OTC 25474

Design and Development of a Greenland Ice and Metocean Geoportal

Ross, E., Loos, E., Fissel, D.B., Lapidakis, Y., and Zhang, O., ASL Environmental Sciences, Inc.

Abstract

Completed and planned metocean and ice measurement programs off Greenland's eastern and western coasts result in large and varied datasets characterizing physical phenomena such as icebergs, sea ice, seabed scours, weather, surface waves, ocean currents, and water properties. Future planned measurement programs will expand on the spatial and temporal breadth of these datasets. Other datasets support the analysis of the measurement data including license area locations, bathymetry, glacier calving areas, and notable submarine features.

In order to plan measurement programs, manage the acquired datasets, and use the data for characterization of the physical environment, a web-based geoportal was designed and developed. The geoportal aids scientists and engineers in their discovery and use of metocean and ice data. The geoportal development required balancing two aspects. Firstly, scientists and engineers have extensive needs to upload, organize, search, visualize, analyze, and download large and varied datasets. Secondly, there are inherent limitations of web technologies due to bandwidth, latency, and security constraints. Many design decisions were focused on balancing these issues and are presented here.

Introduction

The scale of environmental baseline and engineering data necessary for offshore exploration and production projects is large. The required data are multi-dimensional spanning extended geospatial regions, long temporal episodes, and numerous multi-attribute physical phenomena. The offshore environment of Greenland requires characterizing distinctive physical features (sea ice, glacial ice, seabed ice scours) and common key features (meteorology, waves, ocean currents, water quality). Measurement and data analysis programs involve a large number of organizations. Additionally, multiple users of varying disciplines from various organizations require timely access to these data.

A geoportal is a software system that provides web-based access to geospatial data. An effective geoportal must provide tools for performing simple to complex data and metadata search queries, mapping and listing of search results, data visualization and analysis functionality, and the ability to download search results in convenient file formats. ASL Environmental Sciences designed and implemented a geoportal to provide these functionalities.

This paper presents some of the lessons learned while developing the geoportal, including: 1) outlining the main geoportal requirements, 2) identifying the challenges involved in bridging the disciplines of physical marine environmental field data collection, information system design, and the science and engineering of offshore oil and gas exploration and production, and 3) introducing relevant modern web-based technologies.
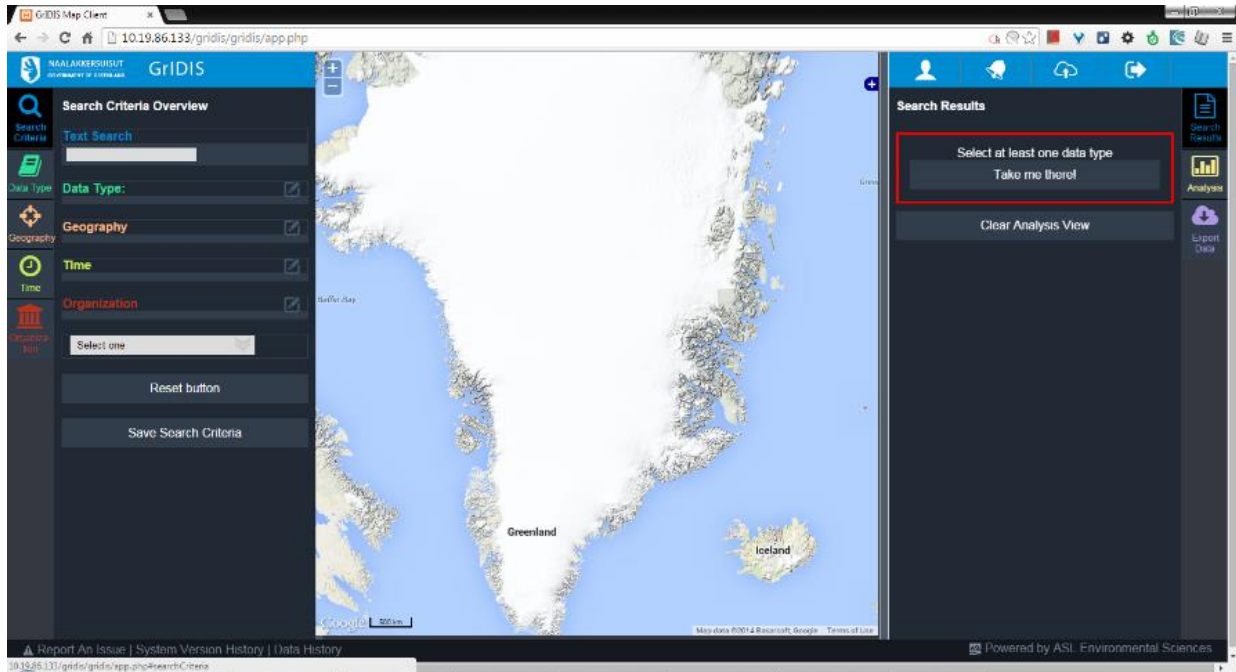
The paper is organized in the following sections:
- Main Functionality – The key functional requirements of the geoportal.
- Other Functional and Non-functional Features – Elements of the geoportal that support the main functionality.
- Innovations and Highlights – A summary of geoportal features that were developed to address constraints inherent to web applications.

