Canadian Integrated Ocean Observing System

INVESTIGATIVE EVALUATION: VISUALIZATION

2017-11-30

Principal Investigator

Claude Tremblay (St. Lawrence Global Observatory)

Authors

Alexandre Brassard Desjardins (St. Lawrence Global Observatory)
Richard Dewey (Ocean Networks Canada)
Eric Guillemot (Ocean Networks Canada)
Stéphane Lapointe (St. Lawrence Global Observatory)
Dwight Owens (Ocean Networks Canada)
Benoit Pirenne (Ocean Networks Canada)
Arianne Plourde (ISMER)
Germain Sauvé (St. Lawrence Global Observatory)
Anne-Sophie Ste-Marie (St. Lawrence Global Observatory)

TABLE OF CONTENTS

SUMMARY

Canada is surrounded by three major oceans and has the world's longest coastline spanning over 244 000 km. The country's 8500 rivers and 2 million lakes represent 8.9% of its total area. Activities related to the marine and aquatic environment provide a substantial contribution to our economy. The country's world renowned scientific experts, extensive science infrastructure as well as a thriving ocean technology industry are amongst its most valuable resources. The importance and the need to better coordinate regional efforts in ocean observation has led to realization of three investigative evaluations (IE) to support the Canadian government towards the implementation of a cohesive Canadian Integrated Ocean Observing System (CIOOS). Those three evaluations are: IE Observation and Data, Cyberinfrastructure, and Visualization.

This present report focuses on the results of the Visualization IE, as well as integrating cross-cutting issues with the two other IEs. The scope of this report encompasses not only visualization general tools and products, but also, evaluating various tools to enable it, while keeping in focus the goal of meeting end-users needs. This report sought to make recommendations that will maximize data discovery trough metadata and data visualization in ways that are useful to the widest range of end users, from the scientific community to the Canadian public. A review and assessment of existing tools to search, access and display data from the ocean observing systems in Canada and abroad is thus provided, as examples of best practices and starting point from which to build upon.

The report makes recommendations for three implementation services models (low, medium and high), as well as proposes additional products that could be added in a phased approach and the associated costs and resources those products would require. Since the use of open source software is recommended, costs are mostly influenced by human resources needed to develop visualization tools, implement existing tools (e.g. data catalogue) and for maintenance and optimization.

Minimally, each Regional Association should foster a data catalogue interface and a standardized harvesting data access point to feed the CIOOS national portal. In turn, the CIOOS portal will be enabled to host a national metadata catalogue interface, as well as a sensor map interface and a geospatial extent based metadata interface to ensure data discovery from coast to coast. With increased service, the other models offer more data discovery and visualization interfaces, with more functionalities to mine through data. The three service models are followed by a phased approach implementation process, aligned with the one of IE Cyberinfrastructure, to ensure the smooth implementation of Regional Associations and the national CIOOS portal.

Along with constant search for best practices compliance and sustained funding of its operations, CIOOS tools must be proven necessary to increase reliable sources of ocean information for end-users (ocean researchers, decision-makers, public at large, etc.). As such, tracking the uses of CIOOS portal visitors will be a must, and tools to reinforce CIOOS identification in the Canadian ocean landscape will help its mission. Therefore, this report also recommends visitors tracking and survey tools, as well as a generic homepage branding to increase the sense of belonging between CIOOS and various RAs, and recognition of the network for visitors browsing from one portal to the other.

RÉSUMÉ

Le Canada est entouré de trois grands océans et possède un littoral de 244 000 km, le plus long au monde. Les 8500 rivières et les 2 millions de lacs du pays représentent 8,9% de sa superficie totale. Les activités liées au monde marin et aquatique apportent une contribution substantielle à notre économie. Les experts scientifiques du pays, de renommée internationale, les infrastructures de recherche, ainsi que l'industrie florissante des technologies marines comptent parmi ses ressources les plus précieuses. L'importance et la nécessité de mieux coordonner les efforts régionaux en matière d'observation des océans ont mené à la réalisation de trois évaluations d'investigation (EI) pour aider le gouvernement canadien à mettre en place un système intégré d'observation des océans canadiens (CIOOS). Ces trois évaluations sont: Observations et Données, Cyberinfrastructure et Visualisation.

Le présent rapport se concentre sur les résultats de l'EI Visualisation, ainsi que sur l'intégration des questions transversales avec les deux autres EI. La portée de ce rapport englobe non seulement la visualisation des outils et des produits de visualisation, mais aussi de leur évaluation, avec comme objectif de répondre efficacement aux besoins des utilisateurs finaux du CIOOS. Ce rapport visait à faire des recommandations qui maximiseront la découverte des données grâce à la visualisation des métadonnées et des données, de manière à être utile pour le plus grand nombre possible d'utilisateurs finaux. Un inventaire et une analyse des outils existants de recherche, d'accès et d'affichage des données, des systèmes d'observation des océans au Canada et à l'international, sont fournis autant à titre d'exemples de bonnes pratiques que de points de départ sur lesquels construire le CIOOS.

Ce rapport formule des recommandations pour trois modèles de services de mise en œuvre (faible, moyen et élevé), des produits supplémentaires, une approche d'implémentation par phase, ainsi que les coûts et ressources associés à ces produits. Étant donné que l'utilisation de logiciels libres est recommandée, les coûts sont principalement influencés par les ressources humaines nécessaires pour développer certains outils de visualisation, mettre en place des outils déjà existants (par exemple, le catalogue de données) et enfin, pour la maintenance et l'optimisation du système.

Au minimum, chaque Association Régionale devra mettre en place un catalogue de métadonnées et un point d'accès aux données standardisé pour alimenter le portail national CIOOS. À son tour, le portail CIOOS sera en mesure d'héberger un catalogue de métadonnées national, ainsi qu'une interface de carte de senseur et une interface de métadonnées basée sur l'étendue géospatiale. Les deux autres modèles de services offriront davantage d'interfaces de découverte et de visualisation de données, et des fonctionnalités supplémentaires plus spécifiques de découverte de données. Les trois modèles de service pourront être mis en place selon un processus en trois phases, aligné sur celui recommandé par l'EI Cyberinfrastructure.

Avec la recherche constante de meilleures pratiques et de financement soutenu de ses opérations, il sera nécessaire de démontrer l'efficacité des outils du CIOOS à devenir une source fiable d'information sur les océans pour les utilisateurs finaux, des chercheurs océaniques au grand public en passant par les décideurs publics. À ce titre, le suivi des utilisations des visiteurs du portail CIOOS sera nécessaire, et les outils pour renforcer la reconnaissance du CIOOS dans le paysage océanique canadien aideront sa mission. Par conséquent, ce rapport recommande également des outils de suivi d'utilisation par les visiteurs, ainsi qu'une image de marque générique des pages d'accueil pour augmenter le sentiment d'appartenance entre le CIOOS et les divers AR, et la reconnaissance du réseau pour les visiteurs naviguant d'un portail à l'autre.

1. INTRODUCTION

1.1. GENERAL INTRODUCTION

Canada is an ocean nation. Its extensive coastline of 244 000 km, the largest of any nation on earth, spans from the temperate North Pacific Ocean, through the Arctic, and down the Atlantic seaboard to the USA. The Gulf of St. Lawrence and Hudson's Bay, two of the earth's great inland seas, are wholly contained within Canada's land mass. The Great Lakes are another shared coastline between Canada and the USA.

About 40% of the Canadian population lives within 100 km of these coastlines (Manson, 2005). Both historically and currently, Canadians have turned to the ocean for their livelihoods and well-being. Canada's ocean economy is diverse, and includes transportation, offshore energy, marine technology, defense, tourism, conservation and fisheries. The ocean economy accounts for about $26 billion, approximately 5% of Canada's annual GDP, and provides employment to more than 315 000 workers (Fisheries and Oceans Canada, 2009). The marine environment and the Great Lakes are also critically important for Indigenous People's subsistence, social and ceremonial uses, and are the backbone of the socioeconomic well-being of Canada's coastal communities.

Advances in marine technology are providing unprecedented access to the ocean and are spawning a myriad of new economic and scientific activities. New, well-paying employment opportunities will bring many more people out to work on the ocean as this "Blue economy", or "ocean industrial revolution" accelerates (McCauley, et al., 2015). It will also add pressure to the ocean systems that provide essential ecosystemic services and that support the existing fisheries, tourism, and other sectors that are major engines of the Canadian economy. To understand and sustainably manage this development, Canada needs an ocean observing capacity that will provide integrated information needed for high-quality research as to inform policy management decisions.

The ocean drives planetary systems such as weather and water cycles, and while the environmental characteristics and fauna of the ocean may differ considerably among regions, the ocean is still an interconnected whole, as exemplified in the One Ocean concept (O'Dor, Fennel, & Berghe, 2009). What happens in one part of the global ocean can have important impacts on other, distant regions. The species on which our fisheries depend are mobile and not constrained by national borders. Interconnectivity applies to environmental threats such as oil spills, invasive species, or rising sea level, temperature, and acidity. Consequently, humans have the shared global burden to provide the ocean science needed to plot a sustainable future. One mechanism that coastal nations are addressing this challenge is by signing international agreements to collect and exchange ocean data and knowledge, and to mutually address shared problems.

In Canada, ocean science is conducted by government, academia, industry, non-governmental organizations (NGOs), and the general public through citizen science. Fisheries and Oceans Canada (DFO) has by far the greatest investment and capacity for ocean science; while DFO's science sector pursues fundamental science, it is responsible of stock assessment for best fisheries management, as well as providing advice in support of its other programs related to ocean protection, such as marine protected areas and species at risk and aquaculture. Its work also supports development of economic opportunities, such as aquaculture, and guides operations including search-and-rescue for the Coast Guard. The Canadian government is strongly committed to ocean science, as evidenced by the Mandate Letter issued by the government for the Minister of Fisheries and Oceans and the Canadian Coast Guard, which directs the Minister to:

(1) “Restore funding to support federal ocean science and monitoring programs”,

(2) “Ensure that decisions are based on science, facts, and evidence, and serve the public interest”, and

(3) “Work with the provinces, territories, Indigenous Peoples, and other stakeholders to better co-manage our three oceans.”

In parallel and supported by Canada’s national academic funding agencies, our university and college sectors also have strong capabilities in ocean science. Academics undertake a variety of research, ranging from individual investigator, narrow-focus, short-term projects to large national networks (e.g. Canadian Healthy Oceans Network, Ocean Networks Canada, Ocean Tracking Network, MEOPAR, ArcticNet) that have the supporting infrastructure to sustain interdisciplinary research and the associated data management for longer (~ 5 years) periods.

Industry, Indigenous governments, NGOs and the public undertake more limited ocean research which is generally tied to specific interests or values of their organization. Many of these programs frequently address issues in which the public has a strong interest. With the advent of user-friendly ocean observing sensors, these groups can rapidly generate large volumes of high-quality data from geographic areas of great interest.

To meet efficiently Canada’s needs in ocean science, it is essential that Canadian investigators from all sectors coordinate their data collection efforts to avoid duplication or lost opportunities, and ensure that data collected is discoverable, usable and shareable by Canadians to the benefit of all Canadians. This issue, highlighted in the Mandate Letter to the Minister of Fisheries and Oceans, was the subject of two reports commissioned by DFO and its partners (Fisheries and Oceans Canada, 2010; Ocean Science and Technology Partnership, 2011) and was a key finding of two reports issued by the Canadian Council of Academies (CCA) which examined the Canadian Ocean Science Sector (Canadian Council of Academies, 2012; Expert Panel on Canadian Ocean Science, 2013). In addition, the Canadian Council of Academies identified 40 priority ocean sciences questions for Canada. Of those, two questions specifically addressed Canada’s ocean information needs:

> *#24 How can a network of Canadian ocean observations be established, operated and maintained to identify environmental change, and its impacts?*
>
> *#25 What indicators are available to assess the state of the ocean, what is the significance of changes observed in those indicators, and what additional indicators need to be deployed?*

The Expert Panel on Canadian Ocean Science (2013) examined how Canadian ocean science research is currently structured and concluded that the country faced three primary gaps:

> **<u>Vision</u>**
> *Canada lacks a national vision and strategy for the oceans.*
>
> **<u>Coordination</u>**
> *We need to pool efforts from the local to the international scale to address our ocean science needs.*

<u>Information</u>

We lack information about the scale and scope of ocean research being carried out nationally, and on the availability and comparability of our existing research activity and of the data being generated.

Canada requires a coordinated integrated ocean observing system to meet the national ocean information needs of government, academia, industry, and the public. Such a system will directly support our international ocean commitments, and permit Canada to play a global leadership role in multidisciplinary ocean science. An ocean observing system it will help coordinate the collection of ocean data, be capable of adaptation in the face of changing needs and a changing environment, and will provide access to data currently not discoverable, especially the extensive holdings of the federal government.

The international and national context both offer favourable conditions for the establishment of a national ocean observing system in Canada. Internationally, a growing number of countries and organizations worldwide have well-established ocean observing systems. Canada's positive global reputation has us well-positioned to sustain our engagement in international efforts (UNCLOS, Espoo, CBD, OSPAR, MARPOL 73/78, GEOOS, GOOS, etc.). Nationally, the amount of information and data generated by Canada's existing ocean observing assets distributed across the country (provincial and federal ministries, research organizations, universities, Indigenous Nations, NGOs, etc.), is already considerable and provides a solid foundation for establishing regional associations within an overarching Canadian Integrated Ocean Observing System (CIOOS) to address Canada's national priorities.

Such a system will require engaging in pan-Canadian efforts to achieve shared standards and practices among the existing organizations (Wilson, Smit, & Wallace, 2016). Each operates at its own level of sustainability, maturity, scope, and funding and will require investment and support in different areas. For example, the St. Lawrence Global Observatory, established in 2005 by a network of provincial and federal department and universities, integrates multidisciplinary data from multiple partners, and in many ways, is a model for future regional associations. To this end, in 2016, Fisheries and Oceans Canada re-initiated a consultation process with stakeholder groups across the country to continue past discussions and move forward with the creation of CIOOS. In 2017, it commissioned three Investigative Evaluations (IEs) to make recommendations regarding the structure of a national observing system. The three IEs addressed issues within the topics of *Data and Observations*, *Visualization*, and *Cyberinfrastructure*.

1.2. INVESTIGATIVE EVALUATION VISUALIZATION

This present report focuses on the results of the investigative evaluations on visualization, as well as integrating cross-cutting issues with the other IEs. The scope of this report encompasses not only visualization general tools and products, but also tools to search, access and display data, while evaluating various tools to enable it. Other generic themes such as branding will also be addressed.

The data life cycle typically includes planning of infrastructure, data acquisition, quality assurance and quality control, utilization, archiving, etc. However, this cycle is only completed when data are made accessible: through both discovery and actual access. For that purpose, there is an increasing need to maximize the efficiency of data visualization in ways that are useful to the widest range of end users, from the scientific community to the Canadian public. To collect data once and reuse it as many time as possible, as well as to limit duplication of effort, or lost opportunity in oceanographic research, it is crucial that CIOOS datasets are easily discoverable to users who do not have prior knowledge that these data might already exist.

Effective data visualization capabilities are key to answering the various end-users needs. Those needs include i) discovering metadata and data through data catalogues, ii) accessing data (download or direct use), iii) using value added applications. User needs span a spectrum of purposes that include: i) integration in scientific research activities, ii) improving public safety (search and rescue, navigation tools). For example, a value-added product such as a map showcasing multiple stations (scientific buoys, tide gages or high-frequency radar), allows their associated data to be easily discoverable, by using various filters, while data download can still be available in a wide range of formats.

Visualization of oceanic and coastal data, by nature, brings a lot of challenges but innovative opportunities as well. In addition to the increasing volume of data, marine data are collected by a wide variety of instruments and institutions, from ministries to NGO's to citizen scientists, and serves a variety of disciplines. Visualization tools must then be able to adaptively present a widely differing collection of ocean core variables such as water temperature or annual fishing survey and stock evaluations. The dynamic boundaries of marine data and the 3D marine environment make the quality control more important to ensure high quality data that can empower decision-makers and inspire the audience more generally. There is an outstanding opportunity for observation systems such as CIOOS to meet these challenges and offer innovative and efficient visualization tools. For example, Marine Spatial Planning (MSP), an ecosystem-based management approach, allows the integration of a wide variety of spatial data coming from diverse applications such as fisheries and aquaculture, traffic management, marine protected areas, etc. MSP tools offer an adaptive and anticipating strategy, and its visualization products allow the diagnostic examination of both future and actual territory. MSP rests on visualization tools such as:

- Inventory and maps of ecologically and biologically important sectors,
- Inventory and maps of human activities and pressures,
- Identification of conflict zones or possible complementarities between usages

The present investigative evaluation first presents an inventory of existing visualization products for both metadata and data (tables, graphs, georeferenced, video, audio, 3D, etc.) that address the needs of ocean observation community (section 2). Following is an assessment based on a short listing from a preliminary inventory of different software suits and tools (section 3). This report also evaluates and details the end-users needs that can be expected from a CIOOS national portal and the RAs, both in terms in metadata discovery, thematic metadata visualization, data previsualization or data access and download (section 4). More importantly, this report recommends a generic CIOOS identification, including branding recommendations, formats and services, and an incremental service model approach, from low-service to high-service (section 5). These models of services are followed by a recommended phased approach, from a prototype to the optimization of all interfaces (subsection 5.4). Finally, section 6 estimates expenses for one regional association for a year (salaries, training, software, etc.), while section 7 ends this report by opening to a few potential additional products and next steps that could be considered in the CIOOS implementation.

1.3. DEFINITION OF SELECTED TERMS

Web Application	Client-server computer program that runs in a web browser. E.g.: SLGO's Marine Conditions, ocean forecasts, EMODnet's GEOVIEWER, Ocean Networks Canada OCEAN 2.0
Product	Transformation of input to create a useful artefact. A data product turns a set of data into a different form using computation. E.g.: a plot, a table in a different format, etc.
Protocol	"A set of instructions for transferring data from one computer to another over a network" (World Wide Web Consortium, 2013)
Regional Association	A regional organisation is in charge of coordinating the integration of multidisciplinary data from multiple data producers in a specific region. It is also in charge of maintaining the required cyberinfrastructure, a website with data discovery to ensure open access to these data.
Server	A computer program that manages access to a resource or service in a network.
Standard	Established norm allowing compatibility and interoperability between software, systems, platforms and devices. E.g. OPeNDAP[1], WMS
Tools	Any software, application or library enabling standard or protocol implementation. E.g.: OpenLayers, Geoserver, THREDDS
Variable	A representation of an information container, whose content will change from time to time. A variable will be characterized by metadata such as units. E.g. CIOOS Core Variables
Visualization type	Specific visualization category. E.g.: time series, maps, etc.
Web Services	A service offered by an electronic device to another electronic device, communicating with each other via the Internet using the http protocol.

[1] NASA Earth Science Data Systems. *The Data Access Protocol — DAP 2.0.* Retrieved from OpenDAP Web site: https://www.opendap.org/pdf/ESE-RFC-004v1.1.pdf

2. INVENTORY OF EXISTING DATA ACCESS AND VISUALIZATION PRODUCTS AND THEIR FUNCTIONALITIES

The implementation of a Canadian integrated ocean observing system implies two courses of actions. First, it will need to aim for best quality of its services by constant search for optimization and innovation. Second, it must rely on what already exists to build from it rather than duplicating past efforts. In this perspective, this first section of the IE Visualization report details an inventory of currently existing data access and visualization products, with an overview of their respective functionalities.

2.1. METADATA

Information used to describe a piece of data, metadata, makes it possible to facilitate its discovery and to evaluate its relevance. The user will access the metadata according to a title, summary or detailed display depending on access points (map, catalogue, API, etc.). The choice of metadata to display will depend on the chosen vocabulary scheme.

A summary display will provide sufficient insight to uniquely identify the dataset (e.g. unique identifier, data provider, title, etc.), while a detailed display will include all the formal elements of the chosen data vocabulary schema, to which may be added the extended properties. Metadata can be displayed on a map using an interactive tooltip or in a static table.

2.1.1. METADATA STANDARDS AND VOCABULARY

To facilitate the visual comparison of similar elements, the use of recognized standards, terminologies or vocabularies must be favored. To read more on metadata standards and which ones should be favored in the context of CIOOS, more information on these topics are addressed in the IE Data and Observations report.

In visualization interfaces, the label associated with a metadata field usually is the same as defined by the standard for the corresponding field. For example, the [dct: title] property of the DCAT3 schema will be represented by the *Title* label (World Wide Web Consortium, 2014).

2.1.2. SPATIOTEMPORAL BOUNDARIES

Temporal metadata makes it possible to locate data in time by displaying in textual form the numeric representation of the date and time. This property can be represented by an interval (1990 - 1995) or a specific time and date (June 2010), with various levels of time scale precision (minutes, hours, days, months, seasons, years, decades, etc.).

A date metadata representation respecting the standard ISO 8601 format will be visualized according to different level of granularity, from the less precise (AAAA) to the most precise (YYYY-MM-DDTHH: MI: SS.MSZ)

By default, the metadata display format usually respects UTC formats. Optionally, the option to override these preferences using drop-down menus or any other selection element could be implemented.

Spatial metadata makes it possible to locate data in space, according to its latitude, longitude, and elevation. Visual representation of spatial metadata can be done using a map symbiology or using textual metadata. The map identifiers could be either a point, a series of points, a polygon, a series of polygon or a centroid, while textual representation of geospatial metadata is usually done in GeoJSON (RFC 79464) format (Internet Engineering Task Force, 2016).

2.1.3. Free form text

The type of metadata whose content is unstructured and cannot be framed by a standard is usually visually represented through free form textual display. The preferences associated with the user's browser should not affect the display of the metadata. For example, a date included in a free form text description will not be interpreted nor transformed.

However, the display of the metadata can be enriched according to a standardized markup language that can be displayed in a browser (i.e. HTML) or according to a simplified markup language (i.e. LML).

2.2. Data

Ocean science data comes a multitude of sensors and observing platforms, is collected in diverse regions and climatic conditions, and serves multiples purposes from fundamental research to decision-making. This wide diversity in data sources and uses is translated into a wide diversity of data products. This subsection lists the most common ones along with their associated features and functionalities.

2.2.1. Tables

Tables are a simple yet effective way to visualize data. Tables can range from an elementary static representation of information to a full-fledged interactive web visualization tool. Numerous data types may be included in a table. Tables can be created in a number of formats following a number of public and industry standards such as .json, .xml[2], .csv; but also Excel and Matlab formats, etc.

2.2.1.1. Static tables

Simplicity can sometimes be the most efficient solution. Static tables can be visualized as a regular image or directly generated from raw data. While they do not allow for direct data manipulation, they are computationally inexpensive, easily understood by the end-user, and do not require complex implementation.

2.2.1.2. Interactive tables

Interactive tables have grown in popularity in recent years and, nowadays, constitute an integral part of web visualization. They can offer a wide variety of features such as presented in **Error! Reference source not found.** . Features increase flexibility for the end-user, but may also increase implementation complexity. It is worth noting that several of these features imply computation to re-render the table.

TABLE 1. DESCRIPTION OF INTERACTIVE TABLE FEATURES.

Feature	Description
Sorting	Order the table entries according to the values of a specific column and a specific algorithm (e.g. sort using a “City” column, in alphabetical order).

[2] World Wide Web Consortium. *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. Retrieved from W3C Web site: https://www.w3.org/TR/xml/

Filtering	Only display values that comply to a specific condition applied to a specific column (e.g. only display entries with year greater than 1970)
Paging	Display data in multiple pages, according to a given page size.
Hiding/showing column	End-user can choose which columns to be displayed.

2.2.2. 2D Graphs

Graphs have long been a popular and powerful high-level visualization tool. Seeing graphically depicted information provides insight that goes beyond the raw numbers; allowing end-users to witness trends, patterns, evolution through time and space, variable dependencies, correlation and much more.

There are multiple types of graph and each one excels at conveying a specific message. To convey relationship or correlation line graph, bar charts or even Venn diagrams are often used. Line graphs and bar charts are excessively good at comparing data within well defined parameters while Venn diagrams can quickly display relationships in a very evocative manner. To visualize the distribution of data, histogram and scatter plots can be quite useful.

2.2.2.1. Line diagram

The line diagram below, in Figure 1 shows one month of temperature record in Gascoyne Inlet, NU, by 21-meter depth, a common example of static line diagram. In Figure 2, a more sophisticated graph shows a map with superimposed vectors of surface currents direction and velocity. Usually, the map is animated, showing changes over time.

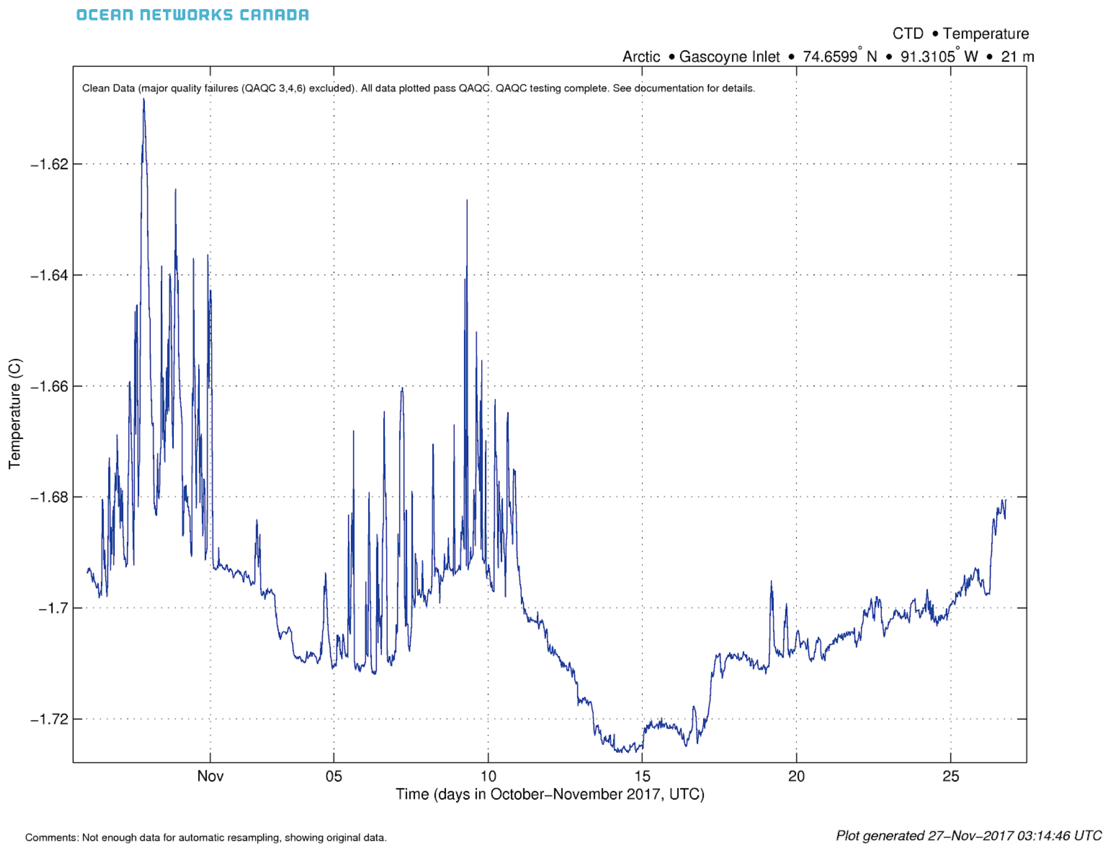


FIGURE 1. LINE DIAGRAM EXAMPLE OF RECORDED TEMPERATURES FROM OCTOBER TO NOVEMBER 2017 IN GASCOYNE INLET, NU, AT 21-METER DEPTH. COURTESY OF OCEAN NETWORKS CANADA.

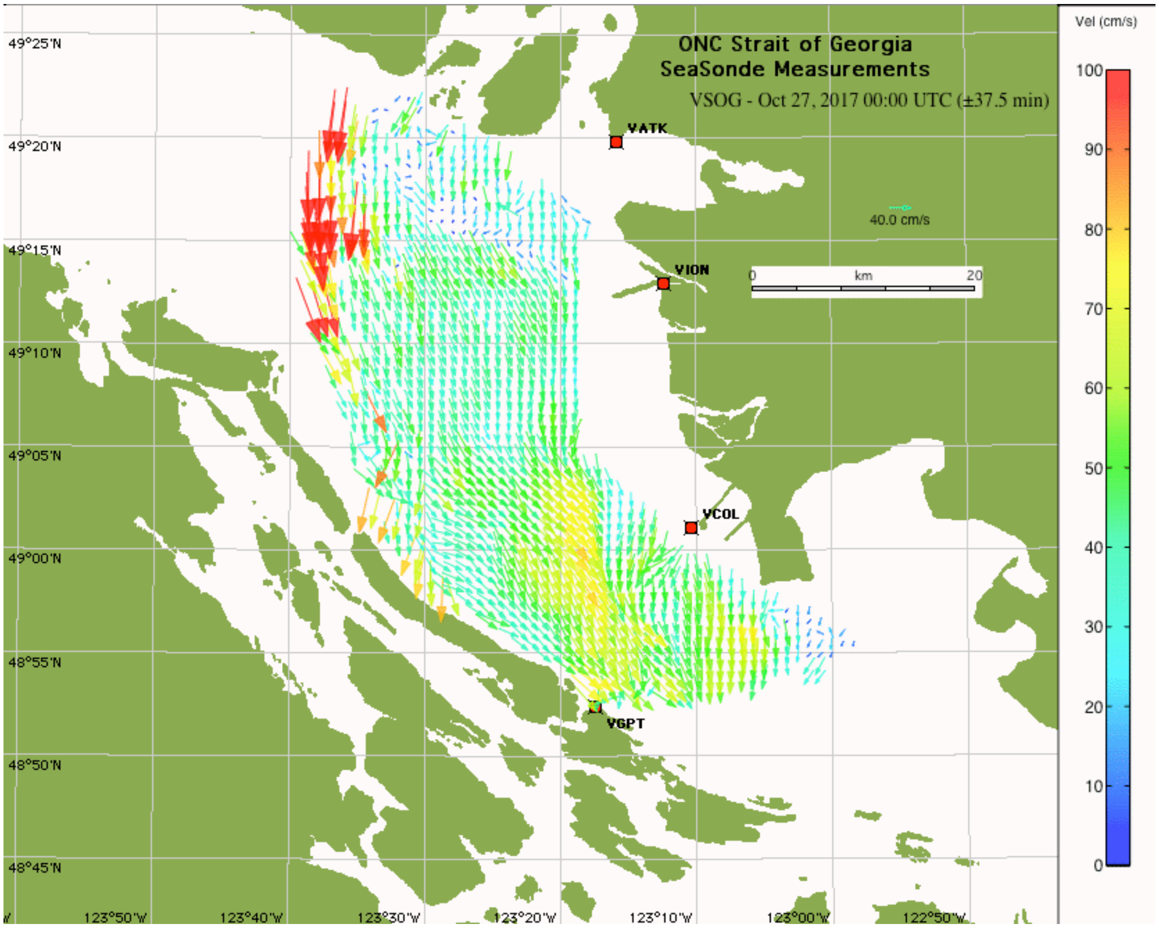


FIGURE 2. GRAPH EXAMPLE REPRESENTING A MAP WITH SUPERIMPOSED VECTORS INDICATING SURFACE CURRENT DIRECTION AND VELOCITY. *The map is animated to represent changes over time.*

Similar to tables, graphs may also be presented in static as well as interactive form.

2.2.2.2. Static graphs

The trade-offs of static graphs versus interactive ones are similar to the ones of static and interactive tables. Despite a loss in flexibility and analysis power, they are simple to implement and maintain and only require the computation necessary for initial rendering.

2.2.2.3. Interactive graphs

Interactive graphs are becoming an increasingly common data visualization format. They provide features that can help the end user dive deeper into the analysis of data or make visual adaptations to better suit their needs. Common features are described in the table below.

TABLE 2. DESCRIPTION OF INTERACTIVE GRAPHS FEATURES.

Feature	Description
Clickable or hoverable points	Reveal the value of the current data point
Zoom	Focus on a section of the graph
Dynamic scale	Modify the scale of any axis
Dynamic symbiology	Change the symbiology used to represent data points
Live update	Progressively render the chart as new data are acquired

2.2.3. Georeferenced data

Georeferenced data are one of the most prevalent types of data used by the ocean community. It is of utmost importance to provide an accurate, yet intuitive, visualization of this type of data, especially considering that the information only holds true significance when associated with a precise location.

Georeferenced data are a great way to reach out to the public. Maps are an intuitive approach to see and understand data that have become more and more established since the mainstream adoption of mobile tools like GPS and Google Maps.

Spatialized information is also a very useful and powerful tool for scientists and decision makers. For example, coastal and marine spatial planning - is an approach used and recommended by UNESCO in the context of sustainable development.

CIOOS spatial information will play an essential role in improving data discoverability and providing integrated data.

2.2.3.1. Geographic and projected coordinate systems

For proper visualization of georeferenced data, it is important to consider both geographic coordinate systems and projected coordinate systems. Geographic coordinate systems are based on a three-dimensional ellipsoidal model of the earth. The current standard (used in cartography, geodesy and navigation) is the World Geodetic System (WGS) 84. While this system is extremely useful for making measurements, projected coordinate systems are the ones used to visualize the data. Projected coordinate systems, as the name entails, are projected on a two-dimensional plane. 2D maps cannot be created without such a transformation.

There are numerous projections and each of them distorts some area of the map to a varying degree. Most web interfaces use EPSG:3857 which is optimised for web use. Other projections are optimised for Arctic and Antarctic representation. Support for various projections is a question for the CIOOS visualization as well as for other IEs, especially for older datasets.

2.2.3.2. *VARIABLE VISUALIZATION STANDARDS*

For data of given ocean variables, specific standards regulate their georeferenced visualization. While beyond the current scope of this investigative evaluation, it is important to consider using well known visualization standards for a given variable. For example, for sea surface currents the International Hydrographic Organization (IHO) provides the S111 standard (International Hydrographic Organization, 2016).

2.2.3.3. *VISUALIZATION FEATURES*

Georeferenced data visualization can range from simple to exceedingly complex. To better understand what is available to the end-users, the following table lists some of the most common features.

TABLE 3. DESCRIPTION OF GEOREFERENCED DATA VISUALIZATION FEATURES.

Feature	Description
Zoom	Bring a section into closeup.
Pan	Drag the map to change the center of focus.
Rotate	Rotate, from a predefined center, the view of the map.
Tilt	Move the view up or down on its vertical axis (i.e.: if we compare the view to a camera, the action would be to keep the same point of focus while increasing or lowering the camera's altitude).
Identify	Click on the map and, if available, obtain information on the feature(s).
Reproject (on the fly)	Change the current projection of the map visualization.
Zoom to extent/feature	Automatically zoom and pan the view of the map so that the given point or area is centered and fully visible on the screen.
Dynamic styling	Change the layer/features opacity, color, etc.
Dynamic symbiology	Change the symbiology used to represent features on the map (e.g. circles could become triangle, etc.).

Layer management	Available layers can be turned on or off in a map to show more or less information pertaining to the same area under consideration. Different layers typically represent different types of georeferenced information.

Visualization style

The styling of visual features on a map can be an overlooked aspect of georeferenced data visualization. There are multiple facets that need to be considered, most notably standard variable visualization styles, standard style definitions and coloring. As previously mentioned, some variables may be associated to a given representation standard (e.g. surface currents and S111), in which case specific steps must be taken to ensure appearance fidelity. To define geospatial styling, there are multiple possibilities. Styled Layer Descriptor (SLD) is an Open Geospatial Consortium (OGC) standard and one of the most commonly used options. Other options include YSLD (SLD equivalent based on YAML), map-box style (style definition based on JSON) and multiple Cascading Style Sheets (CSS) variants. Each style definition has its own set of advantages and disadvantages and may be more adapted to a particular situation. Lastly, the color of various features on the map can be a critical aspect of conveying information. A proper gradation of colors is necessary so that end-users can clearly get a sense of where the value fits in the overall spectrum. It is also recommended to offer a gradation that is adapted to colorblind individuals. At minimum a color palette that is relatively friendly towards the 3 types of anomalous trichromacy could be adopted. A better option would be to allow users to select between various palettes; this would include some designed for deuteranopia, some for protanopia and some for tritanopia. Ultimately, there could be the possibility for users to make their own custom palettes.

