International Journal of the Commons
Vol. 9, no 2 September 2015, pp. 595–618
Publisher: Uopen Journals
URL:http://www.thecommonsjournal.org
URN:NBN:NL:UI:10-1-117247

ISSN: 1875-0281

Building a diagnostic ontology of social-ecological systems

Ulrich J. Frey
Martin-Luther-University of Halle-Wittenberg, Germany
ulrich.frey@landw.uni-halle.de

Michael Cox
Dartmouth College, USA
michael.e.cox@dartmouth.edu

Abstract: This article argues that research on social-ecological systems could profit from the use of ontologies. Ontologies, i.e. formalized conceptualisations of a domain in a computer-readable format, allow making progress in different areas. In particular, a diagnostic approach would be facilitated, in turn addressing the complexity problem (analysing the complexity of social-ecological systems adequately), the panacea problem (the overreliance on simplistic policy prescriptions that do not account for this complexity) and the scatter problem (lack of integration of many research findings into a cohesive set of theoretical statements). Ontologies offer several advantages, e.g. they structure and formalize domains, unify knowledge, decrease terminological confusion, reduce incomparability, reduce redundancy of efforts and allow automated reasoning. We demonstrate the practical use of ontologies by converting the SES framework into an ontology. This leads to several suggestions on how to improve the framework.

Keywords: Institutions, ontology, social-ecological systems

1. Introduction

Humanity currently faces large environmental challenges, including declining resources such as fish, soil, forests and fossil fuels, and pollution problems such as acid rain and climate change. The difficulties of analyzing these problems result in large part from their complexity. Environmental and natural resource problems occur in highly complex social-ecological systems (SESs), where variables do not contribute linearly to an outcome. Instead, variables interact in webs of causation, and the effects of each variable are highly contingent on the scale of analysis and the presence of many other factors. As a result, producing theories relating causes to effects in these systems is extremely difficult, as outcomes emerge from unpredictable interactions. The effects of a variable may appear clear in one system, but contribute to widely divergent effects in others. We call the challenge of constructing theories of such complex systems the *complexity problem* (see Folke 2006 for a more complete treatment of the dimensions of complexity in social and ecological systems). One typical example of complexity in SESs are heterogeneous actors with divergent interests interacting with other stakeholders within an ecosystem that itself behaves in a non-linear and dynamic way.

Due to this complexity, it may be unclear where problems are located within a system. To some (Rittel and Webber 1973), it seems that the problems involved are "wicked", which means that concepts will never lose their ambiguity or fuzziness, that answers necessarily remain unclear or uncertain and that each case is unique. Therefore, there is little learning from past cases, thus little trial-and-error process.

In our view, this is too pessimistic. *Ontologies* are one attempt to clarify concepts. Even if this is not always successful, less ambiguity can definitely be reached as hundreds of working and productive ontologies in "fuzzy" domains demonstrate convincingly (see below). Moreover, SESs are comparable, as shown by large databases of similar systems (Gibson et al. 2005), meta-analyses (Gutiérrez et al. 2011) and the fact that a few frameworks are advantageously used by thousands of different researchers. However, it is indeed difficult to construct meaningful classes in an ontology that do not overlap substantially and which help to cut the domain at its joints.

In addressing this problem, the study of environmental management and policy faces several other challenges. One of these has been an historic overreliance on simplistic policy prescriptions and management strategies that do not adequately account for this complexity. This reliance has led to disappointing results in many areas around the world (Ostrom and Cox 2010). We refer to this overreliance as the *panacea problem*. Much work has been done in turn to address the panacea problem, including research programs in common-pool resource (CPR) management (Ostrom et al. 1994; Ostrom et al. 2002), and social and ecological resilience (Folke 2006; Gallopin 2006). These programs have now examined a multitude of variables that affect outcomes where natural resources are managed. As such, they do represent progress in addressing the complexity and panacea problems by distinguishing between more general patterns across systems and specific patterns to a particular system.

However, this progress has led to another challenge, which we refer to as the *scatter problem*. This is a lack of integration of many research findings into a cohesive set of theoretical statements that explain how relevant conditions interact to produce success or failure over time. The few theories that exist to explain the many findings and concepts are not enough. Unless we assume linearly additive and non-contingent causality, the current state of knowledge falls short of the kind needed to provide useful governance advice across a range of social-ecological settings. Overcoming this scatter problem is complicated by several others. First, data sets of different research groups tend not to be shared or constructed in ways that facilitate integration. For the most part they are not designed to be shared due to investment in time, money or other resources by the research teams. Related to this, the specification and measurement of variables is frequently idiosyncratic to a particular study area, and too informal to facilitate comparisons of findings across studies. Conceptual validity is often unclear and external validity low. This situation mirrors a dilemma found in the ecological sciences (Madin et al. 2008).

The argument set forth in this paper is that a diagnostic approach, when implemented with an ontology, could potentially address all three problems. In this paper we will first discuss the benefits of a diagnostic and ontological approach, and then introduce components of an ontology we have constructed to begin the implementation of these approaches.

1.1. Diagnosis, complexity, and panaceas

The process of diagnosis involves asking a series of questions of a system, where subsequent questions are based on the answers to previous questions and are increasingly specific to the system being analyzed. This can be facilitated by a multilevel *framework* that orders a set of variables based on how likely they are to be relevant to a broad range of systems. A framework is a classificatory structure designed to "identify relevant variables“ (Ostrom 2009, 420), in this case for studying complex SESs. It is multilevel, shows subsystems, relationships and a "common set of potentially relevant variables" (Ostrom 2009, 420) embedded within different contexts (e.g. S or ECO in Figure 1). More broadly relevant variables occur at higher, more primary levels in the framework. Each question addresses a variable and its potential values at a particular level in the framework. An analyst begins at the first level, and then asks increasingly specific questions at lower levels of the framework.

Diagnostic thinking in environmental management was popularized by Young (2002) and then Ostrom (2007, 2008, 2009). Ostrom (2007), whose work is the basis for our ontology, initially proposed a multilevel SES framework, arranging the relevant variables within four primary components at the first level, as shown in Figure 1.

These components are: (1) the governance system; (2) the actors; (3) the resource system; and (4) the resource units. Each of these components contained multiple variables at the framework's second level. Subsequent levels have also been proposed, but have not so far been formalized.