- Development Tools and Methods – Highlights of the project team, project execution methodology, and software tools used to successfully complete the project.

Main Functionality

In order to access the geoportal, the user must sign in and be authenticated. After login, the user accesses all geoportal functions via a single webpage (Figure 1). This single interface provides all of the functions necessary to perform data and metadata search, visualization, analysis, and export.

Figure 1. Landing page after authentication. This main interface is divided into three main areas: the user sets search criteria in the left-hand panel, the center panel displays search results on an interactive map, and the right-hand panel lists search results and provides data analysis tools.

Search

When interacting with a large and varied data archive, finding records of interest is often the first challenge. A search utility is often needed when simply browsing is insufficient given the quantities and variety of possible results. The geoportal search engine (left-hand panel of Figure 1) handles simple and exhaustive criteria and displays results immediately as the search criteria are updated. A summary panel shows the defined search criteria. The user can also load or save a set of search criteria. The other tabs on the left-hand panel provide grouped search criteria (Figure 1) as follows:

- Data Type (Figure 2a) – The user must select one or more of fifteen categories of physical data items, which are grouped by physical medium as listed in Table 1. A final category provides access to local knowledge study results that do not involve quantitative data as in the listed data type categories.

Table 1. Selectable data types for inclusion in search results.

Physical medium	Data type	Description
Ice	Grounded icebergs	Iceberg targets extracted from satellite imagery and classified as grounded. Time-series of image target sets acquired at intervals.
	Iceberg aerial density	Time-series of spatial maps of iceberg occurrence density.
	Iceberg drifts	Position track lines and time-series of iceberg attributes.
	Ice draft	Time-series (equispaced in time) and spatial-series (equispaced in distance) profiles of ice draft acquired by upward-looking sonar (ULS) on fixed subsurface moorings.
	Ice velocity	Time-series of ice velocity acquired by ULS on fixed subsurface moorings.
	Individual icebergs	Iceberg targets extracted from satellite imagery. Time-series of image target sets acquired at intervals.
	Sea ice concentration	Time-series of ice concentration maps, i.e. regions classified as 1/10ths of total ice coverage.
	Sea ice mobility	Time-series of ice mobility maps, i.e. regions classified as one of five mobility values.
Metocean	Meteorology	Time-series of wind velocity, atmospheric pressure, air temperature, humidity and other attributes at fixed metocean buoys and from mobile vessels.
	Ocean currents	Time-series profile of currents acquired by ULS on fixed subsurface moorings.
	Waves	Time-series profile of currents acquired by ULS and metocean buoys on fixed moorings.

	Water quality	Time-series of salinity, temperature, and pressure at fixed locations and from mobile vessels.
Seabed	Scour areal density	Time-series of spatial maps of seabed ice scour occurrence density.
	Individual scours	Ice scours extracted from side-beam and multi-beam sonar with numerous geometric and compositional attributes.

