Application of the MINOE Regulatory Analysis Framework: Case Studies

Julia A. Ekstrom
University of California, Berkeley
2087 Addison Street
Berkeley, CA 94704-1103
jaekstrom@gmail.com

Gloria T. Lau, Kincho H. Law
Engineering Informatics Group
Stanford University
Stanford, CA 94305-4020
glau, law@stanford.edu

Matthew Hardy
Fisheries and Oceans Canada
343 University Ave.
Moncton, NB, E1C 9B6
matthew.hardy@dfo-mpo.gc.ca

ABSTRACT
In this paper, we describe a tool to help holistically understand, research and analyze the relationship between an ecosystem model and the relevant laws. Specifically, a software, MINOE, is being developed to address the needs to identify gaps, overlaps and linkages in the increasingly fragmented set of ocean-related laws. MINOE requires two pieces of information from the users, namely an ecosystem model, and a set of laws and its associated metadata, to perform the analysis. The output from MINOE is a searchable collection of laws organized by ecosystem relationships. Additionally, various visualization modules have been developed to help users synthesize the results for gap and overlap analyses. Two current usage examples are documented to illustrate the potential use of MINOE on legislation and management research.

Categories and Subject Descriptors
I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems – *law.*

General Terms
Management, Experimentation, Legal Aspects.

Keywords
Ocean management, ecosystem, co-occurrence analysis, visualization, law.

1. INTRODUCTION

1.1 Background
Food safety, transportation infrastructure, water supply, coastal ecosystem health are all examples of valued resources that are composed of inter-related sources and sectors. These resources, like many others that are valued in our society, can be viewed as systems to better understand the complexity inherent in the resources. Scientists have long been doing just that – developing conceptual or quantitative models that represent the system's components and associated linkages – to understand the behavior of the system [19]. For example, ecologists document how species interact with one another and to what degree they affect and depend on certain habitats [24]. When these resources are threatened (such as a fishery declines, toxins are found on spinach, or there is insufficient groundwater for remote towns), developing solutions to such problems calls for an analysis of existing governance practices. This paper presents a tool developed to assess governance using a systems perspective.

Analysis of governance practices, especially in developed countries, is commonly conducted through an analysis of laws, regulations, and/or other documents representing policies, programs or other efforts by government agencies. Such legal and management analyses are conducted relying on digital information, such as the laws provided by the local, state, or federal government offices, generally accessible via a keyword-based online informational retrieval system. However, in our research of legal information retrieval systems, we have found a need for providing a systems perspective within the analysis of governance [13]. A comprehensive suite of digital government tools should provide a systems perspective of laws, regulations and other management documents so as to help reveal where there are potential gaps in management, overlapping jurisdictions, and opportunities for coordination. In this paper, the focus is on ocean governance; however, the tool presented can potentially be applied to analyze governance of any issue of concern from a systems perspective.

1.2 Systems Perspective on Ocean Governance
Marine and coastal ecosystem health is perceived to be in decline worldwide. The fragmented nature of governance systems with which oceans and coasts are managed is a major contributor to the decline in ocean health [8, 18]. Governments have traditionally managed ocean and coastal issues and resources within sectors and on a case-by-case basis [8, 31]. Without adequate coordination, the sector-based management has created a system of governance in which linkages between activities and natural resources – including those manifesting as adverse environmental effects – are not deemed to be adequately managed for by existing rules, rights, and decision-making procedures [4, 23].

As a result of this fragmentation, there is a growing need to transition out of the sector-based approach into an ecosystem-based management (EBM) system [17, 26, 28]. A major roadblock to the implementation of EBM is that it requires coordination and communication among sectors within and between levels of government [5, 15, 16]. Therefore, the identification of management gaps and overlaps is a necessary first step toward achieving the transition to EBM.

Dg.o'11, June 12–15, 2011, College Park, MD, USA.

MINOE, which stands for Management Identification for the Needs of Ocean Ecosystems, is a software tool developed to facilitate the linkage between a given ecosystem model and the laws issued by different government agencies. Specifically, MINOE allows users to build an ecosystem model using a matrix representation, from which it automatically attributes linkages between the ecosystem elements and regulations, statutes, management plans and other documents that users might be interested in. Agency authority data can also be recorded as metadata per document, which allows for an agency overlap analysis to be performed by the software. Since both ecosystem models and legal data can be large, several visualization modules are developed to help users synthesize the relationships between large data, such as the gaps, overlaps and linkages in the law in relation to the ecosystem model.

This paper is organized as follows: Section 2 briefly reviews relevant work on the development of EBM tools. Section 3 discusses the inputs to MINOE, namely the ecosystem model and the document corpus, which is often comprised of regulations, statutes and management plans. Output from MINOE is described in Section 4, where the visualization modules are discussed. Two current use cases of MINOE are presented in Section 5, and future extensions are discussed in Section 6.

2. RELATED WORK

The increasing interest and need for applying the ecosystem concept in management and decision-making have generated a growing number of scenario building tools and research programs focused on valuing ecosystem services [20]. Most of these tools thus far focus on the natural science of ecosystems and some are beginning to integrate scientific data with economic information. MINOE could be used in conjunction with these ecological and economic EBM tools to provide baseline law and regulation data in relation to the ecological and economic components.

For example, Ecopath with Ecosim (EwE), built by a team of scientists and software engineers at University of British Columbia, is an ecosystem modeling software used to assist in identifying optimal management strategies for fisheries management [3, 25]. EwE creates quantitative ecosystem models of direct and indirect linkages between species, habitats, and other ecosystem elements. Using a suite of management scenarios, one module of the EwE can help scientists and managers identify what aspects of the economic supply chain will be affected by various ecosystem changes. MINOE could be used in conjunction with EwE to provide users with useful information about the existing legal framework in which the use and protection of ecosystem elements are being regulated, whether the ecosystem linkages are accounted for in any regulations, and what agency or suite of agencies potentially should or could be involved in implementing the scenario [12].