Multiple levels of analysis, arranged in accordance to how specific the variables at each level are to a subset of systems, help analysts think diagnostically and address the complexity problem. Each time an analyst asks a question of a particular variable, this potentially eliminates many other variables from consideration. For example, if an urban groundwater system is analyzed,

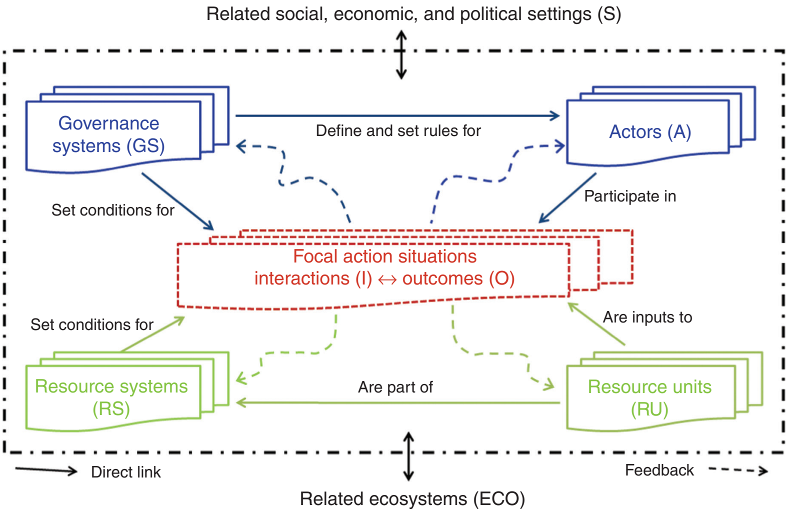

Figure 1: Tier 1 of the SES framework
Source: McGinnis and Ostrom 2014.

questions pertaining to ecological and biophysical attributes of a forest ecosystem are generally not relevant. A diagnostic approach supported by an ontology would formalize this branching process for any particular analysis, and would also emphasize the production of the type of information needed to decide whether the value that one variable takes on determines the potential relevance of another, more specific variable. Since ontologies are a conceptualization of a domain (e.g. of social-ecological systems), they are a kind of *operationalized framework*. As such, they are a tool for the diagnostic process, helping the analyst to choose the right questions at the right level and not miss important ones.

Diagnostics can also help with the *panacea problem*. The panacea problem occurs when an analyst constructs a theory that is either too specific or too general to be useful for a system to which it is applied. A theory may be defined as one or more general propositions used as explanation for a class of phenomena. Good theories may be falsified, are not circular, are internally and externally consistent, have a high explanatory power and are testable (Lakatos and Musgrave 1970; Popper 2002). States and markets, and the regimes of public and private property with which they are assumed to correspond, are frequently proposed solutions to environmental problems. However, these two solutions are much too general to be of much use in any particular system. The variability in the practical effects on outcomes *within* these categories is much too great in comparison with the variability *between* them (see Agrawal et al. 2008 for an application to forestry). Therefore, by themselves they are not very useful ways of categorizing complex institutional arrangements. They would require substantial taxonomic decomposition in order to be of more practical use to those who might implement them. The other extreme is when a highly specific theory is applied to a range of cases for which it is not accurate. Each of these is an example of a theory that is too simple to capture important variation across the systems to which it might be applied (see also Cox 2008).

With a multilevel framework, theories can be constructed at multiple levels of specificity as warranted by the similarities and differences among and between systems across levels. More specific theories will be accurate for a smaller subset of cases, but more useful for that subset. The reverse is true for more general theories. For example, some type of monitoring is likely to be beneficial for almost all systems where natural resources are managed. A theory positing this would be accurate for the great majority of systems, but would not provide much information regarding what type of monitoring should be implemented, or how it should be done in a particular system. Different types of monitoring, based on various configurations of rules and the types of actors carrying them out, will be better suited to different types of natural resource systems and social, economic, and political contexts. Theories of more specific factors complement the more general theories and are more useful for a smaller subset of systems. By constructing theories at the level of generality or specificity that is warranted by the data and information available, analysts can use a multilevel framework and diagnostic thinking to avoid the panacea problem.

Determining which level of specificity is warranted depends on the research question. If we are interested in a comparison of success of different SES like forests or irrigation systems, the first tier is probably the most relevant. If, on the other hand, a researcher is interested in differences in monitoring in irrigation systems, third and fourth tier variables are more of interest (Governance system (1) → Compliance (2) → Monitoring (3) → implementations of monitoring (4)). One particular way in which the diagnostic approach might be implemented is to embed it in a domain-specific ontology. Besides being able to address all three of the main problems discussed here, ontologies are well suited to implement the diagnostic approach. We now discuss the ways an ontology can improve upon the SES framework.

2. Ontologies

2.1. Ontologies and diagnostics

Ontologies are a kind of representation. One of the earlier and more famous representations is the taxonomy of Carolus Linnaeus. He presented the available knowledge of animals and plants at that time in a unified and structured way in one place. The scientific advantages of such an effort are obvious: scattered information is collected in one place and may serve as a common starting point for domain experts by providing them with a unified and consistent vocabulary. Assumptions of other researchers working with it can be checked easily and double efforts can be avoided. However, this system is *not* an ontology but a *classification*. Ontologies have richer relationships between entities than taxonomies, where usually only one relationship is defined (Gruber 1995). One example for this is the relationship of "tree" and "branch". A tree "has" branches. Trees themselves are "a kind of" plants. This relationship is different and could be "is" or "is a kind of", but not "has".

Recently, ontologies have been most widely used in the field of bioinformatics as a means of *structuring* the vast new quantities of genetic data that have become available (Baclawski and Niu 2006). Ontologies have also been developed by various governmental organizations, including the U.S. National Library of Medicine (http://lhncbc.nlm.nih.gov/project/medical-ontology-research or http://www.openclinical.org/ontologies.html), the World Health Organization in the construction of its most recent International Classification of Diseases (ICD 11) and the Food and Agriculture Organization (FAO) (http://aims.fao.org/website/Fisheries-ontologies-/sub2). In addition, there are a few so called "upper ontologies" (e.g. Cyc, Wordnet, SUMO) which are domain-independent and represent upper and more abstract concepts. Domain-specific ontologies may be latched onto these top level ontologies, but usually are used as stand-alone products. There is less research on ontologies and SESs, but see Khazai et al. (2014) and Kumazawa et al. (2014).

Ontologies in general are at the very core of any knowledge-based system, since inferences of any kind of artificial intelligence have to be based on some information about the world.[1] The concept of ontology has been historically used in two completely different and separate ways and domains. The first use is in philosophy where the term was originally taken from. In philosophy it has a distinct history as the study of the nature of existence and reality. In this article, we do not use it in this first philosophical sense, but in a second sense, that of computer science. A common definition of an ontology in computer science is "an explicit specification of a conceptualization," while a conceptualization is "an abstract, simplified view of the world that we wish to represent for some purpose" (Gruber 1993, 1). So, a conceptualization is for example a certain domain like a forest. All concepts (terms) are specified by assigning an unambiguous meaning to them and define their relationships precisely (e.g. forest "has" trees). In the end, this process allows computers to automatically derive new relationships by using logic. For example, if a forest has trees and some trees have leaves, then computers can deduce that there must be leaves in every forest. Therefore, creating an ontology means taking a conceptualization of a domain of knowledge, or a simplified view of what is relevant in the world, and operationalizing it by formalizing it. In the remaining sections of this article we use some technical

---

[1] For an introduction into knowledge-based systems see Akerkar and Sajja (2010); for an account of the development of knowledge representing ontologies and software tools, see Gennari et al. (2003); and for the most recent spectacular success of a knowledge-based system able to beat humans in a knowledge retrieval task, Watson by IBM, see www.ibm.com/smarterplanet/us/en/ibmwatson/.

terms of ontology design. These terms, a short definition and an example are summarized in Table 1 and Figure 2.

Ontologies formalize in a very particular way. In order to construct an ontology of SESs, we would first construct the world (SESs) as *classes* of objects (e.g. resource units), the members of which are called *instances* (e.g. fish species A, tree species B) and share a common set of *attributes* via their common membership (e.g. shape, mobility).