- Geography (Figure 2b) – Four methods allow the user to add geographic constraints to a search query:
  - Custom Polygon – The user enters coordinates of the vertices of a geographic region (polygon).
  - Map Control – The user defines a polygon by clicking points on the map.
  - License Area – The user selects from a pre-defined list of exploration license areas to use as a geographic search filter.
  - Saved Polygons – The user can save a polygon to be used in a search by uploading a standard GIS-format file containing the polygon. The upload is saved in the user's profile for future use.
- Time (Figure 2c) – The user can temporally restrict the search results to a single episode using start and end dates. Alternatively, search results that occur in a periodic time sequence (e.g. every January) can be obtained.
- Organization – The user can restrict the search results to those associated with study procurer and contractor organizations.
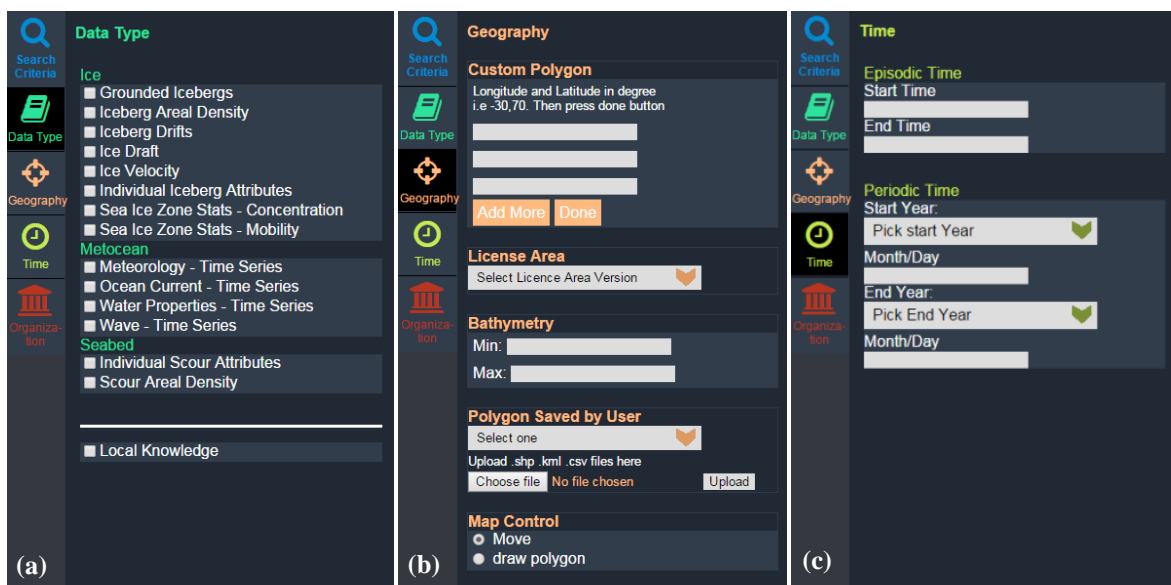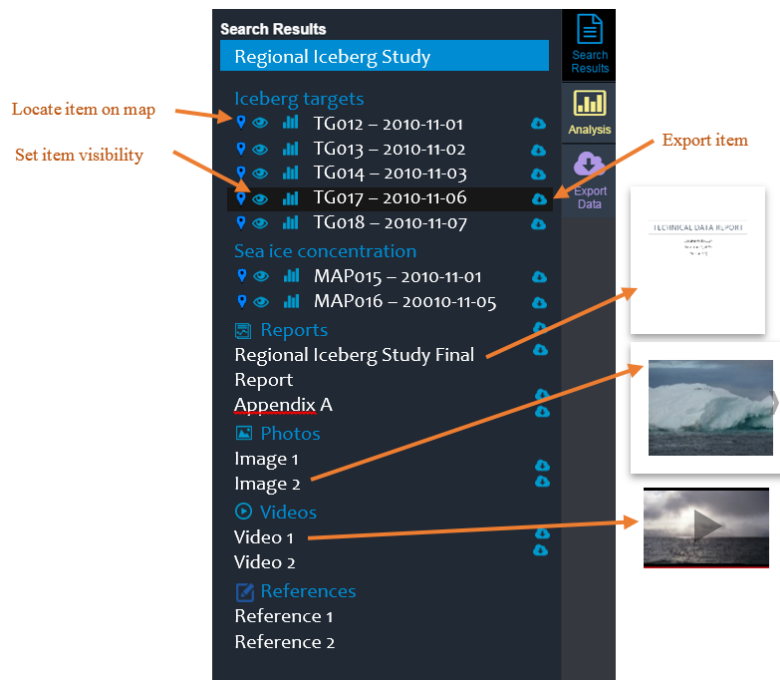

Figure 2. The geoportal search options from left to right: (a) Data Type, (b) Geography, (c) Time.

Visualization

Figure 3 shows the search results panel on the right-hand side of the main interface after a search query has been submitted.