3. USER INPUT TO MINOE

MINOE requires two main pieces of information to conduct an analysis: 1) a model of a system of interest or concern, e.g., ecosystem, socio-ecological model, or any other system of elements linked in some way to one another; and 2) a collection of text documents. Essentially MINOE helps to highlight where links and component in the ecosystem model are discussed in the given set of documents.

3.1 Ecosystem Model

The ecosystem model is represented in MINOE by a symmetrical matrix. Column and row headers are made up of a term or multiple terms that represent each element of the ecosystem. In the domain of ocean management, examples of elements may include “fishing”, “watershed” and “ballast”. Cells contain a “1” if the corresponding elements are linked in the model and a “0” if they are not linked. The definition of a link is determined by the user based on his/her needs and interests. For example, “fishing” might be connected to “watershed” but not “ballast”. Section 3.1.1 describes the details of the ecosystem elements, while Section 3.1.2 outlines the process to build linkages between elements.

3.1.1 Ecosystem Elements

The ecosystem elements are the key concepts that users define to represent the ecosystem model that they are interested in. In our previous work, 46 issues related to ocean and coastal management were manually selected to represent key ocean topics [10]. Examples include transportation, fishing, port(s), and discharge. A basic frequency count of the term or the phrase is implemented to model the concept importance in various statutes and regulations. For instance, the term “transportation” has 43 mentions in 55 units of Federal statutes in our repository, where a unit is defined as a chapter in the U.S. Code (USC).

Based on beta testers’ feedbacks, a single term or phrase alone is generally too limiting in defining an ecosystem element. Domain experts can usually define a list of related terms to a concept, as shown in Figure 1, where term set A represents the relevant concepts for the user and term set B represents the related terms [9]. As a result of user feedback, we have since incorporated synonyms defined by users. As illustrated in Figure 2, users can defined a list of synonyms for a single concept, e.g., “construction”, “community”, “development”, and “waterfront land”, via MINOE’s user interface.

The synonym functionality allows the users much control over the vocabulary used in the ecosystem model; however, the drawback is that it puts a heavy load on the users as well. Domain experts need to be consulted to craft a meaningful set of synonyms, which is not reusable for other EBM tools if the knowledge is not properly captured and stored. On the other hand, topic taxonomies are widely available for various domains, and many taxonomies capture the collaborative effort among working groups of experts in the field. Because of the recent semantic web movement, an increasing number of taxonomies are developed and stored in the Resource Description Framework (RDF) format, which facilitates easy importing and exporting of data.

As such, we have experimented with including taxonomic data into MINOE [11]. Ekstrom, in consultation with domain experts, constructed a domain specific terminological taxonomy using the California Aquatic Invasive Species Management Plan [2]. This document contains an extensive description of the individual pathways of aquatic invasive species in the State of California and a full species list (with vernacular and scientific names). In Table 1 below, the ecosystem element “recreational boating” is defined by two levels of taxonomy nodes. L1 represents a higher level and a more abstract definition of the concept, whereas L2 (e.g., water skiing, sailboat, and etc) represents the specific terms which are generally found in more detailed management documents.

	GROUP	COMPONENT	TERM SET A	TERM SET B (additional terms to represent concept)
IMPACTS	Dependant Activities	Urban residential	urban residential	construction community development "waterfront land" housing streets municipal services neighbourhood backyard building setback septic sewage
		Tourism recreation	tourism recreation*	marina wharf pleasure beach travel fishing sports vacation leisure cottage boating sailing camping accomodation park
		Agriculture	agriculture	farm* cattle food crop manure fertiliser farmer farming farmland harvest
		Forestry	forestry	logging logger lumber forest woodland "tree harvest" sylviculture
		Fisheries aquaculture	fisher* aquaculture	seafood "fish farming" "fish farm" mariculture fisherman
		Energy mining manufacturing industries	mining energy	manufactur* industry industries industrial refining resources mineral hydrocarbon "strip mine" tailings "oil and gas" well exploration extraction
		Land transportation	transportation	truck* highway street bus train railway bridge road culvert paved unpaved
		Marine transportation	shipping "marine transportation"	port shipping transoceanic trade transatlantic cargo ferry "ocean liner" wharf
	Services	Water quality	"water quality"	"clean water" "water pollution" "water pollutant" "water contamination" sewage colliforms closure
		Water source	"water source"	spring headwater* reservoir "ground water" well aquifer
		Aesthetics	aesthestics	beauty view vista tourism values
		Coastal access	"coastal access"	"beach access" "public access" right-of-way tourism riparian recreation
		Fish resources	fish	seafood protein fishery aquaculture subsistence harvesting
		Navigation	navigation	shipping passage channel transportation sealane obstruction