Data types and data serving

Georeferenced data can be expressed by two main data types; raster and vector. Each data type has its own set of advantages and disadvantages, file formats, standards and serving protocols.

Raster data

Raster graphics are digital image represented by a matrix of pixels, or cells, that can be reducible or enlargeable. Raster data type is therefore comprised of rows and columns of cells, each storing single or multiple values. A common example is the bitmapped computer display, where each pixel on the screen is associated with a small number of bits, the color of each pixel being determined by those bits.

This type of data is heavily used for map visualization and, as such, is stored in a wide variety of formats. Some of the commonly used formats are JPEG[3], PNG[4], GeoTIFF[5], netCDF[6], etc. While it is possible to visualize georeferenced data using a simple, static PNG image, this approach does not afford the end-user any

[3] ISO/IEC. *Information technology – Digital compression and coding of continuous-tone still images*. Retrieved from ISO Web site: https://www.iso.org/standard/18902.html

[4] World Wide Web Consortium*. Portable Network Graphics (PNG) Specification (Second Edition).* Retrieved from W3C website: https://www.w3.org/TR/2003/REC-PNG-20031110/

[5] GDAL. *GTiff -- GeoTIFF File Format*. Retrieved from GDAL Web site: http://www.gdal.org/frmt_gtiff.html

[6] Unidata. *Network Common Data Form (NetCDF)*. Retrived from Unidata Web site: https://www.unidata.ucar.edu/software/netcdf/

flexibility, nor significant power of analysis. As previously discussed, an interactive map is often the more preferred option.

Although there are several ways to serve raster data to an interactive map, the most widespread standards are the Web Map Service (WMS) and Web Map Tile Service (WMTS[7]) from the Open Geospatial Consortium (OGC). The WMS protocol serves only map images and does not natively support tiling (Open Geospatial Consortium, 2017). The WMTS protocol may serve data in different formats (topojson, protobuf, geojson, etc.) and was conceived specifically to handle tile support (Open Geospatial Consortium, 2017).

The raster data type has some drawbacks. Since datasets record a value for all points of a designated area, this will often lead to significantly more storage space requirements. The raster data model may also bring issues of transparency and aliasing when overlaying multiple images and ill-defined boundaries with low resolution files. Another point of contention is updating and maintaining: raster images must be completely reproduced with every update. Despite these drawbacks, raster data are not expensive to render, which can be quite a compelling advantage when dealing with complex interactive web maps.

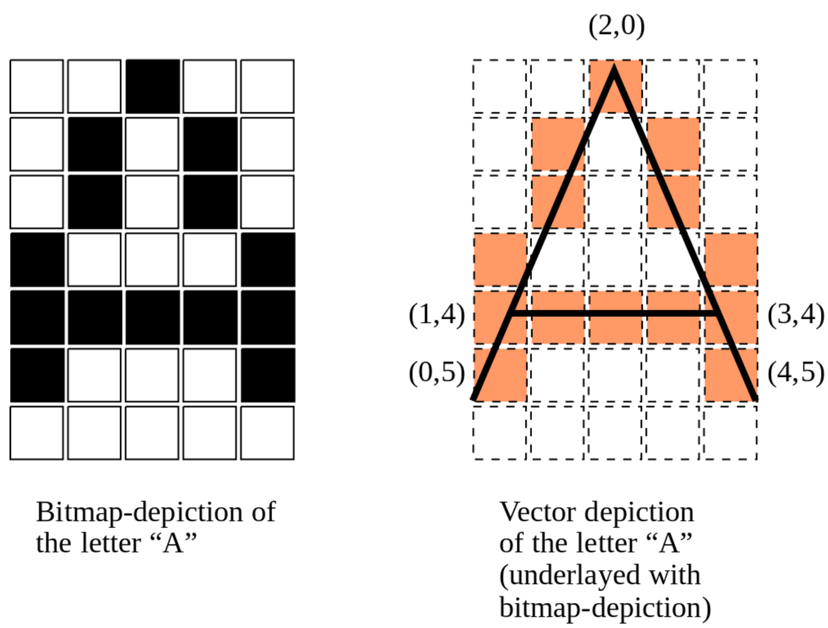


FIGURE 3. BITMAP AND VECTOR DEPICTION OF THE LETTER "A". (WIKIMEDIA COMMONS, 2010)

Vector data

Vector data use geometrical shapes, expressed as vectors, to represent georeferenced features. Different geographical features will require different types of geometry. The main types are points, lines (or polylines) and polygons. For example, a point could represent a buoy, a line a boat trajectory, and a polygon an essential fish habitat.

[7] Open Geospatial Consortium. *OpenGIS Web Map Tile Service Implementation Standard.* Retrieved from OGC Web site: http://www.opengeospatial.org/standards/wmts

The use of vector data for map visualization has grown significantly in recent years. Commonly used formats include GeoJSON, shapefile, Geography Markup Language (GML), and so on. Vector data must be processed (rendered) before they can be visualized.

As is the case for raster data, there are numerous ways to serve vector data. The most standard practice would be through Open Geospatial Consortium's Web Map Tile Service (WMTS), although a simple endpoint providing data in the required format can sometimes be equally sufficient. For example, a simple web service can be developed to return GeoJSON, which can then be easily rendered onto an interactive web map.

The vector data type offers its own set of advantages and disadvantages. It is much easier to implement overlay operations and the result is often visually smoother. Vector data are better adapted to relational databases, as they can be associated to a standard column and several databases provide a multitude of operators. Vectors are easier to scale and re-project. They also, usually, have smaller file sizes and can be updated efficiently, by simply editing the raw data. They provide enhanced analysis capabilities, most notably for networks such as subsea cables, surface drifter trajectories, etc. Despite all these advantages, it is important to stress the fact that they are more computationally expensive to render.

FILTERING AND DIMENSIONAL MANIPULATION

There may be times where the end user only wishes to visualize a small subset of what is currently displayed. For example, within a large georeferenced biodiversity dataset, the user might only wish to see occurrences of right whales. Filtering a dataset in such a way can enhance user analyses, however, as the number of different filters increase, so does the implementation, maintenance complexity and, potentially, the cognitive load for the end user. Another interesting way to manipulate the visualization of georeferenced data is through dimensions. Time and depth are the two most commonly used dimensions in ocean data and manipulating them provide multiple benefits. Manipulating time allow users to visually witness the evolution of a given phenomenon; whether it is the trajectory of a glider, the sea level rise or the global temperature. Modifying depth is also beneficial as several variables relevance depends on it (e.g. water temperature, sea currents, etc.).

2.2.4. VIDEO AND 2D IMAGERY

Video data, for the purposes of this assessment, will be confined to digital video (as opposed to analog video), for which there are several standards and formats. Digital video is the sequential display of visual images in the form of encoded digital data. There are numerous formats and digital resolutions for the visual images, often determined by the camera and original encoding process, and a second encoding layer/stage which optimizes the format for the display of the image sequence, depending on the selection of numerous efficiency requirements (i.e. supported streaming data rates, compression levels, etc.). Ideally, digital video is also georeferenced, if not frame by frame, at least by video segment. At a minimum, the georeferencing includes a timestamp, the latitude and longitude, and depth of the video. Additional georeferencing data could include camera angle/orientation, as well as any artificial lighting employed during the video capture. Most digital video formats have "tracks" to accommodate other coincident streaming information, most notably, an audio track. These additional tracks can, however, be used to store image content information, or annotations, thereby making the video content self-describing. It is also possible to display human or machine-generated log entries or annotations associated with specific video time segments. Specific types of video exist for visualizing 2-D phenomena in a convenient form. For instance, time-lapse videos are a succession of still frame images taken at regular intervals, using the same pointing and the same imaging equipment. Other forms of video include may animations, typically to visualize a time component in a 2-Dimensional environment (e.g. sea-surface currents changes in the last 24 hours).

Video players or viewers for research purposes need to be able to display the full image resolution stored in the digital video file, stop/pause and select individual video frames (single images), advance and play-back video at various speeds, and provide view/edit access to the available annotation or additional tracks. Automatic or manual cataloguing of video content and annotation tagging is important for research video, enabling users to search video annotation archives for specific content without having to replay the video, either in real-time or fast-forward. Video archives integrate numerous files containing the digital video, annotations, georeferencing, and all associated metadata related to the video capture system and any video processing applied.

A key component of any CIOOS centralized or distributed marine video archive will be the need to have a well organized and searchable annotation catalogue, identifying content along with the georeferencing and metadata. In other words, the archive contents need to be discoverable, without downloading and viewing actual data. This may require investment in either manual or automatic content annotation, or a requirement that only “annotated” video will be integrated into a discoverable CIOOS.

2.2.5. Passive acoustic

Passive acoustic data, as with video data, refers to streaming, sequential digital audio data. But instead of an ordered sequence of digital images, a digital audio stream consists only of sequential digital numbers, of some specified bit resolution (i.e. 16 or 24 bits per number) and representing high-frequency digital samples of ambient acoustic pressure. Marine audio is recorded by a hydrophone (underwater microphone). Analog hydrophones require a digitizing process, while modern digital hydrophones include all the A/D conversion elements, including fully calibrated digital conversion to sound pressure.

Raw digital audio is often stored in uncompressed files as a simple sequence of numbers, known as WAV files. There is no additional information stored with each digital audio number (audio level), and the time sequencing is assumed to be at a predefined sample rate, such as 44kHz. Audio compression (with various levels of information loss) is common among Internet audio formats, such as MP3, to minimize file size and download times. Lossless audio compression codecs also exist (e.g. Free Lossless Audio Compression/FLAC), and are being used increasingly to deal with large volume of research data.

Audio content is often represented in one of two forms: by playing and listening to the audio (often with speed control), or by visualizing the audio content by displaying spectra of the audio signal. Mean spectra can be used to assess audio levels within specified frequency bands, or short duration spectra can be stacked vertically into an image format as a spectrogram. Spectrograms are images with time along the x-axis and frequency along the y-axis, and audio intensity/power displayed as a colour intensity. Longer-period spectrograms and long-term spectral averages can enable users or detection algorithms to rapidly evaluate long periods of audio data, without the need for listening. Spectrogram viewers need to facilitate viewing short, medium, and long spectrogram intervals, while offering the ability to play (and download) the audio once a signal of interest has been identified.

As with video, any centralized (or distributed) discoverable archive of audio data must have a searchable catalogue of content annotation, including the georeferencing and metadata. This may require investment and standardization in manual and/or automatic content assessment and identification.

2.2.6. 3D AND VOLUMETRIC DATA

Multiple datasets are defined in more than two dimensions. These datasets are structured as point clouds, 3D polygonal objects or structured and unstructured data in 3 or more dimensions. Most point clouds are generated by 3D scanning instruments such as multibeam echosounders or LIDARs. Point clouds comprise a list of points in 3D space defined by X, Y and Z values. Polygonal objects comprise vertices linked by edges ices in 3D space, forming triangles or more complex polygons in Euclidean space. They are generally produced by triangularization with post-processing of point clouds or isosurfaces extracted from volumetric data. N-Dimensional structured or unstructured data partition space in addressable dimensions. Outputs of weather, climate and ocean forecast models and satellite data are the most common uses of N-Dimensional data.

While 3D point cloud rendering and 3D polygonal model rendering tools are the most common and mature, structured and unstructured volumetric model rendering is still a niche, for which no Web visual standard tools or library exists at a mature level. 3D acceleration is now available to Web browsers and libraries like three.js are starting to emerge for 3D polygonal rendering and animation, 3D point cloud libraries are still lacking and need more custom development for proper rendering. Multiple volume rendering techniques are suitable to certain types of visualization, manipulation and hardware configurations; no one size fits all. Since most data for the marine community are either point cloud or volumetric datasets, the rendering and manipulation of these kinds of dataset are more in the realm of desktop applications than Web interfaces. The computationally expensive requirement for 3D visualization is still mainly reserved for desktop computers so mobile devices are mostly excluded. This might change in a relatively near future since mobile and desktop 3D acceleration hardware are evolving at a rapid pace with technologies like Virtual Reality and Augmented Reality attracting significant research & development capital.

For most current applications and users, only 2D projections of 3D data are viewable. One example of such a 2D projection is a bathymetry point cloud generated by multibeam echosounder data that are post-processed into gridded 2D maps that can be displayed through standard protocols like WMS. Structured volumetric data can be reprojected and accessed via standard WMS where time and depth represent the additional dimensions. The size of volumetric data can be prohibitive for users to download in their entirety, data extraction through specialized protocols like OPeNDAP is better suited to visual rendering on desktop applications.

As with audio and video, any centralized (or distributed) discoverable archive of 3D and volumetric data will need a searchable catalogue of content annotation, including the georeferencing and metadata. More work is required for metadata and data standardization and manual and/or automatic content assessment/identification.

2.2.7. HIGHER DIMENSIONAL DATASETS

Oceanography is frequently collecting data about currents at different depths using Acoustic Current Doppler Profilers (ADCP). They produce complex matrices representing a 3-directional velocity vector for e.g. 100 different depths once per minute. The resulting dataset has essentially 5 dimensions. Ad hoc methods to represent and visualize the data on a 2-D screen typically involve a 3D projection of vectors, using length for velocity and animation for the time dimension. Those are very frequently used data types in ocean sciences and will need to be considered in CIOOS initiatives.

3. Software suites and tools

To provide accurate, meaningful and efficient visualization, the tools of the trade must be carefully examined. In pursuit of this objective, the pros and cons of open source and proprietary software is initially discussed. A similar debate on all included suites versus custom stacks follows. Finally, an inventory of potential tools is presented, along with an assessment of their strengths and weaknesses.

3.1. Open source and proprietary

The debate on whether to use open source or proprietary software has been going on for quite some time now. Before delving into it, the proper definition for each term must be made sufficiently clear. Open source software is "*computer software with its source code made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose*" (St-Laurent, 2008). Proprietary software is defined by Wikipedia as "*computer software for which the software's publisher or another person retains intellectual property rights—usually copyright of the source code, but sometimes patent rights*".

Closed source software offers several advantages. Commercial vendors often focus on a specific market and attempt to satisfy its users. The result is usually greater usability due to a more coherent interface, ease-of-use or functionality. The overall product is usually oriented toward less tech-savvy users and accompanied by strong and targeted technical support. Nonetheless, the facts remain that this type of software can be expensive, that users are dependant on the vendor for the implementation of specific features and, obviously, that there is no opportunity to look over the code and implement any bug fix.

Open source has its own set of advantages. Although not all open source software packages are free, a good portion are and most of them are of much lower cost than their proprietary alternatives. Making the code available to anyone means that it can come under heavy scrutiny, which will generally result in better code quality and improved security. As soon as a potential flaw is detected, users can report it or even fix it themselves. Another compelling reason to use open source is to avoid vendor "lock in". Since there is no dependence on a single vendor for continued improvements, bug fixes or feature implementation, there is a much smaller risk of having orphaned software. Lastly, open source means having a community of engaged individuals working toward a common goal. This community will often provide significant levels of free support to those interested in using their product-. The drawbacks of open source are its lesser usability, orientation toward more skilled users and no targeted, quick support.

While it is important to acknowledge the pros and cons of open source and proprietary software, the context in which they are considered can be the decisive factor. Open source products are generally favoured as long as certain explicit conditions are met. Ideally, the product should be popular, have active development, meet expected features and have a minimum age. The popularity of open source software has been growing steadily in recent years: they are now used by 98% of enterprise-level companies (Murphy, 2010). Microsoft, previously a strong advocate of closed source, is now investing significantly in open source. The White House website has moved to an open source Content Management System and the International Space Station is using an open source operating system. Some prime examples of open source products are the apache http server (most popular web server), the Linux operating system (99.6% market share of supercomputers, host for majority of web servers), Mozilla Firefox, and so on.

3.2. All included and custom stacks

Tools can be provided as is or as part of a whole suite (or stack) that encompasses commonly complementary features. Some examples of suites that aim to fulfill the role of a complete Geographic Information System (GIS) are MapMint (comprised of tools such as MapServer, GDAL, OpenLayers, LibreOffice) and Boundless Suite (comprised of PostGIS, GeoServer, GeoWebCache and OpenLayers).

The aim of these suites is generally to facilitate the implementation of common use cases. Most of them provide support to end users and an additional layer of abstraction over their components. Since authors of the suites are deeply invested in a given set of tools, they strive to ensure that they remain compatible and complement each other. Despite these advantages, the use of individual tools is mostly favored. Even if a given suite exclusively uses free software solutions, they generally charge a significant fee for their services. Adopting a suite is also notably detrimental to flexibility. If a new and improved tool comes along, it is much harder, if not impossible, to use it instead of the one provided by the suite. Likewise, it can be difficult to use software solutions outside of the suite to complement its features. Lastly, it is worth noting that if the technical expertise with the different tools is already acquired, the benefits of ease of use and simplicity are significantly mitigated.

3.3. Inventory of existing visualization tools

The following section lists a selection of existing visualization tools, and some suites that implement them. A more detailed list can be found in **Error! Reference source not found.** in Visualization Tools InventoryAppendix A.

3.3.1. Map Servers

For this document, map servers refer to servers whose main objective is to serve geospatial data for eventual visualization through a client map viewer. Only solutions that adhere to basic map visualization standards like Web Map Service (WMS) and that have some degree of popularity have been considered.

3.3.1.1. Geoserver

Geoserver is a Java software package using Open Geospatial Consortium open standards, allowing users to display spatial data on maps with various output formats. These standards support serving, sharing and editing of data. Geoserver uses OpenLayers as a mapping library and Geotools as a GIS toolkit.

3.3.1.2. MapServer

Similar to Geoserver, MapServer is an open source software package using OGC standards such as WMS[8] and WFS[9], which is used to publish spatial information as "geographic image maps".

3.3.1.3. ArcGIS Server Entreprise

ArcGIS Server is the geographic information system (GIS) server software developed by the private company Esri. ArcGIS Servers allow users to create and manage GIS web services, applications and data.

---

[8] Open Geospatial Consortium. *Web Map Service.* Retrieved from OGC Web site: http://www.opengeospatial.org/standards/wms

[9] Open Geospatial Consortium. *Web Feature Service*. Retrieved from OGC Web site: http://www.opengeospatial.org/standards/wfs

3.3.2. Catalogues

Catalogues are servers that systematize the presentation and access to data and metadata. They usually provide aggregation and acquisition capabilities and might offer the possibility to federate instances.

3.3.2.1. CKAN

CKAN (Comprehensive Knowledge Archive Network) is an open source data management system that provides tools to store, publish, share, find and use open data. Already used by the Canadian Open Data portal, it supports multiple standards, metadata visualization, and data previsualization.

3.3.2.2. GeoNetwork

GeoNetwork is a catalogue application, providing powerful metadata editing and search functions, as well as providing an interactive web map viewer.

3.3.2.3. ESRI Geoportal

ESRI Geoportal is an open source data discovery tool that users can use to manage, publish, search and access geospatial resources and their metadata. Geoportal Server allows users to enter geospatial resources (locations, descriptions) in a central repository, which can then be published publicly or through a private intranet.

3.3.3. Map Viewers

Map viewers are client software solutions that can display maps and provide interactive features.

3.3.3.1. OpenLayers

OpenLayers is an open source mapping library used to create interactive maps on the web. It can display map tiles, vector data and markers from any web page, using standard formats and protocols (WMS, WFS, KML, etc.). It is a powerful GIS-oriented library (as opposed to Leaflet, which is more web and mobile oriented)

3.3.3.2. Leaflet

Leaflet is an open-source JavaScript library for interactive maps that can be extended with plugins. Leaflet has a well-documented and easy-to-use API, and an open source code. Leaflet is designed to create mobile-friendly maps.

3.3.3.3. ArcGIS Javascript API

ArcGIS JavaScript API version 4.5 is Esri's next-generation API that builds full-featured 2D and 3D applications with information layers (terrain, base maps, imagery, integrated mesh layers, 3D objects).

3.3.4. Data Servers

Data servers are web servers that focus on providing convenient access to data and metadata.

3.3.4.1. THREDDS

THREDDS Data Server (TDS) is a web server using remote data access protocols to give end-users access to both metadata and data. TDS uses protocols such as OPeNDAP and HTTP, and OGC standards (WMS, WCS). TDS can also use NetCDF Markup Language (NcML).

3.3.4.2. ERDDAP

ERDDAP is a data server using Data Access Protocol (DAP), allowing download of subsets of gridded and tabular datasets, in various file formats. ERDDAP also offers a powerful search tool, format translation, graphing and mapping, and time standardization.

3.4. ASSESSMENT OF EXISTING VISUALIZATION TOOLS

The assessment of the various visualization tools and suites is made on a per category basis. Each subsection only evaluates the software solutions with the most potential for the given category.

3.4.1. MAP SERVERS

Map Server	**ArcGIS server**	**Geoserver**
Open Source	Closed Source	Open Source
Free	Expensive	Free
Cross platform	Cross platform, proprietary formats	Cross platform
Most common standards (No WPS, no direct WFS-T)	Most common standards (+ SOS, SWE)	Most common standards
Configuration through Mapfiles or Mapscript	REST API configuration	REST API configuration
Caching support	Direct customer support, More appeal for less tech-savvy users	Caching support
Documentation not up-to-date	Good documentation	Good documentation
	Web interface for administration	Web interface for administration
Good performance	More custom features than other map servers	Most active Community

3.4.2. CATALOGUES

GeoNetwork	**ESRI Geoportal**	**CKAN**
Open Source	Open Source	Open Source
Free	Free	Free
Used by many large organizations	Not as popular as the other alternatives	Used by major open governments
REST API (not mature)		REST API
Supports CSW,	Supports CSW, INSPIRE, ISO 19115-19, GEMINI, Dublin Core, WMS, ISO 23950	Supports CSW, INSPIRE, Dublin Core, FGDC, DCAT, etc.
Can harvest with CSW, OAI-PMH, Z39.50, THREDDS, Webdav, Web accessible folders, ESRI GeoPortal, etc.	Can act as a client for ArcIMS, OAI-PMH, ISO 23950, WAF, WMS, WFS, GML, KML	Offers harvesting interface. Extensions are already developed for CKAN, CSW, WAF

		and Single spatial metadata document
Oriented toward geospatial data, less adapted for non-geospatial data	Federated searching	Geospatial search with Solr backend (no multi-polygon geospatial search)
Can be harvested by CKAN	Integrates well with ESRI products	Well documented, strong and active community
Supports multiple languages	Interface not modern, not as user friendly as other options	Support access control
Allow visualization of geospatial data on a map, but no table or graph		Expandable via plugins

3.4.3. Map viewers

Leaflet	**ArcGIS Javascript API**	**OpenLayers**
Open Source	Closed Source	Open Source,
Free	Free for NGO and not-for-profit	Free
Strong and active community	Excellent ESRI support, but less popular than other options	Strong and active community
	Good documentation, wide range of examples	Well documented, wide range of examples
Core support of GIS formats, no out-of-the-box support for projections and GIS formats	Large number of standards and formats (…) as well as proprietary formats	Large number of standards and formats (…)
Not GIS oriented, less tech-savvy users oriented	GIS oriented	GIS oriented
Most advanced features provided by plugins	3D support, several custom features	Open source library with most features for geospatial data manipulation

3.4.4. Data servers

THREDDS	**ERDDAP**
Allow for consuming data in chunks	Processes data in chunks
Easy access for analysis of raw data	Built-in data conversion and data filtering
Efficient data transfer	Great for interoperability, RESTful web services, used by oceanographic community
Supports standards like OPeNDAP, WMS, WCS, Dublin Core,FGDC/DIF, ISO 19115, ADN	Support standards like OPeNDAP, WMS, WCS, WFS, SOS, OBIS, etc.
Supports data formats like NetCDF, HDF5, NcML, GRIB, NEXRAD	Wide range of formats support

Few visualization possibilities	Allow for map, tabular and graph visualization
Weak navigation and search capabilities	Consistent and strong search options
	Support for federation of instances
No i18n support	No support for multiple languages, no support for audio, low support for data with multiple dimensions
Well documented. Supported by Unidata	Documentation not well structured, coded by essentially one individual

4. End-users and desired functionalities

CIOOS will be used by many users spanning a wide range of interests. Although most tools and software applications, used to gather and process raw data, are developed mainly by and for scientists, many type of users and interest groups desire and require access to the ocean data for professional and personal reasons. Specific groupings, by data format, theme and visualization are possible.

4.1. End-users

As CIOOS governance plan aim at targeting users from various sectors and areas of interests, many possible user groups can be defined with each a set of typically used data visualization products. However, needs often also varies within the same sector or user group. For this reason, among others, an iterative collaboration with partners is necessary to ensure that evolving needs are met with adaptations in the products and services offered by CIOOS. Tools to implement such collaboration will not be detailed in this report, however, a user experience survey on the CIOOS website could be a way to track satisfaction and gather cues for improvement directly from all CIOOS website visitors. Other communication tools, such as invitation to join a mailing list, could be useful for track and analyse user statistic, as well as allow to reach out to the user community for further feedback along with sharing network initiatives and other news.

In addition to the differences in products and services need, user groups might have their own specific constraints such as: low-bandwidth connections, mobile access needs, government firewalls, limited computing facilities or resources, limited technical expertise. As such, support to users should be made available through clear and easy communication channels such as an email help desk or contact form. However, this line must be adequately equipped in a way that rapid and efficient feedback and help are provided to users.

Note. Ocean observing systems (OOS) not only help to better understand and manage ecosystems, their contribution can be measured in term of societal benefits such as human health protection, emergency response & preparedness, climate change assessment & forecasts, sustainable resource management (water, energy, agriculture, etc.), and economic development. OOS create value from environmental data and their impact is worth several billion dollars annually. Nevertheless, the sustained investment required to maintain a functional CIOOS justifies the necessity to be able to identify end-users, downloaded datasets, and data products used and their outcomes, from the start of CIOOS implementation. It is recommended that CIOOS optimizes the uses of tools such as Google Analytics, for each Web sites (CIOOS and RA's), to be able to measure and describe in a comprehensive way the use of their products. This will offer CIOOS insightful information, support its legitimacy, help target end-users needs in terms of data and products, and allow to continuously improve its interfaces. Other tracking or feedback mechanisms, always in the respect of end-users' privacy, such as tracking of downloaded databases or visitors targeting surveys, should be explored, considered and implemented before CIOOS and RAs Web sites launches.

TABLE 4. EXAMPLES OF VARIOUS CLIENTELES BENEFITS IN USING CIOOS DATA PRODUCTS AND SERVICES.

	CIOOS Products & Services	
User Groups or Areas of Interests	**Real-Time Data & Archives**	**Web Services & Applications**
• Governments • Decision-makers	Official data sources to complement own datasets	Web services feeding decision-making processes ex: disaster response
• Navigation • Marine Transportation • Port management	Bathymetry - Currents - Waves – Tides - Vessel Traffic - Marinas	Real-time water levels & weather conditions - Web services feeding electronic charts - Navigational aids locations
• Energy Development (Oil & Gas, Wind, etc.)	Physical/biological/geological	Visualization tools, ex: air/water circulation
• Resource Exploitation (e.g. Fishing)	Physical/biological (temperature, salinity, dissolved oxygen, etc.)	Resource mapping applications - Species identification guides
• Resource Conservation & Management	Water quality - Biodiversity - Species at risk habitats	Resource mapping applications - Protected area maps - Species identification guides
• Coastal Zone Preservation & Management	Ecosystem monitoring data - invasive species presence	Visualization tools, ex: Advance warnings (disasters)
• Scientific Research	Physical/biological /geological time-series	Interactive mapping tools
• Tourism Industry	Site locations/accesses (lighthouses, beaches...) - Demographics	Interactive mapping tools
• Environmental Consulting	Official data sources	Web services feeding analytical processes
• Education & Outreach	General interest information	Crowd-sourcing applications - Marine protected areas
• Research & Development • S&T Industry	Official data sources supporting innovation & development projects	3D imagery tools - Interactive mapping tools
• Social Sciences	Applied science datasets	Visualization tools
• Human Health	Pathogen distribution (air, water, soil) - Demographics	Interactive mapping tools, ex: threats (pollution, diseases...)

4.2. METADATA CATALOGUE

Metadata discovery is very important, and some datasets don't require complex visualization interfaces. Below are section detailing the needed tools for end-users, catalogue editors and system administrators.

4.2.1. END-USERS FUNCTIONALITIES

In order to allow various communities (scientific, media, academic, general public), to explore, search and manipulate datasets, the implementation of a visualization tool to discover data and metadata must be integrated while setting up a national portal such as CIOOS. This section lists available useful functionalities of a cataloguing tool.

4.2.1.1. SEARCHING

Minimally, a dataset must be discoverable through browsing in the catalogue, where all datasets are listed without filters. The following tools allow users to refine their data search and exploration.

4.2.1.1.1. FREE TEXT SEARCH

A partial or exact text search function allows users to find a first listing of datasets. Text search can be applied to all textual attributes in a dataset, including metadata. This type of search is similar to what is offered by known web search engines. Pangea Data Publisher is an example of a catalogue that provides a powerful free-text search capability.

Although text search is a must, it is not always ideal to facilitate data discovery. Other search modalities, detailed below, should be considered to support more research scenarios.

4.2.1.1.2. SPATIAL SEARCH

Spatial search is based on the discovery of geo-referenced data and it is performed by the means of a map in which a user-defined polygon makes it possible to limit the datasets to the chosen area. If the research tool makes use of an API, the spatial search can be performed using coordinates defining the polygon. For example, the GeoJSON specification (RFC 7946) could be used to define polygon points when querying the API. This search also obviously must take depth and altitude into account. For example, users may be able to define depth ranges along with geographic bounding boxes and time ranges when selecting data sources, as shown in the following example user interface.

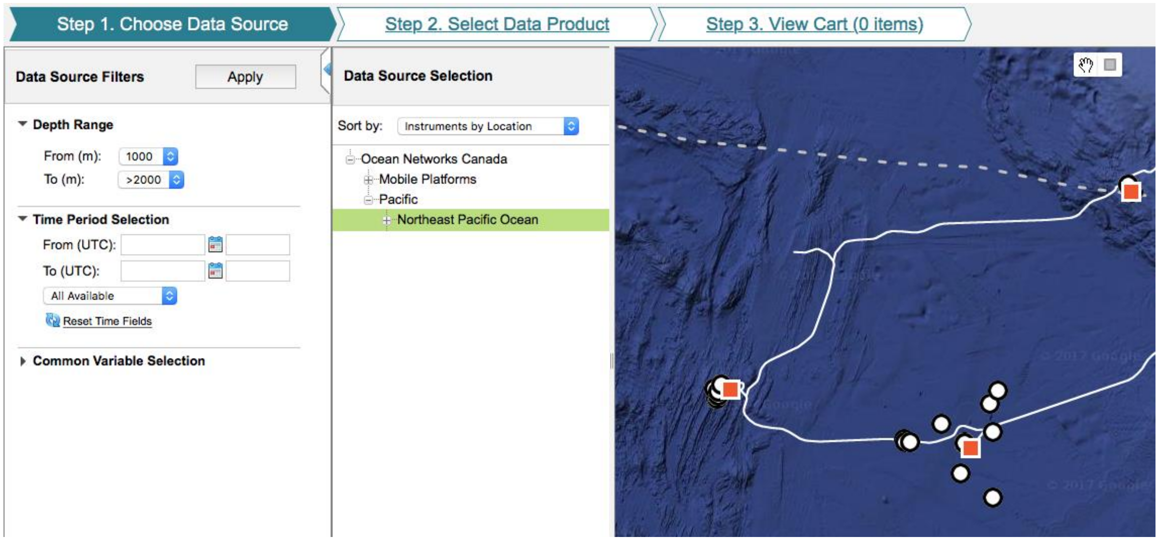


FIGURE 4. DATA SEARCH TOOL BY OCEAN NETWORKS CANADA (HTTP://DMAS.UVIC.CA/DATASEARCH).

4.2.1.1.3. TIME SEARCH

Time search is used to define the period within which a dataset was generated. This interval may be bounded or unbounded. To ensure interoperability, time-based search must meet a recognized standard, for example ISO 8601, a standard recommended in the IE Data and Observations report. Time and spatial searches can often be combined effectively within a single interface.

4.2.1.1.4. FACETED SEARCH

Faceted search allows users to filter a collection of datasets by one or a series of criteria, other than time and space which were discussed above. Generally, users carry out this search by successively adding criteria to refine the scope. To implement this type of search, a controlled vocabulary must be adopted to associate elements of the datasets with criteria, or keywords, that will facilitate their discovery. As an example, this feature could offer taxonomic filtering.

4.2.1.1.5. ADVANCED SEARCH

To meet the needs of advanced users, the catalogue should enable search using multiple combined criteria: free text, keywords, facets, spatial and time search in a single step.

Without being exhaustive, other advanced search approaches using logical operators could be considered, such as:

- Regular expressions: A type of text search used to find strings of characters that meet criteria, using a precise syntax. As an example, *bio.*y* will discover the datasets containing the words biology, biochemistry, without however including the words biochemical.
- Boolean functions: A type of text search that allows the union (OR), intersection (AND) or exclusion (NOT) of search keywords. Combining these operators allows for a multitude of possibilities. For example, a search containing (biology AND freshwater) NOT fish would find results that apply both to biology and to freshwater, but would exclude fish from the results

4.2.1.1.6. CIOOS-DEFINED CATEGORIES SEARCH

A type of search based on categories defined especially for CIOOS and by CIOOS could be based on the specific terms such as the CIOOS core variables[10].

4.2.1.1.7. Protocol search

Protocol search allows the user to discover the available datasets using recognized protocols such as OPeNDAP and WMS. This type of search is particularly useful for users desiring to reuse the datasets in other tools such as MatLab or Ferret.

4.2.1.2. *Summary description, data and metadata*

Discovery of a dataset will enable the user to obtain relevant information, namely: a summary description, standardized metadata and access to one or more resources. The metadata may include a persistent identifier, such as DOI, as well as versioning information. Finally, it must provide standardized provenance information and citation guidelines. For further details on the metadata, see the Data Investigative Evaluation report.

4.2.1.3. *Content syndication*

Content syndication allows the user to subscribe to content of a dataset and in that way, be notified of any modification to it. RSS and Atom are two systems allowing content syndication. Similar approaches, such as newsgroup, mailing lists, followers, etc. can also offer to keep the user updated of any changes in a selected dataset.

4.2.1.4. *Activity stream*

Activity stream is a function allowing the users to obtain the chronological list of any changes made to the content.

4.2.1.5. *Simple workflow*

Without being a feature as such, the data catalogue visualization tool should be user-friendly. To use the tools listed here as well as to quickly access datasets, the workflow should be based on familiar Web-based interactions: point, click, view, download.

4.2.1.6. *Multilingual capability*

The data catalogue user interface, as well as selected metadata fields, must be bilingual (French and English) without, however, imposing linguistic constraints on the data content itself, data-associated information (e.g. contextual information in a PDF[11] document), or back-end metadata standardized information (e.g. controlled vocabularies).

4.2.1.7. *Access to raw resources*

Resources associated with a dataset should be available for download in a variety of formats. For example, a NetCDF file could be available as a full download via HTTP or using the OPeNDAP protocol. OPeNDAP makes it possible to consider offering the user a partial download, when applicable. To improve the browsing experience, the user should be able to download a set of resources associated with a dataset in a single operation. When the resource is hosted elsewhere than on the national portal, the user accesses the resource through an external link.