The geoportal groups its archive into studies or projects. Each study contains data items (with geospatial, temporal, and other attributes) and supplementary information such as reports, images, and video. The search results are similarly grouped. The data items are also displayed on the map interface as points, vessel and drifter tracks, target positions and shapes (icebergs and ice scours), and density and probability maps. As the user hovers a listed search result, the corresponding map item is highlighted and brought to the front of any overlapping map features. A mapped item can be quickly located through clicking an icon next to the corresponding listed item, which will immediately pan the map to that item's geographic location. The visibility of any mapped item can be set via an icon next to the listed item.

Study and contents metadata are immediately retrieved and displayed when the user hovers over any item in the search results list. The supplementary study information in the form of documents and multimedia files (photos, images and video) can be viewed by clicking the item name (see the flyout examples in Figure 3). The export icon next to any data item allows the user to download the data in a GIS-compatible format.

Figure 3. Search results panel. The results are grouped by studies. Each study lists individual datasets grouped by their data type (e.g. iceberg targets). These search result items are also represented on a map. Icons beside each item allow the user to locate the item on the map, toggle the item's visibility on the map, and export the item in GIS-compatible format. Selecting an entry in the reports, photos, videos and references sections will display that item on the screen. Photos are displayed as a navigable gallery. Videos can be played.

Analyze

Fundamental statistical analysis tools can be used to explore the search results. The icon immediately preceding each data item in the search results example shown in Figure 3 selects that item for further analysis. The Analysis panel is immediately displayed with an attribute of the data item selected by default (e.g. iceberg width). The selected item represents either a distribution type (a geospatially extended dataset with a population of samples) or a time-series type (a set of time-series measurements acquired on a fixed mooring, vessel or drifter).

An example of a distribution type is a set of iceberg targets acquired from an image. The user may select from the multiple iceberg attributes and obtain plots of frequency and exceedance distributions and a tabulation of summary statistics for the selected attribute (Figure 4a). The user can also select two attributes to view a bivariate distribution scatter plot (Figure 4b). These plots are rendered promptly for a quick survey of the selected data. Multiple data items from the search results can be selected and combined in the analysis results (Figure 4c).

An example of a time-series type is a profile of ice draft obtained from a fixed subsurface mooring (Figure 4d). The user may scroll and zoom a plot of any of the selected item's attributes. As the user navigates the dataset, summary statistics of the full dataset and the portion selected to view are updated.

Export

The Export panel (Figure 5) serves two purposes. Firstly, a history of recent downloads is listed providing quick access to these links. Secondly, a listing of snapshots of analysis results is displayed. These snapshots can be acquired at any time in the Analysis panel allowing the user to save interesting features discovered in the datasets. The full snapshot listing can be downloaded as a PDF document containing the saved analysis results.
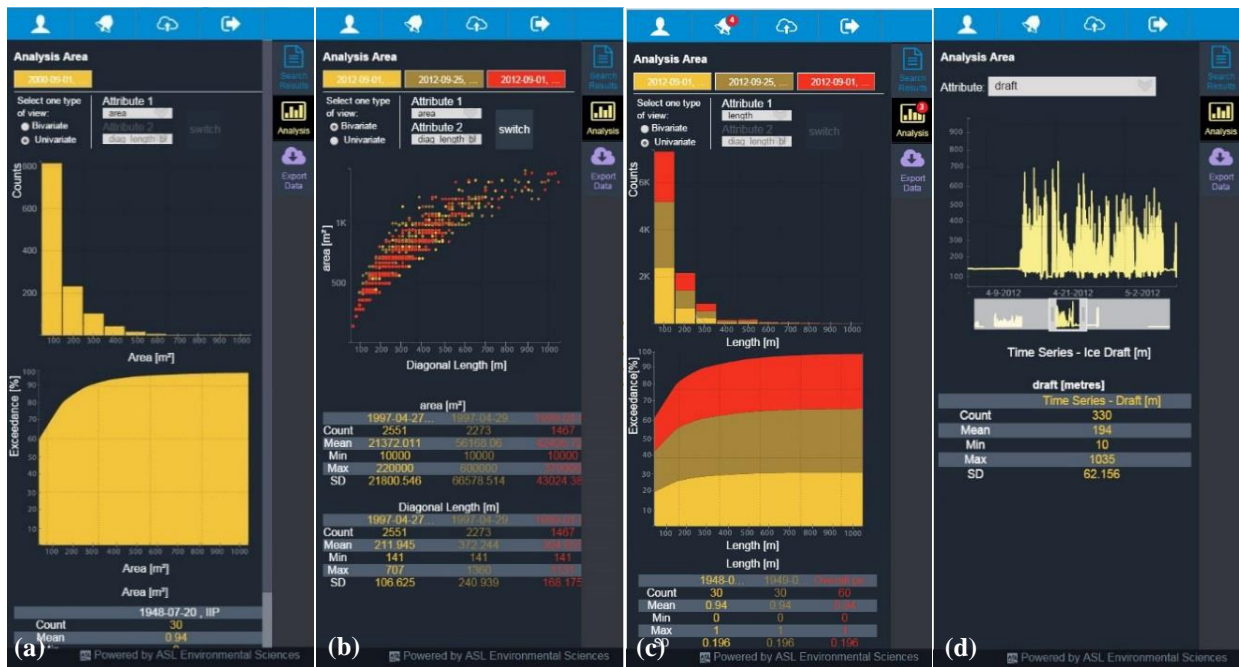
Figure 4. Data analysis panel provides further insight into the search results: (a) frequency and exceedance plots and summary statistics for a selected distribution type data item, (b) bivariate, (c) multiple distribution type items selected, (d) navigable (scroll/zoom) plot and summary statistics of selected time-series type data item.
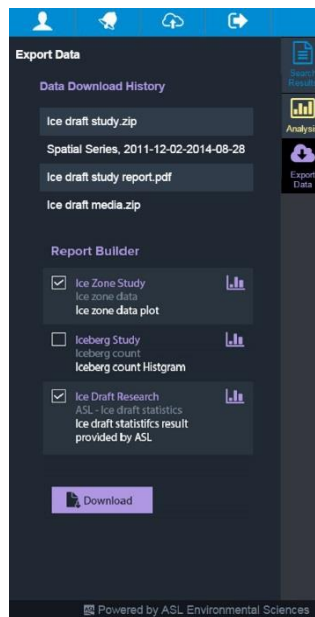