Classes can have several relationships with each other. The most common is the "is a" relationship, where one class is a subclass of another, which is then called the superclass (e.g. fish and trees are subclasses of the superclass resource units). Therefore, any class can be a subclass to other classes (tree to resource unit) and be a superclass to others at the same time (tree to conifers).

If class A (tree species) is a subclass of superclass B (resource units), this means that any instance that is a member of A is also a member of B. This is most often a one-to-many relationship, with one superclass having many subclasses. Subclasses with a common superclass are called sibling classes. Subclasses can have their own subclasses.

Table 1: Definitions of key terms in ontologies (adapted partly from the OWL Web Ontology Language Reference 2004, http://www.w3.org/TR/owl-ref/#Class-def).

Term	Definition	Example
Ontology	A formal framework for representing knowledge; an explicit specification of a conceptualization, while a conceptualization is "an abstract, simplified view of the world that we wish to represent for some purpose" (Gruber 1993, 1).	http://www.openclinical.org/ontologies.html)
Class	Concepts, types of objects, or kinds of things	cars; resource units
Subclass	A set of individuals in a class that are a subset of the set of individuals in another class. A class is by definition a subclass of itself.	electric cars; fish
Instance	Individuals or objects that are members of a class	Tesla S; herring
Relationship	Way in which classes and individuals relate to each other	"is a"
Attribute	Property, characteristic or parameter that classes and objects possess	red; length of 20–50 cm
Component	Part of a class	rear bumper; fin
Restriction	Description of what is or is not true (i.e. a class of all individuals for which all values of the property under consideration are either members of the class or are data values within the specified data range)	A car cannot be an animal.
Logical rule	Statement describing logical inferences that can be drawn from assertions (usually by automatic reasoning)	All fish can swim. A herring is a fish. Therefore, herrings can swim.
Disjoint class	If class A and B are disjoint, then no members of class A can be members of class B.	An animal either flies or cannot fly.

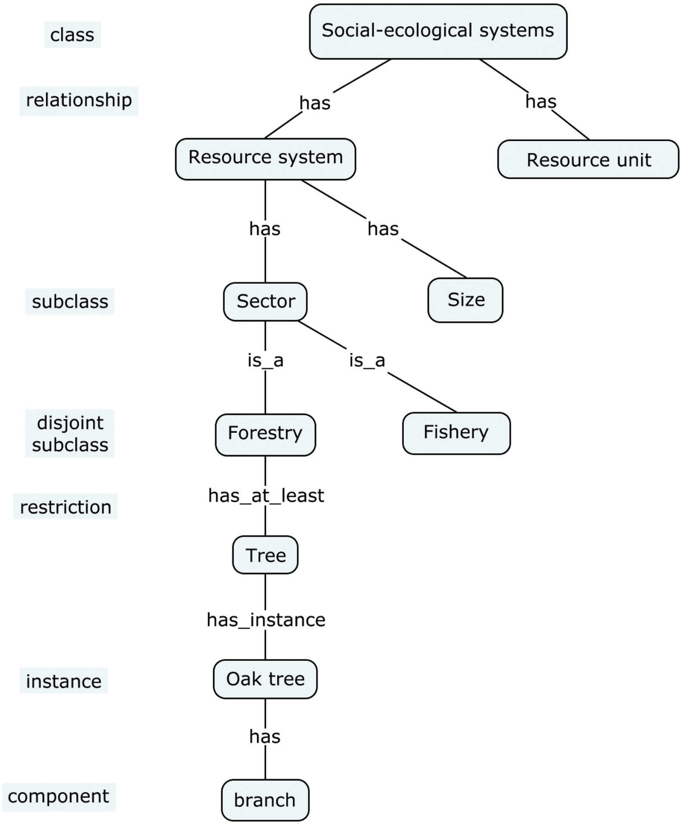

Figure 2: Example for key ontology concepts used in the text.

Each class has a set of attributes associated with it. The term attribute is sometimes used to mean two things (Baclawski and Niu 2006): a variable, or placeholder for several different values, (such as having a color) and the values themselves (such as being the color blue). Here we are referring to the placeholder concept. Subclasses inherit the attributes of their superclasses. For example, if we assign the attribute *biodiversity* to a class we call *ecosystems*, then every subclass of ecosystem (e.g. terrestrial or aquatic) would also have this attribute. It is also possible for a class to derive attributes from more than one superclass, which is called multiple inheritance.

Another very common relationship is the “has a” relationship mentioned earlier, where one class is a component of another class. Trees *have* leaves, for example. Subcomponents do not inherit attributes of their supercomponents, although subclasses would inherit the subcomponents of their superclass, just as they inherit their attributes. So, while the component “branch” would not have genus or species as its superclass “tree species” has, any genus or species of tree would have branches and leaves as subcomponents, because these are components of the more general class of trees. Ultimately, an ontology can have as many types of relationships as a modeler cares to define as relevant to the system of interest.

One important question about ontologies concerns their flexibility. In creating classes, the conceptualization of the domain seems to be very rigid at first. What happens if new cases add variables that do not fit into the ontology? If they cannot be accommodated within the existing concepts (classes), the ontology loses its value. However, constructing an ontology is an ongoing process – new cases are valuable test instances for the ontology. If a new concept is missing, it is added by adding a new subclass, thus making the ontology richer.

Ontologies have additional advantages in being able to address the complexity, the scatter and panacea problems. We will turn to each problem in the following three sections.

2.2. Addressing the scatter problem

Today, the wealth of information available within a given scientific field has become problematic in itself. No researcher is able to be aware all of the relevant information in his or her lifetime, let alone structure, organize and use it in regard for his or her research focus. Ontologies facilitate many tasks that would be impossible without them, and which may help address the scatter problem and panacea problems. Specifically, ontologies are able to:

1. *Decrease terminological confusion*
   Ontologies are a suitable tool to provide a common ground for definitions of concepts. In order to develop an ontology, concepts used in it have to be stripped of their inherent semantic heterogeneity; this helps avoid inaccuracies, and pinpoints one meaning to fuzzy concepts (or forces the developer to abandon the term or to split it up in more precise subconcepts).

2. *Help unify knowledge*
   Constructing ontologies forces the ontology engineer to not only disambiguate semantic heterogeneity, but to provide formal specifications for the terms used. The relationships between concepts have to be logically correct, or internally consistent (e.g. X implies Y or A is a part of B must be true for all concepts for which these relationships are defined – or a computer-based reasoner applied to the ontology will not able to deduce new relationships correctly). This formalization process in turn helps the research community to work together on a set of specified and formalized concepts.

3. *Reduce incomparability*
   Data incomparability is a common problem, because research data usually comes from many different sources. Different research goals and data formats prevent large-scale integration of research. It is important to point out that this is not a minor technical inconvenience, but is a real impediment in many disciplines (e.g. proteomics and genomics). Often, data in separate databases

simply cannot be compared or integrated, because there is no common ground in terms of common data fields and metadata, although the content may be very closely related. The huge amount of data being produced prohibits any "manual" or human-based solution. Ontologies have been put forward to solve this problem, since they provide an underlying structure to the data that would allow the automatic filtering and sorting of different data sources into identical records, since their content is related even if their format is not.