Figure 1: Expanded list of terms to represent concepts in the ecosystem model

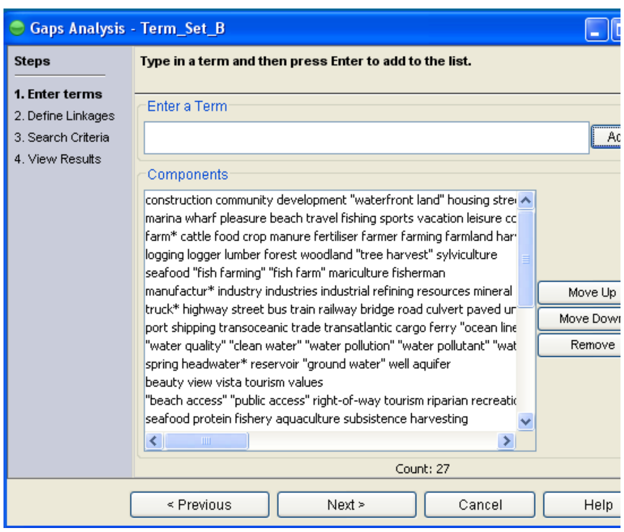


Figure 2: GUI to define synonyms to represent concepts

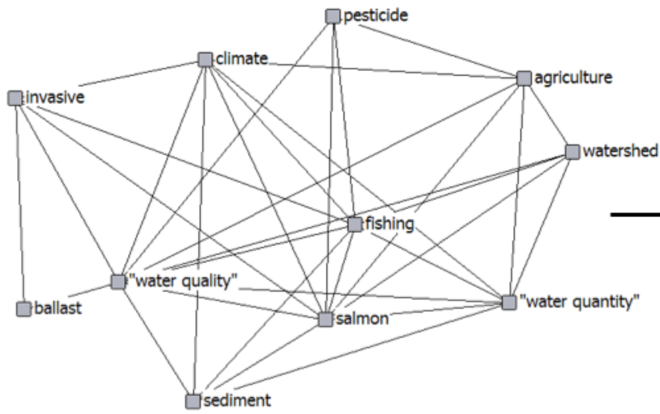
The technique for the evaluation and the benchmarking test with taxonomic data has been described in details elsewhere [11]. The result of comparing the usage of L1 and L2 terms in identifying agency involvement in environmental regulations and statutes was inconclusive. For example, "recreational equipment" without the use of the terminological taxonomy did not occur in any federal law or regulation. But the use of the lower level terms did reveal government agencies involved. In other situations, as with "recreational boating", the rank of agency involvement remained consistent with and without the taxonomy. In general, while related taxonomic terms could help in certain situations, there was not notable improvement in results from the use of the taxonomy.

Although there is not sufficient conclusive result from our preliminary experiment with taxonomic data, we believe that the

Table 1: Excerpt of a taxonomy used to represent ecosystem elements

Concept	Taxonomic Level	Term(s)
Recreational boating	L1	recreational boating
	L2	water skiing
	L2	sailing
	L2	sailboat
	L2	hull fouling
	L2	anchor
	L2	boat trailer
Recreational equipment	L1	recreational equipment
	L2	Asian overbite clam, Corbula amurensis
	L2	water craft
	L2	scuba
	L2	snorkel*
	L2	jet ski

ability to capture, import and export taxonomies from MINOE is important. MINOE is envisioned to be a component in a long suite of EBM tools that policy makers and other interested users might need in order to do their research, and thus the capability to interoperate with other tools is crucial. With taxonomies defined and stored as RDF documents, users can reuse their hierarchical representation and synonyms of ecosystem concepts. Therefore, we plan to provide a plugin and an Application Programming Interface (API) to import and export taxonomies in and out of



	salmon	fishing	sediment	watershed	water quality	water quantity	pesticide	agriculture	ballast	invasive	climate
salmon		1	1	1	1	1	1	1	0	1	1
fishing			1	1	1	1	1	0	0	1	1
sediment				0	1	1	0	0	0	0	1
watershed					1	1	0	1	0	0	0
"water quality"						1	1	1	1	1	1
"water quantity"							0	1	0	0	1
pesticide								1	0	0	0
agriculture									0	0	1
ballast										1	0
invasive											1
climate											

Figure 3: An ecosystem model represented graphically and in a matrix format

MINOE so that we can further examine the effects of including related taxonomic terms in the analysis.

3.1.2 Ecosystem Linkages

Matrices are a common way to organize the complexity of stressors, activities, and other elements and their associated linkages to inform scientific research and management [21, 30]. As an example, state and federal agency personnel, NGOs, and university scientists conducted a series of workshops to organize and prioritize impacts of wave energy parks off Oregon's coast [1]. The reports from the workshops published a series of matrices to organize impacts of various stressors and other linkages that need to be considered for implementing and monitoring such wave parks. These matrices, representing conceptually modeled system of interest, can then be used in MINOE to evaluate the laws, other management documents, and government agencies that deal with the issues defined in the matrices.

As such, a Boolean matrix of ones and zeroes are used to represent the presence and absence of linkages between ecosystem elements. An example is shown in Figure 3, where a 1 in "fishing" and "watershed" represents an ecological linkage between the two elements, while a 0 in "fishing" and "ballast" represents the lack of such a relationship between the two elements. Currently, the strength of linkages between concepts is not captured in MINOE. The strength of the ecological relationship can be represented with a non-Boolean matrix; however, further study will be performed to better understand the benefits and the impacts of a non-Boolean representation.

3.2 Documents and Metadata

MINOE indexes a corpus of text documents in order to link the ecosystem elements to the texts and their metadata. Users can import their documents and metadata into MINOE, or they can start with a prebuilt corpus. We have collected, cleaned, parsed and meta-tagged a list of regulations and statutes in several jurisdictions. As illustrated in Figure 4, the prebuilt corpus contains a large number of Federal and State laws for users to choose from.