Figure 5. The export panel displays the history of recent downloads and allows the user to download generated analysis results as a PDF document.

Other Functional and Non-functional Features

In addition to the main functions of search, visualize, analyze and export, a number of secondary functions are accessible through a control panel (Figure 6). The control panel is a tabbed overlay that is visible over the main interface, which is disabled and faded, i.e. the control panel is modal. Tabs on the left-hand side of the control panel access the different functions:

- Personal Information – The user can review and edit login and contact information.
- Reported Issues – The user can review any issue (bug fix, feature request, etc.) that they have submitted and the dialog thread with a system administrator. The user can also submit a new issue report.

- Provided Data – The user can review the list of datasets that they have uploaded. The uploaded datasets (data files and metadata) are quality controlled and prepared for insertion into a database. After insertion the web-based interface is tested to ensure the inserted data are visible to the search and mapping engines. The progress of this process can be tracked through the list of provided datasets. The user can also upload a new dataset.
- Saved Search Criteria – The user can review the list of search criteria sets that they have saved, edit their labels, and delete saved search entries.
- Uploaded Polygons – The user can review the list of uploaded polygons, edit their labels, and deleted polygon entries.

Figure 6. The control panel allowing the user to edit their personal profile information, maintain their uploaded data, search criteria and polygons, and track the status of reported issues.

An alert icon in the top right-hand area (Figure 4c) will be overlaid with a count of the number of unread alert messages for the user. Alert messages include system maintenance announcements, progress updates on the insertion of new study data, and feedback on reported system issue. The alert messages are shown when the user hovers above the alert icon. Some messages may lead to further user action, e.g. a response to a requested feature enhancement is received and the user would like to respond further. In these cases, clicking the message will open either the Reported Issues tab or the Provided Data tab of the control panel.

A number of non-functional features ensure that the geoportal is performant. The entire geoportal functionality is available through the main interface. There is only one "screen" in the geoportal. Less used functions (such as the control panel described above) are accessed through overlays on top of the main interface. Keeping the user interaction to a single page ensures that their progress on searching and navigating results is not lost while accessing secondary functions.

As a single-page application, the only event that can upset the interface is when the user's web browser window closes. This may happen accidentally (e.g. power loss, computer failure) or purposely (e.g. user manually closes web browser). In any case, the goal is to display the same state of the interface to the user upon their return to the geoportal. Restoring the interface state includes setting search criteria, panning/zooming the map, displaying search results including selected items for analysis, and selecting the left-hand and right-hand tabs.