---

[10] See Table 3 and Table 4 of the Section 2.6 the IE Data and Observations report.

[11] International Organization for Standardization. *ISO 32000-2:2017*. Retrieved from ISO Web site: https://www.iso.org/standard/63534.html

4.2.1.8. *Metadata visualization*

The previously described features and search tools will be used to facilitate data discovery. When searching, datasets will appear listed, displaying at least the dataset's name, description and data producer. Some additional information such as associated resources file formats, keywords, or timescale can be displayed. Once a dataset has been selected, all available metadata are displayed. Depending on the type of metadata available, certain geospatial visualizations, such as the associated geographic boundaries, can be displayed on a map. Since a dataset can contain multiple elements or resources, each element can be accessed with more details, including a set of metadata specific to it.

4.2.1.9. *Data previsualization*

Data can be previewed according to various scenarios. Minimally, as stated above, the user will be able to view metadata information. Metadata standards and recommendations are covered extensively in the IE Data and Observations report.

The display of metadata should ideally be accompanied by preview features. File formats of dataset-associated information (PDF, text, JPEG, etc.) can be previewed using a discovery tool without leaving the portal. Tabular data (CSV, XLS, etc.) can also be previewed and may provide users additional features such as sorting, filtering, etc. Depending on the format of the source, spatial data can be displayed using points on a map (KMZ, GeoJSON, etc.) or inside a zone (Shapefile, XML, etc.). In some cases, different preview options for the same data source might be available. For example, a tabular file may also be previewed using a map if it contains geospatial information such as coordinates.

Data preview should be done regardless of the storage location. Remote data must be accessible using the appropriate method (WMS, web service, etc.) to be previewed in the search tool.

4.2.2. Administration user functionalities

Even if the primary users of the discovery tool are the end-users, the catalogue environment will have to provide a minimal set of functionalities to system administrators. Some of which have to do with access control.

4.2.2.1. *Access control*

While access to metadata is freely available for end-user, the system must provide restricted access control for administrative tasks.

On the national portal, the administrator should be able to define location and source type of a machine-readable data source that can be harvested (pull technology).

At regional nodes, the administrator will be able to assign permissions for data providers to add datasets through an API or a direct and manual access (push technology), or alternatively define location and source type of a machine-readable data source that can be harvested.

4.2.3. System administrator functionalities

The needs of system administrators differ from the needs of end users. These needs will have to be considered when setting up the catalogue.

4.2.3.1. *White box system*

The system will have to give access to its internal operating features. As a result, the system will be fully extensible through the API, adding plugins, or adding third-party modules.

4.2.3.2. *Web-based*

The search tool must be accessible using the Internet with a web browser to centralize the work: maintenance, deployment, backup, storage, security. The system must be adaptive and responsive, i.e. independent of the platform or browser used.

4.2.3.3. *API net-centric*

To ensure interoperability between CIOOS and the regional node systems, the tool will expose an API allowing access and exchange of data according to different levels of security. The exchange format and syntax shall be defined to ensure interoperability. The format of the metadata provided by the API will be transmitted according to the necessary protocols. In addition to the exchange format, the services offered by the API will have to adopt a controlled vocabulary. Thus, the syntax of the metadata needs to respect a standard such DCAT or QDC. For details on controlled vocabulary recommendations, see the Data Investigative Evaluation.

4.2.3.4. *Active project*

The chosen cataloguing solution should be able to show continuous activity from the developing community to ensure its longevity. Various clues indicate levels of active development, such as:

- Name and categories of organizations currently using the product
- User feedback
- Number and quality of contributors to the project
- Support level of the community
- Maturity of the project

4.3. Thematic metadata visualization interface

In conjunction with the general-purpose metadata interface or catalogue, thematic metadata discovery and visualization interfaces are needed. These interfaces limit the kind of datasets available yet give a more flexible set of filters and visuals tailored to different needs. EMODNet implement different portals where specific data are viewable and discoverable. IOOS has different portals and viewers. Portals can be defined as data access websites offering a collection of services and interfaces to where specific data reside, including data viewers tied to specific data types.

Since CIOOS will have one centralized general catalogue interface for metadata, different viewers will be made available to allow users access to datasets of different types. The centralized general catalogue will allow users to navigate through multiple sites to discover and visualize datasets of common interest. Combination of visual products and preconfigured search filters can be implemented in a different, more specific interface.

A survey of the main ocean observatories has identified common trends and styles of interfaces, which may, ultimately, shape user expectations and needs.

4.3.1. Geographic Information System style interface

The most common interface style is data layers that can be displayed on a map. This interface is the one used by Geographic Information Systems (GIS). A list of available datasets is displayed, and the user selects the desired layers to display on the map. When the number of datasets becomes too large for a simple list, either filter or grouping options, or a data catalogue can be used to add datasets on the map. When the catalogue solution is used, the user must navigate between the map and the catalogue.

When an interface is associated with one specific variable such as bathymetry, a common legend can be used and displayed permanently. However, when different layers representing different parameters and variables can be stacked, a flexible legend describing the symbiology for all active layers must be available. The visibility of the selected layers can be changed either by a simple check box or some interfaces can change the transparency level.

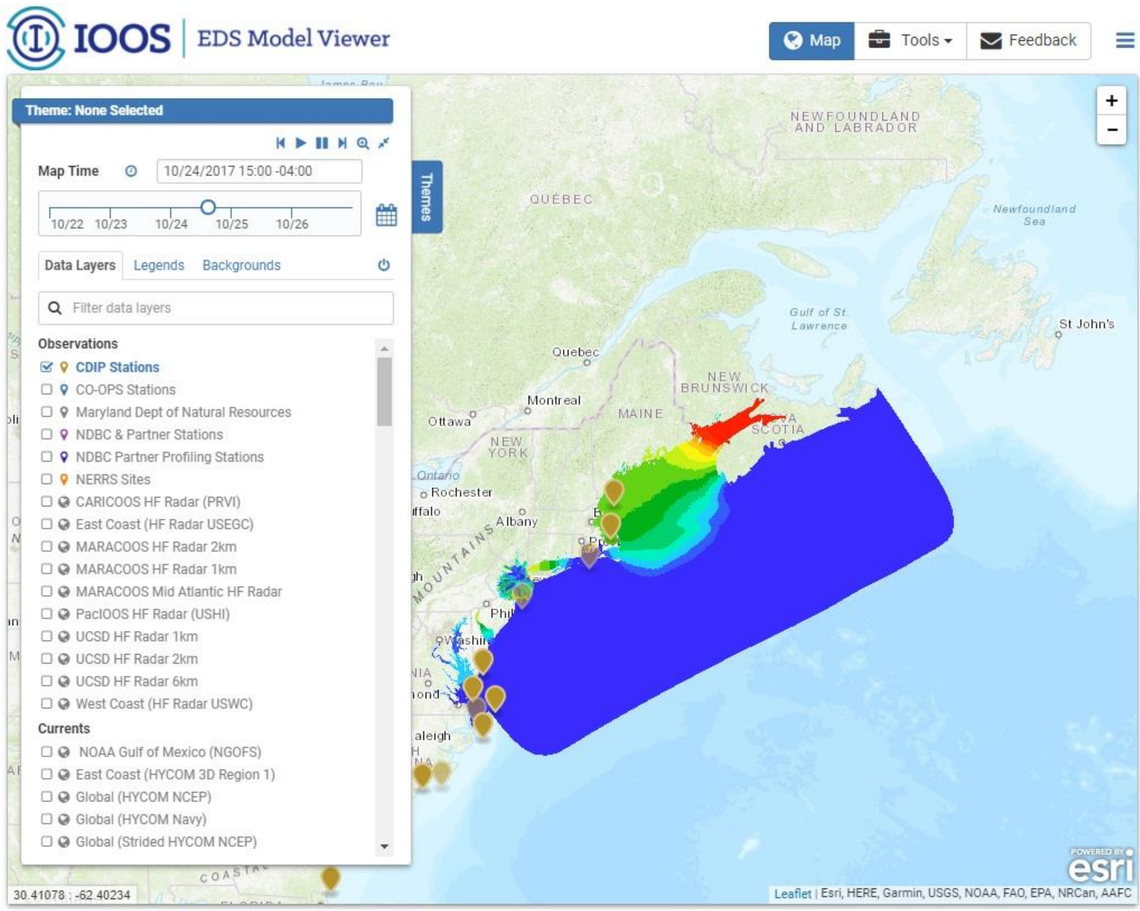


FIGURE 5. EXAMPLE OF A STANDARD GEOGRAPHIC INFORMATION SYSTEM VIEWING TOOL WITH MULTIPLE SELECTABLE LAYERS IN THE IOOS EDS MODEL VIEWER WEB APPLICATION.

More data information is normally available when clicking on any feature on the map. When added layer types differ greatly and multiple layers may overlap, normally only one information, either the last added layer or an amalgam of available data or layers, is displayed.

The filters for layers and layer-related information links vary according to the homogeneity of the layers available. An interface dedicated solely to bathymetry can afford to offer a single link to a general context. It is also possible to offer the same filters and options than those of the catalogue if the interface is flexible enough, although it is rare that a GIS interface integrates all the data available in the catalogue. A same interface can be used for different themes, by limiting the available layers of data and the search and filtering features available.

It is necessary to have access to the data to apply client-side symbiology. Otherwise, it is possible to use mapping services to provide generic symbiology according to the layer information.

It is also possible to use only metadata, if it includes the geospatial extend of the data. However, this option limits the interface to a basic geographic representation of a catalogue rather than a full-fledged visualization tool.

4.3.2. Sensor map visualization interface

Since several types of data are collected by instruments installed on fixed stations, or sensors, mapping these instruments' location and giving access to their data is another popular interface. To represent the sensors, the use of a client-side symbiology is most often used. Since the rendering is done on the browser, it is possible to change the visual used when selecting a sensor or to cluster the icons according to the actual zoom level.

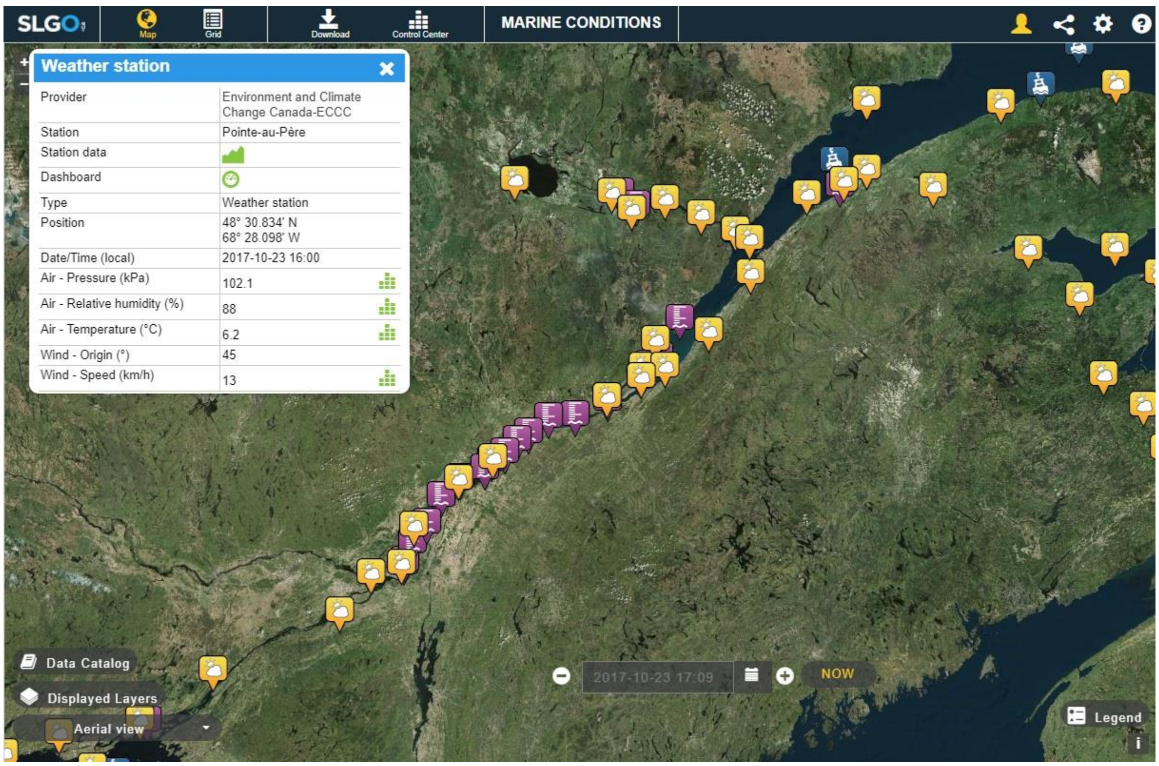


FIGURE 6. EXAMPLE OF TIDE GAUGES AND WEATHER STATIONS ON THE ST. LAWRENCE WITH LATEST INFORMATION FROM POINT-AU-PÈRE STATION ON SLGO'S MARINE CONDITIONS WEB APPLICATION.

Access to the station's data is done either directly in a pop-up window when it is clicked on the map, or following a link towards data access in that same pop-up window. Displayed data could be archived data or near real-time, but rare are the interfaces displaying real-time data directly on the map. Although possible, it is more demanding on the computer and software infrastructure side. To overcome this, it is best to collect all data in a single server in a way to provide a single access point rather than call every data provider. This approach quickly becomes necessary with a bigger number of sensors and variables to display.

Common available filters are sensor type (ex. CTD) and collected variable type (ex. water temperature).

4.3.3. Moving instruments (wave gliders, trackers)

The map interfaces presented above are a good fit for fixed instruments and stations, however another kind of interface is necessary for instruments such as wave gliders and biotrackers, whose location move over time and their trajectory is relevant information to display. Whether with only position or more parameters, the main filter dimension is time. It is also often possible to select the project or mission associated. The trajectory performed by the instrument can be filtered by its departure or arrival location, and time. As location at a given time is collected, the interface can present either the complete path of the instrument or the trajectory for a specified time range.

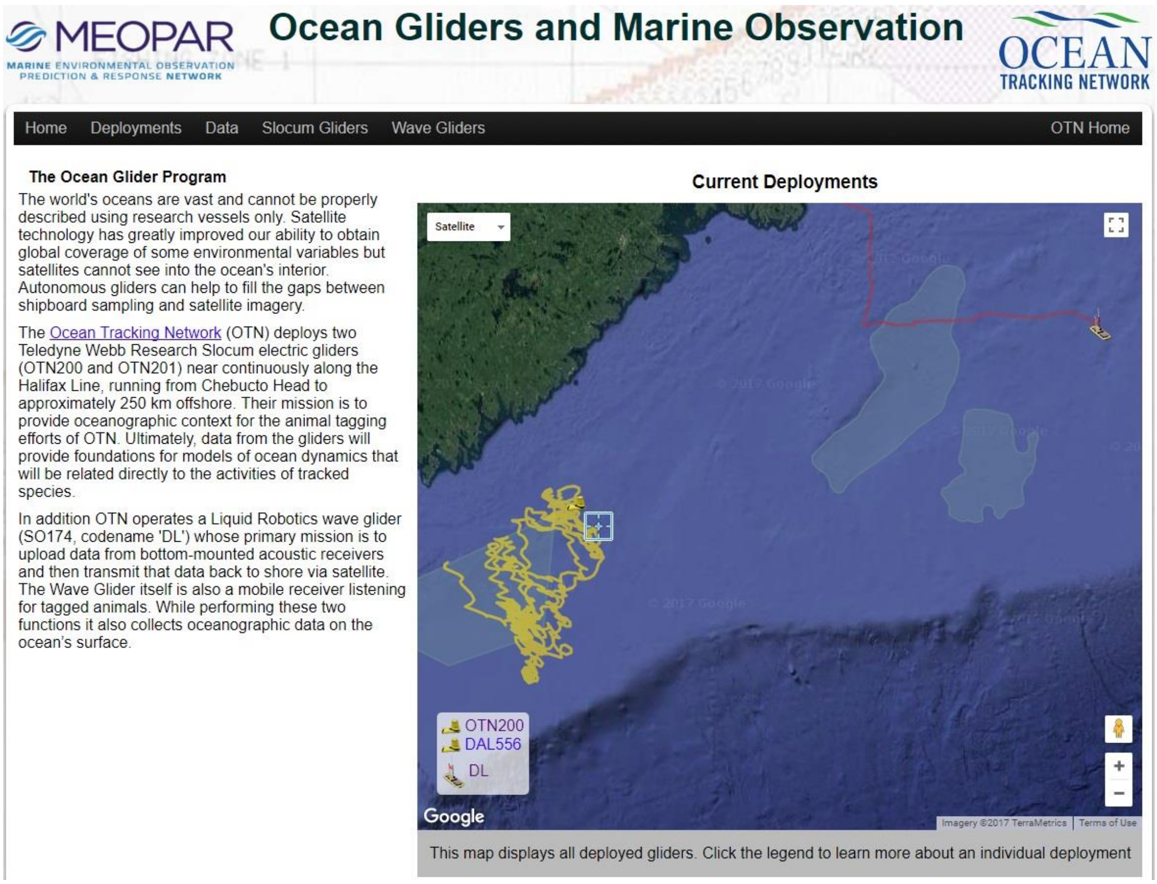


FIGURE 7. EXAMPLE OF A WAVE GLIDER BASED INTERFACE WITH MEOPAR OCEANTRACK.

In the case of biotrackers, species is an important filter since a single project usually contains several trackers on more than one species. It is uncommon that additional data are available other than the position according to time. The trajectory or dot observations are usually used, for a specific time range, to represent the evolving tracker's position. When a simple emitter is used to detect the specimen via fixed receiving stations, a list of points is preferred since the trajectory can't be derived.

4.3.4. Real-time data visualization interface

Real-time data visualization interfaces tend to allow the user to choose an instrument and display all its variables. Seldom, some interfaces allow to choose a set of instruments and their associated variables. This former interface style is simpler to implement because of its less demanding software infrastructure: a single call is made for only one instrument and its variables, this smaller amount of data doesn't require to be merged to other datasets from potentially other data providers, and thus may be kept decentralized.

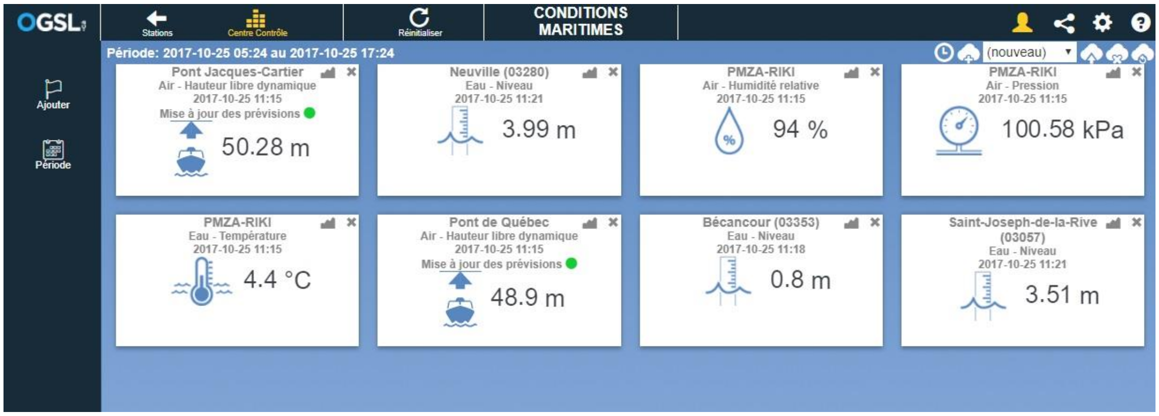


FIGURE 8. DASHBOARD FOR SPECIFIC REAL-TIME DATA OF SENSORS DEFINED BY A USER IN SLGO'S MARINE CONDITIONS WEB APPLICATION.

This interface can be a sub-interface of a principal sensor map interface, where clicking on a sensor allow to open a real-time data visualization dashboard. It could otherwise be an independent interface where selection of instruments is available via a menu. Since the infrastructure demand is heavier for real-time support, few interfaces display data for all the available sensors at the same time, but rather display a selection of by-default sensors from the complete list.

The use of real-time data, or near real-time data, allows for a more operational use of the visualized data. Pilots, port facilities, shipowners or recreational boaters are often very interested in this type of visualization, usually for restricted geographic areas and variables, which can guide the interface style details to meet their needs. For navigation purposes, forecast data can either be consulted in parallel or within the same interface.

Observation data can often be used to qualify and correct forecast data. Web-based data access is popular with niche application developers such as the Portable Pilot Unit developers, electronic navigation map or decision support application. These applications can meet much more constraining operational requirements and impossible to achieve in Web application.

4.3.5. Multidimensional model viewer

Interfaces could allow the extraction of volumetric data from oceanic prediction systems. However, we are not aware at the time of writing this report of other existing interfaces than the Ocean Navigator's, illustrated in Figure 9. Ocean Navigator allows end-users to define transects to extract data of the entire water column. It is also possible to define a section vertically by defining the desired depth and area. Choices of parameters and time are available filters.

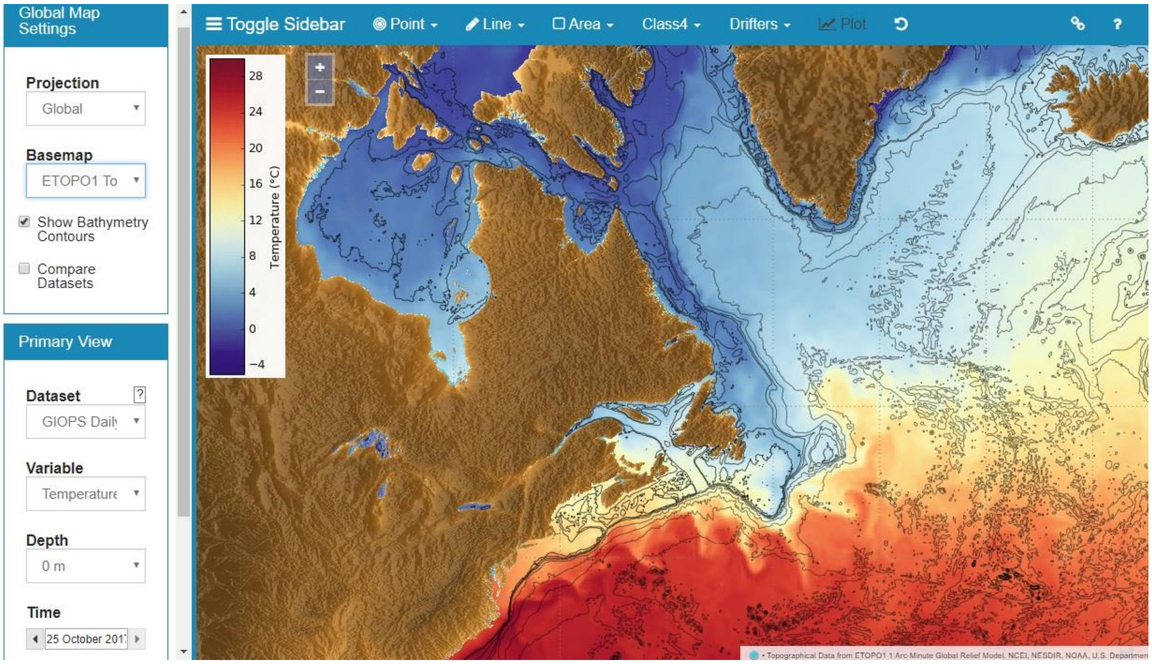


FIGURE 9. FISHERIES AND OCEANS CANADA'S OCEAN NAVIGATOR DATA EXTRACTION MAIN VIEW, ACCESSIBLE AT HTTP://NAVIGATOR.OCEANSDATA.CA/PUBLIC/

Since server-side demand is expensive for accessing and retrieving data of such large datasets, it is impossible, without a high-performance infrastructure, to allow some form of interactivity. Since rendering is done on the server side, static images can be returned by the visualization service. It is however possible to extract and download data in several visual or digital formats. Although the OPeNDAP protocol allows access to volumetric data, there is no mature and standardized protocol for visual access.

4.3.6. BIODIVERSITY WITH DATA ANALYSIS

A few interfaces display the results of a summary analysis on all the datasets. Parameters, filters and tools of the analysis are often specific to the scientific domain related to the interface's thematic. For example, this type of minimal analytic interface is often used in biology and biodiversity themed interfaces. Predefined parameters, such as the individual observation counts, the period (specific time bin), choice of species, sampling method, various habitats extend, are necessarily identified by the targeted end-user community, to ensure the presence of essential filters. Some biology portals give access to the database and let more tech-savvy users create their own query.

To be able to implement this kind of interface, that requires analysis through many datasets of many providers, the data needs to be centralized and formatted in a very standard way, so filters can be applied uniformly. Since the visualization rely on the server to analyse the data, interactivity is harder to implement since bandwidth or server load must be considered.

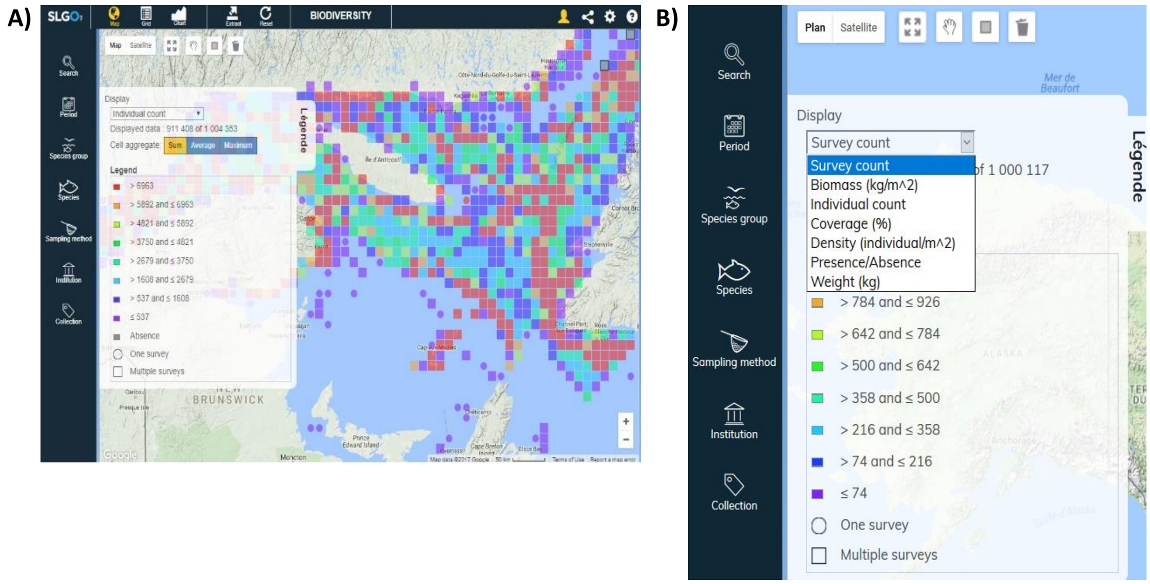


FIGURE 10. EXAMPLE OF DATA CLUSTERING AND ANALYSIS SERVER SIDE WITH SLGO'S BIODIVERSITY (A) WEB INTERFACE AND (B) SPECIFIC FILTERS ADAPTED TO END-USERS NEEDS.

4.3.7. Audio

Exploring large audio archives is a technical challenge. There are two proven techniques utilized at a few sites hosting and serving marine and ocean audio data. The first approach is to display spectrograms of the audio signal, where time is the x-axis, frequency the y-axis, and colours represent variations in sound intensity. Spectrograms can display data over seconds, minutes, hours, or even longer time periods, allowing a user to visually inspect the audio data for particular spectral signatures, thereby "finding" signal patterns of interest. The second, more sophisticated technique requires that the audio be pre-processed by passing it through a detection algorithm, which then catalogues the identified content in an annotation database. The database can be searched with standard query techniques. Both approaches are means for compressing the audio data into a manageable size for either screen display or content search queries.

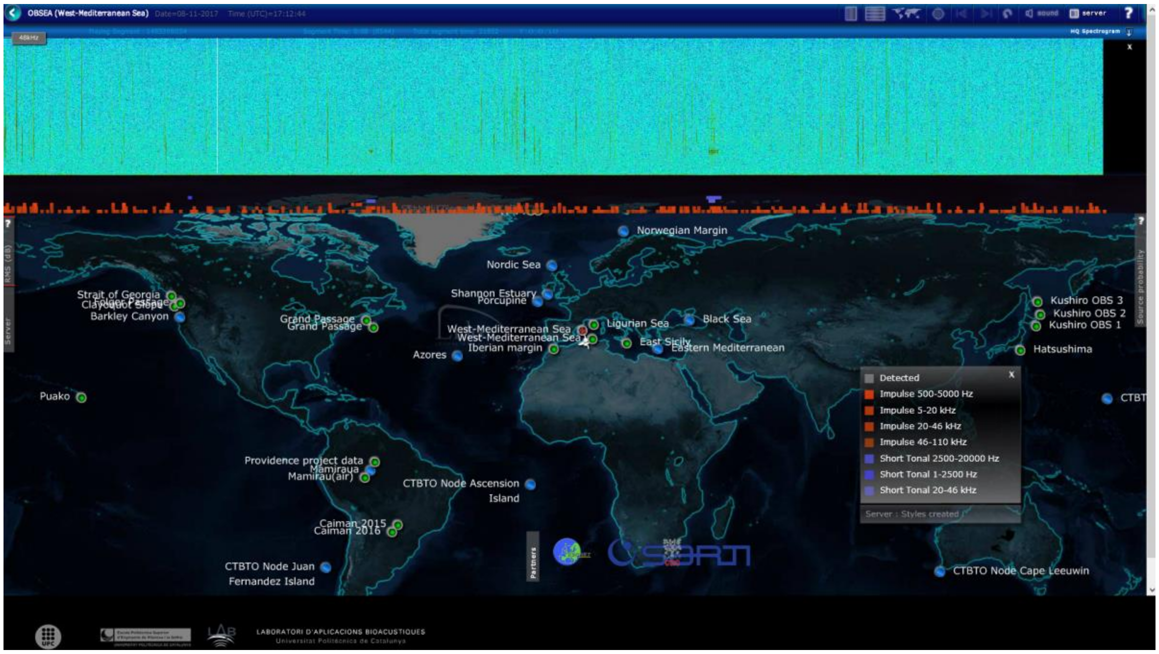


FIGURE 11. A SCREEN-GRAB OF THE LISTENING TO THE DEEP OCEAN (LIDO) WEB SITE, SHOWING A MAP OF AVAILABLE HYDROPHONES SITES, A COLOUR CODED SPECTROGRAM (UPPER PANEL), AND THE LOCATIONS OF DETECTED EVENTS FROM THE ANNOTATION CATALOGUE (THE RED BARS, MIDDLE PANEL).

4.3.8. VIDEO

Video is another challenging data type to support exploration and content discovery, since the data are extremely voluminous, and the content may be diverse. Shown below, in Figure 12, is a screen-grab of Ocean Networks Canada's SeaTube interface, where video has been ingested into an archive along with two critical ancillary data streams. First is the video metadata, which includes such information as the camera and video formats, but also the location (latitude, longitude, depth) and time of the video. This information is displayed both numerically and graphically (map) to orient the user to the location from which the video was collected. The second critical data stream is the associated annotation catalogue, (displayed centre-bottom here) which is either created during the video collection or afterward. The annotation catalogue describes the current content, but can be textually searched to skip ahead to video segments containing other occurrences of the desired content. Users can also contribute their own/updated annotations for extended content cataloguing and future discovery.

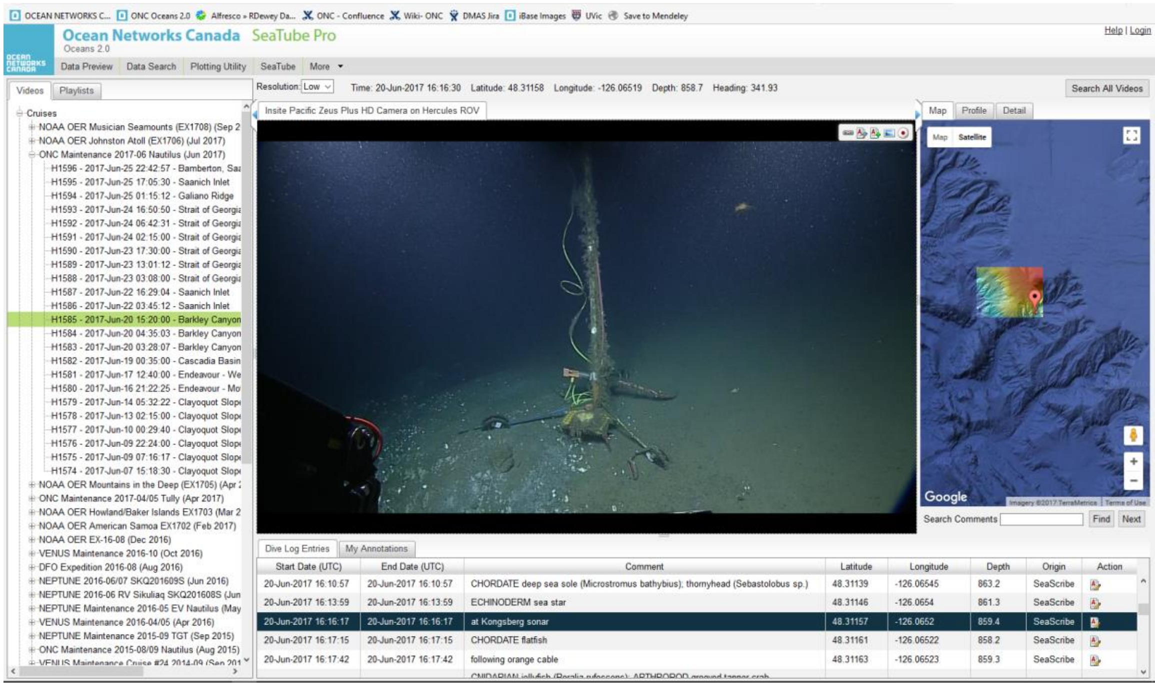


FIGURE 12. OCEAN NETWORKS CANADA'S OCEANS 2.0 SEATUBE INTERFACE, PROVIDING A MEANS FOR VIEWING, EXPLORING, AND SEARCHING EXTENSIVE ANNOTATED VIDEO ARCHIVES.

4.3.9. Animated and interactive interface

The technical capabilities of browsers and mobile devices have evolved rapidly, and some mainstream interfaces now use animation and interactive tools. They now include graphical features formerly mostly seen on native desktop applications. However, to optimize their display, they currently limit data resolution and require a greater server-side data manipulation. The aesthetics are a major asset for general public end-users and citizens, despite current necessary data alteration. An example of such an interface is illustrated in the Figure 13 showing the Ventusky interface.

While the quality and resolution of data aren't adequate for Ocean Observing system, some rendering techniques like current streams and client side and interactive rendering of almost every data visualization are noteworthy. Those technique will be expected by the users sooner rather than later. They are easier to implement and facilitate the development of responsive layout for mobile devices.