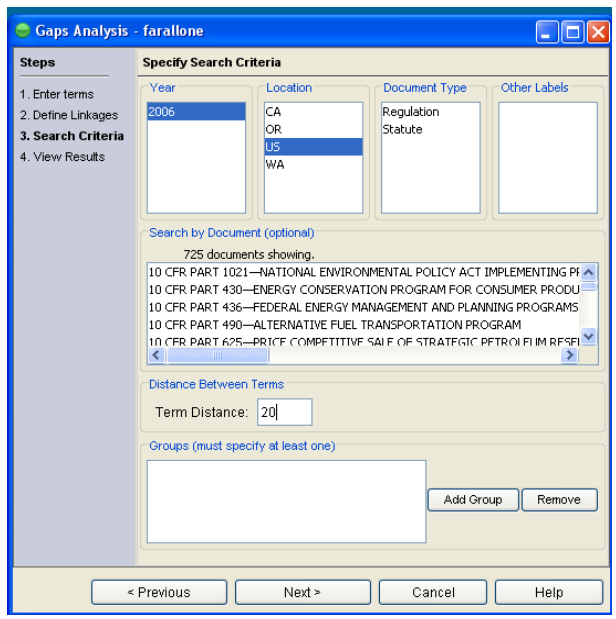


Figure 4: Precompiled corpus of Federal and State regulations and statues in MINOE

Meta-tagging here refers to the agency authority for each regulation or statute, which is a crucial piece of information for MINOE to perform agency overlap analysis. For instance, the Environmental Protection Agency (EPA), the Army Corps of Engineers (ACE), and etc., have authority to implement the Clean Water Act (33 USC 1251 et seq.). As such, the Clean Water Act is tagged with agencies such as the EPA, the ACE, and etc. Tagging is currently a manual process by compiling the agency authority list from the National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center in addition to skimming laws. The Federal Register publishes an agency metadata field which specifies "the agency or sub-agency responsible for the rule, proposed rule, or notice." MINOE can potentially benefit from interfacing with these externally available metadata so that agency authority tagging can be automated. As such, further research is necessary to determine the availability, consistency and coverage of such data.

Metadata is not limited to agency authority only. For instance, if the user is interested in analyzing the financial impacts of legislating certain ecological elements, he or she can tag the cost per legislation and MINOE will build the linkage between the ecosystem concept and the financial impact via the tagged documents. MINOE currently reads the metadata from a Boolean matrix that captures the agency-document relationship. This data structure is designed to readily accept various metadata so that users can define the end goals of their analysis and not be restricted to agency authority only.

4. OUTPUT FROM MINOE

4.1 Searchable Collection of Law

MINOE indexes the collection of law and any other documents that the user provides. The most basic functionality of MINOE is capturing the occurrence of ecosystem elements, defined by either a single term or a list of synonyms, in various sections of the law and its metadata. As suggested in Section 3.1.1, a taxonomy can be used to define terms, related terms and synonyms to represent a concept. MINOE keeps track of the frequency of all terms, and computes the concept frequency by summing up all terms used to define a concept.

Figure 5 shows an example of the searchable collection of law, where users can search for keywords such as "ballast", and matching keywords will be highlighted in the corresponding sections of the law. Since the metadata, namely the agency authority information, is also indexed in MINOE, users can also visualize the correlation between the metadata and keywords as shown in Figure 6. This is computed by counting the total

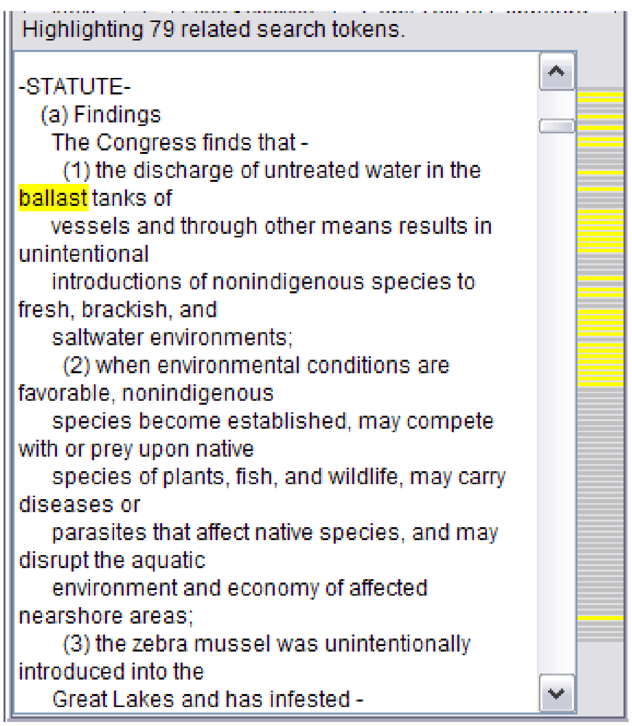


Figure 5: Searchable Collection of Law with matching keywords highlighted

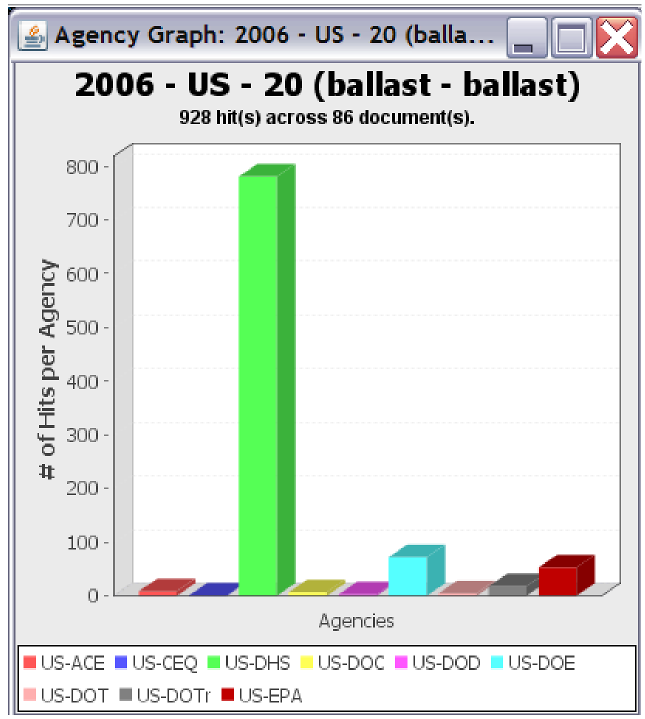


Figure 6: Aggregated count of agency involvement for the ecosystem element "ballast"

occurrences that each concept appears in laws tagged with each agency. For instance, as shown in Figure 6, the occurrence frequency of the ecosystem element "ballast" is charted against each agency, such as the Department of Homeland Security (DHS). According to the chart, the term "ballast" occurs most frequently in those laws that are governed by the DHS.