Many layers of security are integrated throughout the geoportal. All users must be authenticated with a username and password to access the geoportal. Dataset providers determine who can gain access to uploaded data and metadata by setting permissions for individuals or all members of particular organizations. Every interaction between the web browser and the web server include checking the user's authentication. When the user enters information into the web interface, the entered values are checked for valid ranges and format. Modern known common threats are sanitized from any user input. When this information is bundled into a request and transmitted to the web server, the server itself sanitizes the information and checks for valid values. Web server requests may include a read query executed on the geospatial database. These read queries are

performed by a limited-access system user to prevent non-read-only interactions with the database. Finally, encrypted access to the geoportal is accomplished through the secure HTTPS protocol.

Responsiveness of geoportals is a key indicator of performance. Upon any user interaction, the web interface must respond quickly with cues to the user of the results of their actions. A long execution time between submitting a search query and seeing the mapped results on the screen can render the system unusable, for example. Responsiveness was built into the geoportal through two mechanisms. Firstly, after an interaction the user interface responds by guiding the user to the expected following action. For example, selecting a search result item for analysis displays the Analysis panel and begins loading the selected item and item attribute analysis results. Other cues include visually highlighting mapped items when the user hovers above corresponding search result items and using overlays on tabs and links to indicate contained information (e.g. a count of the number of unread alert messages overlaid the alert icon, a count of the number of selected data items for analysis overlaid the Analysis tab).

Secondly, numerous strategies were used to cope with the bandwidth limitations inherent in web applications. All of the information to be displayed in the web interface must be transmitted from the web server. This constrains the amount of information that is feasibly transferable while retaining responsiveness in the interface. In particular, this challenge is significant when displaying a large number of search result items in the search results list and map and when analyzing and navigating large datasets. Scrolling and zooming of a time-series dataset may require navigating a dataset containing millions of records. Mapping a set of imaged icebergs may require displaying thousands of individual complex polygons on the map. The objective was to handle these situations while maintaining a timely response regardless of the scale of visible data. Solutions to accomplish this objective are discussed in the following section.

Innovations and Highlights

As a web application, the geoportal is easily accessed using any recent version of the major web browsers. Unlike desktop applications, a web application does not require installation or configuration and any application updates are immediately available to all users. The geoportal does not require the use of any web browser plugins so the user will never be prompted to install software.

A key innovation objective used throughout the geoportal design was to mitigate the negative impact of limited bandwidth on the system performance. Several geoportal functions can lead to this. The following lists these functions and the mitigation strategies used to ensure their performance.

Performing a search
After editing an individual search criterion, the user receives results in their web browser within a few seconds. This includes updating the list of search result items (Figure 3) and the map. A search can produce a considerable amount of varied map content. The larger map items include: thousands of iceberg polygon outlines, long and complex drift tracks, and diverse density and probability maps. It is not feasible to transfer the complete representation of these items due to their sheer data size. Therefore, the web server creates a compressed version of a map item greatly reducing the amount of data transferred to the web browser. The compression is performed such that the item eventually rendered in the map is visually equivalent to a rendering produced by the uncompressed data. As the user zooms into a region of the map dense with search results, more detail of the mapped results are exposed through the server compression process.

Navigating the map
Search results can appear anywhere in the map including regions that are out of the field of view. The geoportal only loads the map content that would appear in the map extent that is currently viewable. As the user pans and zooms the map, the web server provides additional content to display.

Locating an expected search result on the map can be challenging if it is not within the current map extent. The user can automatically pan the map to the location of a search result item by clicking the pin icon adjacent to the search result item. Additionally, as the user hovers over a search result item, the corresponding mapped item appears in the foreground of any overlapping features and is highlighted causing it to be immediately locatable.

Analyzing selected search results
When a search result item is selected for analysis, the results are displayed within a few seconds. The analysis results are computed by the web server resulting in a small data packet size to transfer to the web browser. This includes computing probability distributions and summary statistics.

Scrolling and zooming a navigable time-series plot must provide renderings of datasets up to millions of records. Transferring the raw data to plot in the web browser is unachievable due to the size of the dataset. Therefore, multiple

versions of each time-series are created when inserted into the geoportal database. The versions capture the complete time-series at various resolution levels. These are used to maintain a low data packet size transferred to the web browser effectively via a lossless compression scheme.

Development Tools and Methods

The success of the geoportal development required a team with diverse technical expertise who represented all stakeholders, a project execution methodology that suited the envisioned system scope, and a set of tools suited to the challenges of the project. These elements are described in more detail in the following sections.