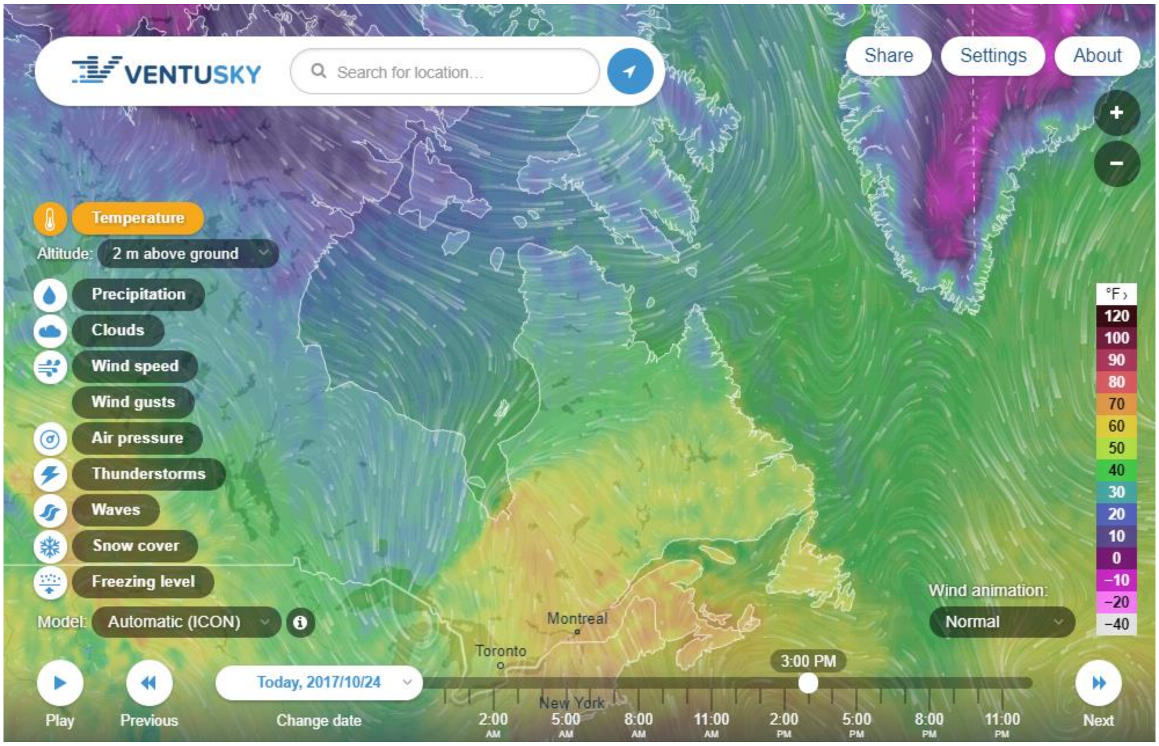


FIGURE 13. EXAMPLE OF A FULLY ANIMATED AND INTERACTIVE WEB INTERFACE WHERE EVERYTHING IS RENDERED ON THE CLIENT BROWSER (ACCESSIBLE VIA HTTP://VENTUSKY.COM/).

4.4. DATA PREVISUALIZATION

Four main previsualization types encompass all the variables: tables, 2D line graphs, gridded maps and static images. The data format and the interface used to visualize the data dictate which pre-visualizations are available. In some cases, two variables can be combined to form vectors. Only video, audio and multidimensional data requires other special interfaces and/or hardware, but they can have some form of previsualization via images. For example, video previewing can be done using still frames, compressed versions or time-lapses. For other complex structures or multi-dimensional datasets, 2D projections or slices of data from a selected transect can be used and served via static images.

4.4.1. TABULAR DATA

On an interactive tabular interface, the user expects to be able to sort the columns and be able to copy and paste the data, or generally download the table in various formats. In some cases, the dynamic addition or removal of columns is available. Pagination of data can be explicit for a fixed length table, or implicit as more rows appear while scrolling. For data preview purposes, pagination might not be necessary as only a subset of data would be displayed. Sorting while having explicit pagination is more complex to implement depending on the volume or structure of the data.

Date/Time (local)	Depth (m)	Water - Temperature (°C)	Water - Salinity (PSU)	Water - Level (m)
2017-10-22 13:15	6.09	5.3	26.18	2
2017-10-22 13:18	6.13	5.3	26.18	2.04
2017-10-22 13:21	6.16	5.3	26.18	2.07
2017-10-22 13:24	6.19	5.3	26.18	2.1
2017-10-22 13:27	6.23	5.31	26.18	2.14
2017-10-22 13:30	6.26	5.31	26.19	2.17
2017-10-22 13:33	6.31	5.31	26.19	2.22
2017-10-22 13:36	6.35	5.31	26.19	2.26
2017-10-22 13:39	6.39	5.31	26.21	2.3
2017-10-22 13:42	6.43	5.31	26.2	2.34
2017-10-22 13:45	6.48	5.31	26.2	2.39
2017-10-22 13:48	6.51	5.31	26.2	2.42
2017-10-22 13:51	6.55	5.32	26.18	2.46
2017-10-22 13:54	6.59	5.32	26.18	2.5
2017-10-22 13:57	6.63	5.32	26.18	2.54
2017-10-22 14:00	6.66	5.32	26.18	2.57
2017-10-22 14:03	6.69	5.32	26.16	2.6
2017-10-22 14:06	6.73	5.33	26.15	2.64

FIGURE 14. EXAMPLE OF A TABLE VIEW OF MULTIPLES VARIABLES INCLUDING WATER TEMPERATURE FOR A SENSOR IN SLGO'S MARINE CONDITION WEB APPLICATION.

4.4.2. 2D LINE GRAPH

The combination of one or more variables on the same interface can be represented by 2D line graphs. Users often have the choice of the variables to graph. Colored lines as well as interactivity allow to identify the variables that may overlap. X-axis zooming, value visualization on click or hover are other features commonly available (see Figure 15). Another way to present multiple variables is vertically stacking them (see Figure 16). In that case, a static image is used, and may include some interactivity.

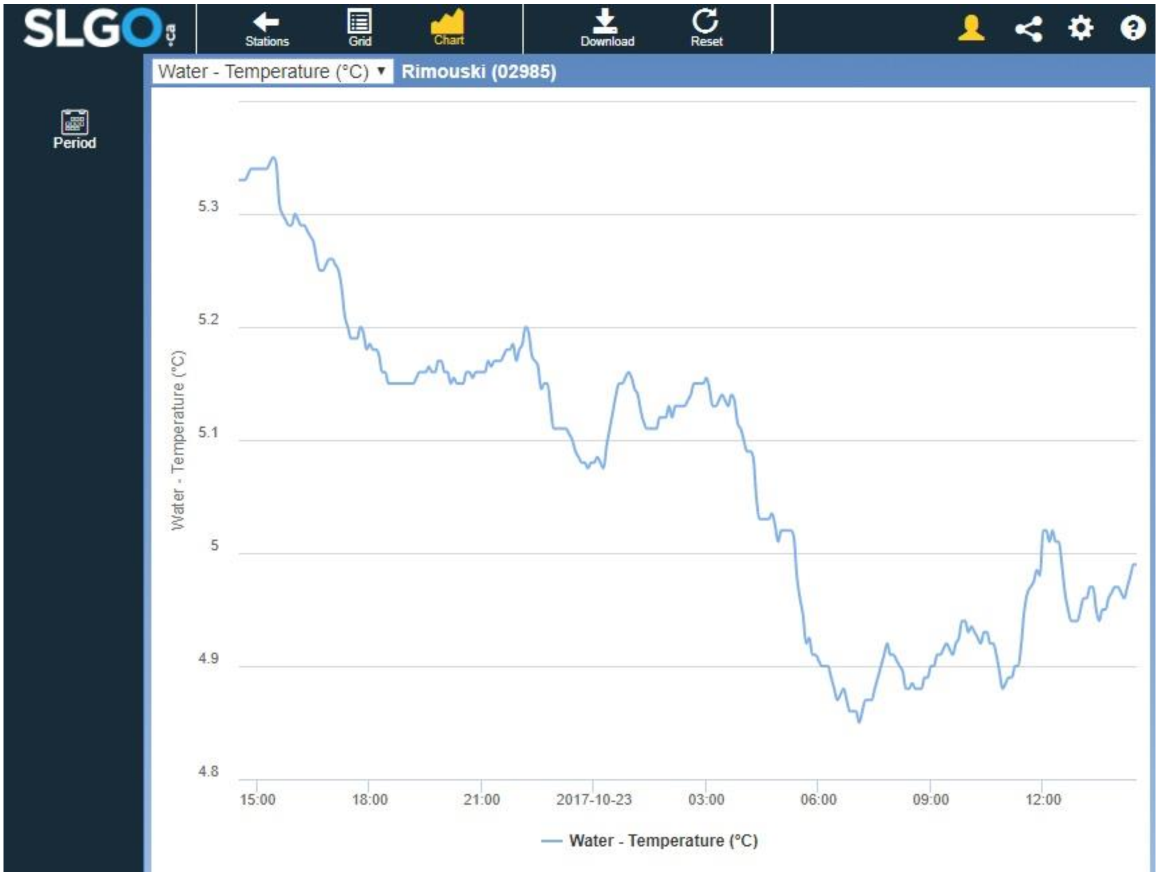


FIGURE 15. EXAMPLE OF A 2D LINE GRAPH (TIME SERIES) OF WATER TEMPERATURE FOR A SENSOR IN SLGO'S MARINE CONDITION WEB APPLICATION.

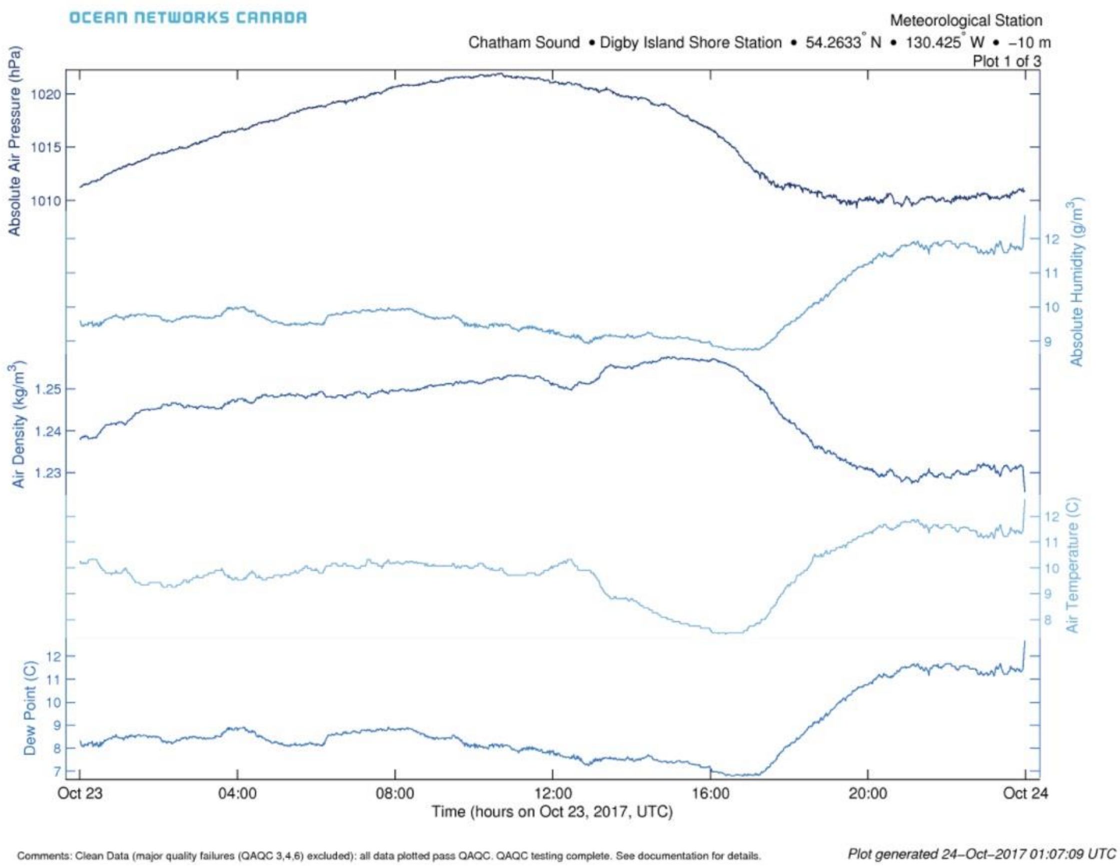


FIGURE 16. EXAMPLE OF MULTIPLE LINE GRAPH ON THE SAME X AXIS (TIME) ON ONC'S OCEAN 2.0 DATA PREVIEW.

4.4.3. GRIDDED AND VECTOR MAP

Most interfaces displaying a map are interactive. Moving and zooming are basic available features. The base map, or background map, can often be changed with a few different topographic options, or satellite imagery. Satellite background base map and political and street maps are normally available. Rarely, the choice of projection is available, although this feature is more commonly found on interfaces with specialized in polar data.

Some interfaces allow to modify other parameters such as time and depth. These parameters are standardized in the Web Mapping Service (WMS) protocol. Finally, when multiple layers can be displayed at the same time, layer transparency control is usually available.

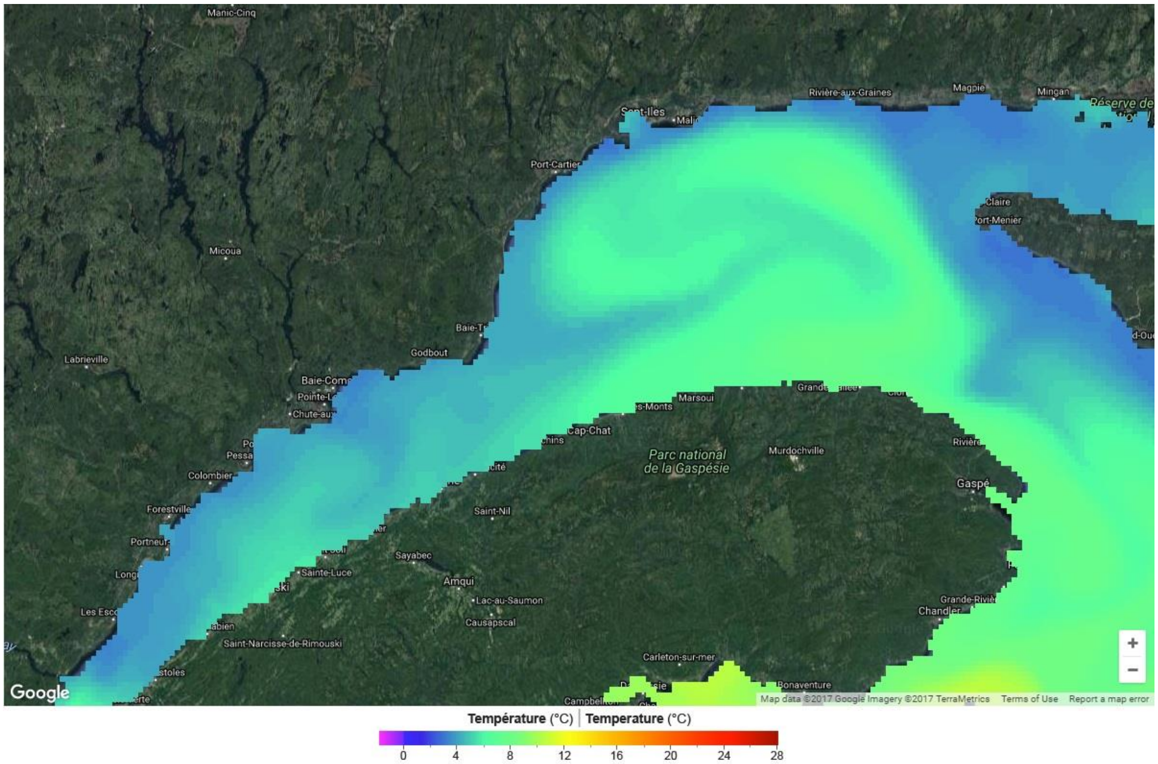


FIGURE 17. GRIDDED 2D MAP OF WATER TEMPERATURE FORECASTS FROM FISHERIES AND OCEANS CANADA AND ENVIRONEMENT AND CLIMATE CHANGE CANADA IN SLGO'S OCEAN FORECASTS WEB APPLICATION.

In most maps, more information is displayed upon clicking on any item of a layer (extend, symbiology, etc.), while clicking on any spot not linked to any activated layers won't do anything. The format of the data displayed in the former case differs greatly depending on data types, interface style, and users needs.

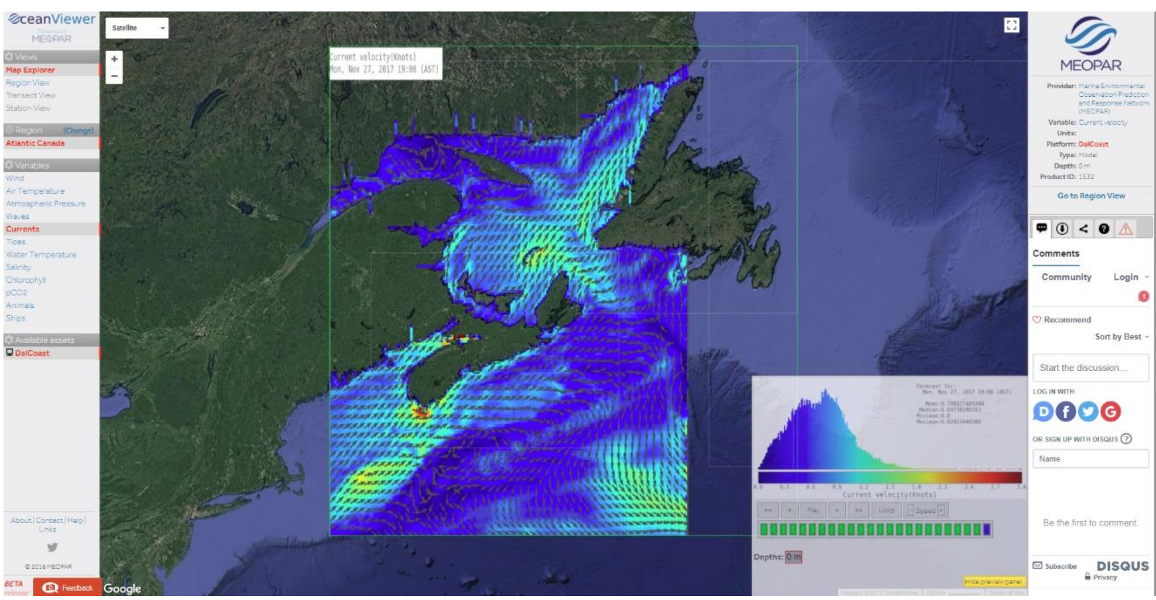


FIGURE 18. EXAMPLE OF WATER CURRENT DIRECTIONS AND SPEEDS REPRESENTED AS VECTORS ON MEOPAR'S OCEANVIEWER WEB APPLICATION.

Legend display depends on the number of layers that can be activated simultaneously. If only one layer of data is available, the legend is on or off the map. In the case of multiple layers, a dedicated window might list the active layers and a succinct version of their respective legends.

Some interfaces allow the direct download or print of current view of the displayed map. When available, all captions are then included in addition to certain metadata of the selected layers, as opposed to, for example, taking a screenshot of the view.

4.4.3.1. Client versus server-side map

When users can select a specific point on a map or interface, a client-side rendering allows to change the symbiology of the selected item. This technique is often used for sensor selection or when clustering is needed. Since all the data are transferred only once, this technique requires a limited amount of data in order not to overload the client browser. Some interface mixes maps with client-side rendering and server-side rendering.

4.4.4. Static image

If the rendering time is unacceptable for viewing data in an interactive way, it may be easier to generate a static image on the server side, when libraries or rendering software are not available on the client side, or other constraints make it impossible to render the data on the client side. In this case, an image is transferred containing the rendering of the visualization. Obviously, this static image does not allow interactivity. In other cases, the data itself is an image (picture, screenshot, satellite image, etc.).

Some basic interfaces use only static images that include legend, symbiology, attribution, or other relevant information related to the data. This method simplifies the development of the visualization interface and ensures that all data can be visualized at minima in static images, but at the obvious expense of interactivity. It should be noted that newer interfaces tend to favor interactive tables and graphics instead of images.

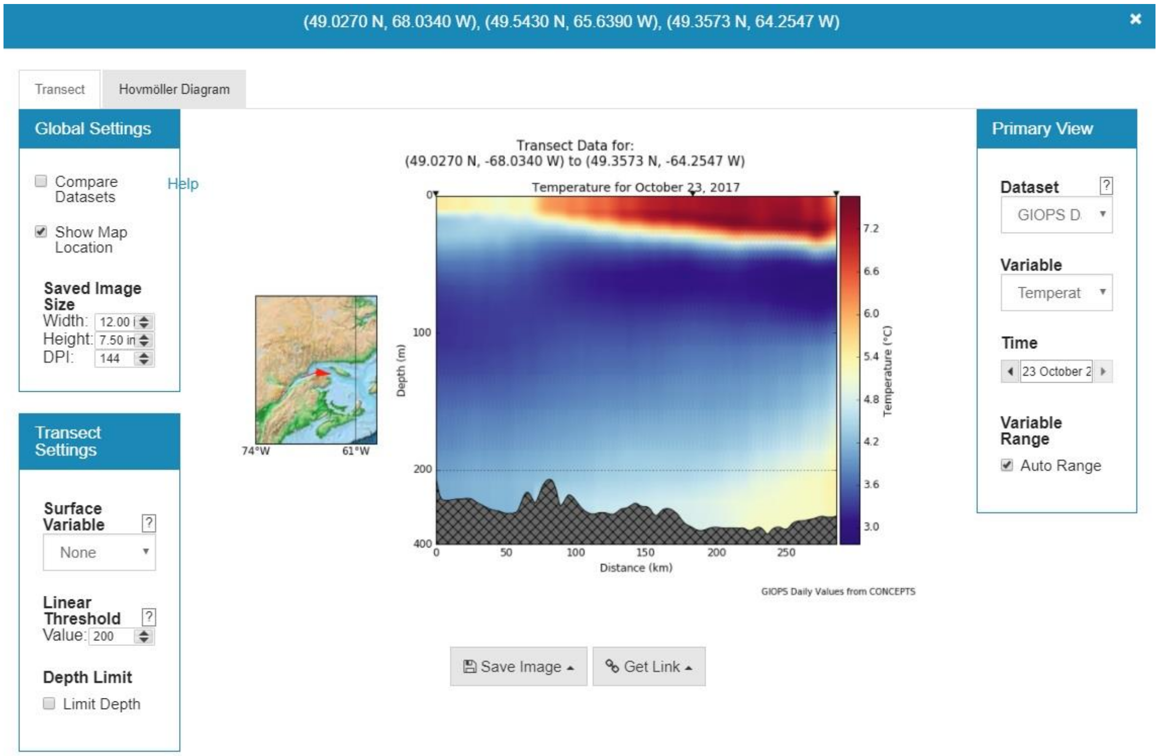


FIGURE 19. EXAMPLE OF WATER TEMPERATURE FROM A CUSTOM TRANSECT OR SLICE IN A 3D VOLUMETRIC DATASET FROM CONCEPT OCEAN-NAVIGATOR. THE DATA ARE RENDERED ON THE SERVER AND SENT TO THE CLIENT AS AN IMAGE.

In some cases, mainly where the information could be printed or shown on a high-resolution device, PDF files could be sent instead of an image. Since PDF has good support for vector-based rendering, file sizes can be kept small. It should be noted that PDF won’t be shown inside the interface but in another window or even application.

4.4.5. METADATA AND ATTRIBUTION VISUALIZATION

In conjunction with the visual representation of the data, some of the metadata such as attribution, quality flag, type or file format, reference or available download options can be displayed. An image (organization logo, protocol icon) is sometimes used in addition to the text, as illustrated in Figure 20. Depending on the interface and the manipulations required to display data, the metadata displayed is the only place to view these metadata and the data simultaneously.

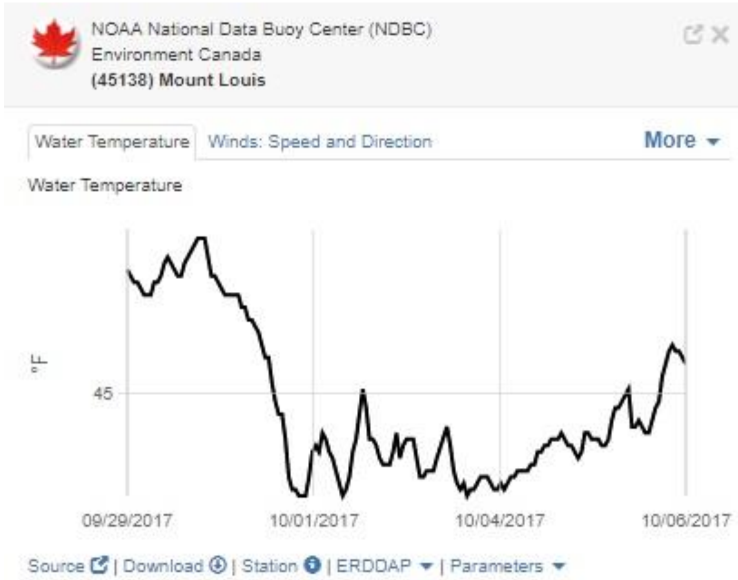


FIGURE 20. EXAMPLE OF PREVISUALIZATION OF TEMPERATURE DATA ON A 2D GRAPH WITH LOGO OF DATA PROVIDER ATTRIBUTION ON IOOS SENSOR MAP.

4.5. DATA ACCESS AND DOWNLOAD

There are several possibilities for accessing data. As with visualization, the data type, size, and the requirement manipulations on them dictate the most efficient and effective way of transferring information. It is possible to offer several types of access according to various user needs.

For smaller datasets, under 1 GB, and in widespread formats, they can be placed in a standard Web server, where a simple HTTP link will allow browsers to download it. If the file type is known by the browser, it can be opened and viewed directly within the browser. This method limits the download data, and increase the importance of dataset granularity. This URL link can be easily added to documentation or on web pages, and is especially appropriate for static data (images, pdfs, etc.).

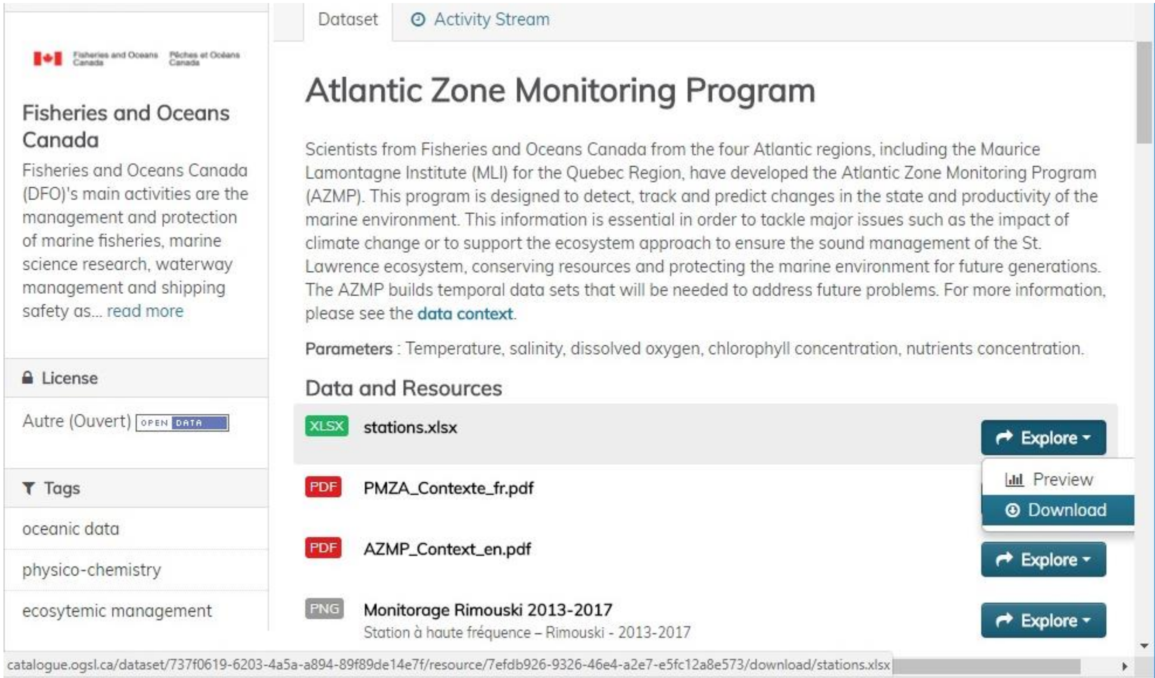


FIGURE 21. EXAMPLE OF A LINK TO DOWNLOAD A DATASET RESOURCES IN SLGO'S CKAN CATALOGUE.

An interface allowing the filtering of data to be downloaded is often necessary, rather than automatically downloading the complete dataset. An interface, such as ERDDAP's, illustrated below, can allow the user to limit the dimensions (time, coordinates) and extracted elements (variables), as well as to specify the file type of the extraction. The result of such a query is normally a direct download when the harvesting is done.

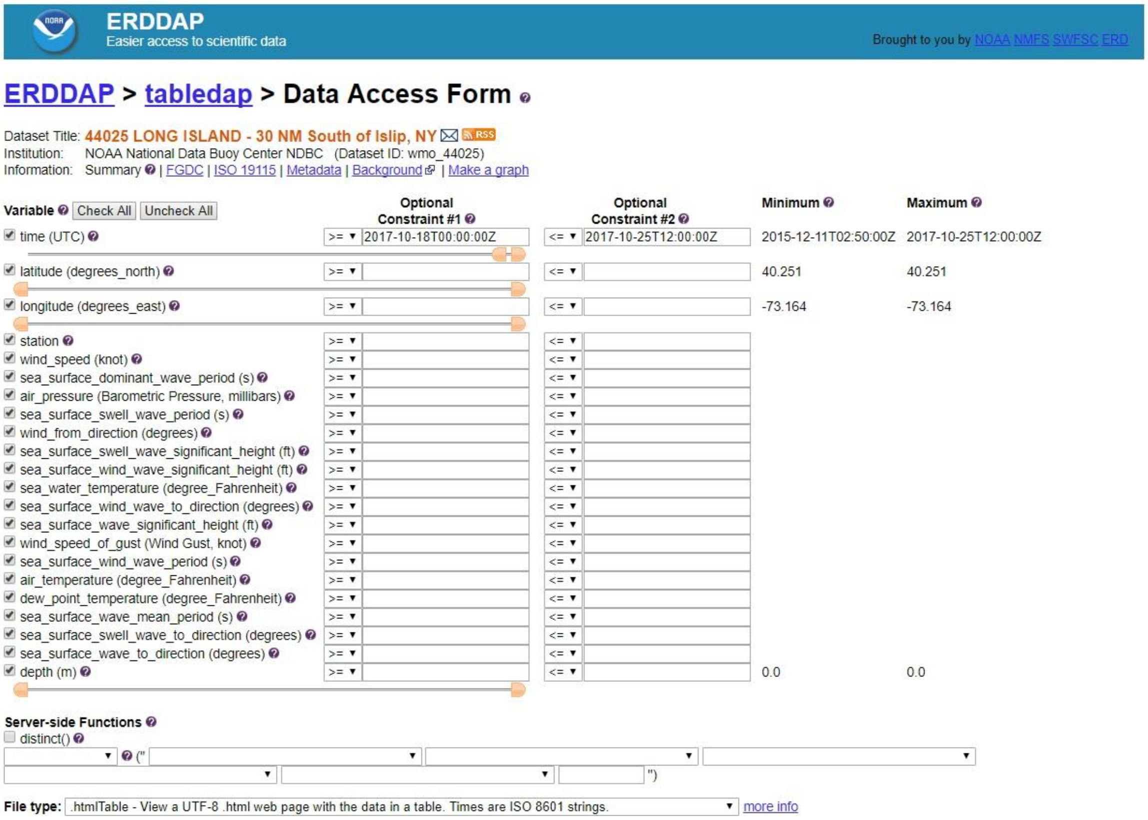


FIGURE 22. EXAMPLE OF ERDDAP DATA ACCESS FORM DATA OF A BUOY, FROM IOOS'S SENSOR MAP WEB INTERFACE.

To allow users to select multiple data sets and the servers to retrieve and format the data efficiently, an interface with a shopping cart logic can be used. The user can choose according to a list of available variables and parameters, for every dataset of interest. The 'shopping cart' concept allows for hiding a long and complex assembly and allows to further optimize the data assembly and download (see Figure 23). The user doesn't have to follow multiple links to download the chosen data.

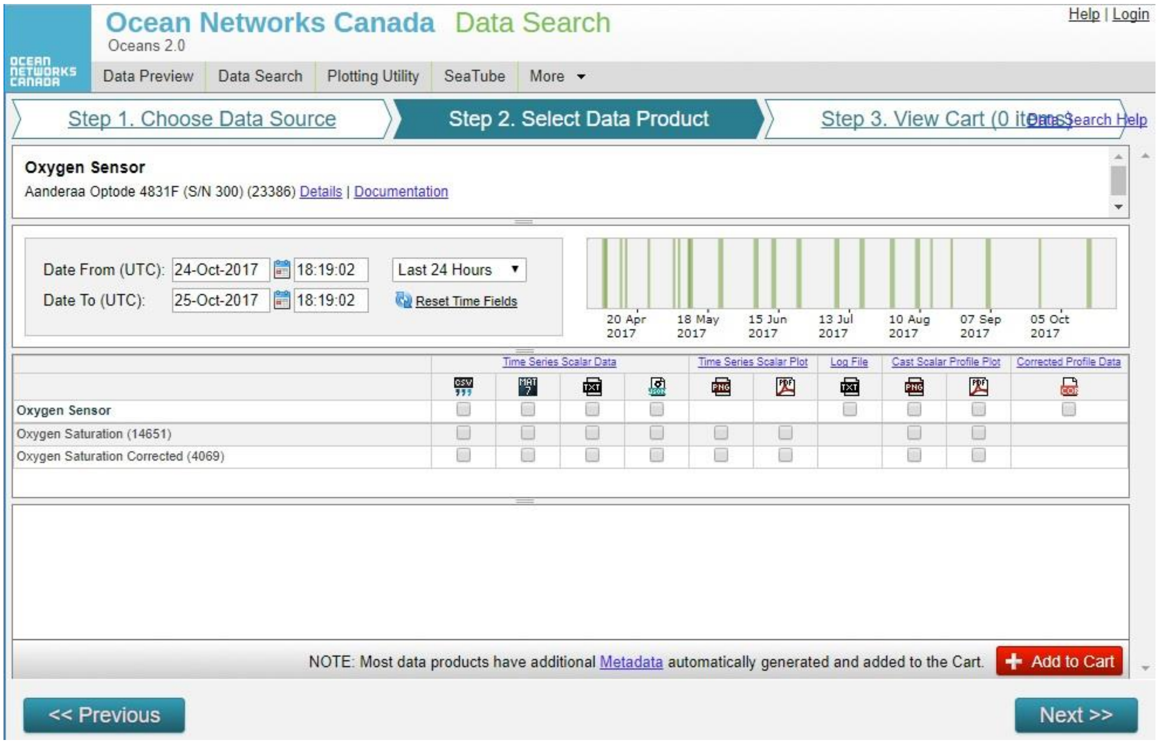


FIGURE 23. EXAMPLE OF A DATA SELECTION INTERFACE WITH A CART SYSTEM FOR DATA DOWNLOAD ON ONC'S OCEAN 2.0 DATA DOWNLOAD TOOL WEB INTERFACE. THE TOOL ALLOWS USERS TO SELECT MULTIPLE VARIABLES FROM MULTIPLE LOCATIONS AND AT THE SAME TIME SPECIFY THE DESIRED DATA FORMATS ARE OFFERED.

4.5.1. Link to extract or access protocol

For more tech-savvy or experienced users, for giving an access to a continuous stream of data or for a data access consumable by other computer systems, data access needs to be available through a standardized protocol such as OPeNDAP, or Web services. Obviously, the visualization available and the interface functionalities depend on the constraints of the service and data. It is important to visually represent this type of link and provide relevant, updated and complete documentation associated with the protocol or service.

4.5.2. Visual representation of types of link

Since several data extraction points may be available for the same dataset, it is important to correctly represent the type of access. It also allows to quickly indicate to the end-users which file formats are available to download. This applies both for extraction interfaces and direct data access interfaces. A standardization on the visual representing the formats, whether file formats or protocol types, should be implemented.