4. *Reduce redundancy of efforts and increase reusability*
   As soon as one concept has been described formally (and satisfactorily to the research community) it can be used over and over again. It is then easy to integrate into other ontologies. Reusability is a major issue in software engineering, and many structural features of programming languages themselves are designed in a way to help this (e.g. object-oriented programming with capsuled classes and functions; modularity at the project level, etc.). Ontologies would in fact provide that important advantage to research as well. However, reaching consensus between experts about key concepts is difficult.

Our approach suggests to use a generally accepted and broadly useable framework – the SES framework – as starting point and use the process of ontology building to reach broader consensus, however imperfect.

2.3. Addressing the panacea problem

We now discuss the ways in which ontologies can help address the panacea problem:

1. *Ontologies allow automated reasoning*. Sometimes even structured queries do not suffice to extract important information because the amount of data is just too large. This is where another advantage of ontologies comes in: they allow automated reasoning. On the basis of ontologies and with the help of formal logic, new knowledge can be inferred from existing data. For example, if a particular kind of monitoring – say monitoring by hired guards from the state government – is effective in helping produce the sustainable management of a tropical forest ecosystem, this produces several findings at once. It supports a very specific relationship between a certain type of monitoring in a certain type of resource system. But in the context of an ontology, where each of these classes are defined as subclasses of more general superclasses, this observation supports more general findings as well, which computers can automatically infer. A more general relationship that is supported is simply that monitoring helps produce sustainable outcomes in forest ecosystems – or more general yet, in any type of resource system. It is precisely this mix of general and specific theories that we need to move beyond the panacea problem, and ontologies can help us automate this important task.

2. *Ontologies are designed in a goal oriented way*. For well-designed ontologies this implies that the concepts are at the right level for the problems at hand. In addition, the class hierarchies within ontologies reaching from the most general concepts to very specific instances enabled researchers to choose the right level of specificity.
3. *Ontologies formalize and operationalize existing guiding structures like frameworks*. By formalizing diagnostic frameworks which are in turn especially suited to address the panacea problem (as discussed above), ontologies extend these particular strengths of frameworks.

2.4. Addressing the complexity problem

The very structure of ontologies – a rigid hierarchy of classes that are clearly defined – helps to address complex issues in SESs: The parent-child relationship of a superclass and its subclasses is the primary way in which an ontology can implement a diagnostic approach and thus address the complexity problem. When analyzing a particular system, each class, for example a class A, represents a set of diagnostic questions. First, there is a potential question about the value of each attribute of class A – e.g. does a high mobility of resource units influence manageability? Second, if class A has subclasses, (say classes B and C) the analyst can ask to which of these subclasses, B or C, the instance they are studying belongs (e.g. listing all fish species that are harvested under subclass fish species). The answer to this last question leads to further questions based on the attributes and subclasses of classes B and C. If classes B and C are disjoint, meaning that an instance cannot be a member of more than one of them (e.g. coastal marine system vs. open-ocean system), then we only need to ask questions associated with one or the other, but not both. If a subclass of the class of resource systems is the class of coastal marine systems, then for those systems we do not need to ask about attributes that are only relevant for open-ocean systems, assuming these two subclasses are disjoint.

Additionally, the subcomponent relationship can play an important role in the diagnostic process. If a class has a subcomponent, this leads to an additional set of questions that need to be asked, based on the attributes, subclasses, and subsequent subcomponents of that subcomponent class. Subcomponent relationships affect the diagnostic process differently than do subclass relationships. This is because subcomponents of a class are not disjoint from each other, since e.g. a branch is a component (part) of several different trees: unlike subclasses, where a user may select a single subclass and ignore other disjoint subclasses, a user must examine each subcomponent of a class.

In an ontology there are many hierarchies along different types of classes and their subclasses and subcomponents. Instead of unpacking one single hierarchy then, as may have been implied in the original SES framework, the analyst would proceed along several scales in order to explore the relationships between different types and subtypes of important factors.

2.5. Practical use of ontologies

Apart from these advantages, much of the potential of ontologies lies in the fact that the knowledge coded in them can be used by and with computers. This allows them to accomplish the following useful tasks:

1. *Improve communication* between humans and between humans and computers. Information can be linked or grouped together that had been separated into different groups for various reasons (e.g. sections in a library, information on different websites, similar articles in a journal). Ontologies provide a common understanding of concepts and are in a fundamental way language independent. Content (e.g. in scientific databases, on websites in different languages) can be extracted, analyzed and used in more general ways, if the same ontology is used.
2. *Enable and facilitates searches and queries*. The enhanced interoperability of content is one fundamental prerequisite for searches and queries across data. Structuring data according to search terms is perhaps one of the most important tasks in modern science. Ontologies facilitate this, because the representation of knowledge is structured, uniform, reusable and formalized.

However, there are some tasks that ontologies are not suited for. They are not sophisticated enough to be stand-alone products without human intervention. In practice, this means that the construction, text-mining and merging of ontologies and the generation of new knowledge have to be constantly supervised and checked. Furthermore, to be of use they have to be updated and treated as dynamic learning systems with new content continuously added. Although reasoners in ontologies do have inferential capabilities, no ontology solves logical puzzles on its own. For them – and all other computer systems – the quality of human input is decisive in determining their usefulness. Last, but not least, ontologies are never goal free. They are always built towards a specific function; this in turn limits other uses any one of them may have.

3. Methods – Constructing a SES ontology

3.1. Two approaches and their limitations

Roughly speaking, constructing an ontology consists of first choosing the domain and the purpose of the ontology, then selecting the classes, their attributes and relationships (subclasses and subcomponents primarily) and the instances (for a more detailed description, see Gomez-Perez et al. 2004). The logical formalisms for drawing correct inferences have been worked out in the first half of the 20th century and are available today as part of software packages like Protege (http://protege.stanford.edu/).

There are two basic ways of constructing ontologies. First, researchers may enter information manually. Second, automated text mining and information extraction can be used to build knowledge into an ontology. Both have their problems. The first method is work-intensive, slow and in most cases leads to incomplete projects, due to restrictions in resources and time. The second method often leads to poor quality, because text sources are unstructured, messy, vague and inconsistent. In addition, much uncertainty remains concerning the quality in regard to the unchecked parts of the ontology (which are the majority, since vast amounts of data are usually concerned).

There have been many attempts to overcome the limitations of both approaches. One example is the use of text sources structured by humans which follow a certain format like scientific articles or Wikipedia. Another idea is to "crowdsource" to many human reasoners the validation of inferences of the ontology reasoner, like Cyc does. Often both are used in combination, where automated text-mining is checked by human domain experts. Since the ontology used here begins with several key concepts of SESs which are well known from the literature, its first tiers are manually entered into the ontology. Its extensions, however, particularly at lower levels, might in the future be produced in part by text-mining, community-based extensions, and the integration of other ontologies.