The Team

The conceptual design, planning, implementation, and testing of the geoportal were conducted by a team of individuals with different expertise in web application software engineering, database design, metocean and ice science, oceanography, cartography, and geography. Initially, the conceptual design of the geoportal was drafted according to the client's requirements. Familiarity with geospatial systems and oceanographic data enabled translation of the client's needs into the basic requirements of the geoportal. In addition to incorporating the latest functionality available in common geoportal systems, new data visualization and analysis tools were developed. The system implementation required programming expertise in the very different languages, platforms and code libraries available for the system's front-end (the web browser interface) and back-end (the web server and database).

Clear communication between the project team and the client was essential to fulfill the project requirements. A steering committee was created by the client to oversee the progress and compliance of work to the proposed requirements. The steering committee was comprised of individuals with diverse backgrounds and expertise in offshore oil and gas, nautical and ice engineering, and metocean science and provided feedback relevant to the development of the geoportal through: the evaluation of monthly status reports, review of preliminary and detailed design documents, and system prototype demonstrations.

An in-person workshop was organized by the project team after early system design elaboration to allow for user interface testing. This proved to be vital to the development of the geoportal system as the project team had the opportunity to reconcile the user's needs against a prototype and to obtain detailed information on particular dataset characteristics relevant to the design of the database. The interactions during the two-day workshop yielded valuable information that was incorporated into the geoportal throughout the subsequent development phases.

The Development Methodology

An object-oriented analysis and design approach (Satzinger, Jackson & Burd 2005) was taken from the start of the project in order to iteratively elaborate the system requirements. The first iteration began by examining the results of a system engineering technique called "use-case analysis" that was performed while creating the proposed system conceptual design. This analysis of the geoportal requirements led to an initial definition of the main system functionalities (use-cases):

- Insert datasets,
- Search datasets,
- Visualize search results,
- Analyze search results,
- Export search results,
- Update geoportal component,
- Backup/restore geoportal system,

and the user roles and external systems that would interact with these functions:

- Data provider,
- Data analyst,
- Data manager,
- System administrator, and
- Web application hosting service.

Iterations of system modeling through use-case analysis significantly expanded the steps of each use-case and identified multiple scenarios under which each use-case could occur. Software engineering techniques were used to translate the use-case details into distinct concepts within a system model. Each modeling iteration adhered to a principle of stepwise refinement, i.e. a top-down approach to progressively elaborating system functionality by dividing a use-case into tasks and sub-use-cases. The Unified Modeling Language (UML) (Rumbaugh, Jacobson & Booch 2004) was used to represent the system components and the interactions between components. UML defines a number of diagrams that the software engineer
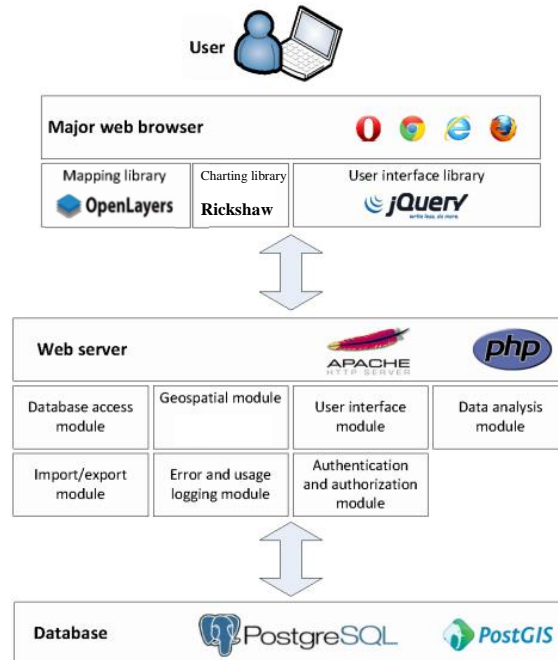
can use to describe system software components and the messages that must flow between these components. Coupling these system model diagrams with mock-ups of the geoportal's graphical user interface (i.e. the web browser application) produced a complete design of the geoportal that was readily documented and distributed amongst the project team and steering committee.

Development Tools
The UML diagramming completed while modeling the geoportal system became very useful. The diagrams were used in design review documents, administrator documentation, and as communication tools among the project team. Additionally, some diagramming software allows the user to develop a complete software system model that is linked to all UML diagrams. The system modeling software became an invaluable tool to capture engineering decisions and implement these changes across the entire model.