5. RECOMMENDATIONS: CIOOS IDENTIFICATION

5.1. REGIONAL ASSOCIATION AND DATA NODES

Each Regional Association (RA) will be responsible for managing the interactions with one or more data nodes. This includes data connections, permissions to disseminate, but also the necessary server infrastructure, cloud-based or not, to host and maintain metadata and in some circumstances the data. The RA is also in charge of providing web services for the CIOOS central node to harvest metadata and services for data pre-visualization and visualization. Each RA will manage the provision of their node's data, with curated metadata conforming to standards defined by the CIOOS. For metadata and data visualization, protocols and standard formats are required. Multiple paths or tools to implement protocols, should/must be made available, depending on the infrastructure already in place.

Aside from the data node management by the RA, other data nodes already exist within different organizations such as Fisheries and Oceans Canada or the Federal Geospatial Platform and these data nodes can feed the CIOOS central node directly or through one or more of the RAs.

5.1.1. REGIONAL ASSOCIATION DATA NODES

Regional associations (RAs) have the responsibility to curate and validate metadata so there will be minimal processing demands on the CIOOS central facility. RAs will manipulate the data and implement all the services required for previsualization. Since RAs will each represent a set of nodes, they are responsible for helping regional data producers implement the set of standards required for CIOOS, or if need be, to manipulate the data and metadata to make their integration with CIOOS compatible with required standards. For previsualization of rapidly changing data like real-time data, the RAs must implement automated ways to harvest, transform and conform the data to access standards. To enhance user interface consistency, RAs should not only give access to the data via CIOOS, but adopt the same CIOOS interface for data discovery and visualization. While not a prerequisite for the low-service model, this is recommended for RAs that do not already have a visualization interface that can be leveraged. However, RAs are encouraged to develop any other metadata interfaces and visualization services that can better serve their data and user base, as fully functioning RAs will unlikely be limited by data types and variables defined by CIOOS.

5.1.2. OTHER DATA NODES

Since multiple federal departments already have infrastructure in place for data discovery and extraction, they can easily become a data node if their infrastructure implements the set of standards and services required for CIOOS integration. With possibly minor adjustments, they could extend existing working services, or develop interfaces and services mapping already in place, while working either as proxies or through one or more existing Regional Association. With vast amounts of data already available in partial nodes such as DFO's or FGP's, their reusability through CIOOS needs to be encouraged.

Should certain types of data require too much work to achieve partial CIOOS compatibility, these federal department nodes may partially or totally transfer their responsibilities to a node making data available to one or more regional association(s).

5.2. General properties

Properties and features of the national CIOOS portal are described here in terms of generic branding orientations, and recommendations on data formats and services. These subsections are defined as general properties, that should be implemented notwithstanding the service model prioritized from section 5.3.

5.2.1. Branding

CIOOS will be a national and regional collaboration with three regions to begin and many partners who will contribute in different ways. This has the potential to complicate visual identity and branding. This justifies the need for branding and finding the right balance between specific RA look and feel, and a strong CIOOS branding.

Having a strong identity to CIOOS for RAs will contribute for CIOOS general recognition as a collaborative organization, provide a sense of belonging and credibility, both nationally and internationally, and eventually will facilitate access to future funding.

In the scope of this report, a reflection has been initiated around the question of what the visual branding of the CIOOS portal would look like and on which model would it be based. Indeed, a few scenarios have been identified partly based on some existing models, in the following ascending order of uniformity:

- Landing page only, followed by independent RA websites
- Similar header and footer
- Similar header, footer and main navigation menu
- Similar header, footer, main navigation menu and general layout structure of the page's content
- Fully identical

The first and the last options were rejected from the start as the former would not allow enough sense of branding and general sense of belonging to the CIOOS, while the later would not support enough customization of RA pages as well as preventing a sense of belonging for regional partners within a single RA.

Of the three remaining options, the third was favored as it would build a strong sense of belonging to CIOOS, build a common look and feel, while keeping a certain flexibility in the design and the content, allowing specific themes for each RA. This will in turn build a sense of belonging within each RA and allow to illustrate cultural specifics or fields of scientific expertise. However, it has been identified that it might be challenging in terms of reaching that common look, and that it might leverage less of existing work than other options.

5.2.1.1. Header

The similar header across RAs and CIOOS would allow recognition for a unified network of observing systems. For RAs, a map of Canada with the region highlighted could be an example on how to 1) identify the region and 2) point out that other regions can also be explored (see Figure 24). The header could also more specifically point out to other RAs, or back to the CIOOS webpage when on a RA webpage.

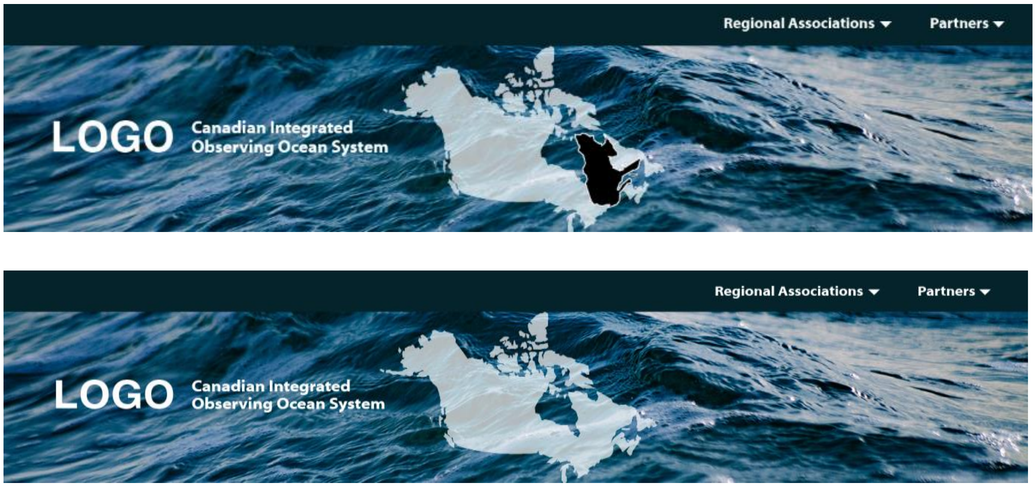
1)
2)


FIGURE 24. HEADER EXAMPLE FOR A (1) RA'S WEBPAGE WITH THE REGION HIGHLIGHTED AND (2) CIOOS WEBPAGE.

5.2.1.2. *Footer*

Footer structured on the same layout allow for quick and easy points of references, for common links tools that are usually located in that area:

- *(when on RA)* About us section with summary, mission, data producing members, composition of various managing bodies such as board of directors, scientific advisory committee, etc.
- *About CIOOS* section
- List of links to the other regional associations
- Contact form or links, social medias and newsletter subscription

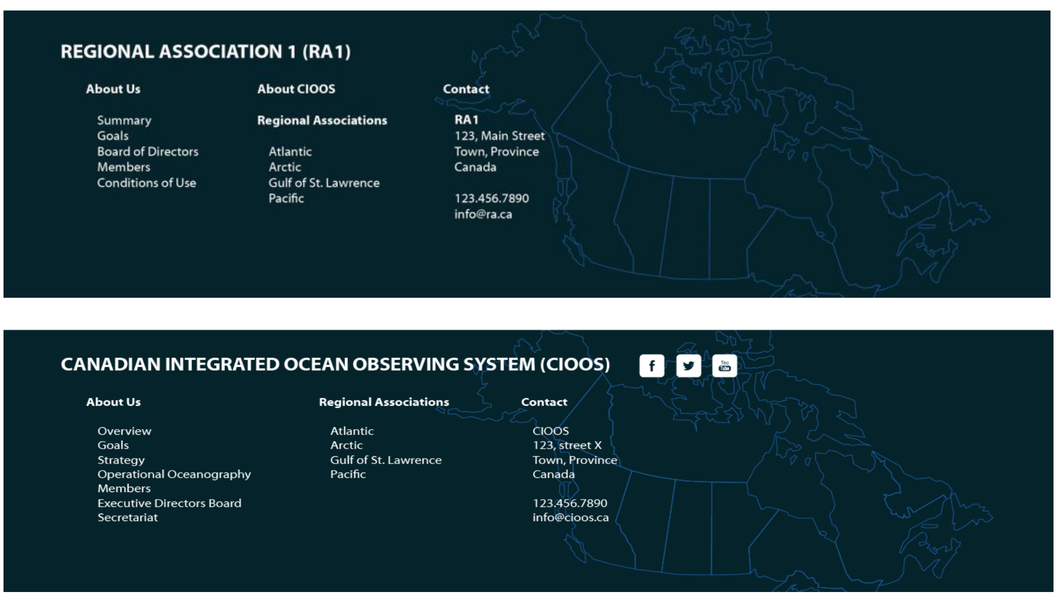


FIGURE 25. FOOTER EXAMPLE FOR A (1) RA'S WEBPAGE WITH AN ADDITIONAL "ABOUT CIOOS" LINK AND (2) CIOOS WEBPAGE.

5.2.1.3. *Main navigation menu*

The main navigation menu has been identified as a tool to contribute to the branding, while staying flexible enough to adapt to regional specificities. The whole menu is recommended to be located and using the same design across the RAs and the CIOOS main webpage to allow a baseline of uniformity. Basic sections (about us, news, data catalogue, language choice) are also recommended to be placed in the same order across the RAs and CIOOS. However, specific sections and pages, defined in each RAs, are free to added in the order they see fit. In the case of data viewers, according to the service model (see Section 5.3) that would be implemented, it could be possible that a common visualization tool would exist across the RAs and CIOOS. In that case, it should also be placed in the main navigation menu, at the same location to 1) be easily identified across the webpages and 2) lower the number of clicks before visualizing data. The same reasoning can be applied to the link for the data catalogue.

Note. A lot of visually pleasing portals displaying geospatial information now commonly have an interactive map directly on their homepage. As this method minimize the number of clicks to actually reach data, it should be looked into while designing CIOOS and the RAs homepages.

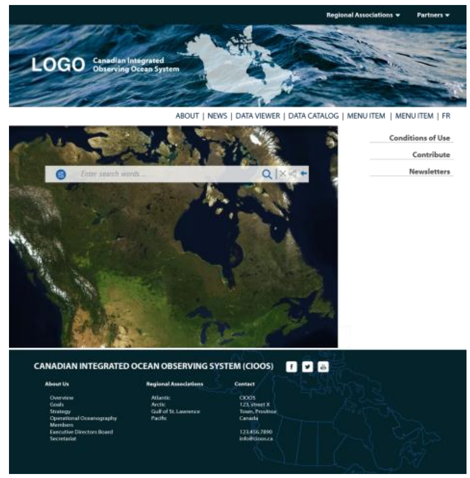
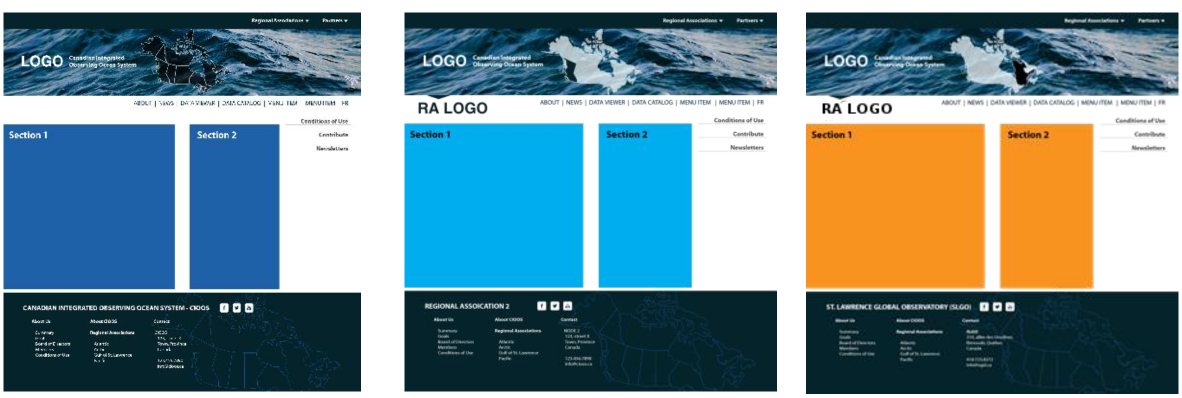


FIGURE 26. LAYOUT EXAMPLE FOR A CIOOS HOMEPAGE WITH AN UPFRONT INTERACTIVE MAP.

5.2.1.4. *HOMEPAGE CONTENT LAYOUT*

Finally, this scenario recommends using a generic similar structure in the layout of content for all the homepages of the RAs and CIOOS webpages. While this report doesn't specify what and how the details of the design should be, the purpose of this additional piece of uniformization is to maximize the sense of belonging. The content of the section however should be defined by each RAs according to their regional expertise or scientific focus, as well as being graphicly design in a way to meet their cultural regional identity, colors and symbols.



(1) (2) (3)

FIGURE 27.LAYOUT EXAMPLE FOR A (1) CIOOS HOMEPAGE AND (2)(3) RA HOMEPAGES SHOWING SIMILAR LAYOUT PAGES.

5.2.1.5. *COMMON VISUALIZATION TOOL*

For visualization tools that would be shared across RAs and the CIOOS national portal, a common branding should be implemented to acknowledge its belonging in the network. As such, CIOOS logo should be clearly identifiable near the RA's logo, in the same pattern for all RAs duplication of the tool.

5.2.2. Metadata catalogue

Because of the diversity of the audiences the CIOOS portal will have to serve, the interfaces will need to be flexible enough to accommodate the different users and their corresponding specific needs. From oceanographers to the general public, CIOOS interfaces should remain globally easy to use and intuitive to discover. Both interfaces and metadata must be available both in French and English, and any CIOOS help desk established must be able to respond to inquiries in these two languages as well. To enable data discovery, they should be easily explorable, for example with a thematic classification, where datasets may be tagged in more than one category. For example, SLGO's data catalogue organizes metadata according to a selection of six themes, identified after surveying their partner organizations as well as their Scientific Advisory Board. These six main themes were associated with the following subtopics, in a way to cover most of the fields of interest of their partners as well as to facilitate data discovery through subtopic keywords:

- **Biogeochemistry**; Nutrients, organic matter, pH, contaminants, primary production, dissolved oxygen, carbon sequestration
- **Biology and ecosystems**; Biodiversity, habitat, physiology, genetic, invasive species, species at risk, environmental stressors
- **Physics**; Currents, water level, bathymetry, temperature, ice, salinity, waves, submarine acoustic
- **Marine geology**; Sediments, sediment transport, granulometry, benthic imagery, minerals, dating
- **Coastal environment**; Coastal erosion, coastline, coastal habitat, conservation, shoreline restoration, river access
- **Human activity**; Maritime traffic, fisheries, aquaculture, aboriginal uses, tourism, climate changes adaptation

Other themes may be applicable for RAs in different oceanographic settings, or adapt more closely the core variables listed at the CIOOS level.

Finally, as a recommendation for all free-form text of the catalogue, the character encoding must correspond to the UTF-8 format.

5.2.3. Thematic visualization

Having custom visualization interfaces opens the possibility to add other visualization types more easily in the future. This is a mean to create interfaces that give answers and not just show data.

While basic access to data and visualization are possible with ERDDAP, for non-scientific users, a more user-friendly interface is required. Nearly all Ocean Observing Systems and organisations have a custom interface or have modified one already available. Most custom interfaces revolve around OpenLayers for the display and custom web services for layer selection. This kind of interface is a feature required by many casual or non-scientific users. While accurate scientific graphing can be available via ERDDAP, interactive 2D graph may be more suitable for many users. Different client-side libraries like chart.js and highcharts are available to implement simple interactive data manipulation. Web technologies and libraries are rapidly evolving, and so are end-user expectations.

Other visualization types for interactive exploration and manipulation of audio, video, multi-dimensional data, require custom interfaces where special software libraries may be in use. In particular, spectrogram viewers exist and are successful for exploring and searching audio archives. Audio annotation catalogues also accelerate searching audio archives. Similarly, video archives require integrated means for viewing and searching annotated video segments. There are existing examples of visualization interfaces for both audio and video data types. Only custom interfaces can fulfill some of CIOOS users' current and future visualization needs.

User expectations evolve in concert with Web technologies. As such, interfaces should be developed for modern browsers. This will allow CIOOS interfaces to stay relevant, use recent features and reduce implementation complexity. At the time of writing, modern browsers support HTML5, CSS3 and, at the very least, ECMAScript 5. In the future, a guiding, yet not all-encompassing, principle would be to use a market share cut-off percentage (e.g. consider browsers with market share below 3% as potentially not supported candidates).

The number of custom interfaces to develop and maintain is a main distinguishing feature between different service models (see section 5.3).

5.2.4. Data types and associated visualizations

There are numerous ways to visualize information, however only a few of them apply in a given situation, for a given data type. Considering the broad range of ocean core variables that the CIOOS aims to cover, it is important to recognize which circumstances calls for a specific visualization. The following subsections will briefly go over the common data types and their possible visualizations, as well as more unique ones and their specialized visualizations.

5.2.4.1. Time and date

Although UTC time is well understood by scientific users, local time or custom time zone is often more useful especially for users looking after real-time data. An advantage of rendering the interface visualization on the client side is the possibility of standardizing the display, while implementing the internationalization and custom time display depending on the device used or the local time. This is the responsibility of the interface's developer to clearly show which time zone the time and date are converted to if local time is used. Users should be able to select time display in UTC or not.

5.2.4.2. Fixed instruments

This is one of the most common forms of data. A typical example would be a buoy equipped with various sensors, making measurements at regular intervals and leading to time series measurements. This data can be easily visualized in tabular form, where the user can see the time of the measurement as well as the values of the associated sensors. A line graph showing the evolution of a sensor value through time (e.g. x=time, y=temperature) should be another simple visualization aid. As for map visualization, this type of data is usually represented through points on the map. Each point will have a color associated to a value on a given scale.

5.2.4.3. Moving instruments

There are use cases where sensors are attached to live animals or a moving platform (see Figure 28 for an example of shipboard mounted thermosalinograph). This type of data can be associated with gliders, drifters, vessels, animals, etc. As with fixed sensors, this can be visualized in simple tabular form, where time, position and other sensor values are displayed. For graphical representation, sensor values can be plotted against time. On a map, trajectories should be provided as a line for a predetermined time interval. Lines should have a clear start or end points, so users can infer the direction.

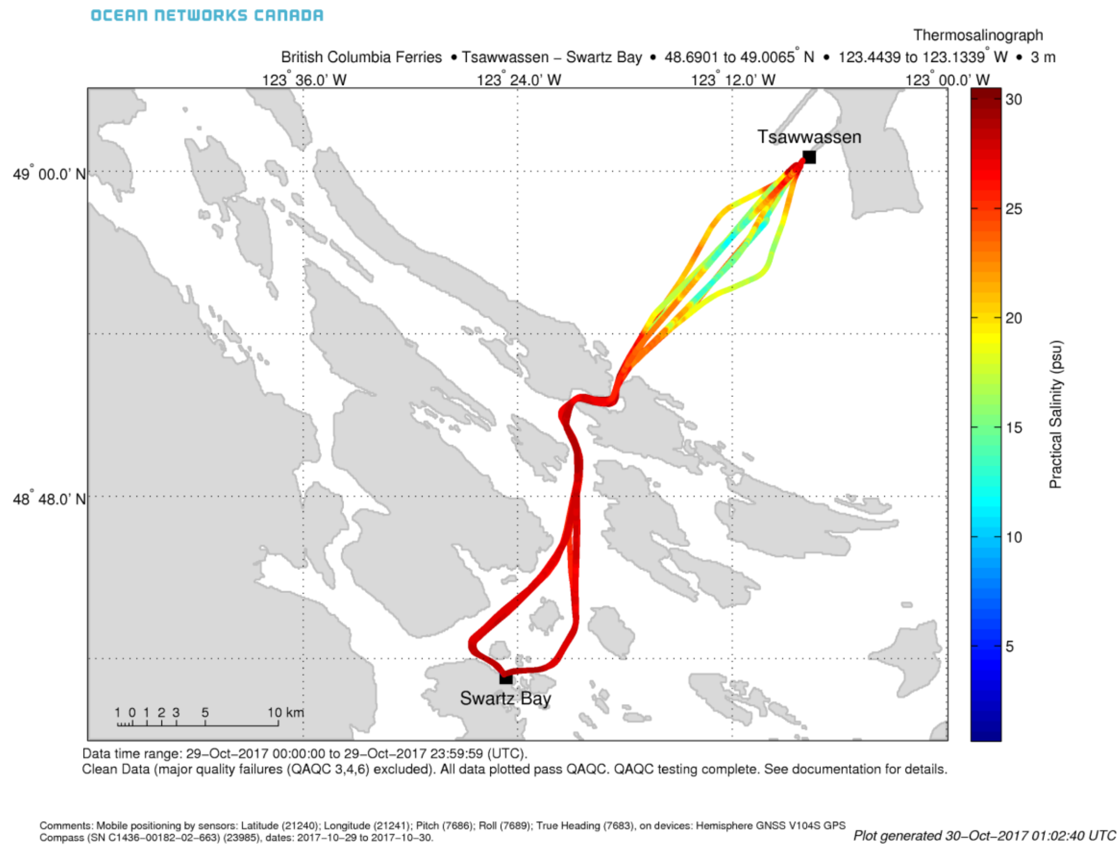


FIGURE 28. PLOT SHOWING THE EVOLUTION OF THE SALINITY VALUE IN THE SOUTHERN STRAIT OF GEORGIA ALONG THE TRANSIT PATH OF A BC FERRY GOING FROM VANCOUVER ISLAND TO THE MAINLAND SEVERAL TIMES DURING A DAY. PLOT COURTESY OF OCEAN NETWORKS CANADA'S OCEANS 2.0 SYSTEM.

5.2.4.4. MULTIDIMENSIONAL

Multidimensional data are becoming more and more prevalent and, as such, proper visualization must be provided. A flat representation of the different values can be presented in tabular form, although other means of visualization are necessary to extract more adequate insight. For 3-dimensional data, a 2-dimensional graphical representation may still be used. The graph provides 2 dimensions through classic axes and the third dimension is provided using a specific color from a given scale (see Figure 29 for an example).

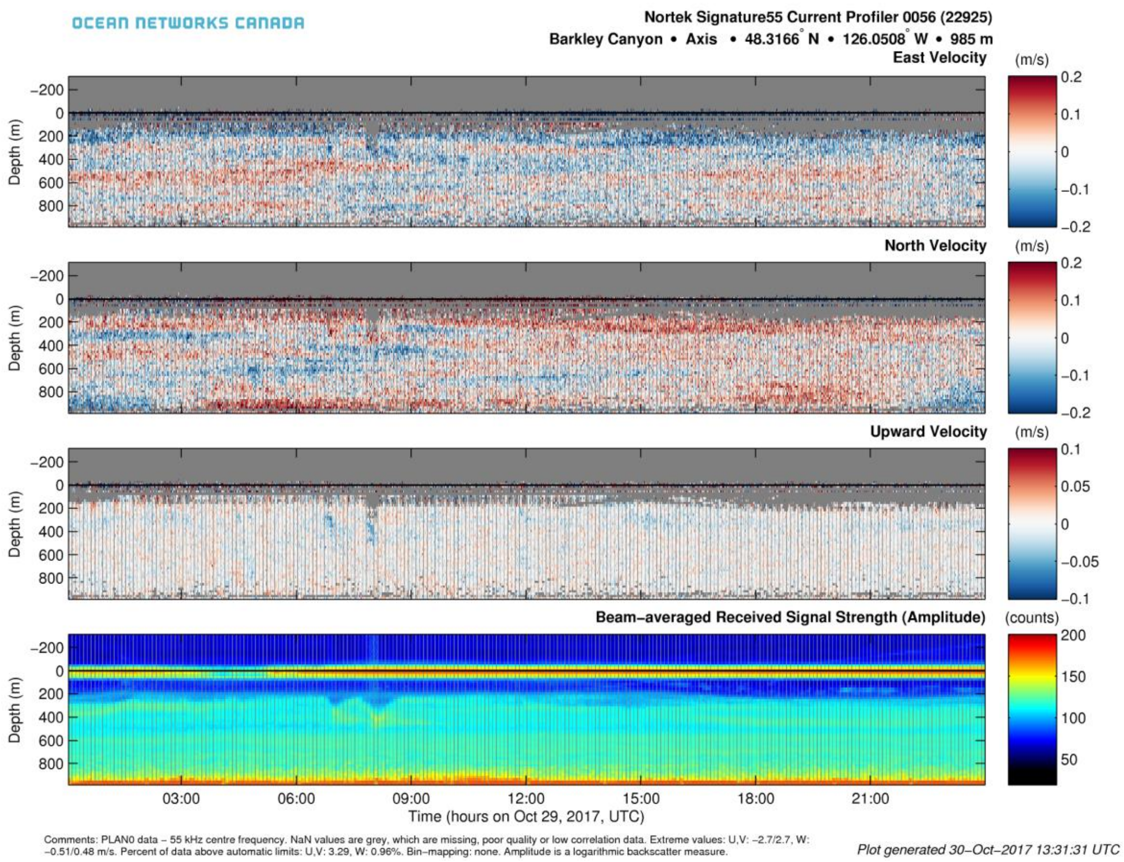


FIGURE 29. THE ABOVE PLOTS REPRESENT DATA FROM A SINGLE DAY OF OBSERVATIONS OF AN ACOUSTIC DOPPLER CURRENT PROFILER (ADCP) MEASURING 3 VELOCITIES (E, N, UP) IN ABOUT 130 DIFFERENT BINS BY ALMOST 1,000M DEPTH WEST OF VANCOUVER ISLAND. THE COMPLEXITY OF THE MULTI-DIMENSIONAL DATA SET PRODUCED BY THIS SINGLE INSTRUMENT EVERY 10 MINUTES REQUIRES THE USE OF MULTIPLE PLOTS AND COLOUR CODING ON TOP OF LINES AND VECTORS IN AN X-Y PROJECTION. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

Pure 3-dimensional representations of data on a map is currently not mature and standardized enough to be considered. What is possible, however, is to offer users the possibility to select or draw a transect on a map, and then obtain the graphical representation of that 3-dimensional data (see Figure 30 for an example).

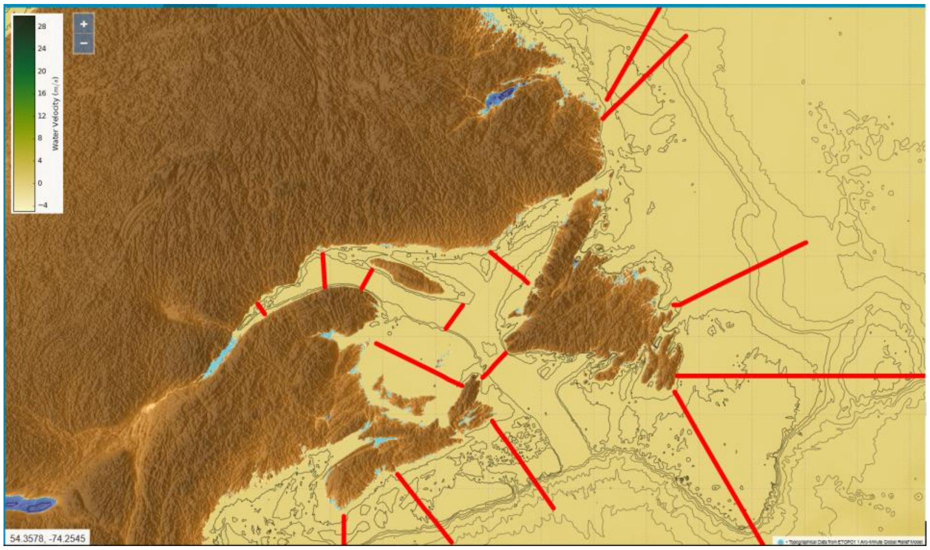


FIGURE 30. THE ABOVE IMAGE SHOWS VARIOUS TRANSECTS FOR THE ATLANTIC ZONE MONITORING PROGRAM (AZMP) IN THE OCEAN NAVIGATOR APPLICATION, ACCESSIBLE AT [HTTP://NAVIGATOR.OCEANSDATA.CA/PUBLIC/](http://navigator.oceansdata.ca/public/).

A 4-dimensional graphical representation is not as common. A potential example is when the additional dimension of direction (wind, currents, etc.) is involved, as illustrated by the black arrows in Figure 31.

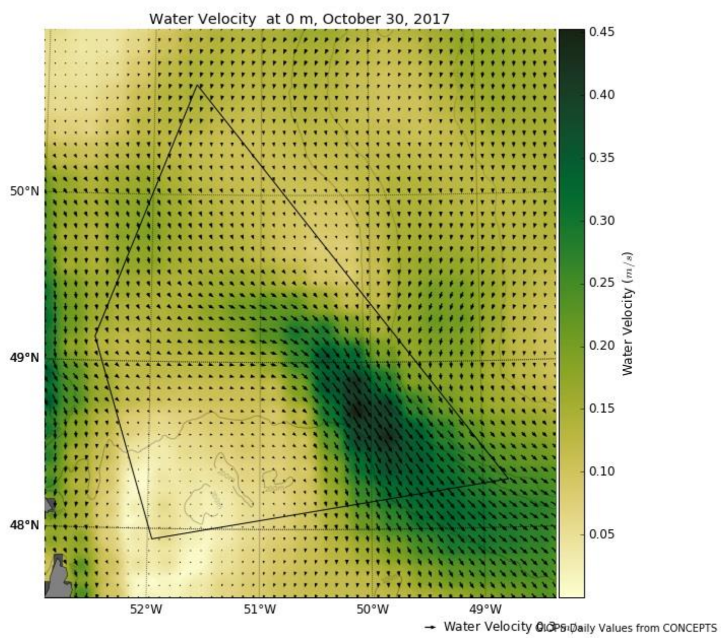


FIGURE 31. 4-DIMENSION VISUALIZATION OF WATER CURRENTS DIRECTION AND VELOCITY. THE 4-DIMENSION DISPLAYED ARE LONGITUDE (X-AXIS), LATITUDE (Y-AXIS), DIRECTION (BLACK ARROWS) AND VELOCITY (COLOUR CODING). ILLUSTRATION FROM OCEAN NAVIGATOR APPLICATION.

5.2.4.5. Gridded

Data points can be structured in regular fashion within a well-defined grid. Such gridded data are usually rich in information and better suited for visualization on a map. Common use cases are forecast models, satellite remote sensing, etc. The visualization should be done through a simple raster format and allow the possibility of clicking on any point to get its value.

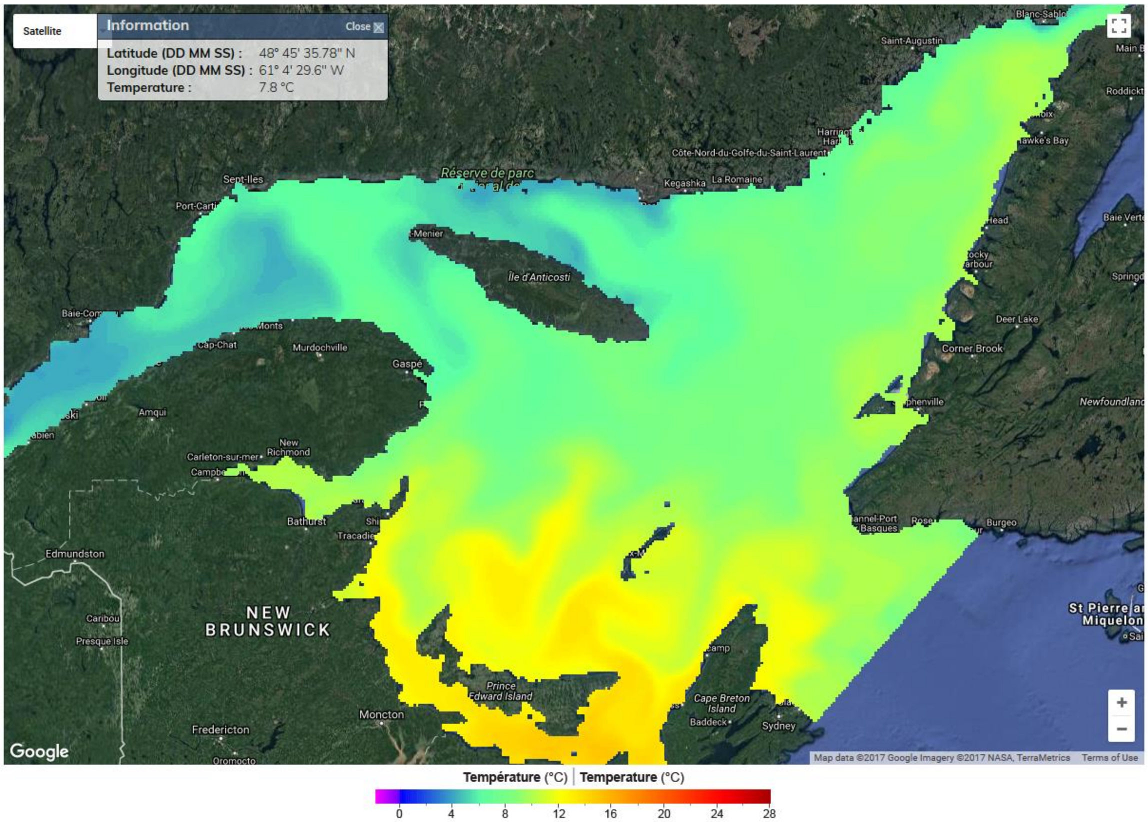


FIGURE 32. EXAMPLE OF GRIDDED MAP VISUALIZATION OF WATER TEMPERATURE IN THE ST. LAWRENCE GULF, FROM SLGO'S OCEAN FORECASTS APPLICATION.

The above example (Figure 32), from SLGO's Ocean Forecasts application, provides a good visualization of the water temperature within the St. Lawrence River and Gulf region. Any point in that region can be clicked to display accurate values and associated coordinates.

5.2.4.6. DIRECTIONAL

Some values are only truly meaningful when accompanied by the direction in which they act (or broken down into its directional components). Typical examples include wind speed and direction, current speed and direction, current x velocity and current y velocity, etc. As is customary, the various values may be displayed inside a table. For graphical representation multiple options are available; magnitude and direction can be separated into different time dependant line graphs and so can the vector components (e.g. velocity in x against time and velocity in y against time). However, this does not provide as much insight to end users and is not the recommended route. Better options would be to display both vector components as distinct lines in a time dependant graph or, similarly, display both magnitude and direction.

FIGURE 33. EXAMPLE OF LINE GRAPH FOR DIRECTIONAL DATA OF SURFACE CURRENTS SPEED (BLUE LINE) AND DIRECTION (GREEN LINE), FROM SLGO'S MARINE CONDITIONS APPLICATION.

The above line graph (Figure 33) from SLGO's Marine Conditions allows users to visualize both the current speed (blue line) and its direction (green line).

Representation of directional data on a map should remain simple and intuitive. Direction should be expressed using arrows, while magnitude should be represented by the size and color of the individual arrows (see example in Figure 34).

