it is now, all direct measurements are lumped into the output.
6. IHO DCDB Multibeam and IHO DCDB Singlebeam Existing Data – These are map service layers that show the density of soundings per unit area for each sensor type. As with many of the layers described above, the web service type is not amenable to

programmatic calculation of data coverage. If these same data were presented as image services, they could be consumed directly in the proposed No Data Planning tool. Digging a bit deeper into a related layer that is available by the IHO DCDB but is not utilized in the MACHC Web Application, the IHO DCDB/NOAA NCEI Multibeam Survey Footprints are simplified polygons footprints for multibeam bathymetric surveys archived at NCEI and are made available as a Feature Layer. By virtue of it being multibeam survey footprints, there isn't a need for source type attribution, so this layer could be utilized in the Data/No Data layer. The proposed processing tool would simply query the feature service for features of coverage intersecting the user-defined area of interest, providing an input to the combined overlay with other data/nodata image and feature services.

7. GEBCO 2014 Un-mapped, GEBCO 2019 Un-mapped, GEBCO 2019 Depths and GEBCO 2020 Depths – These group layers also contain representation of data presence and absence in a format not compatible with dynamic analysis as required for the proposed tool. That said, the inclusion of the bathy data and the sensor type as separate layers is in line with the goal of this study. Digging deeper into the REST folders for the GEBCO data, we discovered an image service for 2020 bathymetry data with depth values along with a corresponding map service for 2020 sensor types. While the image service is helpful, we cannot combine the two services to generate a simple data / no data raster as desired. Again, we recommend that an image service is provided for the sensor type coverage to support the proposed objective. This content could (and perhaps should) also be made available as raster services by Seabed 2030.

**Discussion**

Based on the data assessment presented above, we must note that the current repository of web services available for the MACHC area are not amenable to dynamic, programmatic

delineation of a no data planning layer. That said, both the ArcGIS and OGC platforms used for delivery of these services could be leveraged to provide the type of outputs required to achieve the current project objective. For each bathymetric coverage dataset, the availability of an ArcGIS image service or an OGC WCS service that provide integer values indicative of the type of sensor used to collect the data would enable the use of desktop GIS or web-based scripting/geoprocessing tools to extract the information necessary to provide a comprehensive no data planning layer. Using a sequence of queries, one could assemble a suite of temporary outputs that can in turn be combined to delineate the output dataset. First, one would query for all pixels that have a specific sensor type (where TID value is a direct measurement). Second, use that initial output as a mask to generate a surface representing where data is present or absent.

The following diagram illustrates a sample geoprocessing workflow for generating a single data/nodata raster from a set of image services with TID values as described in this report. Only two input services are shown in the graphic but the actual tool would iterate through all of the required input rasters (some generated from Feature Service, WFS and/or WCS source data) and might have slightly different filtering expressions to yield binary output rasters. An alternative approach using raster functions would follow a similar logic.

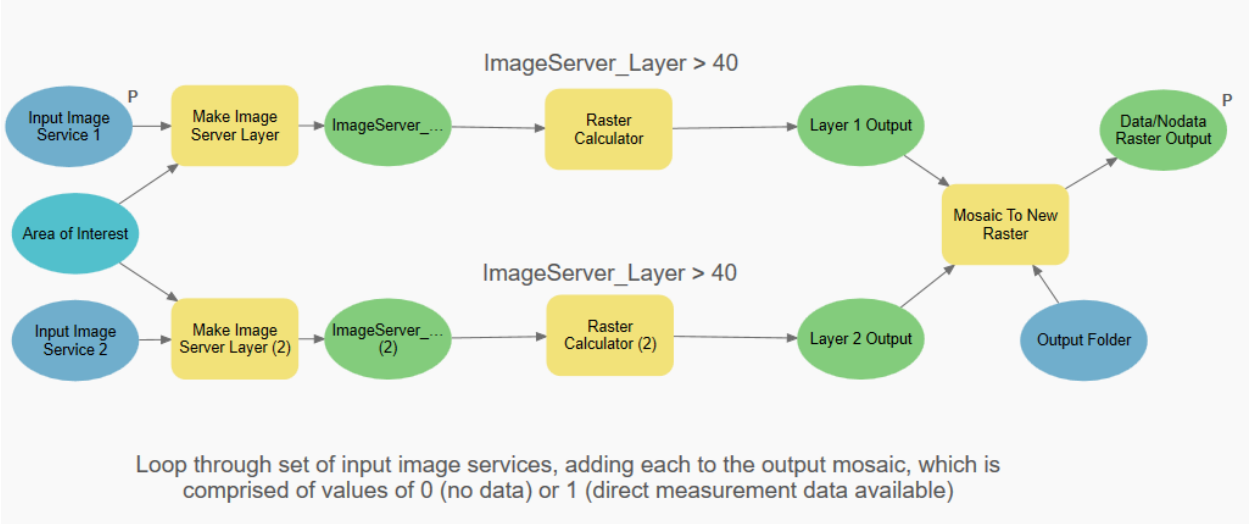



*Figure: ArcGIS Geoprocessing Model Example*

One important consideration for these services is the limit set on the maximum request size (number of pixels). Given that we only need access to integer values, these request limits could be set at a relatively high level because the size of the output is significantly smaller than would be required for floating point data. Another consideration is the potential use of multiband images in image (or WCS) services. By including the raw elevation values in one band and the sensor type in a second band, one could write a simple query that would select only those pixels with direct measurement source type to yield an output showing data / no data. Furthermore, this type of image service would enable a sequential query to obtain an output bathymetry raster with no data pixels in all zones that are not based on direct measurement types. Multiband image services such as these could be used to dynamically present results based on raster functions (on-the fly, lightweight calculations that take advantage of the pixels on the screen at runtime) or static geoprocessing functions that extract full resolution outputs as downloadable datasets.

For organizations that already have the resources to publish ArcGIS image services or open source WCS layers, the level of effort and required resources to publish and maintain the sensor-type information with the capability to query and extract data values for large areas is not

substantial. If we can encourage and support these organizations in this effort, perhaps even providing direct technical assistance or resources during updates, the result would be significantly positive. For organizations that do not have the time or resources to take on this extra step, perhaps a third-party entity already involved in hosting GEBCO data (e.g. CCOM) could take on this effort, publishing an independent image service.

To complement the goal of creating an authoritative data/no data delineation tool using image services, we suggest that when organizations publish an update, they run a simple analysis that compares the previous data/nodata surface against the current (updated) one, yielding a raster (ideally published as an image service or feature service) that shows areas of change. This, in turn, could be used to identify and access the new source data. Again, this task is something a third-party might take on as part of the effort to publish an aggregated result/solution.

To facilitate and automate this combined workflow in a web-based tool, an expanded version of the model shown above could be published directly as a “Geoprocessing Service” (and/or python script), which in turn could be embedded in other web applications. The tool would take a user-defined area of interest as an input, generating a simple binary raster (data/nodata) for that extent as a georeferenced image file (i.e., geotiff). Importantly, detailed processing metadata listing the steps, tools, and datasets used for each tool run (usage episode) are captured and provided as messages included with the raster output. On this note, the tool can be configured to point to different source web services that follow a standardized schema and metadata guideline.

Using readily available ArcGIS Online or ArcGIS Portal web apps, this type of tool could easily be added to an app builder template to provide access to the service. Alternatively, authorized users could access and run the geoprocessing service from a simple REST API request, a Python Notebook, or in ArcGIS Desktop/Pro.

While the results of this cursory review reveal limitations on the project objective, we hope the assessment and related recommendations provide a basis for collaborative discussion among all stakeholders on next steps.