Baseline information such as the bar chart in Figure 6 allows users to explore what agencies need to collaborate and in what capacity, where this type of basic data are previously hard to come by and would need to be manually compiled. To maximize reuse of data, we plan to provide an API in the future to export information from MINOE so that other EBM tools can perform rollup and aggregate reporting to further the analysis.

4.2 Visualization of Gaps and Linkages in the Ecosystem Model

A searchable collection of the law and its metadata is essentially a text-based user interface to MINOE's underlying data and analysis. When the dataset gets large, which is often the case with regulations, statutes and management documents, a graphical user interface is a lot easier for users to visualize the output and to explore and synthesize the data.

One of the most interesting features in MINOE is the visualization module that allows users to explore concept co-occurrences in their document collection. In the context of ocean law, the co-occurrences and the lack of co-occurrences of ecological elements defined in the ecosystem model can be visualized through a connected graph. The co-occurrences are shown as linkages, and the lack of co-occurrences, or the gaps, are shown as broken linkages.

We have built a small ecosystem model consisting of 11 elements to demonstrate the visualization of gaps and linkages; please see Figure 3 for the matrix representation of the ecosystem model. As shown in Figure 7, the ecosystem model, which was originally inputted by the user using a matrix format as discussed in Section 3.1, is now displayed as a connected graph. Terms representing ecosystem elements are sized relative to the frequency of occurrence in the document collection. For instance, the term "fishing" is the largest in size because this term occurs more times in this collection of U.S. laws from 2006 (4,201 occurrences) than any of the other terms used to construct the ecosystem model. Dotted lines indicate gaps (those ecosystem modeled links that do not co-occur within the law), and solid gray lines indicate those modeled links for which ecosystem elements do co-occur closely together in one or more law. Thickness of line varies with the frequency of elements' co-occurrence. For example, a thick line connects "salmon" and "watershed" because there are 200 co-occurrences of these elements in the U.S. federal laws, while a relatively thin line connects "invasive" and "ballast" because there are only four co-occurrences of the two ecosystem elements. In this ecosystem model, "salmon" and "climate" should be connected but do not co-occur in any law in this collection, and therefore a dotted line is drawn to represent the gap.

We see tremendous potentials in assisting users to visualize gaps and linkages in the law. With the knowledge that a strong link between certain ecosystem elements exists, users can ask for specific sections of law where the concepts co-occur and further their analysis. We believe that the ability to visualize the linkages prior to further analysis in the law is a significant time saver for users. Likewise, the ability to visualize the gaps is also important. For instance, isolated islands of ecosystem elements could be observed from the graph, where appropriate links to the elements are missing. The element "climate" is a good example; it is found in the law (69 times in this compilation) but is surrounded by dotted line gaps from several other elements in the ecosystem model. This means that the user has defined linkages between "climate" and other ecosystem elements, but MINOE did not find co-occurrences between "climate" and these elements in the law.

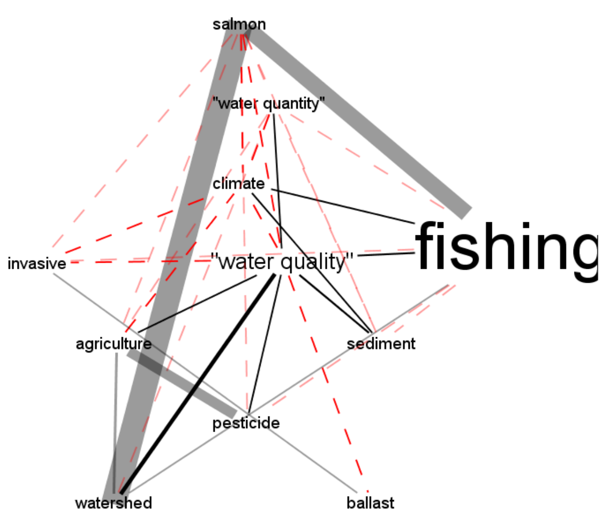


Figure 7: Graphic of US federal laws and regulations representing the ecosystem model in Figure 3

This holistic observation helps users refocus their research on why such an isolated island is formed in the law, rather than spending time on pairwise analysis to verify the presence or the absence of the linkages.

4.3 Visualization of Overlapping Laws and Agencies

The second visualization module – currently under development – in MINOE is the overlap analysis module. As shown in Figure 8, this visualization helps users to identify overlapping agencies on different ecosystem elements. Each agency is shown as connected to the laws where it has authority to implement. The bigger nodes represent more frequent occurrences of the concept in that particular section of the law. For instance, the ecosystem element "ballast" appears to be fragmented across many agencies, where the Department of Homeland Security (DHS) and the Environmental Protection Agency (EPA) have the largest number of laws. For more details of an agency overlap analysis performed using MINOE, please see [10].

5. CURRENT USAGE

MINOE was designed and developed iteratively with the help of many beta users. In this section, two current use cases are described to illustrate the reach of MINOE.