3.2. Making the SES ontology work

In order to take full advantage of an ontology's promise for implementing the diagnostic approach, there are several rules that should be followed. To best understand these rules, it helps to imagine having a set of objects of interest in a particular domain of knowledge. Each of these has a set of attributes associated with it (e.g. forests have the attributes of average tree basal area and a level of biodiversity). In creating an ontology we are mainly interested in three things: (1) taking these objects and placing them into hierarchies of classes and subclasses; (2) specifying the important attributes that each class has; and (3) specifying other important relationships between classes, primarily the subcomponent-relationship. Here, we use a top-down approach, because the SES framework is already in place and can be used as a structure for the ontology.

None of these processes are automatic or unambiguous. Identifying key classes and attributes depends heavily on the goal of the ontology. For example, two groups of researchers focusing on a forest may have quite different questions that are best answered by the measurement of different variables. To allow for potentially different uses of an ontology, it is helped if it follows the principles of *minimal encoding bias* and *minimal ontological commitment* (Gruber 1995). This essentially means that it should not formalize and specify more than necessary, or in ways that are overly specific to the needs of a particular research group.

Gruber (1995) also emphasizes the importance of clarity, coherence, and extensibility for ontologies. Clarity is important for human users, while coherence supports valid inferences by both people and computers. Extensibility means that additional material can be added to an ontology without fundamentally altering existing material. We have taken each of these principles into account in the construction of our ontology.

One of the most fundamental decisions to make in constructing an ontology is when to create a subclass (or subcomponent) for an existing class. If a new subclass is not created, then the defining attribute to be associated with that subclasses remains as an attribute of the original class. For example, we could use the resource system *sector* variable (e.g. pastures, forests, fisheries) from the original SES framework as either an attribute of the resource system class or as a means of creating a set of subclasses defined by the different values of that attribute. The class resource system would then become a superclass, i.e. the parent class. Here we list the guidelines that we used for making this decision as we constructed our ontology with one example for each (see also Noy et al. 2001).

1. *Do not create single classes*: We avoid creating just one subclass of a superclass.
2. *Do not create too many subclasses*: If there are many subclasses of a superclass, we create intermediary classes between the two levels.
3. *Create disjoint classes*: We preferably create subclasses that are disjoint from each other (e.g. open-ocean vs. coastal systems).
4. *Create meaningful subclasses*: We preferably create a new subclass if its own attributes or subcomponents are not shared with the superclass or sibling subclasses.
5. *Create meaningful superclasses*: If a superclass of a set of subclasses has no unique attributes or subcomponents that the subclasses share through inheritance, we drop it from the ontology.
6. *Create similar classes*: We preferably create subclasses whose instances share common sets of values for their attributes.
7. *Create meaningful components*: We preferably create a subcomponent of a class if part of the definition of the class is that it is composed of multiple instances of the subcomponent.
8. *Create classes with a key property*: We preferably create subclasses or subcomponents of a class if these classes will take part in a relationship (e.g. be a subcomponent of) another class.
9. *Use the expertise of domain experts*: We are aware that the overall structure should reflect common usage of the terms by experts in the domain (here: we used the SES framework by Nobel Prize winner Elinor Ostrom as the starting point).

None of these rules are absolute. The first guideline is fairly intuitive. It is usually not very meaningful to create only one subclass of a class. There would be no other subclass in which to locate those instances that are a member of the superclass but not the subclass. The second guideline responds to the opposite situation, where there are too many subclasses, which is not usually very helpful either. The most useful hierarchy is generally neither to tall or too flat.

Next, if two classes are disjoint (rule 3), then a particular instance cannot be a member of both classes. This helps address complexity. If subclasses are disjoint, then an analyst knows that he or she is only dealing with one subclass, and does not have to worry about the attributes of its sibling classes. If two subclasses are not disjoint, which is often the case, then an analyst has to take both into account and deal with the ensuing complexity.

The fourth guideline is critical. If a class has, say, 10 attributes associated with it and we are considering if any of them should be used to create subclasses, we should use the one that leads to subclasses with their own, non-overlapping, sets of attributes (which would be subsets of the original 10). If we create subclasses that don't do this, we have not helped the analyst exclude from an analysis those variables that are not essential for his or her case. The most useful subclasses have attributes that are not shared with sibling classes or the superclass. Applying guidelines three and four to the resource sector example, we could use the sector variable (e.g. fishery, forestry, irrigation) as a means of breaking down resource systems into subclasses, which in this case would be disjoint from each other and have their own sets of attributes, along with some they have in common and inherit from their shared superclass.

The fifth guideline is directly related to the fourth. Just as we want subclasses with their own attributes, a useful superclass retains some attributes that are, through inheritance, shared by each one of the subclasses. If there are no such attributes, there is less need for the superclass, and the subclasses do not in fact have any attributes in common. For example, each resource system has a spatial or geographic extent, even if the sector subclasses are disjoint like in the examples shown above.

The sixth guideline reflects a different type of classification. In this situation, subclasses are not created because they have their own sets of variables or attributes associated with them. Instead, they may have the same set of attributes, but are still identifiable as different classes because they exhibit patterns in the values those attributes take on. This is the standard, well-known exercise of classifying objects based on their similarities across a standard set of variables, and it is commonly done in many disciplines.

The seventh rule regards subcomponents exclusively. As noted earlier, these play a different diagnostic role than do subclasses. If a class has a subcomponent class, this means that in order to understand the class we *must* look at the relevant attributes of at least one, but quite possibly several, instances of the subcomponent. In order to understand a particular tree, for example, we may need to examine many of its branches, roots, and leaves. If this is the case, then we should create these classes as subcomponents of the tree class.

The eighth rule amounts to saying that if a particular attribute value ends up being important in helping to define classes other than its own class, then it is probably important enough of a distinction to warrant its own subclass.

The final rule simply states that we should organize the ontology to be recognizable by, and therefore useful for, domain experts. One way we have done this in our ontology is to use superclasses that, while not necessarily having their own attributes (breaking rule 5) are commonly used to organize their subclasses in the relevant literature. While these are useful guidelines, it has to be kept in mind that although there are many best practices in designing an ontology, there is no single "right" way to do it. The structure, scope and level of detail of an ontology are heavily dependent on its goal.

4. Results – Converting the SES framework into an ontology

So far, we have demonstrated that ontologies may help with the complexity, the panacea and the scatter problem by describing their potential. In order to make the usefulness of ontologies more concrete, we will now show how formalizing the SES framework into an ontology improves diagnostic thinking and the analysis of SESs. As already stated above, the design of an ontology depends heavily on its purpose (goal). In our case, the purpose is to facilitate the analysis of SESs regarding social, economic and ecological outcomes. In order to do so, the ontology should be able to accommodate all concepts of potential interest holding all variables that have been mentioned as potentially relevant in the literature.

In order to achieve this, the overall structure of the SES ontology is identical to the SES framework, i.e. all first level concepts including interactions, related ecosystems and social, economic and political settings are converted to superclasses (Figure 3).