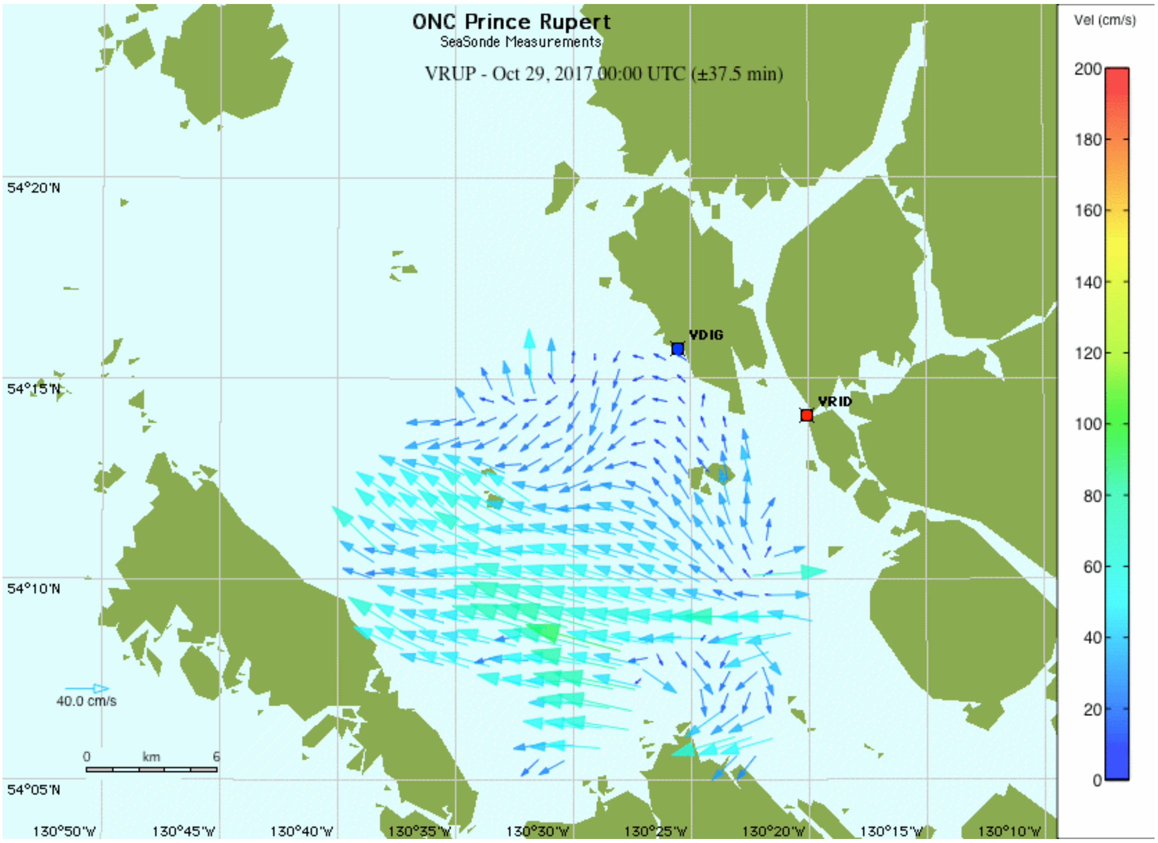


FIGURE 34. MAP VIEW OF SURFACE CURRENT EVOLUTION OVER A 24-HOUR PERIOD ENCODES DIRECTION WITH ARROWS AND INTENSITY OF CURRENTS WITH COLOUR CODING AND ARROW SIZE FOR EACH OF THE CELLS MEASURED BY THE TWO HIGH FREQUENCY RADARS (HFR) NEAR THE PORT OF PRINCE RUPERT, BRITISH COLUMBIA. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA'S OCEAN 2.0 SYSTEM.

5.2.4.7. POLYGONAL AND AREA

Whether it is to identify a protected area, delimit an aquaculture zone, or display administrative regions, polygons can be excessively useful (see Figure 35 for an example of a polygon delimited in a map interface by SLGO). While it is possible to present the polygon definition along with the associated values in tabular form, in map visualization this type of data takes on all its meaning. Vectors can be easily used to express polygons, which can then be rendered on an interactive map. Clicking anywhere within the defined region should provide the user with the associated values.

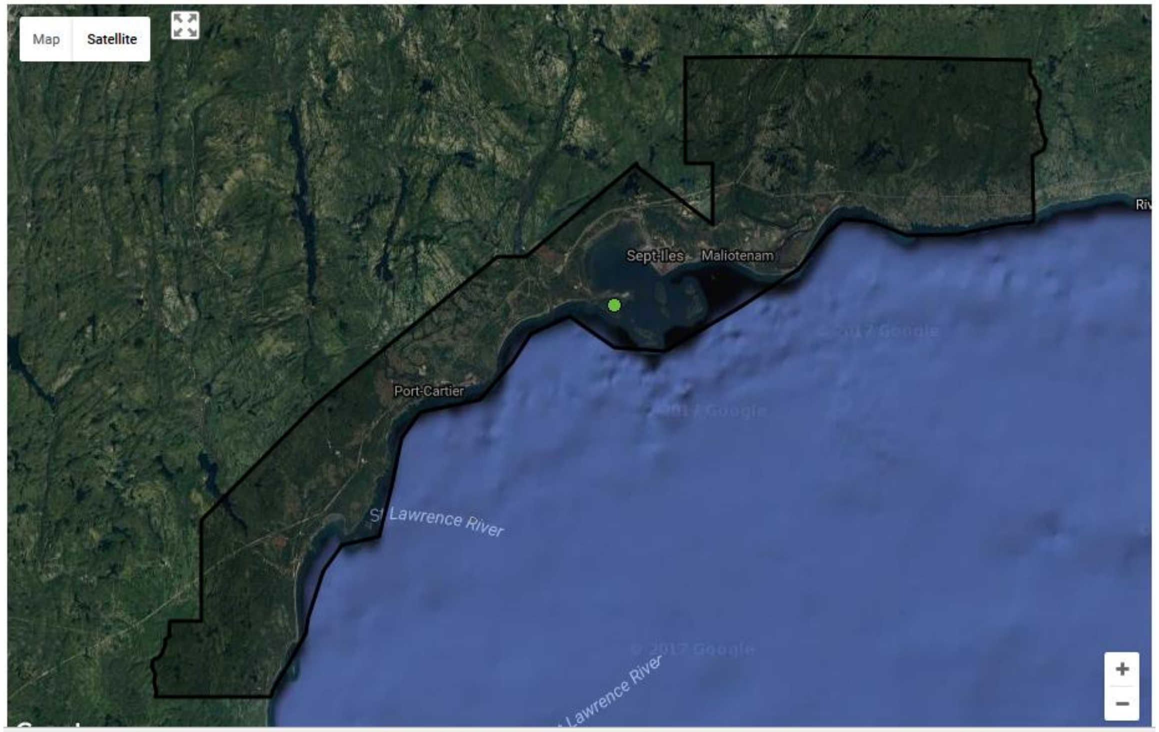


FIGURE 35. MAP INTERFACE SHOWING A POLYGON, DELIMITING A SPECIFIC REGION IN SLGO'S "ENVIRONMENTAL REFERENCES" APPLICATION.

5.2.4.8. AUDIO

Audio data, more frequently referred to as passive acoustic data, is difficult to search and visualize on its own. Phenomena of interest in each hydrophone stream are often rare and difficult to find. Listening to each stream is not only difficult due to the sheer volume of data but more importantly literally out of reach as many of the possibly interesting snippets in the signal are beyond the audible reach of the human ear. Two distinct approaches can help with identifying interesting data content. One recommended method is to employ software algorithms involving neural networks trained to detect well known phenomena to quickly scan many hours of recordings, identify expected signals and describe them in annotations, which can in turn be easily searched using standard database search methods.

Resource Type: Device Data

Resource: Ocean Sonics icListen HF Hydrophone 1

<< first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 30

ID	Campaign	Resource Type	Resource Name	Annotation Summary	Start Date (UTC)	End Date (UTC)	User ID	Flagged	Shared	Modified Date (UTC)
233080		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Humpback whale	01-Jan-2014 03:01:59	01-Jan-2014 03:01:59	21759	false	true	29-Jun-2015 21:50:15
233079		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Fin whale	01-Jan-2014 02:51:59	01-Jan-2014 02:51:59	21759	false	true	29-Jun-2015 21:50:15
233078		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Humpback whale	01-Jan-2014 02:51:59	01-Jan-2014 02:51:59	21759	false	true	29-Jun-2015 21:50:15
233077		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Fin whale	01-Jan-2014 02:41:59	01-Jan-2014 02:41:59	21759	false	true	29-Jun-2015 21:50:15
233076		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Humpback whale	01-Jan-2014 02:41:59	01-Jan-2014 02:41:59	21759	false	true	29-Jun-2015 21:50:15
233075		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Humpback whale	01-Jan-2014 02:31:59	01-Jan-2014 02:31:59	21759	false	true	29-Jun-2015 21:50:15
233074		Device Data	Ocean Sonics icListen HF Hydrophone 1251 (23157)	| ICLISTENHF1251_20 | Fin whale	01-Jan-2014 02:31:59	01-Jan-2014 02:31:59	21759	false	true	29-Jun-2015 21:50:15

FIGURE 36. EXAMPLE SHOWING TABULAR RESULTS FROM A SEARCH OF HYDROPHONE ANNOTATIONS. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

Another recommended method that will keep the "human in the loop" involves the simple generation of spectrograms representing various time periods of hydrophone data as visual images (see Figure 37). Spectrograms display time versus frequency and a colour coding represents the intensity. This is a convenient and common method, as humans are very good at detecting patterns in images. The images can then be annotated on their own for further search/reference, and linked to the original audio content.

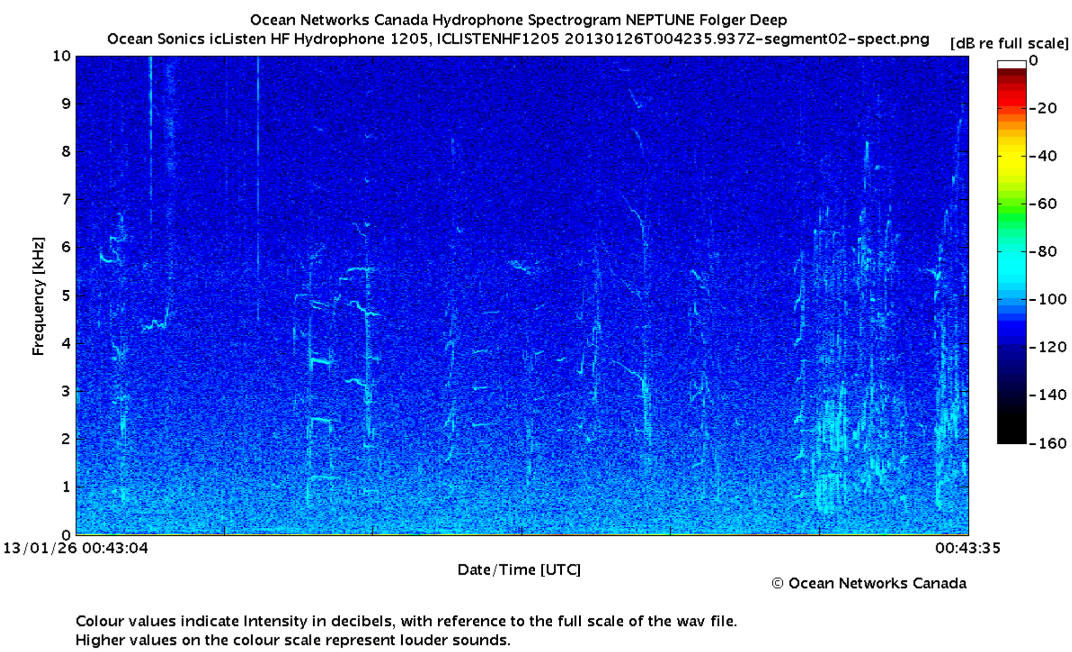


FIGURE 37. HYDROPHONE SPECTROGRAM REPRESENTING 31 SECONDS OF RECORDED CALLS BY NORTHERN RESIDENT KILLER WHALES OF THE G CLAN, RECORDED AT A DEPTH OF 96 METRES, 26 JANUARY 2013. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

A search through a list of annotations or spectra leads to the ability to browse visually through multiple spectrograms in the same web-based tool and to download audio files of interest. An example interface is shown in Figure 38, which allows users to visually browse through spectrograms, with the option to select and download hydrophone data in a variety of formats, including WAV, MP3, FFT, or HYD. Such an interface could also include an embedded player for interactively controlling recording playback.

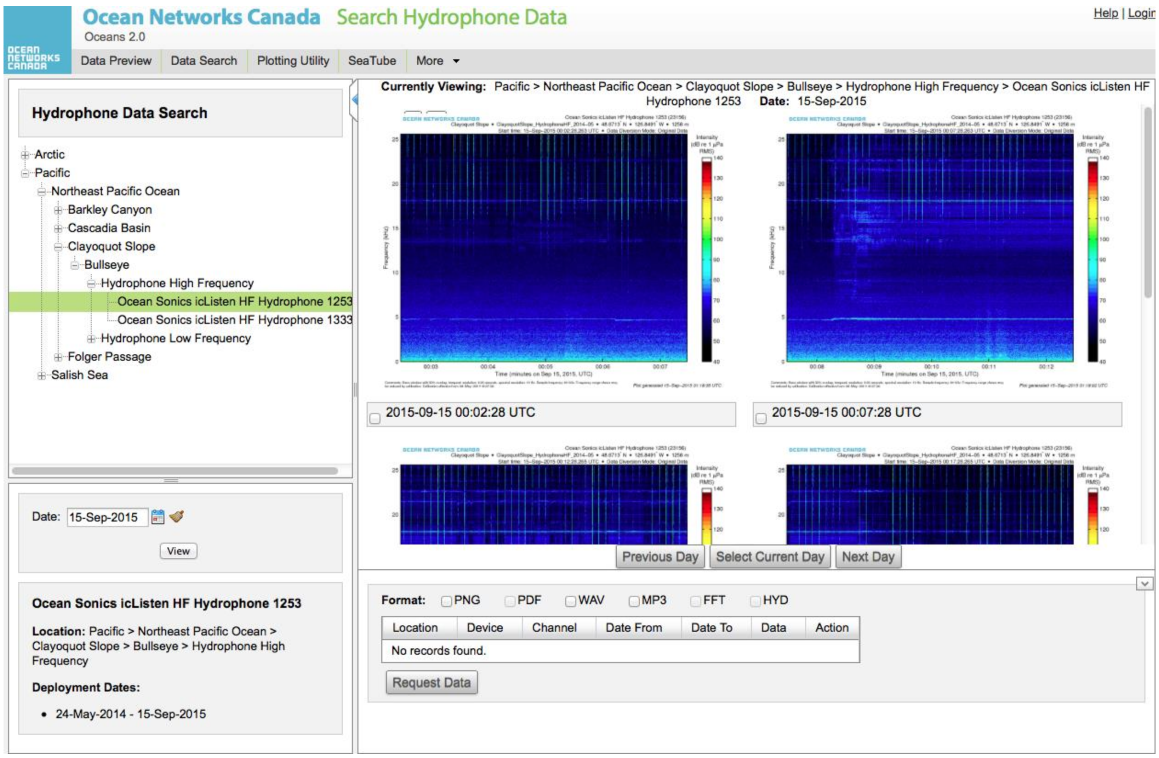


FIGURE 38. GRAPHICAL USER INTERFACE FOR SEARCHING, BROWSING, SELECTING AND DOWNLOADING HYDROPHONE DATA IN VARIOUS FORMATS. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

5.2.4.9. VIDEO

Video data represents an invaluable tool, particularly for marine biologists. The advent of cabled observatories, provided power and communication capabilities have made these instruments more ubiquitous and allowed for more such data to become available. High-capacity battery packs and memory cards also make stand-alone, autonomous cameras possible. Recorded video, flooding data archives by the thousands of hours then represents a challenge for scientists eager to exploit the image content. As for passive acoustic data coming from hydrophones, video content needs to be characterized. This can be done manually, through annotations created by human experts or citizen scientists. Annotations can also be created using automatic analysis systems that will try to identify specific patterns in the images. Once annotations are created, tools must be available to search through them and subsequently automatically play the video for someone to confirm the annotation, download a section of the stream or perform some other visual work.

Figure 39, Figure 40 and Figure 41 show examples of tools that currently exist and that support video data, from the Oceans 2.0 SeaTube interface. Those figures respectively illustrate the application's annotation search, browser-based video player, and its general interface. SeaTube is the comprehensive video search and visualization tool part of Ocean Networks Canada's Oceans 2.0 suite of features. It merges video, annotations, geographical context in three dimensions, together with the ability, among others, to prepare playlists.

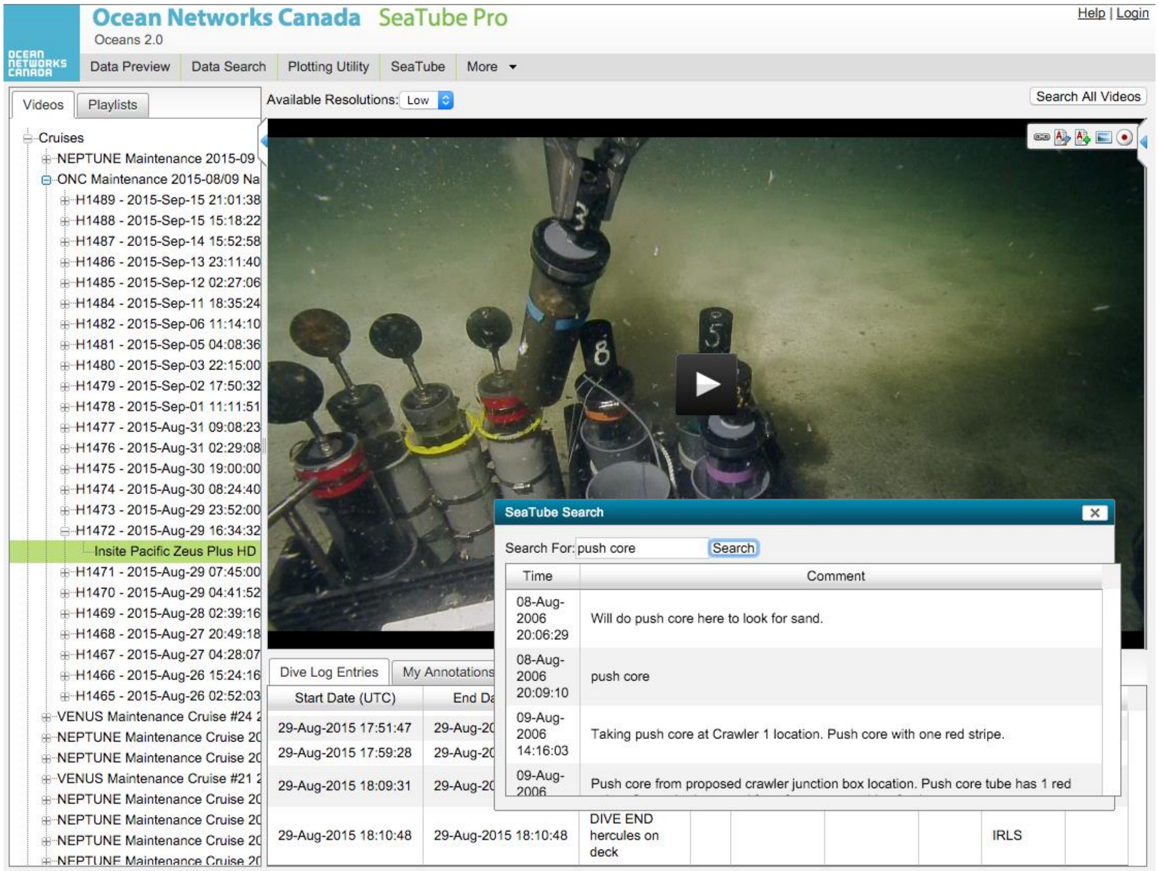


FIGURE 39. TABLE OF ANNOTATION SEARCH RESULTS WITH STREAMING VIDEO, FROM OCEANS 2.0 SEATUBE INTERFACE. ANNOTATIONS CAN BE SEARCHED TEXTUALLY; WHEN SELECTED, CORRESPONDING VIDEO SEGMENT PLAYS. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

FIGURE 40. STANDARD BROWSER-BASED VIDEO PLAYER, WITH PLAY/PAUSE, VOLUME CONTROL, EXPAND/SHRINK AND PLAYHEAD POSITION CONTROLS. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

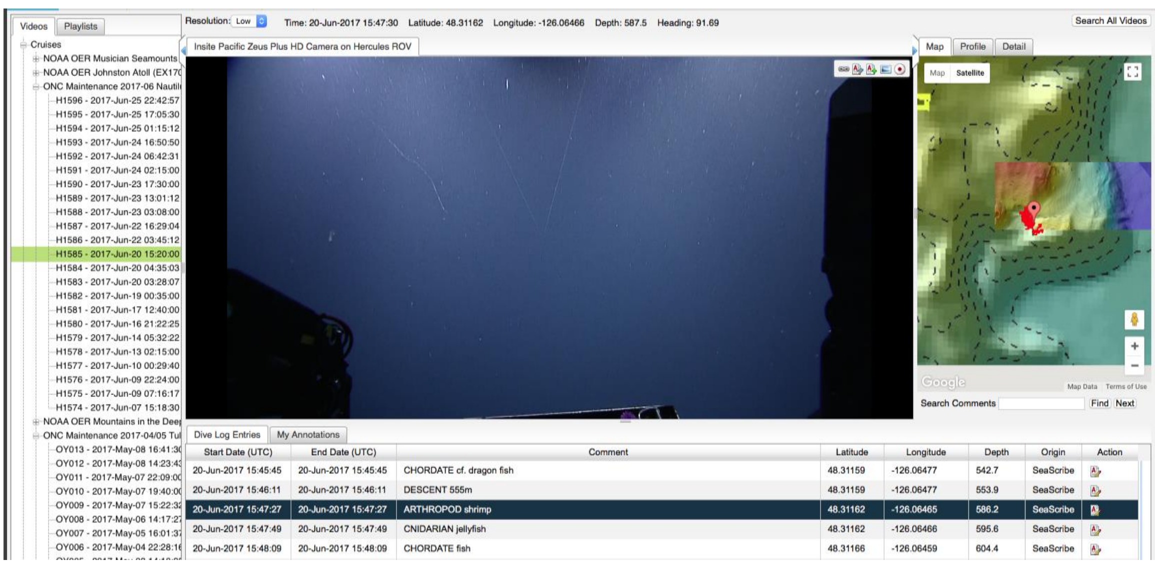


FIGURE 41. SEATUBE BY OCEAN NETWORKS CANADA'S OCEANS 2.0 SUITE OF FEATURES. ILLUSTRATION COURTESY OF OCEAN NETWORKS CANADA.

5.2.5. REQUIRED DATA FORMATS AND SERVICES

Every service model, recommended in section 5, is built with the same set of data formats and services. While Web Mapping Service (WMS) isn't required for regional associations at the low-service model, it would be a welcome addition since many data providers already host WMS and WFS services.

Since all visualization interfaces are web based, every standard and protocol need to be efficient and well supported by client-side JavaScript libraries. JSON and REST services are preferred over XML and SOAP[12] for these reasons.

5.2.6. Formats and services for metadata previsualization

Metadata standards are closely linked to the catalogue service. Notwithstanding the standards established by the IE Data, the catalogue will have to give access to these data, so they can be visualized in the different interfaces as well as the in the catalogue itself. For the visualization interfaces, as users are interested often in only a subset of metadata, there needs to be an access to the spatiotemporal extend, as well as the different keywords needed to easily identify the themes associated with the data, to facilitate their discovery.

Resources associated with a dataset can be used to identify service points for data access, preview access, or other useful links. The problem of granularity will have to be considered since a dataset can be composed of several resources accessible through several points of service.

5.2.7. Formats and services for data previsualization

Since the different visualization interfaces are implemented as an application run in a web browser, the viewing formats that are already native or easily supported by them are prioritized. Interactivity of the interfaces is also preferable and must be facilitated. As such, interfaces allowing data clustering, or when the information is centralised, and the amount of data is small, should use geoJSON. Since the symbiology must be applied on the client side, the information returned in geoJSON applies more to well-defined topics with a simple and standard symbiology.

It should be noted that the attribution and the type of visualization, that must be displayed at all times, will have to be defined by metadata, as this information will not necessarily be available in the geoJSON tables. Table 5 summarizes the recommended services for each data type, which are detailed in the subsections below.

TABLE 5. STANDARDS AND SERVICES FOR DATA PREVISUALIZATION, BY DATA TYPE.

Visualization services	
Data type	Services
Spatial data	Web Mapping Service (WMS) Web Map Tile Service (WMTS)
Metadata spatial data	GeoJSON
Video	Still frame via static images HTML 5 native support
Time series, graph and table	JSON

[12] World Wide Web Consortium. *Simple Object Access Protocol 1.2*. Retrieved from W3C Web site: https://www.w3.org/TR/soap/

Metadata	JSON via CKAN or other metadata API
Audio	Static image HTML 5 native support
Other	Static image

5.2.7.1. *Table and 2D graph data*

Since metadata are needed for preview rendering, the information returned by the previewing services should be as limited as possible to the data itself, i.e. excluding viewing method information. For example, the type of chart, color of the axes, title size and other visualization details should be defined by the viewing tool. This allows to reuse the same preview service points for multiple viewing types, in addition to limiting data transfer. The Web services will also be easier to maintain and implement. In this case, the responsibility of the visualization standardization, as well as metadata internalization, is transferred to the interface's tool, rather than the service points. For example, the ISO8601 time standardization, or controlled vocabularies describing the units and variables, allow the visualization tool to translate the information according to end-users needs, as well as fitting perfectly with ERDDAP's JSON output.


```
{
 "table": {
          "columnNames": ["time", "WaveHeight"],
          "columnTypes": ["String", "float"],
          "columnUnits": ["UTC", "m"],
          "rows": [
     ["2016-02-27T00:00:00Z", 1.094],
     ["2016-02-27T01:00:00Z", 1.25],
     ["2016-02-27T02:00:00Z", 1.25],
     ["2016-02-27T03:00:00Z", 1.25],
     ["2016-02-27T04:00:00Z", 1.094],
     ["2016-02-27T05:00:00Z", 1.094],
     ["2016-02-27T06:00:00Z", 1.094]
   ]
  }
}
```


FIGURE 42.EXAMPLE OF JSON OUTPUT FOR A 2D LINE GRAPH OF WAVE HEIGHT, IN METERS, OVER TIME, IN UTC.

5.2.7.2. *Web mapping service*

Web Mapping Service (WMS) is the de facto standard developed by OGC to serve georeferenced images for visualizing on Web-based maps. A multitude of tools and software packages for desktop and servers are available. Most, if not all, GIS packages support WMS minimally for visualization. While simplifications, mainly for performance, are frequently seen on the Web, the standard can accommodate any visual requirement need.

Web Map Tile Service (WMTS) should be permitted as long that the returned type is an image. For vector tile in mapbox[13] or protobuf[14] format, will require discussion of standardization of color grading and symbiology since this need to be done in the client side, i.e. by the visualization interface.

5.2.7.3. *STATIC IMAGES*

For any other type of visualization, a static image will be returned allowing to pre-generate the rendering or to realize it as the request is made on the server side. Although limiting the possible interface interactivity, this option makes it possible to reuse several systems already in place. For transfer optimization purposes, JPEG format should be used unless quality problems related to compression or transparency are identified, in which case PNG format should be favored. Obviously for a chart or table, a JSON access point is prioritized. Finally, a PDF file should not be the only type of image available since it can not easily be displayed inside a preview window.

5.2.7.4. *AUDIO*

Besides the visual analysis of audio through spectrogram or FFT, listening to the audio is available with the native HTML 5 audio control. As of now, file format supported are WAV, MP3 and OGG.

Since control is dependent of the browser's implementation, the quality and efficiency aren't guaranteed. That being said, audio support is relatively mature in modern browser. Standard control let user play, pause, stop and change the current playback position.

5.2.7.5. *VIDEO*

Video support is done through the HTML 5 video control. Since HTML 5 video support is different from browser to browser, it is done in an opportunistic and not mandatory manner. Web video technologies are changing rapidly and support for new video codec like H.265 and HEVC seem to be implemented more consistently in browser than current codec.

Basic video control like play, pause, stop, changing the playback position and sound volume should be available on all video supported by the browser control. Since control is dependent of the browser's implementation, the quality and efficiency aren't guarantied.

5.2.8. FORMATS AND SERVICES FOR DATA DOWNLOAD AND EXTRACTION

The interoperability enabled by services allows global portals to provide seamless access to distributed datasets housed in data nodes. Users get seamless access to the datasets via the global portal. Along with such easy and centralized access, data providers and regional systems require recognition in the form of a visible indicator (e.g. wordmark or logo) associated with the resources that they are providing. It is important that the metadata associated with any data access result contain the origin and attribution references.

A dataset need not be restricted to one endpoint or service for data extraction. Some users might want a more efficient file format to download the entirety of the dataset while others may access subsets of the

---

[13] *Mapbox vector tile specification*. Retrieved from Mapbox's Github repository: https://github.com/mapbox/vector-tile-spec/tree/master/2.1

[14] *Protocol buffers.* Retrieved from Google Developers Web site: https://developers.google.com/protocol-buffers/

data in a more interactive format. Ocean Networks Canada's Oceans 2.0 system, illustrated in Figure 43, implements this feature whereby any data stream can be obtained in any of format.

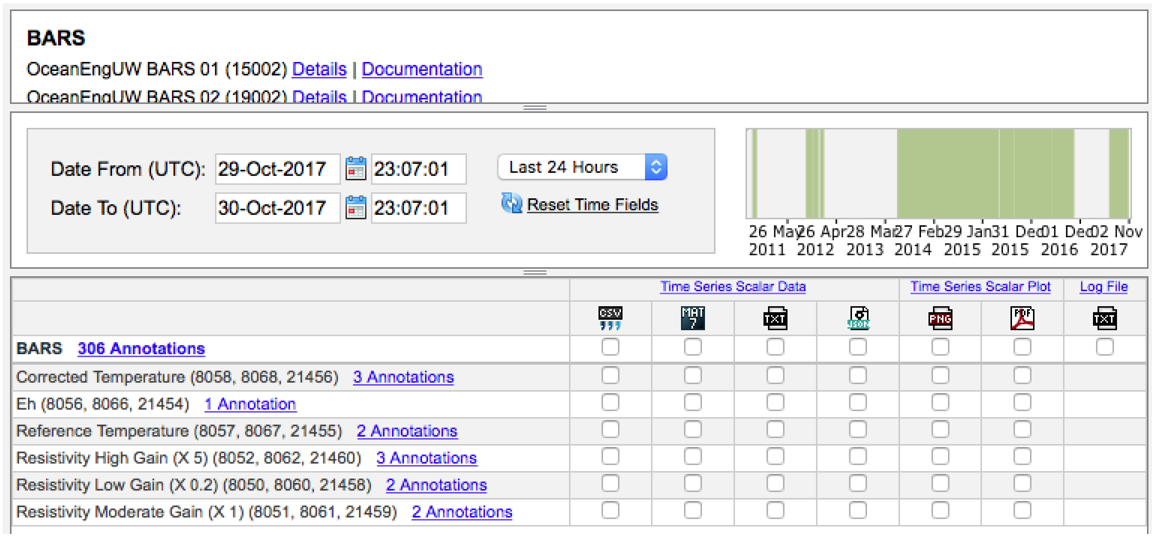


FIGURE 43. EXAMPLE OF DATA ACCESS CHOICES FROM OCEANS 2.0, WHICH OFFERS USER-SPECIFIC FILE FORMATS FOR THEIR DATA DOWNLOAD. IN THIS EXAMPLE, A SPECIFIC TEMPERATURE PROBE CONTAINING A NUMBER OF SENSORS CAN BE OFFERED TO USERS AS EITHER CSV, MAT, TXT, JSON TABLES, OR AS PNG, OR AS PDF PLOTS.

The support of multiple endpoints (or file formats) for data access will be used to implement versioning of standards in the future. It is important for any visualization of those endpoints that the associated protocol and file format is visually recognisable via standardised icons (e.g. Excel file) and logos (e.g. ERDDAP logo), as illustrated in Figure 44.

FIGURE 44. VARIOUS LOGOS AND ICONS ILLUSTRATING THE AVAILABLE FILE FORMATS (CSV, MAT, TXT, JSON, PNG, PDF) FOR DATA DOWNLOAD, IN OCEANS 2.0 DATA SEARCH APPLICATION, BY OCEAN NETWORKS CANADA. NOTE THAT THE SAME FILE FORMATS ARE AVAILABLE FOR VARIOUS DATA PRODUCT OUTPUTS (E.G. TXT AVAILABLE FOR TIME SERIES SCALAR DATA AND LOG FILE).

The list of data type accessible via those endpoints can be comprise but not limited to the table below.

TABLE 6. DESCRIPTION OF TYPES OF LINK FOR DATA DOWNLOAD VIA VARIOUS ENDPOINTS.

Type of link	Description
OPeNDAP	Link to an OPeNDAP service URL
ERDDAP webpage	ERDDAP Web page for data extraction
Formatted numerical data	CSV, XLS or NetCDF file

Static Image	JPEG, PNG and GIF[15] file
Other Website	Link to another website
Document	PDF or Word file
Web accessible folder (WAF)	Link to a Folder (HTML listing or FTP access)
Compressed archive	Zip or GZ file

Data access for time series, graphs and tables can be query-based using OPeNDAP automated scripts and development of software. ERDDAP can be used to manipulate and translate OPeNDAP and other data access protocols in a Web interface or another application.

5.3. SERVICE MODELS

The following set of recommendations have been detailed according to three levels of service: low-service, medium-service and high-service. These levels of implementation will allow to present various options in a scalable approach, while highlighting priorities that should not be overlooked at any cost, as well as desired approaches to maximize CIOOS capabilities. For all the three levels, a generic metadata and data visualization pathway is described in Figure 45. In this figure, regional associations host metadata and data from their regional data producers. They can also do it, when necessary, for other data nodes, for example DFO, NRCAN or FGP data, or it can be directly being made accessible through services for both the RA local interfaces and CIOOS interfaces. In the three shade levels, the figure illustrates the various visualization tools available for each service levels, which are described below.

For the low, moderate and high service recommendations that follow, it is assumed that low-service criteria are included at the moderate-service level, and low and moderate-service criteria are included at the high-service level. As stated above, the general properties listed in section 5.2 are to be included in each of the following service models.

[15] World Wide Web Consortium. *Graphics interchange format Version 89a*. Retrieved from W3C website: https://www.w3.org/Graphics/GIF/spec-gif89a.txt

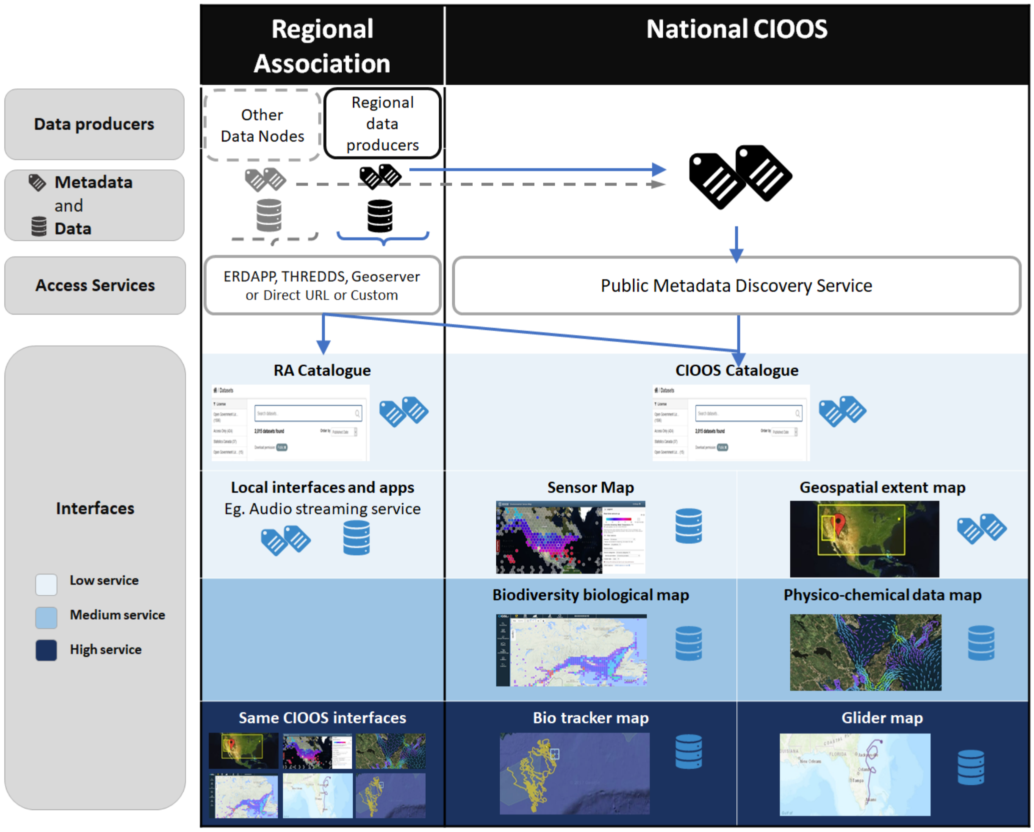


FIGURE 45. OVERALL STRUCTURE OF SERVICES AND INTERFACES, FOR REGIONAL ASSOCIATIONS, CIOOS, AND OTHER DATA NODES (E.G. FISHERIES AND OCEANS CANADA, FEDERAL GEOSPATIAL PLATFORM). THREE LEVELS OF SERVICES ARE ILLUSTRATED BY THREE LEVELS OF COLOUR LIGHTNESS: LOW-SERVICE IN LIGHT BLUE, MEDIUM-SERVICE IN MEDIUM BLUE AND HIGH-SERVICE IN DARK BLUE. DASH LINES REPRESENT A NON-MANDATORY PATHWAY.