5.1 California Aquatic Invasive Species

MINOE can be used to assist regulatory and legislative analysis. For example, one of the main goals set by the California Aquatic Invasive Species Management Plan was for the State to conduct an analysis of existing management, identifying what laws and regulations the State already has that pertain to each specific pathway and invasive species [2]. Additionally, one of the plan's primary tasks is to identify which agencies are and should be involved in management of invasive species. Given the complexity and long list of pathways through which non-native species are introduced into the state waters [2, 22, 27], this can be a time consuming project. Tasked with documenting all California State legislation relevant to aquatic invasive species, a

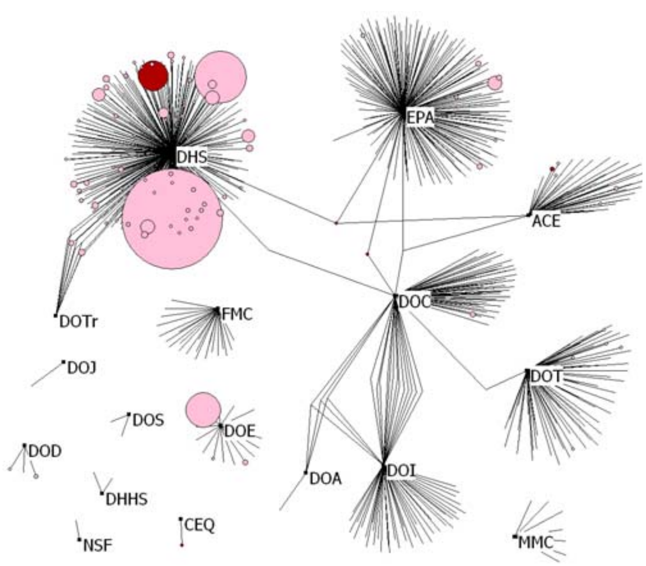


Figure 8: Agency overlap visualization of the concept "ballast"

California State Research Fellow used the beta version of MINOE, in combination with more traditional law databases, to carry out this project.

The user found the MINOE output to be useful in that it provided the same matrix structure so she could investigate and compare the usage of the different invasive species and vector pathways simultaneously, but still independently of one another. This, along with the graphics module (example of output shown in Figure 7), and the clickable agency involvement bar graphs (example of output shown in Figure 6), provide a clear benefit over the traditional database systems she queried, which provided output in the form of simple lists. The other benefit reported was that the user could view multiple jurisdictions through a single query (three states' and federal law), whereas otherwise she had to go to several different websites to investigate her target jurisdictions. The main reported drawback of MINOE relative to the more widely used legal databases was that the compilation of laws and regulations provided through MINOE, which at the time were current as of the year 2006, but her analysis was conducted in 2008. An effort is needed to develop a practical strategy to have an updated compilation (while maintaining a timestamped version of prior legal database).

5.2 Southern Gulf of St. Lawrence, Canada

The Federal government of Canada is transitioning towards an ecosystem-based management approach for the coastal and marine regions through the Oceans Act (1996). Efforts to plan for and carry out activities relating to integrated ecosystem-related goals became a priority with the Canada Oceans Action Plan in 2005 [29]. A pilot project was initiated to assess the use of MINOE in identifying potential gaps in Canada's context of coastal and oceans in relation to adverse environmental effects.

The MINOE software is used here to assist in conducting an analysis to identify potential gaps in a set of Provincial and Federal laws, regulations and best management practices for the jurisdictions adjacent to the southern Gulf of St. Lawrence [9]. Their document collection consists of laws, regulations, and best management practices to represent management from Federal level and three provinces (Nova Scotia, New Brunswick, and Prince Edward Island). The Department of Fisheries and Oceans (DFO) has also developed a conceptual model on the basis of the DPSIR (Driver-Pressure-State-Impact-Response) structure to link activities to cumulative and adverse environmental effects for coastal and marine waters and the associated watersheds [23]. This project served as a pilot, which is currently under evaluation by DFO and partners to determine the utility of extending MINOE's application in the context of a collaborative process for identifying areas where management measures may not be present, enforced or sufficiently effective in mitigating cumulative effects to significant fish and fish habitat.

6. FUTURE WORK

As discussed in previous sections, we envision MINOE to collaborate and interoperate with other EBM tools so that data reuse can be maximized to save users' time. Therefore, we plan to standardize our data structure, add in support for importing and exporting of data by users, and provide an API for programmatic access to the data and to call MINOE as a service by other EBM tools. The development of these features for increased interoperability will be an iterative process where beta testers' feedback can help tremendously.

One of the key issues with MINOE is word sense disambiguation. For instance, we have noticed that searching through a law database using the term "marine" would produce results in two distinct groups – "marine" laws related to ocean and coastal management, and "marine" laws related to the United States Marines (i.e., insurance or retirement regulations not related to ocean law). Another example of an ambiguous concept is the term "well" as shown in Figure 1. The term "well" is used to represent the concept of "water source", but this term can be used in many different ways, such as an adverb, which does not relate at all to the concept of "water source". These two classes of issues together introduce a lot of noise in the data and in the model, where visualization results could be inaccurate.

We plan to employ two different approaches to resolve the two sources of ambiguities identified above. The "marine" example, where the same term is used in different contexts, will require the software to understand contexts. We propose to combine a traditional dictionary-based approach in addition to a training-based approach for this task. For both approaches, our goal is to gather a list of terms that frequently co-occur with the concept for each sense of that particular concept.

WordNet [14] is a popular dictionary used in the field of computational linguistics and natural language processing. It provides an inventory list of different word senses for each word. In our "marine" example, we will be able to associate n-grams of terms that are used in different senses of "marine", such as "soldier … Marine Corps" vs "inhibiting the sea … marine plants" both found in definitions of different senses of "marine" in WordNet. Essentially, we now have the co-occurrence data for the term marine's usage in different contexts. Comparing this identified pattern of co-occurrences with those found in the law will help MINOE to disambiguate between word senses.