Starting with the top entity "thing", the superclass is "SES", with the second-tier variables becoming subclasses. A first difference to the SES framework consists in different relationships as can be seen in Figure 3: they are more precisely defined as e.g. "has attribute" or "has component", which is a distinction

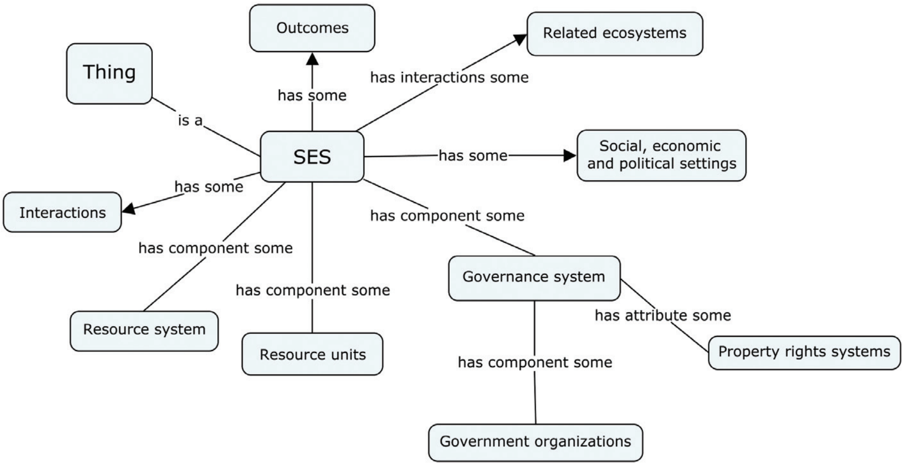

Figure 3: Top tiers of the SES ontology with relationships.

not made in the original framework. A further differentiation is made between "some" which stands for the logical equivalent of "there is at least one" and other logical types like "all". For visualization purposes, the class actors have been left out.

Below, we subdivide two concepts in order to present a more detailed analysis on lower tiers of the framework (which have been spelled out e.g. by Basurto et al. 2013).

If we want to analyze the sustainability of a SES, the superclass of resource units is particularly relevant. A more detailed diagnosis would also need to look beyond its seven subclasses. At this point, the rules for designing good ontologies described above (Section 2.2) lead to the insight that "RU6 Distinctive Markings" is not at the right place in the framework, because there are no further possible subclasses. Resource units either have distinctive markings or not. Therefore, this would suggest to move this attribute further down.

Next, "RU5 Number of units" is very important for any analysis interested in the sustainability of an ecosystem, but it is definitely not a broad superclass. Instead, it is a very specific attribute to many species that are harvested in a particular system and thus should be associated with individual resource units at this level. Therefore, again, we would suggest to move this attribute further down.

Clearly, interactions between actors are important for outcomes. Here, the SES framework conveniently provides the superclass "Interactions". Designing an ontology means to carefully define relationships between classes. So, for example, the interaction "I4 Conflicts among users", relates to different sources of conflicts, like distribution of appropriated units, distribution of duties, rules treating groups differently, etc. However, these linkages have not been spelled out in the framework. Therefore, the ontology design process suggests that it would be important *to add links between top-level interactions*. While not complete for all lower tiers, this is done for the top tiers in our SES ontology with some classes (see Figures 4 and 5 below) developed in detail. The OWL-file is available on request.

One of the strengths of ontologies is the use of automatic reasoning. Here is an example: if the ontology contains two disjoint classes, say broad-leaved trees and conifers, with thousands of instances (tree species) that may e.g. have been imported from existing taxonomy files, two inconsistencies may be found easily. First, if one tree species has the property of having needles, but is sorted into the class of broad-leaved trees, this will pop up as an error. Second, if one tree species appears in both classes, this again will trigger an error, since the classes are disjoint (either – or). The more complex the relationships, the more useful this feature becomes, going beyond the above mentioned examples of finding category mistakes.

Ontologies need lower, fine-grained levels to be of use. Therefore, we show two examples (Figures 4 and 5) on how this could be done in order to expand the SES framework. For this, we choose to decompose concept GS4 under the first tier "Governance systems".

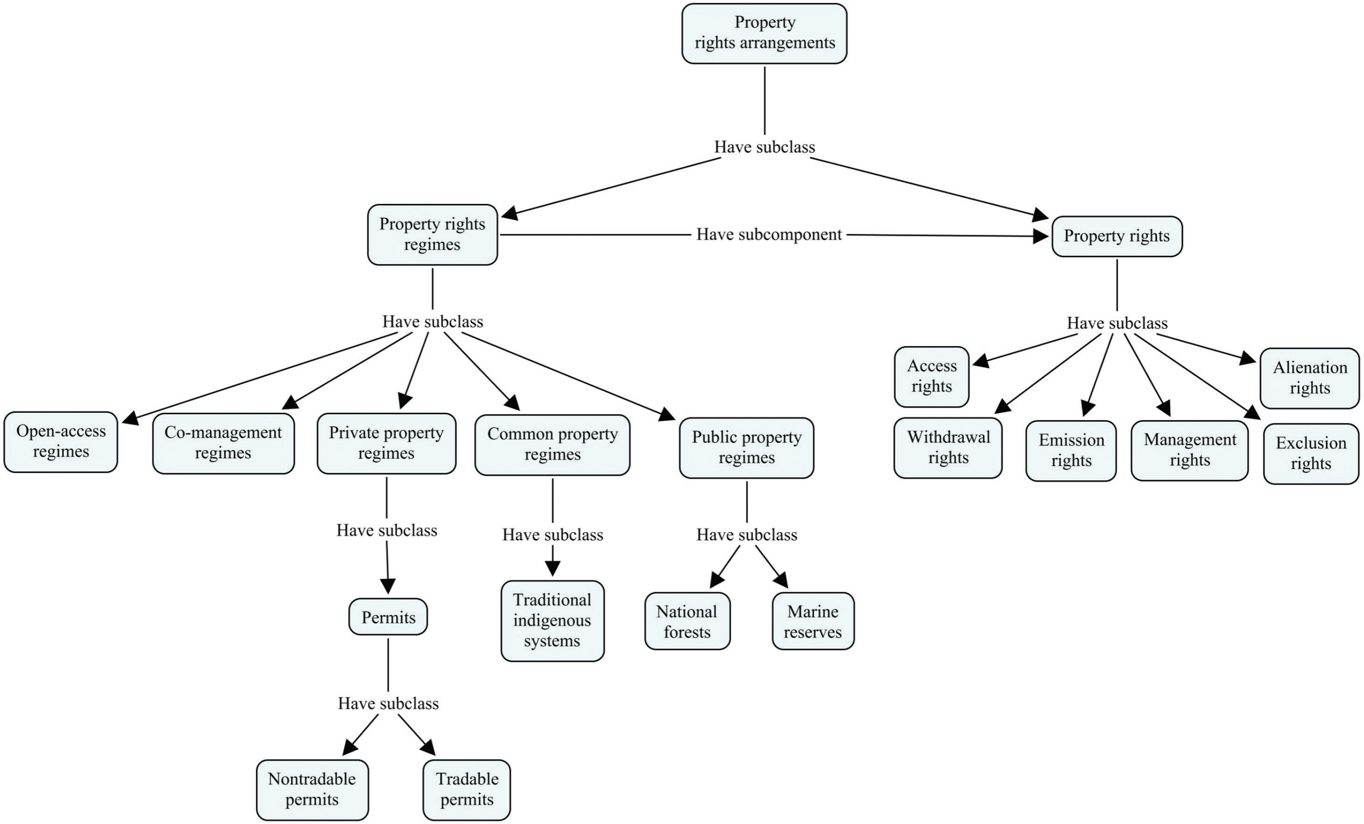

Figure 4: Possible property right arrangements.