5.3.1. LOW-SERVICE MODEL

This first level of CIOOS implementation includes a metadata catalogue and two theme based interfaces allowing metadata visualization and data previsualization. Their visualization elements and functionalities are described below.

The set of visualization interfaces is based on centralized metadata available in the CIOOS catalogue. Since data and visualization services aren’t hosted by CIOOS, the visualization interfaces use the metadata to generate interactive maps with spatialized metadata, accessible to the user. Data visualization services, increasing the information displayed, when available, are not mandatory, although a great addition when possible.

5.3.1.1. METADATA CATALOGUE INTERFACE

The catalogue interface allows the users to easily query the discovery service for every type of dataset included in the catalogue, with a simple free text search, followed by a series of additional advanced search filters, as defined in section 4.2 of this report. Since all datasets are available through this interface, thematic search can help to quickly refine query results in a field of interest. Those themes, represented on the first page, direct the users to a more detailed query interface once the theme is specified.

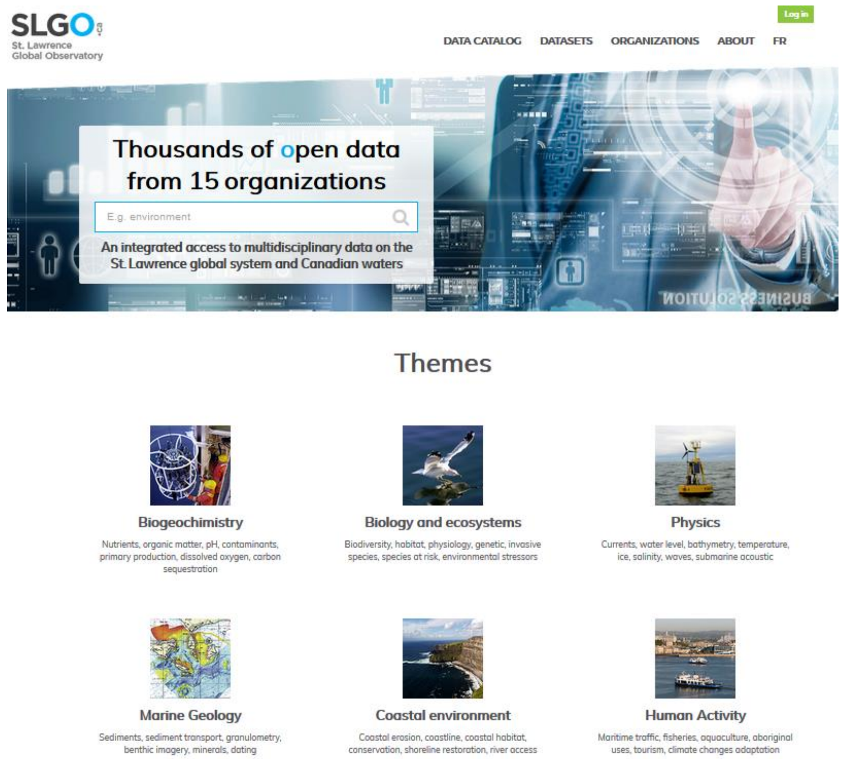


FIGURE 46. THEMATIC CHOICES AND ASSOCIATED KEYWORDS ON THE HOMEPAGE OF SLGO'S CATALOGUE, USED FOR QUICK DATA DISCOVERY. ACCESSIBLE AT HTTP://CATALOGUE.SLGO.CA/EN/.

In addition to the catalogue, two themes based interface have been identified as a minimum for the baseline: data acquired via sensors (buoys, weather stations, etc.) and spatially bounded dataset.

5.3.1.2. Sensor map interface

To visualize the data of various fixed instruments, a sensor map type of interface is essential (see Figure 47 for an example of sensor map). To realize this interface, it is necessary to extract the instruments and the parameters available, by sensors, from the metadata. Sensors position, and available parameters, will be extracted from the base profile defined through ISO 19115 and NetCDF-CF file defined as resources. A task for the metadata server is to periodically query available metadata and to store only the required sensors information in a centralised database in an optimised format. Then, a service accessing this list of metadata in geoJSON format can display it efficiently on the sensor map.

Since the number of sensors for a region or area can be high, when zooming out of a map, clustering should be used, and the number of available sensors or variables displayed. The specific information displayed when the user clicks on the agglomeration remains to be defined.

The interface must basically consist of a map spawning on a full window width, with information overlaid, including transparency levels where needed. This type of layout allows the interface to be included on a home page or to be easily rebranded if a regional association decides to use it on their own website.

Since the display is based on available metadata, access is performed when the user selects a specific sensor. A window appears, listing the access points to the available data, the data source organization, the choice of the parameters from this sensor to be displayed and, depending on the availability of the visualization services of this instrument, an interactive graph or an image, representing the latest available data. The period of time of the data previsualized must be indicated.

The number of available parameters for each sensor has no limit, and so any available data must be listed. As visualization in the sensor map is based on metadata, sensors whose data previsualization can't be done in the same interface (e.g. hydrophone data that should be previewed in an external interface or downloaded as a whole) can still be part of the available sensors shown on the map. This ensures discoverability and accessibility for that kind of data, through displayed metadata. Indeed, if no visualization services are available for a sensor, a link leading at least to an external data visualization interface would be available. In each pop-up window where metadata are displayed, a link to the dataset entry in the metadata catalogue is displayed.

The data format used to define sensors, available parameters and various access points remains to be defined. Everything must be easily consumable by a Web application running in a web browser. For this reason, geoJSON is the recommended encoding format, allowing the interface to locate the instrument while including the rest of the necessary data for previsualization.

The period of available data for previsualization should be configurable to define whether an instrument is displayed or not. This could vary depending on the regional association and CIOOS.

Sensor map interface on RA website

If a regional association decides to use the sensor map interface to give access to the sensors specific to that region, it can simply host and configure the extraction service for their own local catalogue. If their catalogue is in a different format than the CIOOS catalogue, or use different systems to give access to the sensors metadata, whether it is proprietary, SOS or SensorThings, a service extension could support the extraction of these other systems (Bailey & Bochenek, 2016).

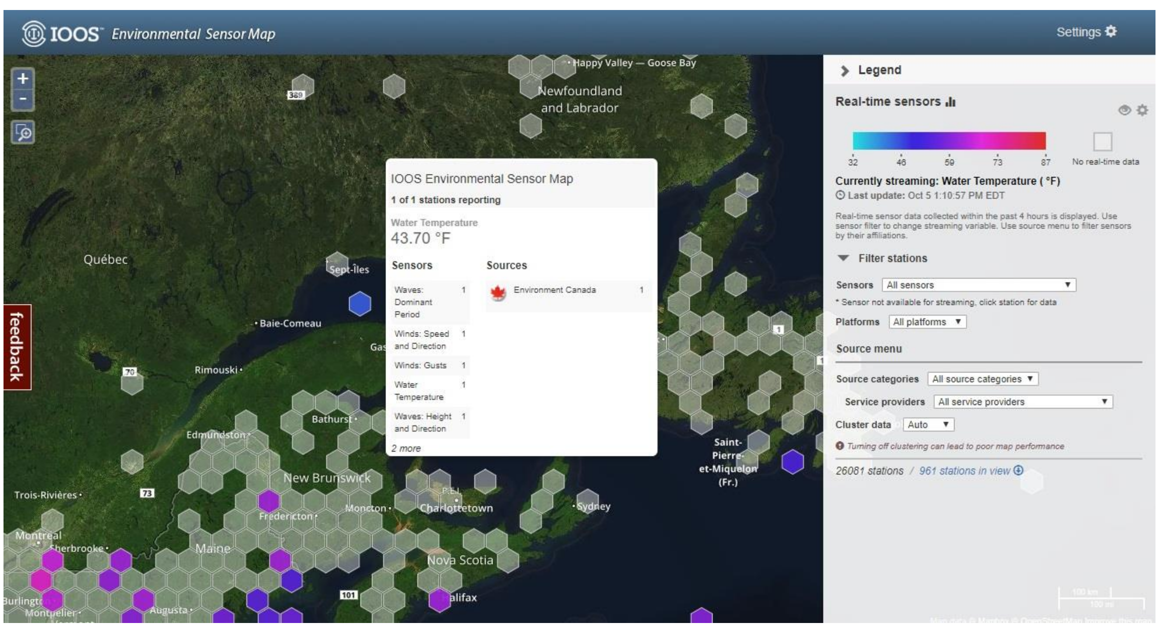


FIGURE 47. SENSOR VISUALIZATION INTERFACE ON IOOS'S 'ENVIRONMENTAL SENSOR MAP' INTERFACE.

5.3.1.3. *Geospatial Extend based metadata visualization*

A geospatial extend based map focuses on the location of datasets. The same search filters as the catalogue ones should be available and applied on any dataset with a specific geospatial extent. A color code could be applied according to the theme associated with the datasets (e.g. for physical vs for biological datasets). In addition to displaying the dataset extend on the map, a window displaying a list of these datasets with some metadata field must be displayed. A useful tool for this window, for each dataset, is an option to zoom to the selected dataset's extent on the map. This is particularly useful for data sets located in small specific areas.

More metadata are displayed when users click on a listed dataset. A subset of the metadata, links to the data access points, and a link to the associated page of the metadata catalogue must be displayed.

Since this interface requires only access to metadata, no preview service is required. As metadata are centralized, it is simpler to optimize the services and database used by this interface. This allows for an effective interactive interface.

As for the sensor map, this interface should take all the available width of the page, with any other information overlaid when needed. See Figure 48 for an example of geospatial extent map interface, by GEOSS Geoportal.

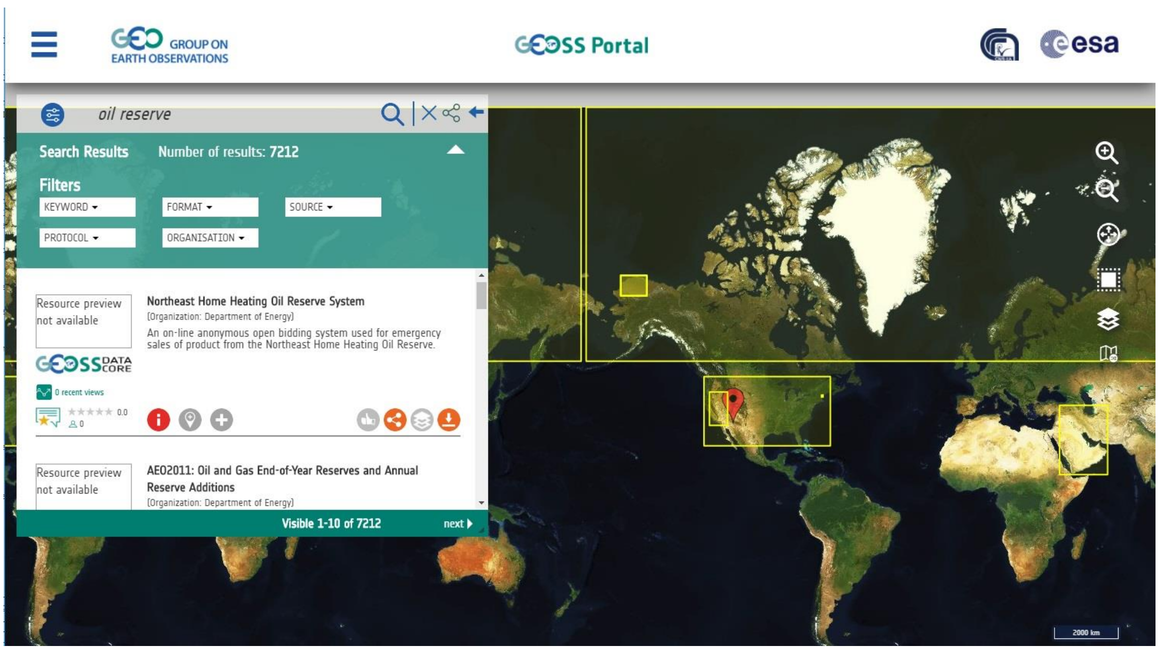


FIGURE 48. METADATA QUERY WINDOWS WITH GEOSPATIAL EXTENT VISUALIZATION A MAP BY GEOSS GEOPORTAL INTERFACE. ACCESSIBLE AT HTTP://WWW.GEOPORTAL.COM

5.3.2. Moderate-service model

This service model includes all the low-service model recommendations, while adding a few more functionalities and interfaces, increasing the service offered to end-users.

A first recommendation for an increase service is to better support Arctic data. We recommend that all interfaces for the moderate-service model must be able to modify map projections, while including at least one projection designed for the Arctic, in addition to the Web Mercator standard.

In addition to the sensor map and geospatial extent map detailed in the low-service section, two new interfaces are here recommended: one displaying gridded observation data and a second one displaying biological and biodiversity data. These two new interfaces need access to WMS map visualization services, which require not only the configuration and hosting of these services on the data nodes, but also a standardization of legends and color codes used by data type for all services.

We recommend that the CIOOS main portal offers multiple projections; at minimum there should be the following commonly used projections:

- EPSG:3857;
- one projection optimised for the representation of the Arctic region (e.g. EPSG:32661);
- one projection optimised for the representation of the Antarctic region (e.g. EPSG:3031).

This is a suggested minimum, but a wider selection could be offered. However, specific projections should not be enforced on the regional nodes, as they should be free to decide which projections better suit their specific needs (see Figure 49 for an example by Polar Data Catalogue). On the other hand, since regional nodes might provide Web Map Services (WMS) to CIOOS's central portal, it is recommended, although not mandatory, to provide such services using EPSG:4326. This recommendation is merely to avoid potentially expensive computation that might occur for complex data layers using peculiar projections.

In addition to the map projection choices, base map layers options should also be available to the users.

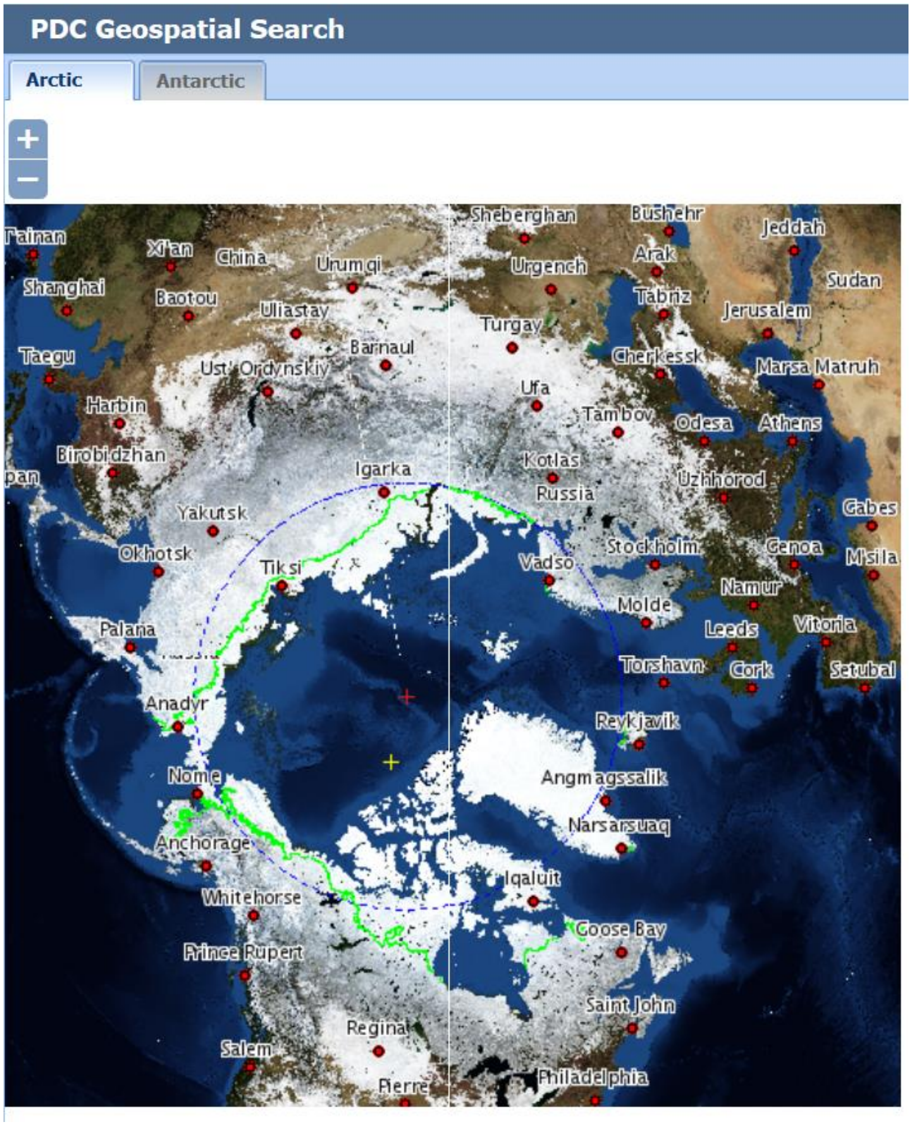


FIGURE 49. PROJECTION OPTIMIZED FOR THE ARCTIC REGION BY THE POLAR DATA CATALOGUE, ACCESSIBLE ON HTTPS://WWW.POLARDATA.CA/PDCSEARCH/.

5.3.2.1. GEOSPATIAL DATA VISUALIZATION INTERFACE FOR PHYSICAL AND BIOGEOCHEMICAL DATA

To visualize data covering several points on a map, either uniformly or not, a Geographic Information System interface style interface with layer selection must be implemented. This interface will display, when selecting in a list of data layers, associated observation layers for physical and chemical data. For examples of variables for each of those two domains, see Table 3 and 4 of the section 2.6 of the IE Data and Observations report. In this interface, the most recent satellite images, bathymetric data, CODAR observations and High Frequency Radar (HFR) data types can be displayed.

A minimum of layer search options will be implemented, to allow end-users to find and select layers from the listed available layers. To identify available layers for this interface, a keyword must be added to the datasets, in addition to specifying the WMS access point for data preview.

The activated and available layers will be listed in their respective windows, overlapping the main map. The legend could be displayed under each active layer in its window, or in a separate window.

To allow the comparison between different layers displaying the same parameter, e.g. water temperature, the different providers will have to use the same legend and color map. This symbiology determined by the CIOOS steering committee will have to be enforced by the regional associations when creating a new layer or reusing a data provider's layer. The legend display for each active layer and layer transparency control will be available, as the zoom-to-extent tool.

On the map, when clicking on an element of an active layer, more information about that layer appears in an overlapping window. This information includes real-time observations (or last updated), the data source, the link to the data catalogue entry, and any other links associated with the dataset. A call to retrieve the information at the given point is performed and the returned information is displayed in addition to the metadata of the dataset.

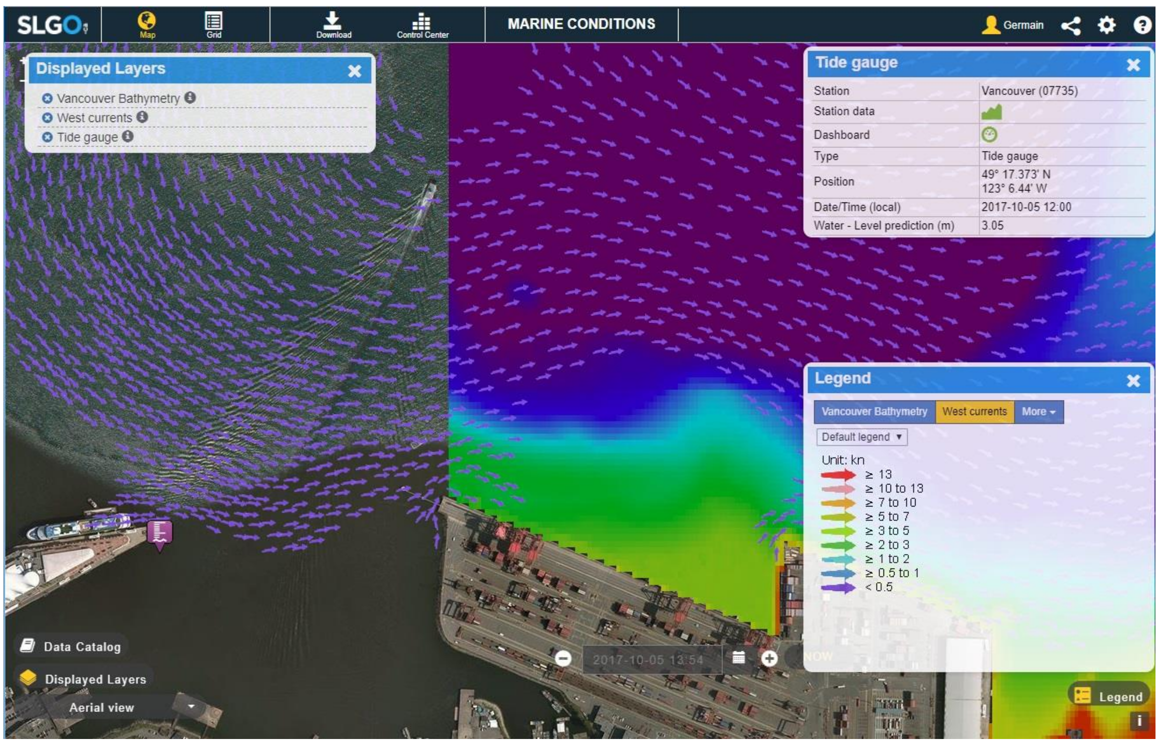


FIGURE 50. MAP BASED DATA VISUALIZATION FROM SLGO'S MARINE CONDITIONS WEB APPLICATION. DISPLAYED LAYERS ARE BATHYMETRY, CURRENTS AND TIDE GAUGE SENSORS. CLICKING ON THE TIDE GAUGE ICON (BOTTOM LEFT QUADRANT) OPENS A 'TIDE GAUGE' WINDOW WITH MORE STATION INFORMATION (ID, LINK TO DATA IN TABLE OR DASHBOARD, TYPE, POSITION, TIME AND VALUE OF CURRENT ASSOCIATED DATA). THE LEGEND WINDOW ALLOWS TO SELECT A TAB FOR EACH OF THE DISPLAYED LAYER; IN THIS CASE, THE CURRENS LEGEND.

5.3.2.2. *BIODIVERSITY AND BIOLOGICAL INTERFACE*

Biodiversity and biological data, in a similar way as the physical and chemical data presented above, are available as gridded layers. Based on the same functionalities of a GIS interface, we recommended implementing, for the mid-service model, this interface allowing to display biodiversity and biological metadata on a map based on layer choices.

As before, a keyword associated with each available datasets of this kind makes it possible to identify them to be harvested for the interface. In addition, the appropriate GBIF code will be added to the dataset to allow searching by species. Other search tools must be available to discover the layers from a list in a window overlapping the main map. The design should be in line, with a similar layout, with the physicochemical interface presented above, in terms of tools placement and features.

It should be noted that this interface, showcasing biodiversity and biological data, does not allow for extensive data analysis. In this case, these layers will have to identify beforehand the desired period or agglomeration (average, presence or absence, species included, etc.).

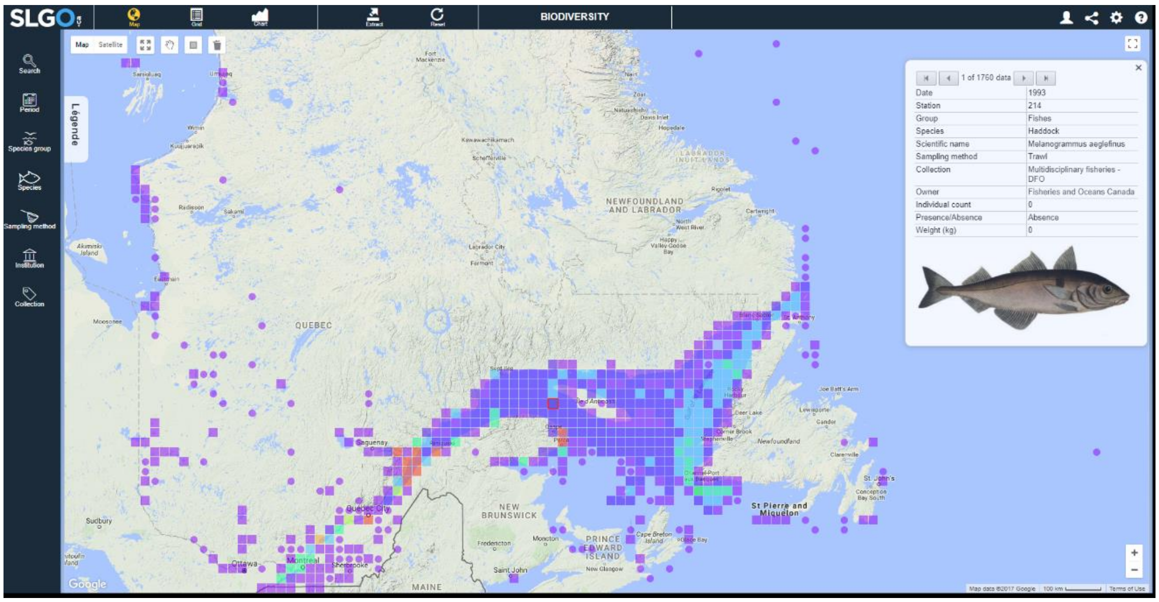


FIGURE 51. BIODIVERSITY DATA INTERFACE ON SLGO'S 'BIODIVERSITY' WEB APPLICATION. FILTERS ON THE LEFT SIDEBAR INCLUDE TIME PERIOD, SPECIES, SAMPLING METHOD, INSTITUTION AND COLLECTION. DATA ARE AGGLOMERATED BY NUMBER OF SURVEY COUNTS (● = 1 SURVEY, ■ ≤ 74 SURVEYS, ■ > 926 SURVEYS). MORE INFORMATION ON A SURVEY IS DISPLAYED UPON CLICKING ON A DOT OR SQUARE ON THE MAP, IN AN OVERLAID WINDOW, INCLUDING SPECIES, SAMPLING METHOD AND DATA PRODUCER ATTRIBUTION. ARROWS ALLOW TO MOVE FROM ONE SURVEY TO ANOTHER FROM THE CLICKED AGGLOMERATION.

5.3.3. High-service model

Similarly to the medium-service model, this high-service model includes all the features and tools implemented in previous models. In this case, it includes all items from both low and medium service, while building on those to implement more features and visualization interfaces.

In this level, time is added as a global parameter to the map interfaces. For the physicochemical data interface, it will be possible to select a specific time, current or in the past, allowing to visualize historical data. Each layer should, if relevant, consider the requested time, otherwise show the most recent data. To add layers with forecast in a subsequent development phase of CIOOS, only the time control would be modified to support the selection of time in the future.

A time filter should also be added to the other interfaces, to enable the selection of time periods. For the biodiversity and biology data interface, as well as the geospatial extend metadata visualization, a time interval section tool must be available. In addition, the interactive charts and tables will have to take into account that time filter and allow pagination of the data. This pagination will make it easier to implement interactive data display, by limiting transfer for extensive time range requests. JSON data access services will need to consider time as a query parameter.

Three other visualization interfaces will be implemented in addition to those presented in the sections above. Two of them focus on mobile sensors: 1) remotely operated underwater vehicle (ROV) and gliders, and 2) wildlife telemetry.

5.3.3.1. GLIDER INTERFACE

Using a similar layer-based interface presented in previous interfaces, the glider visualization interface allows end-users to select a time range (start, end), listing all the corresponding gliders, AUVs and ROVs in service during this period in an overlapping window. The available datasets are identified through their metadata, by using a specific keyword as well as the required start and end time information. The movement of the underwater sensors are detailed in sections 4.3.3 and 5.2.4.3 of this report.

When users click on a point or a path of one of the selected layers, an information window overlapping the main map lists the metadata, including the data source, data download access points, duration of the trip, as wells as the variables and data collected at the specified location.

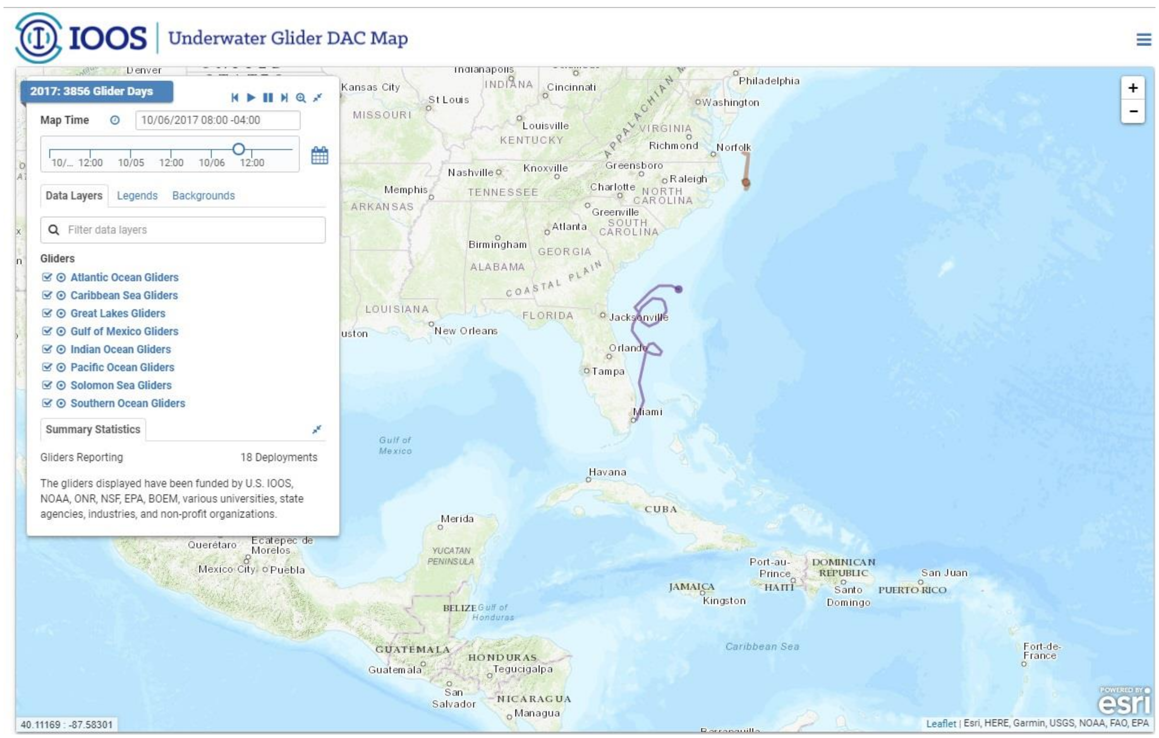


FIGURE 52. GLIDER DATA INTERFACE VIA IOOS 'UNDERWATER GLIDER DAC MAP'.

5.3.3.2. WILDLIFE TELEMETRY INTERFACE

Just like the moving instruments interface presented above, the wildlife telemetry interface allows to visualize layers of data associated to bio trackers. This interface is implemented using the same features of a GIS layer-based interface. Data layers are identified in datasets by a specific keyword, time coverage, as well as GBIF codes for the species. The addition of GBIF codes allow easy and efficient search in datasets. Movement display is done accordingly to sections 4.3.3 and 5.2.4.3 specifications for moving instruments.

Although the trajectory might represent the only parameter collected and information to visualize, the interface should not be limited to it. As for the glider interface, when a user clicks on a point or a path of one of the selected layers, an information window overlapping the main map lists the metadata (data source, data download access point, trip duration) as well as available data and variables collected.

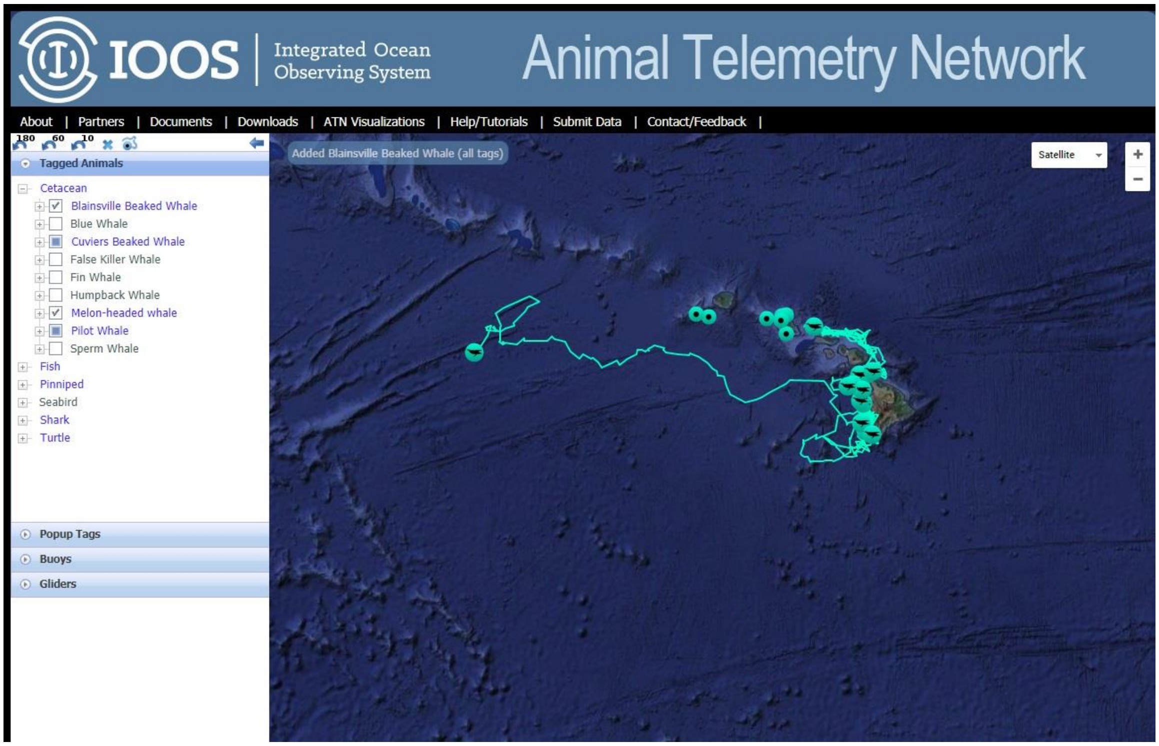


FIGURE 53. EXAMPLE OF TRACKER DATA VIA IOOS ANIMAL TELEMETRY NETWORK WEB APPLICATION

5.4. IMPLEMENTATION PHASES

To implement a system of this magnitude, a phased approached is recommended. This approach allows to validate, measure, correct assumptions and react responsibly to development changes.