Apart from WordNet, there are other sources of descriptions or definitions for each word sense, such as the Wikipedia [32]. For instance, in Wikipedia, there are definitions for marine (military) and marine (ocean), each with a separate page of related content. There is an API to access Wikipedia data, and the definitions and content can be used to learn common terms associated with each word sense similar to the dictionary-based approach. However, we believe that the language use in the law is quite different from that of Wikipedia, and as a result, we plan to start with a dictionary-based approach using WordNet before we reach out to the open web.

Since our goal is to extract co-occurrence patterns from the dictionary to disambiguate concepts, we should not limit ourselves to the dictionary only where its definitions are possibly short and incomplete. A training approach would involve a set of tagged data that define the different senses of key concepts that are commonly seen in ocean laws. For instance, as our domain expert has manually filtered out U.S. marine laws from ocean marine laws, the two piles of laws can be tagged and a classifier can be developed to learn the terms that commonly co-occur with each sense of "marine". Laws from other jurisdictions can be classified using this model, assuming that the language use is similar to that of the training set. Various classifiers can be trained for the task, such as k-nearest neighbor (KNN) [6], support vector machine (SVM) [7] and so on.

The second type of ambiguity, as illustrated in the "well" example, has a simpler solution. We propose to tag part-of-speech by using the Stanford parser [17]. The Stanford parser is a program that works out the grammatical structure of sentences such as noun vs. adverb, which is well suited for this task. Most concepts should be nouns and noun phrases. So for example, if a section of the law contains the term "well" tagged as an adverb, that section should not be linked to the concept "well" as a "water source".

7. CONCLUSIONS

Searching for concepts in regulations, statutes and management documents is a time consuming task, as one might have to run a number of related queries against multiple databases. Searching for agency authority information for multiple concepts is even more difficult, since one has to record agency authority data per section of the law, and correlate that to concepts where they appear in the law. Research on overlaps, gaps and linkages between concepts in the law is the most challenging part, as one would need to run Boolean queries, such as the "near" operator, to find co-occurrence frequencies between concepts. MINOE is a comprehensive EBM tool designed to simplify the above tasks.

The input to MINOE involve two pieces of information: 1) an ecosystem model that specifies the ecosystem elements which can be defined as a list of synonymic keywords, and their linkages represented in a Boolean matrix format; 2) a corpus of documents with optional metadata, which represents the agency authority information in our examples. The second input is optional, since MINOE comes preinstalled with a set of regulations and statutes from various jurisdictions, so users can start analyzing their ecosystem model with respect to this set of laws provided in MINOE. The metadata, namely the agency authority table, allows users to perform analysis such as agency overlaps for each ecosystem element. If appropriate, other metadata can be inputted and analyzed as well, e.g., budget could be a relevant metadata for a financial impact analysis on various legislations.

The output from MINOE is a searchable collection of law, where it indexes the corpus of documents and provides a search interface to it. Agency authority data can also be aggregated per ecosystem element, where one can see which agency has the largest number of matching laws for that particular element. In addition, two visualization modules provide the capability to visualize gaps, overlaps and linkages in the law in relation to the ecosystem defined by the user. For each ecosystem concept, potential overlapping authority between agencies is depicted as two or more laws under the authority of the same or different agencies that reference the same concept. The ecosystem model as a whole can also be visualized as a connected graph of the ecosystem elements of various sizes, in proportion to their occurrence frequencies in the law. Here, the presence and absence of linkages between ecosystem elements can be visualized by users easily.

The architecture of MINOE allows for flexibility in importing and exporting data so that other EBM tools can reuse its analysis rather than recreate it. MINOE can serve as an extension to ecosystem modeling programs, as previously mentioned in Section 2, as complementary to scenario-building tools. Output from such tool can be fed as input to MINOE creating an ecosystem model automatically, alleviating the research load from users. Future work is planned to develop new import/export modules and APIs to MINOE so that it can interact with other EBM tools smoothly. In addition, one of the key challenges in MINOE is word sense disambiguation, where we propose to use a combination of dictionary-based and training-based approaches to resolve ambiguities.

8. ACKNOWLEDGMENTS

Our thanks to the David and Lucile Packard Foundation (Ecosystem-Based Management Tools Initiative Fund) for supporting this research. The authors would like to acknowledge the effort by Mr. Daniel J. Spiteri on implementing MINOE. We are grateful to those who tested MINOE in its beta version and provided valuable feedback at various stages of the project. The authors also acknowledge partial support by the National Science Foundation grant IIS-0811460.

9. REFERENCES

[1] Boehlert, G. W., McMurray, G. R., and Tortorici, C. E. 2007. Ecological effects of wave energy development in the Pacific Northwest. *U.S. Department of Commerce, NOAA Technical Memorandum* NMFS-F/SPO-92 (October 11-12, 2007).

[2] California Department of Fish and Game. 2008. *California Aquatic Invasive Species Management Plan.* DOI= http://www.dfg.ca.gov/invasives/plan/.

[3] Christensen, V. and Walters, C. 2004. Ecopath with Ecosim: methods, capabilities and limitations. *Ecological Modeling.* 172:109-139.

[4] Cormier, R., Kannen, A., Davis, I., Sardá, R., and Diedrich, A. 2010. Policy Fragmentation Implications in Ecosystem-Based Management in Practice. *Proceedings of the ICES Annual Science Conference,* Nantes, France, September 20-24, 2010.

[5] Cortner, H. J., Wallace, M. G., Burke, S., and Moote, M. A. 1998. Institutions matter: the need to address the institutional challenges of ecosystem management. *Landscape And Urban Planning,* 159-166.

[6] Cover, T. and Hart, P. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21-27.

[7] Cristianini, N. and Shawe-Taylor, J. 2000. *An introduction to support vector machines and other kernel-based learning methods.* Cambridge University Press.