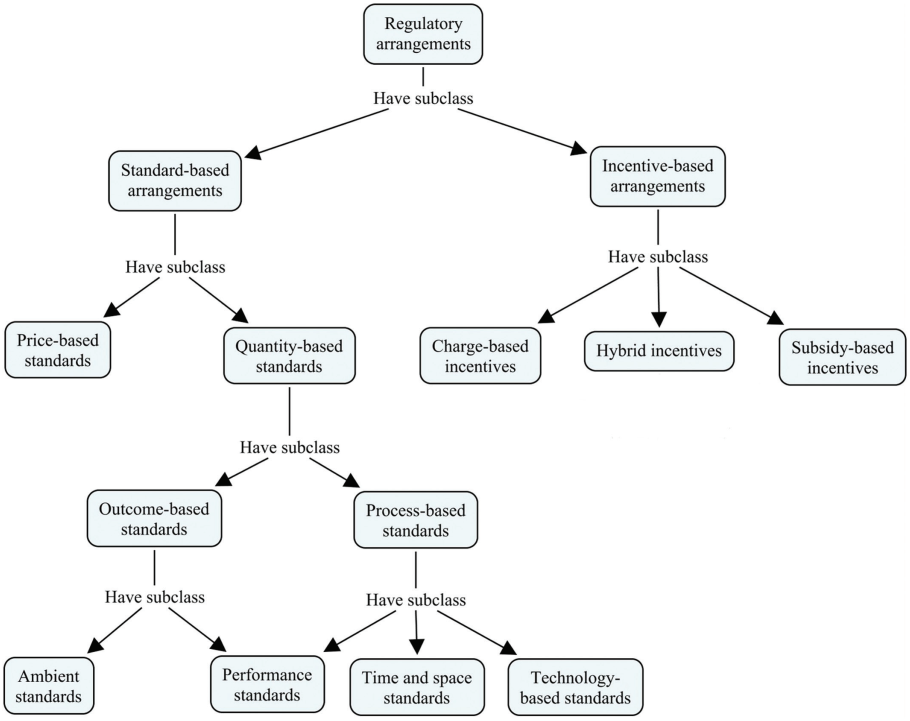

Figure 5: Regulatory arrangements.

As shown in Figure 4, we have two subclasses of property right arrangements: the property regime, or type of actor to which the rights are assigned, and the actual rights that are assigned (see also Starrett 2003). We also specify that property rights are a subcomponent of property regimes. The literature on property regimes has traditionally subdivided them based on whether or not goods are privately owned, owned in common by groups (common property), owned by state governments (public property), or open-access without ownership (Hanna and Munasinghe 1995). There is also a more recently popular category, particularly in fisheries governance. These are co-management property regimes, which are a combination of public property and either common or private property regimes (Jentoft 1989). For illustrative purposes we include some subclasses of these regimes as well.

Much of the policy literature makes rather simple arguments using these general property regimes, but this ontology makes clear that they are just one type of property rights arrangement, and have as subcomponents the specific property rights that give the actors specified in the regimes the rights to take a particular actions with respect to a particular resource system or unit. This points to their complex nature. At the same time, this shows that no ontology can be without some design purpose which may differ according to goal and theoretical background.

The second example is regulatory arrangements (Figure 5), many of which are commonly referred to as command and control policies, the implication being that they involve greater impositions on those who must abide by them. In contrast, market-based instruments are assumed to exert less control. We resist this distinction and terminology here, because (1) we believe that the image of governance arrangements always propagating from a state government to a society that is otherwise entirely free of either constraints or self-governance is a unfortunate myth, and (2) markets ultimately require many rules to function, many of which can exert much control over those who participate in them.

All regulatory arrangements have an attribute *regulatory object*, which describes the unit (not the actor) that is being regulated (e.g. fish, pollutants).

We first divide regulatory arrangements into incentive-based and standard-based arrangements (Figure 5). Incentive-based arrangements do not forbid or require a particular behavior or outcome, but change the incentives involved in doing so. Thus, they are distinguished from the standard-based arrangements in that the make less use of certain rules that prescribe a particular action or outcome on the part of the targets of the policy. There are many words to describe the two most basic types of incentive-based arrangements, which we group into the subclasses of charge-based incentives and subsidy-based incentives.

All of the incentive-based arrangements share a common attribute *incentive amount*, describing the amount of the financial incentive (e.g. $-50 for a charge and $50 for a subsidy). The distinction between charges and subsidies is not based on them having different attributes, but by the value of this financial incentive and whether it is positive or negative (hybrids have both positive and negative incentive amount values). Each incentive also has an attribute *incentive event* that describes the event to which it is attached. For example, *tariffs*, as a kind of charge-based arrangement, are mostly attached to the transport of a good across political boundaries. Many taxes are attached to specific transactions, while many subsidies are attached to the production of a particular good. The third subclass is a hybrid between the other two, with the most common example being deposit-refund schemes.

The second main subclass of regulatory arrangements is standard-based arrangements, which are the most directly prescriptive subclass of the governance arrangements. We initially divide these into price-based and quantity-based standards, the second of which we unpack further here. Quantity-based standards regulate the amount of a particular physical quantity within a unit of time and/or space. We subdivide these into outcome-based and process-based standards. Outcome-based standards define a particular outcome and require that it be obtained without regard to any particular course of action, whereas process-based standards prescribe a particular course of action but not which outcome should result.

These few examples of designing an SES ontology hopefully make clear what could be gained by following this more formalized approach to analysis.

5. Conclusion

As just discussed, ontologies serve as a helpful diagnostic guide, explicitly guiding an analyst towards the questions that need to be asked, based on the components of interest. At the same time, good ontologies address the *panacea problem* by their hierarchical design, different levels of specificity and goal-orientedness. The *scatter problem* is addressed by strict definitions and relationships which may provide a common starting point like the framework does. Then, the formalized class structure with subclasses is well-suited for diagnostic thinking and thus addresses the *complexity problem*.

However, it is important to note that the ontology does not provide the actual diagnostic expertise that is required in order to use it effectively. For example, knowledge about causal relationships is not currently contained in the ontology, and yet causal relationships are central in guiding an analyst to ask certain questions and not others. Indeed, these are probably the most important relationships in determining what is and what is not important for a particular analysis. That an ontology could contain causal relationships is not out of the question, but this has certainly not traditionally been done. However, ontologies do help in discovering logical errors and inconsistencies in a framework. This, and the fact that missing relationships can be automatically added, greatly helps in designing consistent and unambiguous representations of SESs.

We also note that a downside of ontology formalization is the loss of nuances, since ontology design is mostly about standardization. This would not only mean that at least a part of the SES-community would have to agree and be prepared to work with to a certain set of definitions. On the other hand, technical requirements for the use of ontologies are rather low, since numerous tools are already available that allow ontology design for non-experts and provide world-wide availability via the Internet (http://webprotege.stanford.edu/). Researchers working with such ontologies could interact via well-established tools like wikis, comments on the website or mailing lists.