5.4.1. PHASE 1: PROTOTYPE NODE

The first phase implements a simplified version of the systems consisting of only two to three core variables and a limited node of simulated data node and national node. The definition of basic profile metadata profiles, selection of supported tools and protocols needs to be done before any other work can start.

Selected core variables should be visualized as sensors data and as gridded data. Surface current and water level are excellent candidate since they are available for fixed instruments like buoys, mobile sensors like wave gliders and in through gridded format as CODAR output.

Two data producing organisations should be selected to simulate regional nodes. Those organisations will host a metadata catalogue with the required dataset. The more disparate the tools used, the better. For example, a CKAN catalogue and a GeoServer catalogue could be used. By limiting the number of variables and participating nodes but maximizing tools and metadata diversity, this will challenge assumptions and standards selections for interoperability. In conjunction with those two organisations, a national node implementation with the selected catalogue and harvesting service need to be implemented.

The organisations representing simulated regional associations should host web service for visualization of the selected datasets. Those Web services needs to output JSON for tabular and 2D line graph visualization and other supporting WMS or WMTS protocols. If services are already available at the organisations, it should be used and not install or configure another. Like metadata, this will validate and/or document the steps needed to meet CIOOS web service endpoint recommendations.

Prototype versions of visualization interface are developed with minimal effort on aesthetic but with maximum coverage and functionalities on complete metadata display and queries. A simple sensors metadata extraction and caching service need to be implemented.

This phase makes it possible to check the necessary extent of the basic profile to support all the visualization needs at the description level, for search filters, the attribution and the access to the visualization services. Actions needed to standardize metadata and implement, modify or augment existing services are documented. Some metrics may also be implemented such as the response time of the display and the amount of data needed to centralize all the metadata. These metrics will serve as a guide during subsequent phases on the actual need for performance or optimization effort required. This information will correct, re-prioritize some steps in the implementation of the complete system. The experience gained will align the best practices needed when adding datasets.

In parallel, two websites simulating a regional association and the CIOOS are set up. This allows to develop the common branding and ensure the necessary level of work for the implementation of the CMS and other aspects of aesthetics and navigation.

5.4.2. Phase 2: RA and CIOOS metadata implementation

Once the analysis of the first phase is complete, the implementation of the regional associations and the main site can begin. The experience gained during the first phase will provide better guidance for development and facilitate the integration of metadata. In addition, the template required to achieve common branding of regional associations will be refined and implemented.

The implementation or configuration of regional catalogues for harvesting will be implemented at the very beginning. The process of acceptance by the regional association so that the dataset is available for harvesting will need to be established. The CIOOS implementation of the main catalogue, and the configuration for metadata harvesting of the different regional associations, is done in parallel.

A selection of datasets with less data is performed. It is these datasets that will be added during this phase. This allows you to focus on adding metadata and visualization service configurations.

Since the focus is on metadata, development priority is given to internationalization, information on attribution, completeness of information, search filters and interface aesthetics.

Visualization service are implemented or configured accordingly to CIOOS standards and requirement.

The development of the interfaces continues based on the prototypes made during the first phase. The interfaces of sensors and geographic extent is prioritized. Depending on the level of services selected, physicochemical and biodiversity interfaces are also developed.

The setting up of the different sites of the regional associations is realized in parallel with the development of the interfaces.

At the end of this phase, the sites of the regional associations and CIOOS are made available online.

5.4.3. Phase 3: Optimization and final visualization interface

The last phase is devoted to optimizing, patching and adding larger datasets.

Sites and existing interfaces usage based tests are made from the beginning of this phase to identify weaknesses and gaps perceived by users. Any issues raised during the second phase and compiled during the tests are addressed. Subsequently the integration of larger data sets is made to the catalogues of regional associations.

Depending on the level of service chosen, wildlife telemetry interfaces as well as glider or physicochemical and biodiversity interfaces are realized. The optimization of already existing interfaces is done in parallel.

6. Costs per service model

The establishment of a CIOOS, as elsewhere, requires a team of high quality personnel to ensure the dissemination of quality data in a timely manner, and this, in compliance with recognized standards. In this sense, even if they are NGOs, regional associations must have incentive to attract high-demand and scarce staff. They must not only be able to offer competitive salaries, but also to be able to offer contracts of a minimum duration of 5 years. The proposed salaries below address this issue. They are competitive, but may vary from a region to another.

The table below shows the team and an estimate of associated costs required for the data visualization operations, for each of the three models of service presented in section 5.3. This budget also considers costs for staff training and essential software for programming and coordination tasks, some of which generate recurring subscription fees (e.g. Intellij, Adobe Creative Cloud, Gmail for business, Office 365). For any of those software and tools, free open source alternatives could be considered, when possible.

It should be noted that this evaluation is for the regional associations and does not consider the team required to create and maintain the CIOOS portal.

TABLE 7. ESTIMATED EXPENSES AND STAFF REQUIREMENTS FOR LOW-, MODERATE- AND HIGH-SERVICE MODELS IMPLEMENTATION FOR ONE REGIONAL ASSOCIATION, PER YEAR.

	Low-service		Moderate-service		High-service	
Expenses	**Amount (K)**	**# of employee**	**Amount (K)**	**# of employee**	**Amount (K)**	**# of employee**
Salaries						
IT Team manager	$100,0	1,00	$100,0	1,00	$100,0	1,00
Programmer analyst	$180,0	2,00	$270,0	3,00	$360,0	4,00
IT Technician	$150,0	2,00	$225,0	3,00	$375,0	5,00
Data manager	$9,0	0,10	$22,5	0,25	$45,0	0,50
Project coordinator & Outreach	$80,0	1,00	$120,0	2,00	$240,0	3,00
Webmaster	$65,0	1,00	$65,0	1,00	$65,0	1,00
Web integrator	-	-	$50,0	1,00	$50,0	1,00
Communication officer	$6,5	0,10	$13,0	0,20	$19,5	0,30
Sub-total	**$590,5**	**7,20**	**$865,5**	**11,45**	**$1 254,5**	**15,80**
Travel expenses						
IT and data Expert Committee Meeting	$10,0		$20,0		$40,0	
Sub-total	**$10,0**		**$20,0**		**$40,0**	
Equipment						
Software	$5,0		$10,0		$15,0	
Sub-total	**$5,0**		**$10,0**		**$5,0**	
Fixed Operating Costs						
Training	$10,0		$30,0		$60,0	
Sub-total	**$10,0**		**$30,0**		**$60,0**	
Communications & Representation						
Conferences, workshops, registration and travel expenses	$5,0		$15,0		$20,0	
Sub-total	**$5,0**		**$15,0**		**$20,0**	
TOTAL (K$/YEAR)	**$620,5**		**$940,5**		**$1 379,5**	

7. Additional products and next steps

7.1. Forecast models and volumetric data visualization

One important omission from all service levels of CIOOS is forecast model data. To be able to consult both the observations and the historical data, it is necessary to add a time filter. Fortunately, once this filter is available, the inclusion of forecast data is a formality. However, the difficulty of implementing forecasts comes from the high total volume and high rate of data updates, especially for multidimensional forecasting models, on the hardware infrastructure.

The prototype of Ocean Navigator makes it possible to extract and visualize observations and forecasts models from gridded volumetric data. This interface has a high potential, by allowing data access at various depth, which is not possible on surface visualization interfaces. By adopting Ocean Navigator and supporting its development via the CIOOS, the visualization of various forecast volumetric data, available through OPeNDAP, is easily and efficiently achieved with a minimum of effort.

Since this type of interface is relatively unique, the CIOOS would stand out by taking a step ahead in this area. The addition of acoustic parameter is an interesting avenue for the future.

7.2. Audio and video visualization

Since no special audio or video interfaces are developed especially for CIOOS, this lack should be addressed in a subsequent phase. The requirements on the infrastructure to support audio and video data and the lack of standards are the main difficulties to implement specific audio and video interfaces.

A lot of efforts are currently deployed in those fields. This is encouraging us to think that once CIOOS will be operational, the implementation of these future functionalities will be facilitated by increased technical capabilities as well as user expectations.

7.3. Data centralization and product generation at CIOOS

Although the CIOOS infrastructure is based on data decentralization, this infrastructure limits the ability to generate some specific visualization products. Centralization would facilitate support and optimization of the visualization of real-time data as well as some products at the national scale, like biodiversity observations of specific species. The concept of a thematic node partially alleviates this gap. However, when infrastructure requirements become a bottleneck or when multiple datasets outside of a single theme node should be centralised, this partial centralisation should be done at CIOOS.

This requires a much higher hardware infrastructure performance, at the national node level, to store and configure visualization services. Like metadata, a data ingestion service will have to be set up by both the regional associations and the national node. However, if implemented, centralization will allow end-users to download data to one place instead of having to visit several different links, and as such, greatly improve and facilitate data discovery and access.

7.4. Data analysis on CIOOS interface

If data are centralized on the CIOOS portal, it will allow the implementation of tools, not only for visualization, but also for data analysis. Those analysis functions are necessary for the implementation of more advanced tools, such as a biodiversity interface where the data grouping could be done following observation type parameters, or such as an audio data interface allowing spectral analysis.

These features require a powerful hardware infrastructure perhaps bordering on a high-performance computing cluster. In addition, the data storage format will also need to be standardized, in addition to the current recommended metadata format standards. An optimized data format for analysis would have to be adopted.

7.5. True 3D and volumetric data visualization

Volumetric and 3D data are more and more available. As with audio and video, as of now the metadata and data access standards aren't mature enough and the volume of data is very taxing on the underlying infrastructure. One other difficulty is the computing power of the user devices required for 3D visualization and manipulation.

Much like with audio and video, lots of efforts are currently deployed in those fields. Technical capabilities and standards are growing and maturing rapidly. Other emerging technologies like virtual and augmented reality are not only pushing the technology forward but give new means for visualization and navigation through data. Framework like Mozilla A-Frame and WebVR are powerful 3D scene management that will mature or spawn better libraries for the implementation of Web interface in the future.

7.6. Open Source recommendation

Choosing a visualization tool cannot be done without properly considering the associated license.

Acquisition cost of visualization tools with a proprietary license must be carefully studied. Moreover, an in-depth evaluation of the terms of use (End User License Agreement – EULA) must be done so the proper rights and responsibilities can be determined (e.g. warranty and indemnity, intellectual property, limitation as to the number of computers/users, etc.)

Another category of visualization tools license are free diffusion or open licenses:

- Free software (i.e.: public-domain software)
- Free and open-source software
- Open source software

Overall, these license types grant more rights to the user. However, the obligations regarding the redistribution of modified source code and the production of sub licenses must be evaluated.

Finally, a reflection must be made regarding the visualization tools developed by the regional nodes. Although favoring contribution between the partners, free diffusion or open licenses may face issues related to the research and development as well as to the investments required by the development (technology acquisition, labor, etc.).

CONCLUSION

Canada, as a country surrounded by three of the world's oceans, and strong historical, cultural, societal, technological, and economic bonds with its oceans, has a unique opportunity to reinforce its leadership in ocean observing. With growing blue technologies and the rising open data mindset amongst more and more scientists, the timing for a Canadian Integrated Ocean Observing System is at its best. This is reinforced by the many efforts made in the previous years, that have built up to the launch of this current series of three investigative evaluations, initiated by Fisheries and Oceans Canada. Taken separately, these reports are filled with rich technical, structural and strategic information and recommendations. Merging these three reports into one shows the full potential of an upcoming CIOOS and the maturity of its stage, at the verge of its concrete realization. These reports include a concerted phased approach for the implementation of CIOOS, guiding the next steps notwithstanding the models of service that would be aimed for.

This report identifies currently existing visualization products, software suites, and interfaces by type of data, from a wide variety of ocean observing systems across Canada and abroad. Aligned with the other two IEs, the recommendations of this report focus on standards and services rather than specific tools. This ensures maximum autonomy for Regional Associations, providing them the flexibility to adjust their set of tools to their regional specificities and fields of expertise. The focus here is on interoperability and on ensuring recognized pathways for metadata and data to be discovered and visualized from one portal to another.

As an overall consideration, the data producer's attribution is of outmost importance. Indeed, even with the growing movement of open source, a lot of constraints still discourage many researchers to publish and share openly their data. As such, CIOOS visualization interfaces will have to display flawlessly the data producers' attribution and credit, for any service model and any phase of implementation.

Finally, this report widens its scope with a brief overview of additional steps and features that could be included in the implementation of CIOOS. Most recent technologies and high-volume data types (true 3D, higher-dimensional, volumetric, forecasts and video including virtual reality) has been examined to be kept in scope for future developments. Including cutting-edge visualization technologies and answering closely the users needs will without doubt solidify CIOOS purpose and reinforce its position at the forefront of ocean observing systems.

References

Bailey, K., & Bochenek, R. (2016, June 3). Environmental Sensor Map update. Integrated Ocean Observing System (IOOS). Retrieved from https://ioos.noaa.gov/wp-content/uploads/2016/06/E-Sensor-Map_DMAC-Mtg_Bailey.ppt.pdf

International Hydrographic Organization. (2016, November 3). *S-100 Universal Hydrographic Data Model*. Retrieved from IHO Web site: https://www.iho.int/iho_pubs/standard/S-100_Index.htm

Internet Engineering Task Force. (2016, August). *RFC 7946 - The GeoJSON Format*. Retrieved from IETF Web site: https://tools.ietf.org/html/rfc7946

Murphy, D. (2010, August 15). *Survey: 98 Percent of Companies Use Open-Source, 29 Percent Contribute Back*. Retrieved from News and Opinion - PCMag: https://www.pcmag.com/article2/0,2817,2367829,00.asp

Open Geospatial Consortium. (2017). *OpenGIS Web Map Tile Service Implementation Standard*. Retrieved from OGC Web site.

Open Geospatial Consortium. (2017). *Web Map Service*. Retrieved 11 15, 2017, from OGC Web site: http://www.opengeospatial.org/standards/wms

St-Laurent, A. M. (2008). *Understanding Open Source and Free Software Licensing.* O'Reilly Media.

Wikimedia Commons. (2010, February 19). *File:Bitmap vs vector.svg*. Retrieved from Wikimedia Commons: https://commons.wikimedia.org/wiki/File:Bitmap_vs_vector.svg

Wikipedia. (2017, July 24). *GIS file formats*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/GIS_file_formats

World Wide Web Consortium. (2013, June 27). *Linked Data Glossary*. Retrieved from WWWC Web site: https://www.w3.org/TR/ld-glossary/

World Wide Web Consortium. (2014, January 16). *Data Catalog Vocabulary (DCAT)*. Retrieved from W3C Recommendations: https://www.w3.org/TR/vocab-dcat/

APPENDICES

Appendix A. VISUALIZATION TOOLS INVENTORY

The following table was used to make an inventory and a preliminary assessment of currently existing visualization tools. Rows were highlighted in green to signify a particularly interesting option.

TABLE 8. INVENTORY AND PRELIMINARY ASSESSMENT OF EXISTING VISUALIZATION TOOLS.

Name	Reference Url	Type	Description	Standards	License	Supported Formats	Note
GeoServer	http://geoserver.org/	Map server	GeoServer is a Java-based software server that allows users to view and edit geospatial data. Using open standards set forth by the Open Geospatial Consortium (OGC), GeoServer allows for great flexibility in map creation and data sharing. **Visualization types:** Map, tabular	WMS, WMTS, WFS, WFS-T, WCS, WPS, FE, SLD, GML, KML	GNU General Public License : free	**Vector**: Shapefiles, External WFS PostGIS, ArcSDE, DB2, Oracle Spatial, MySql, SQL Server **Raster**: GeoTiff, JPG and PNG (with world file), image pyramid, GDAL formats, Image Mosaic, Oracle GeoRaster, NetCDF	https://live.osgeo.org/en/overview/geoserver_overview.html
MapServer	http://mapserver.org/	Map server	MapServer is an Open Source platform for publishing spatial data and interactive mapping applications to the web. **Visualization types:** Map	WMS, WFS, WCS, GML, GMLCOV, WMC, SLD, FE, SOS, OM, SWE, OWS, INSPIRE View Service compliant, WMTS through MapCache	MIT-style license: free	**Vector**: ArcInfo , ArcSDE, Contour, DGN, ESRI File Geodatabase, ESRI PersonalGeodatabase (MDB), ESRI Shapefiles (SHP), GML, GPS Exchange Format (GPX), , Inline, KML - Keyhole Markup Language, MapInfo, MSSQL, MySQL, NTF, OGR, Oracle Spatial, PostGIS/PostgreSQL, SDTS, S57, SpatiaLite, USGS TIGER, Vector field rendering , - UVraster, Virtual Spatial Data, WFS **Raster**: Formats supported through GDAL, see list here :	https://live.osgeo.org/en/overview/mapserver_overview.html

						http://www.gdal.org/formats_list.html	
QGIS Server	http://qgis.org/en/site/	GIS server	QGIS Server is an open source WMS 1.3 and WFS 1.0.0 implementation which, in addition, implements advanced cartographic features for thematic mapping. **Visualization types:** Map	WMS, WFS, SLD, WCS	GPL : free	**Vector**: Format supported through GDAL OGR library :http://www.gdal.org/ogr_formats.html Also support other formats through data provider plugin, like GRASS and PostgreSQL **Raster**: Format supported through GDAL: http://www.gdal.org/formats_list.html Grass support provided through plugin	https://live.osgeo.org/en/overview/qgis_mapserver_overview.html
Deegree	http://www.deegree.org/	GIS server	deegree is open source software for spatial data infrastructures and the geospatial web. deegree offers components for geospatial data management, including data access, visualization, discovery and security. Open standards are at the heart of deegree. **Visualization type:** Map	WFS, WMS, WMTS, CSW, WPS, GML	LGPL : free		https://live.osgeo.org/en/overview/deegree_overview.html **Low activity**
MapMint	http://mapmint.com/en/Home	GIS server/suite	MapMint is a Geographic Information System (GIS) software running on a web server made to facilitate the administration of a Spatial Data Infrastructure (SDI), publishing cartographic portals and webmapping applications.	Implements WPS, WFS, WMS, WMTS, WCS, WMC, GML, SLD	MIT : free		Use other tools to work. Server side : zoo-project, mapserver, gdal, libreoffice. Client

			Visualization type: Map				side : openlayers, jquery, html5, css3
MapGuide	https://mapguide.osgeo.org/	GIS server	MapGuide Open Source is a web-based platform that enables users to develop and deploy web mapping applications and geospatial web services **Visualization type:** Map	WMS, WFS	LGPL: free		Low activity
Geomajas	http://www.geomajas.org	GIS framework	Geomajas is a collection of free and open source GIS libraries, tools and API's for a complete end-to-end web mapping solution. **Visualization type:** Map	Might be possible through plugins.	AGPL : free		Does not seem to associate closely with OGC standards. Not strongly oriented toward standards support, but more toward consuming standard services. Low activity /popularity
Mapnik	http://mapnik.org/	Map toolkit	Mapnik is a Free Toolkit for developing mapping applications. It is written in modern C++ and has Python bindings that support fast-paced agile development. It can comfortably be used for both desktop map design and web development. **Visualization type:** Map	MVT, WMS	LPGL: Free		Seems powerful and scalable, but not geared toward standards.
Map Surfer	http://www.mapsurfer net.com/	GIS framework	MapSurfer.NET is a free, modern and advanced framework for producing a cartographic product of a high quality. This full featured framework offers a rich set of tools and techniques to automate various	WMS, WMTS	Custom : Free		Does not seem popular

			cartographic tasks. MapSurfer.NET is designed to be fast and flexible for being used both in desktop and web applications. **Visualization type:** Map				
OpenLayers	https://openlayers.org/	Dynamic map displayer / javascript library	OpenLayers makes it easy to put a dynamic map in any web page. It can display map tiles, vector data and markers loaded from any source. OpenLayers has been developed to further the use of geographic information of all kinds. It is completely free, Open Source JavaScript, released under the 2-clause BSD License (also known as the FreeBSD). **Visualization typs:** Map	Can display standard formats: WMS, WFS, KML, GML, WMTS, GeoJSON, WKT	FreeBSD: Free	WMS, WFS, KML, GML, WMTS, GeoJSON, MVT, topojson, EsriJSON, WKT, XML, IGC, GPX, JSONFeature, TextFeature, ArcGISRest	Oriented toward GIS, currently most powerful library.
Leaflet	http://leafletjs.com/	Dynamic map displayer /javascript library	Leaflet is the leading open-source JavaScript library for mobile-friendly interactive maps. **Visualization type:** Map				Less oriented toward GIS, more toward web / mobile friendly
D3JS	https://d3js.org/	Data based document manipulator / javascript library	D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation. **Visualization types:** Map, tabular, graph		BSD 3-clause: Free		Not a general purpose mapping api but has some basic map features (map projections, vector mapping). Powerful charting features. Extremely popular library, but can be complex.

GeoEXT	http://www.geoext.org/	Dynamic map displayer /javascript library	GeoExt is Open Source and enables building desktop-like GIS applications through the web. It is a JavaScript framework that combines the GIS functionality of OpenLayers with the user interface of the ExtJS library provided by Sencha.				Seems to have trouble keeping up with OpenLayers and EXT JS versions
Tangram	https://mapzen.com/products/tangram/	Dynamic map displayer/ javascript library	Tangram is an open-source map renderer designed to grant you ludicrous levels of control over your map design. By drawing vector data live in a web browser or mobile device, it allows real-time map design, display, and interactivity. Using OpenGL, Tangram saddles and rides your graphics card into a new world of cartographic exploration. Animated shaders, 3D buildings, and dynamic filtering can be combined to produce effects normally seen only in science fiction. Map styles, data filters, labels, and even graphics card code can be defined in a plaintext scene file, and APIs permit direct interactive control of the style. **Visualization type:** Map	TopoJSON, GeoJSON, MVT			Seems good for 3D, dynamic styling, shaders and overall control of style.
Googlemap api	https://developers.google.com/maps/	Dynamic map displayer /javascript library	Google Maps is a web mapping service developed by Google. It offers satellite imagery, street maps, 360° panoramic views of streets (Street View), real-time traffic conditions (Google Traffic), and route planning for traveling by foot, car, bicycle (in beta), or public transportation. **Visualization type:** Map	Can display standard formats: WMS, WMTS, KML, GeoJSON	Custom. Standard = free with up to 25,000 map loads per day. Premium = rate	WMS, WMTS, KML, GeoJSON	Not oriented toward GIS, oriented more toward human activities.

					calculation based on type of API request		
THREDDS	http://www.unidata.ucar.edu/software/thredds/current/tds/	Open source data portal (of real-time and archived datasets with data transport architecture and protocol)	The THREDDS Data Server (TDS) is a web server, more specifically a tomcat servlet, that provides metadata and data access for scientific datasets, using a variety of remote data access protocols. Client could be an ordinary browser, although this gives limited functionality. Usually, a client is a graphics program (like GrADS, Ferret or ncBrowse) or web application (like DChart) linked with a library corresponding to the protocol used. **Visualization types:** Map (WMS)	**Metadata standard**: Dublin Core, FGDC/DIF, ISO 19115, ADN **Protocols**: OPeNDAP (DAP2), ncWMS implementing WMS, WCS, HTTP (bulk file access) **Experimental protocols**: CDM RWS, NEXRAD (level 2 & 3)	netCDF library license (MIT-style): free	NetCDF (v. 3 & 4), HDF5, NcML, GRIB, NEXRAD	
CKAN	https://ckan.org/	Open source data portal and data management	CKAN provides a streamlined way to make your data discoverable and presentable. Each dataset is given its own page for the listing of data resources and a rich collection of metadata, making it a valuable and easily searchable data catalogue. **Visualization types:** Map, tabular, graph	**Metadata standard:** CSW, INSPIRE, DCAT, DCIP	AGPL v.3	All formats can be managed/stored. **Viewable formats** (some require plugin activation available in	There are many more view plugins developed by the CKAN team and others which are

						CKAN core): CSV, XLS, HTML, JSON, PDF, RSS, TXT, WMS, XML	hosted on separate repositories.
ERDDAP	https://coastwatch.pfeg.noaa.gov/erddap/index.html	Data server	ERDDAP serves as middleware between providers and consumers of oceanographic data. It aggregates datasets from remote sources and provides services for these datasets such as search, format translation, graphing and time standardization. ERDDAP uses DAP (Data Access Protocol) as the data transport mechanism through which clients can receive formatted data. **Visualization types:** Map, tabular, graph	WMS, UCUM, UDUNITS, OPeNDAP, OBIS, SOS, WCS, WFS, ISO 8601:2004 "extended (Date)	Apache: free open source	**Special types**: .html, .graph, WMS, SOS, OBIS, OPeNDAP **Grid file types**: http://coastwatch.pfeg.noaa.gov/erddap/griddap/documentation.html#fileType **Table file types**: http://coastwatch.pfeg.noaa.gov/erddap/tabledap/documentation.html#fileType **Image file types**: .geotif, .kml, .smallPdf, .pdf, .largePdf, .smallPng, .png, .largePng, .transparentPng	Used by a lot of oceanographic data providers. Provide quite incredible interoperability. Consistent, strong search options. No bilingual support
GeoNetwork	http://geonetwork-opensource.org/	geo catalog	GeoNetwork is a catalog application to manage spatially referenced resources. It provides powerful metadata editing and search functions as well as an interactive web map viewer. It is currently used in numerous Spatial Data Infrastructure initiatives across the world. **Visualization type:** Map	CSW, OAI-PMH, OpenSearch-GEO, Z39.50 Metadata standard: ISO19115, ISO19119, ISO19139, ISO19110, Dublin Core	GPL v2 : free open source	Can upload/download any file format. **Harvesting**: OGC=CSW 2.0.2, OAI-PMH, Z39.50, Thredds, Webdav, Web Accessible Folders, ESRI GeoPortal, Other GeoNetwork nodes **View**: Map viewer based on OGC	Used by many large organisations around the world. CKAN can connect to GeoNetwork. Multilingual interface. Can harvest from numerous other sources. Significant

						can view formats like WMS, WMTS, KML, OWS	UI improvements in recent versions.
OceanNavigator	https://github.com/geoffholden/ocean-navigator	Data visualization tool	Ocean Navigator is a Data Visualization tool that enables users to discover and view 3D ocean model output quickly and easily. The ocean model output is stored in NetCDF files. These files are self-describing and contain the 2D or 3D model output for one or more timesteps and one or more variables. To facilitate reading all these files, we make use of a server called THREDDS Data Server. THREDDS aggregates all the NetCDF files and allows users to query subsets of the files. The server-side component of the Ocean Navigator is written in Python, using the Flask web API **Visualization types:** Map, tabular, graph				
Ocean 2.0	http://www.oceannetworks.ca/innovation-centre/smart-ocean-systems/ocean-observing-systems/oceans-20		A unique and critical component developed to connect the subsea instruments systems, providing the capability for the 24/7 acquisition of extremely diverse and vast amounts of data, quality control and calibration, storage, visualization and access by a potentially global audience, as well as providing a convenient interface to handle otherwise complex tasks associated with the remote monitoring and control of the observatory infrastructure itself. **Visualization types:** time series, profiles, contour sections, spectrograms, maps, scrolling displays, audio and video players				

GeoNode	http://geonode.org/	Geospatial Content Management System	GeoNode is a web-based application and platform for developing geospatial information systems (GIS) and for deploying spatial data infrastructures (SDI). It is designed to be extended and modified, and can be integrated into existing platforms.		GPL v3		
Esri Geoportal	http://www.esri.com/software/arcgis/geoportal	geo catalog	Esri Geoportal Server is a free, open source product that enables discovery and use of geospatial resources including datasets, rasters, and Web services. It helps organizations manage and publish metadata for their geospatial resources to let users discover and connect to those resources. The Geoportal Server supports standards-based clearinghouse and metadata discovery applications. **Visualization type:** Map	CSW, INSPIRE, ISO 19115, ISO 19119, ISO19115-2, GEMINI, Dublin Core, WMS, ISO 23950	Apache 2.0 : free open source	Can upload/download most formats Can act as client for : ArcIMS, ISO 23950, OAI-PMH,WAF, WMS, WFS, GML,KML	Similar to Geonetwork, however seems less active
ArcGIS javascript api	https://developers.arcgis.com/javascript/	Dynamic map displayer /javascript library	The 4.x series of the ArcGIS API for JavaScript is Esri's next-generation JavaScript API that integrates 2D and 3D into a single, easy-to-use, powerful API. Version 4.5 lets you build full-featured 3D applications powered by web scenes that can include rich information layers such as terrain, basemaps, imagery, features, integrated mesh layers, and 3D objects. **Visualization type:** Map	Can display standard formats such as: WMS, WFS, KML, WMTS,CSV	Custom: free for NGO / not-for-profit	WMS,WMTS,WFS,WCS,KML,CSV, GeoRSS, GeoJSON, ArcGIS proprietary formats	Similarities with OpenLayers, however less popular but with 3D features
ArcGIS Server Enterprise	http://www.esri.com/en/arcgis/products/arcgis-	gis server	ArcGIS Server is the core server geographic information system (GIS) software made by Esri. ArcGIS Server is used for creating and managing GIS Web services, applications, and data. ArcGIS Server is typically deployed on-premises within the organization’s service-oriented	WMS, WFS, WCS, WMTS, WPS, SLD, KML, GML + arcGIS	Custom: request quote / expensive	JPIP, KML, databases, NetCDF, shapefiles, text files, SDC files, rasters	

	enterprise/overview		architecture (SOA) or off-premises in a cloud computing environment.[2] **Visualization type:** Map	proprietary formats			
Caris Spatial Fusion Enterprise	http://www.caris.com/products/sfe/	gis server	Spatial Fusion® Enterprise offers robust server tools for managing web applications and their spatial data presentation, as well as a simple yet powerful end user interface. A proprietary framework for data drawing, including map tiling and caching strategies, allows users to seamlessly navigate small to very large datasets. **Visualization type:** Map	WMS, WMTS, WFS, WCS, GML, GeoRSS, KML			

Appendix B. List of Acronyms

ADCP	Acoustic Current Doppler Profilers
ADN	Autodesk Developer Network
API	Application Programming Interface
ArcIMS	Arc Internet Map Server
AZMP	Atlantic Zone Monitoring Program
BC	British Columbia
CBD	Convention on Biological Diversity
CCA	Canadian Council of Academies
CIOOS	Canadian Integrated Ocean Observing System
CKAN	Comprehensive Knowledge Archive Network
CODAR	Coastal Ocean Dynamics Applications Radar
CSS	Cascading Style Sheets
CSV	Comma-separated values
CSW	Catalog Service for the Web
CTD	Conductivity, Temperature, and Depth instrument
DAC	IOOS Data Assembly Centers
DAP	Data Access Protocol
DCAT	Data Catalog Vocabulary
DFO	Fisheries and Oceans Canada
DIF	Directory Interchange Format
DOI	Digital Object Identifier
EPSG	European Petroleum Survey Group
ERDDAP	Environmental Research Division's Data Access Program
Espoo	Convention on Environmental Impact Assessment in a Transboundary Context
EULA	End User License Agreement
FFT	Fast Fourier Transform
FGDC	Federal Geographic Data Committee (USA)
FGP	Federal Geospatial Platform (CAN)
FLAC	Free Lossless Audio Compression
GB	Gigabyte
GBIF	Global Biodiversity Information Facility
GDAL	Geospatial Data Abstraction Library
GDP	Gross domestic product
GEO	Group on Earth Observations
GEOSS	Global Earth Observation System of Systems
GeoTIFF	Georeferenced Tagged Image File Format
GIS	Geographic Information System
GML	Geography Markup Language
GOOS	Global Ocean Observing System
GPS	Global Positioning System

GRIB	GRIdded Binary
GZ	Gnu Zipped file archives
HDF	Hierarchical Data Format
HEVC	High Efficiency Video Coding
HFR	High Frequency Radar
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
HYD	Hydrophone - ONC custom digital file format
IE	Investigative Evaluation
IHO	International Hydrographic Organization
IOOS	Integrated Ocean Observing System (USA)
ISO	International Organization for Standardization
IT	Information Technology
JPEG	Joint Photographic Experts Group
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
KMZ	KML file zipped
LIDAR	Light Detection and Ranging
LML	Lightweight Markup Language
MARPOL	MARine POLlution - International Convention for the Prevention of Pollution from Ships
MEOPAR	Marine Environmental Observation Prediction and Response
MP3	Moving Picture Experts Group (MPEG)-3
MSP	Marine Spatial Planning
NcML	NetCDF Markup Language
NetCDF	Network Common Data Form
NetCDF-CF	Network Common Data Form – Climate and Forecast
NEXRAD	Next-Generation Radar
NGO	Non-governmental organization
NRCAN	Natural Resources Canada
NU	Nunavut
OAI-PMH	Open Archives Initiative - Protocol for Metadata Harvesting
OBIS	Ocean Biogeographic Information System
OGC	Open Geospatial Consortium
ONC	Ocean Networks Canada
OSPAR	Convention for the Protection of the Marine Environment of the North-East Atlantic
PDF	Portable Document Format
PNG	Portable Network Graphics
QDC	Qualified Dublin Core
RA	Regional Association
REST	Representational state transfer web services
RFC	Requests for Comments

ROV	Remotely Operated underwater Vehicle
RSS	Rich Site Summary or Really Simple Syndication
SLD	Styled Layer Descriptor
SLGO	St. Lawrence Global Observatory
SOAP	Simple Object Access Protocol
SOS	Sensor Observation Service
SWE	Sensor Web Enablement
TDS	THREDDS Data Server
THREDDS	Thematic Real-Time Environmental Distributed Data Services
UNCLOS	United Nations Convention on the Law of the Sea
UNESCO	United Nations Educational, Scientific and Cultural Organization
URL	Uniform Resource Locator
USA	United States of America
UTC	Coordinated Universal Time
UTF	Unicode Transformation Format
WAF	Web Application Firewall
WAV	Waveform audio file format
WCS	Web Coverage Service
WFS	Web Feature service
WGS	World Geodetic System
WMS	Web Map Service
WMTS	Web Map Tile Service
XLS	Microsoft Excel file format
XML	eXtensible Markup Language
YSLD	YAML-SLD

Appendix C. CIOOS IE Visualization Statement of Work

Visualization tools & platforms

This IE will create an inventory of what visualization products currently exist that address the needs of the ocean observation community, and evaluate those existing platforms with regards to how they could be best leveraged or built upon to support CIOOS. This IE will also evaluate the pros and cons of different software options, for example the use of open source software, and identify best practices which should be followed. Based on what's available, this IE will identify what the CIOOS presence could and should look like (e.g. IOOS, SLGO, Oceans 2.0, MSDI, Open Canada, etc.). The Federal Geospatial Platform (FGP) is also noted as an important platform, however, it is not visible to anyone outside of the Government and therefore alternative options would be required to support GOOS and CIOOS' role as a GRA. It is recommended that this IE also survey the community to determine their visualization needs and ensure CIOOS can address the priority end-user requirements.

This IE is expected to identify the baseline set of products and tools that CIOOS will provide, including the common visualization elements (maps, models etc.), and required resolution of the data products. The IE should outline the tools and costs required to achieve that baseline, as well as propose additional products that could be added in a phased approach and the associated costs and resources those products would require. For example, the European Marine Observation and Data Network (EMODnet) is being developed in three major phases, from a prototype with coverage of a limited selection of sea-basins, parameters and data products at low resolution. They also use data portals to help organize and visualize their data.

Recommendations will include an outline of what the baseline CIOOS presence should look like, including the products it will provide and what the requirements are to achieve that and the costs involved. The recommendations should also propose potential next steps in a phased approach, and the associated requirements and costs.