[8] Crowder, L., Osherenko, G., Young, O., Airame, S., Norse, E. A., Baron, N., Day, J. C., Douvere, F., Ehler, C. N., Halpern, B. S., Langdon, S. J., McLeod, K. L., Ogden, J. C., Peach, R. E., Rosenberg, A. A., and Wilson, J. A. 2006. Resolving mismatches in U.S. ocean governance. *Science,* 313:617–618.

[9] Ekstrom, J. A. and Hardy, M. 2010. Functional Fit and Gap Analysis between Regulatory Provisions and a Coastal Ecosystem Model. *Gulf Region Oceans Management Series.* 2010/02.

[10] Ekstrom, J. A., Lau, G. T. 2008. Exploratory Text Mining of Ocean Law to Measure Overlapping Jurisdictions. *Proceedings of the 9th Annual International Conference on Digital Government Research (dg.o2008),* Montreal, Canada, May 18-21, 2008.

[11] Ekstrom, J. A., Lau, G. T., Cheng, J. C. P., Spiteri, D. J., and Law, K. H. 2010. Gauging Agency Involvement in Environmental Management Using Text Analysis of Laws and Regulations. *I/S: A Journal of Law and Policy for the Information Society*, 6(2).

[12] Ekstrom, J. A., Lau, G. T., Spiteri, D. J., Cheng, J. C. P., and Law, K. H. 2010. MINOE: A Software Tool to Evaluate Ocean Management in the Context of Ecosystems. *Coastal Management*, 38(5), 457-473.

[13] Ekstrom, J. A., and O. R. Young. 2009. Evaluating functional fit between a set of institutions and an ecosystem.*Ecology and Society* **14**(2): 16. [online] URL: http://www.ecologyandsociety.org/vol14/iss2/art16/

[14] Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database.* Bradford Books.

[15] Joint Ocean Commission Initiative. 2009. *One Coast, One Future: Securing the Health of West Coast Ecosystems and Economies.* Washington, D.C. DOI= http://www.jointoceancommission.com/resource-center/1-Reports/2009-01-15_One_Coast_One_Future.pdf.

[16] Juda, L. and Hennessey, T. 2001. Governance profiles and the management of the uses of large marine ecosystems. *Ocean Development and International Law*, 32, no. 1: 43-69.

[17] Klein, D. and Manning, C. D. 2003. Accurate Unlexicalized Parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 423-430. DOI= http://nlp.stanford.edu/~manning/papers/unlexicalized-parsing.pdf.

[18] McLeod, K. L., Lubchenco, J., Palumbi, S. R., and Rosenberg, A. A. 2005. Scientific Consensus Statement on Marine Ecosystem-Based Management. *Communication Partnership for Science and the Sea*.

[19] Meadows, D.H. (edited by D. Wright) 2008. Thinking in systems: a primer. Chelsea Green Publishing: White River Junction, VT.

[20] NatureServe. 2009. *Ecosystem-Based Management Tools Network*. DOI= http://www.ebmtools.org.

[21] Newton, J., Mumford, T., Dohrmann, J., West, J., Llanso, R., Berry, H., and Redman, S. 2000. A Conceptual Model for Environmental Monitoring of a Marine System. *Puget Sound Ambient Monitoring Program (PSAMP)*, ed. PSAMP, 40: PSAMP.

[22] Office of Technology Assessment (OTA). 1993. *Harmful Non-Indigenous Species in the United States.* United States Congress, Washington, DC.

[23] Ouellette, M. and Hardy, M. 2010. Integrated Coastal Zone Management: Bridging the Land-Water Divide. *Proceedings of the ICES Annual Science Conference*, Nantes, France, September 20-24, 2010.

[24] Paine, R.T. 1980. Food webs: linkages, interaction strength, and community infrastructure. Journal of Animal Ecology 49: 667-685.

[25] Pauly, D., Christensen, V., and Walters, C. 2000. Ecopath, Ecosim, adn Ecospace as tools for evaluating ecosystem impact of fisheries. *ICES Journal of Marine Science*, 57(3): 697-706.

[26] Rosenberg, A. A. and McLeod, K. 2005. Implementing ecosystem-based approaches to management for the conservation of ecosystem services. *Marine Ecology Progress Series*, 300: 270-274.

[27] Schmitz, D. and Simberloff, D. 2001. Needed: A national center for biological invasions. *Issues in Science and Technology*, Summer: 57-62.

[28] Sherman, K., Sissenwine, M., Christensen, V., Duda, A., EHempel, G., Ibe, C., Levin, S., Lluch-Belda, D., Matishov, G., McGlade, J., O'Toole, M., Seitzinger, S., Serra, R., Skjoldal, H. R., Tang, Q., Thulin, J., Vandeweerd, V., and Zwanenburg, K. 2005. A global movement toward an ecosystem approach to management of marine resources. *Marine Ecology Progress Series*, 300: 275-279.

[29] Siron, R., Sherman, K., Skjoldal, H. R., and Hiltz, E. 2008. Ecosystem-Based Management in the Artic Ocean: A Multi-Level Spatial Approach. *Artic* vol 61, Suppl 1: 86-102. DOI= http://pubs.aina.ucalgary.ca/arctic/Arctic61-S-86.pdf.

[30] Thom, R. M., Williams, G. D., and Borde, A. B. 2003. Conceptual Models as a Tool for Assisting, Restoring, and Managing Puget Sound Habitats and Resources. Paper read at PSP.

[31] United States Commission on Ocean Policy (USCOP). 2004. *An ocean blueprint for the 21st century final report of the U.S. Commission on Ocean Policy*. Washington, D.C., USA. DOI= http://www.oceancommission.gov.

[32] Wikipedia: The Free Encyclopedia. 2001. Wikimedia Foundation Inc. DOI= http://en.wikipedia.org/.