Our representation here is far from exhaustive. In the future, we plan to further refine and test the governance system component, as well as begin to unpack the other three main components of the SES framework. Additionally, we need to add at least one class for the large variety of important outcomes that occur in SESs (see Agrawal and Chhatre 2011). As this effort has likely indicated, this will be an inherently interdisciplinary and collaborative undertaking. The ontology introduced here could serve as a starting point – it is in a format that allows users to easily explore it (via Protégé, a tool for building ontologies), and to make their own amendments and potentially new versions. In the future, the ontology discussed may be made available online (via the ontology development environment designed for the web called Webprotégé). This means that the technical implementation to work online with an ontology world-wide should be a minor issue with any university IT-department able to set it up in a matter of days. Given that the SES framework it is based on is very well known, this could also help to bring a wider community together to work with it. However, its usefulness ultimately depends on the input of the community with many successful examples leading the way.

An instructive example for collaborative research with interactive ontologies is http://www.ontologyonline.org. This site combines a browseable ontology with scholarly references and links to sites of other research groups. A good example of the potential for online ontology visualization can be found at the Bioportal site managed by the National Center for Biomedical Ontology (http://www.bioontology.org/bioportal). Other successful examples of the approach we might take include "OBI", which is an ontology for the description of biological and clinical investigations (http://obi-ontology.org/page/Main_Page), and the Gene ontology (http://www.geneontology.org/). This project aims to standardize the representation of genes across species and databases and provides a controlled vocabulary of terms.

Ultimately, how well the ontology is made is determined by how much it helps explore individual cases and, over time, to facilitate theory building across the multiple levels and scales that it contains. Undoubtedly, formalizing the key concepts of the research of SESs and linking them together via logically correct inferences – as the ontology does – is a first and important step in unifying the available knowledge, which has been done in other fields, with proven benefits.

Literature cited

Agrawal, A. and A. Chhatre. 2011. Against Mono-Consequentialism: Multiple Outcomes and Their Drivers in Social–Ecological Systems. *Global Environmental Change* 21(1):1–3.

Agrawal, A., A. Chhatre, and R. Hardin. 2008. Changing Governance of the World's Forests. *Science* 320(5882):1460–1462.

Akerkar, R. and P. Sajja. 2010. *Knowledge-Based Systems*. Sudbury, Mass: Jones and Bartlett.

Baclawski, K. and T. Niu. 2006. *Ontologies for Bioinformatics. Computational Molecular Biology*. Cambridge, Mass: MIT Press.

Basurto, X., S. Gelcich, and E. Ostrom. 2013. The Social–Ecological System Framework as a Knowledge Classificatory System for Benthic Small-Scale Fisheries. *Global Environmental Change* 23(6):1366–1380.

Cox, M. 2008. Balancing Accuracy and Meaning in Common-Pool Resource Theory. *Ecology and Society* 13(2):44.

Folke, C. 2006. Resilience: The Emergence of a Perspective for Social–Ecological Systems Analyses. *Global Environmental Change* 16(3):253–267.

Gallopin, G. C. 2006. Linkages Between Vulnerability, Resilience, and Adaptive Capacity. *Global Environmental Change* 16(3):293–303.

Gennari, J., M. Musen, R. Fergerson, W. Grosso, M. Crubézy, H. Eriksson, N. Noy, and S. Tu. 2003. The Evolution of Protégé: An Environment for Knowledge-

Based Systems Development. *International Journal of Human-Computer Studies* 58(1):89–123.

Gibson, C. C., J. Williams, and E. Ostrom. 2005. Local Enforcement and Better Forests. *World Development* 33(2):273–284.

Gomez-Perez, A., M. Fernandez-Lopez, and O. Corcho. 2004. *Ontological Engineering with Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web*. Heidelberg: Springer.

Gruber, T. R. A. 1993. Translation Approach to Ontology Specifications. URL: http://www.ksl.stanford.edu/KSL_Abstracts/KSL-92-71.html (accessed March 17, 2014).

Gruber, T. R. 1995. Toward Principles for The Design of Ontologies Used for Knowledge Sharing. *International Journal of Human Computer Studies* 43:907–928.

Gutiérrez, N. L., R. Hilborn, and O. Defeo. 2011. Leadership, Social Capital and Incentives Promote Successful Fisheries. *Nature* 470(7334):386–389.

Hanna, S. and M. Munasinghe. 1995. *Property Rights and the Environment: Social and Ecological Issues*. Washington, D.C.: Beijer International Institute of Ecological Economics, World Bank.

Jentoft, S. 1989. Fisheries Co-management: Delegating Government Responsibility to Fishermen's Organizations. *Marine Policy* 13(2):137–154.

Khazai, B., T. Kunz-Plapp, C. Büscher, and A. Wegner. 2014. VuWiki: An Ontology-Based Semantic Wiki for Vulnerability Assessments. *International Journal of Disaster Risk Science* 5(1):55–73.

Kumazawa, T., K. Kozaki, T. Matsui, O. Saito, M. Ohta, K. Hara, M. Uwasu, M. Kimura, and R. Mizoguchi. 2014. Initial Design Process of the Sustainability Science Ontology for Knowledge-Sharing to Support Co-Deliberation. *Sustainability Science* 9(2):173–192.

Lakatos, I. and A. Musgrave. 1970. *Criticism and The Growth of Knowledge*. Cambridge: University Press.

Madin, J. S., S. Bowers, M. P. Schildhauer, and M. B. Jones. 2008. Advancing Ecological Research with Ontologies. *Trends in Ecology & Evolution* 23(3):159–168.

McGinnis, M. and E. Ostrom. 2014. Social-Ecological System Framework: Initial Changes and Continuing Challenges. *Ecology and Society* 19(2).

Noy, N. F. and D. L. McGuinness. 2001. Ontology Development 101: A Guide to Creating Your First Ontology. URL: http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html (accessed March 17, 2014).

Ostrom, E. 2007. A Diagnostic Approach for Going Beyond Panaceas. *Proceedings of the National Academy of Sciences* 104(39):15181–15187.

Ostrom, E. 2008. Frameworks and Theories of Environmental Change. *Global Environmental Change* 18(2):249–252.

Ostrom, E. 2009. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science* 325:419–422.

Ostrom, E. and M. Cox. 2010. Moving Beyond Panaceas: A Multi-Tiered Diagnostic Approach for Social-Ecological Analysis. *Environmental Conservation* 37(04):451–463.

Ostrom, E., R. Gardner, and J. M. Walker. 1994. *Rules, Games, and Common-Pool Resources*. Ann Arbor: University of Michigan Press.

Ostrom, E., T. Dietz, N. Dolšak, P. C. Stern, S. Stonich, and E. U. Weber, eds. 2002. *The Drama of the Commons*. Washington: National Academy Press.

Popper, K. 2002. *The Logic of Scientific Discovery*. New York: Routledge, London, New York.

Rittel, H. W. and M. M. Webber. 1973. Dilemmas in a General Theory of Planning. *Policy Sciences* 4:155–169.

Starrett, D. A. 2003. Property Rights, Public Goods and The Environment. In *Handbook of Environmental Economics.* vol. 1. eds. K. G. Mäler and J. R. Vincent. Amsterdam: Elsevier.

Young, O. R. 2002. *The Institutional Dimensions of Environmental Change: Fit, Interplay, and Scale*. Cambridge, Massachusetts: MIT Press.