The geoportal system model components were programmed into PHP (PHP 2014), JavaScript and other software modules. A test-driven development was used, i.e. unit test cases were created for each software module that checked all possible execution scenarios using expected and error values for input arguments. These test cases were created prior to module programming wherever possible. Software tools were used to automate the testing of software modules against the test cases. In particular, PHPUnit (PHPUnit 2014) and QUnit (QUnit 2014) executed all PHP and JavaScript unit tests, respectively.

In concert with the unit testing tools, a revision control system and continuous integration (CI) server streamlined the testing and deployment processes as the source code for the geoportal system grew and evolved. After updating or creating source code on their local development computer, a developer commits these changes to a central repository of source code accessible by all developers. The CI server automatically detects this commit and operates the unit testing tools to execute all unit tests. Reports on the success and failure of the unit tests are automatically generated. A successful source code commit that passes all tests can be deployed to an operating version of the geoportal system, i.e. a web server and database running on a separate individual computer. The CI server automatically optimizes the source code and updates the web server and database. This leads to system development decisions being constantly integrated into a functional prototype that evolved into the final operational system.

Figure 7. The geoportal's three-layer architecture. Users interact with the geoportal through their web browser. The web browser and web server exchange data to support the front-end functionality. The web server exchanges data with a database to employ persistent and relationally structured data.

System Integration
The geoportal consists of a large integration of custom software modules and open source code libraries structured within a common three-layered architecture (Figure 7) consisting of a presentation layer, logic layer and data layer. The presentation layer is accessible through the user's web browser. All recent versions of major web browsers are supported. The user interface components are a combination of manually created controls and existing controls from the jQuery (jQuery 2014) JavaScript library. The OpenLayers JavaScript library (OpenLayers 2014) was used to render and support all interactions with the geoportal map. The Rickshaw JavaScript library (Rickshaw 2014) was used to create data analysis distribution and

time-series plots. The logic layer handles data requests and submissions from the user's web browser, performs calculations, and acts as a liaison for data flow between the user and the third data layer. The logic layer employs the Apache web server (Apache 2014) supporting scripting via PHP. The data layer provides a persistent storage facility using a relational database management system (RDBMS). PostgreSQL (PostgreSQL 2014) was chosen as the RDBMS for the geoportal with the PostGIS extension (PostGIS 2014) to extend the database to support geospatial data and calculations.

Conclusions

A web-based geoportal was created to support management of and interaction with a large ice, metocean and seabed dataset. Numerous design decisions supported the geoportal's key objectives of responsiveness, usability, security and functional capabilities for searching, visualizing, analyzing and downloading physical environmental study data and metadata. The final result was an intuitive application that executes in the user's web browser requiring no installation or configuration. The geoportal provides an effective balance of ease-of-use and powerful tools for data discovery, analysis and visualization. These tools include familiar functions common to geographical information systems (GIS) and science and engineering desktop applications.

Development of the geoportal will continue. New data analysis tools are expected to be added based on the user's needs. Although the geoportal already employs sophisticated data caching and compression schemes, further improvements in these areas will improve overall system performance by reducing the time necessary to move data between the various system layers. As the need arises, the database architecture will be expanded to include new datasets related to sea and glacial ice, metocean parameters and the seabed.

Acknowledgements

References

Apache (Version 2.4.10) [Software]. 2014. Retrieved from http://www.apache.org/.
jQuery (Version 1.11.2) [Software]. 2014. Retrieved from http://jquery.com/.
OpenLayers (Version 3.0) [Software]. 2014. Retrieved from http://openlayers.org/.
PHP (Version 5.6.3) [Software]. 2014. Retrieved from http://php.net.
PHPUnit (Version 4.4) [Software]. 2014. Retrieved from https://phpunit.de/.
PostGIS (Version 2.1.4) [Software]. 2014. Retrieved from http://postgis.net/.
PostgreSQL (Version 9.4) [Software]. 2014. Retrieved from http://www.postgresql.org/.
QUnit (Version 1.16.0) [Software]. 2014. Retrieved from http://qunitjs.com/.
Rickshaw (Version 3) [Software]. 2014. Retrieved from http://code.shutterstock.com/rickshaw/.
Rumbaugh, James, Ivar Jacobson, and Grady Booch. Unified Modeling Language Reference Manual, The. Pearson Higher Education, 2004.
Satzinger, John W., Robert B. Jackson, and Stephen D. Burd. 2005. *Object-Oriented Analysis and Design with the Unified Process*. Thomson Course Technology.