# Designing Annotation Mechanisms with Users in Mind: A Paper Prototyping Case Study from the Scientific Environment for Ecological Knowledge (SEEK) Project

Laura Downey
Long Term Ecological Research Network Office
University of New Mexico
Albuquerque, New Mexico, USA
lldowney@lternet.edu

**Abstract.** One of the goals of the SEEK project is to research the use of a Semantic Mediation System (SMS) that can 1) help scientists more easily discover highly relevant data and processing components for use in constructing scientific workflows, 2) automate or semi-automate the merging of heterogeneous data sets, and, 3) perform automatic transformation in a scientific workflow. SEEK is a major contributor to Kepler, a research prototype that provides a modeling and analysis environment for scientists. The SMS in Kepler is designed to make use of formal ontologies. In order for Kepler to exploit SMS, users need to annotate data and workflow components with terms from scientific ontologies. To better understand user expectations in annotation activities and to inform the design of annotation mechanisms, a paper prototyping activity was conducted. This paper describes that design activity, as well as future research implications for evaluating users' understanding of, and interaction with, semantic technologies.

**Key words:** user interface, ontology, annotation, evaluation

## 1 Introduction

The SEEK project [1], a National Science Foundation initiative, is focused on 1) creating cyber infrastructure and applications for ecological, environmental, and biodiversity research and, 2) educating the ecological community about ecoinformatics. SEEK is a major contributor to the Kepler Workflow System [2], [3], an open source modeling and analysis tool for creating, visualizing, executing and documenting scientific workflows. A scientific workflow is a collection of data flow and analytical steps that formalizes the research process: a formal model of the way scientists work Kepler is based on state of the art distributed computing technologies (grid computing) and also includes a semantic mediation system [4], [5], [6], to help scientists constructing workflows discover highly relevant data and processing components, to aid in the merging of heterogeneous data sets, and to perform automatic transformations within a workflow.

### Background

SEEK values user input in the design and development of its tools. Scientists (Kepler users) have been involved in the Kepler development process via user representatives on the development team, teams of scientists providing use cases and feedback, and various usability engineering activities including usability testing, user profiling and structured feedback sessions.

Prior to the paper prototyping activity, two iterations of annotation screens existed. The early screen designs were developed with input from user representatives on the SEEK team who had a high level of understanding of the semantic issues involved. The first design was a result of exposing functionality so that annotating data and workflow components was possible. The second design was an attempt to provide users with a centralized place in the user interface where they could make annotation decisions and also access information necessary to that decision making process. These initial designs focused on letting users browse the ontology categories and understand the various relationships between categories. Figure 1 shows the second iteration of the annotation dialog which contains a very rich set of functionality:

- The annotation target is displayed to remind the user which component they chose to annotate and a list of any current annotations for that target is displayed.
- The dialog also provides the ability to query information about a term in the ontology. An ontology term's definition, parents, children and siblings can be toggled for display in one small space to make efficient use of screen real estate.
- A section on what other workflow components contain a selected annotation is provided to aid the user in making a selection by examining how other workflow components in the system are annotated.
- In an effort to support various display modes, besides the standard alphabetical list view, a tree browsing view was added as well as a graphical or visual view of an ontology. This was done to support user preference and varying degrees of user experience, and different learning modes.

The overall design goal of the second iteration of the annotation screen was to offer "one stop shopping" for annotations and to provide supporting decision-making information. The annotation dialog had become rich in functionality. But there was a concern that scientists with no knowledge engineering background might find it overly complex. That category of user had not been previously involved and we wanted to make sure we also obtained their input and feedback.

A small set of formal ontologies [7] was developed by knowledge engineers and subject matter experts for use in Kepler while computer scientists were developing SMS – a "smart" data discovery and integration system performing query processing and reasoning tasks over formal ontologies. "Smart" in this context means the mediation tasks are implicitly used by the system, not explicitly or directly used by the user during discovery and integration, in contrast to some ontology based information retrieval systems [8]. With the SMS system, the formal ontologies, and some workflows in place, the next step was to introduce the notion of annotation using formal ontologies to users with little or no knowledge engineering experience in order to determine how best to design annotation mechanisms for their use. To accomplish this objective within time and resource constraints, the project team selected a formative design approach, paper prototyping, as a way to quickly gather valuable information from a small number of users.

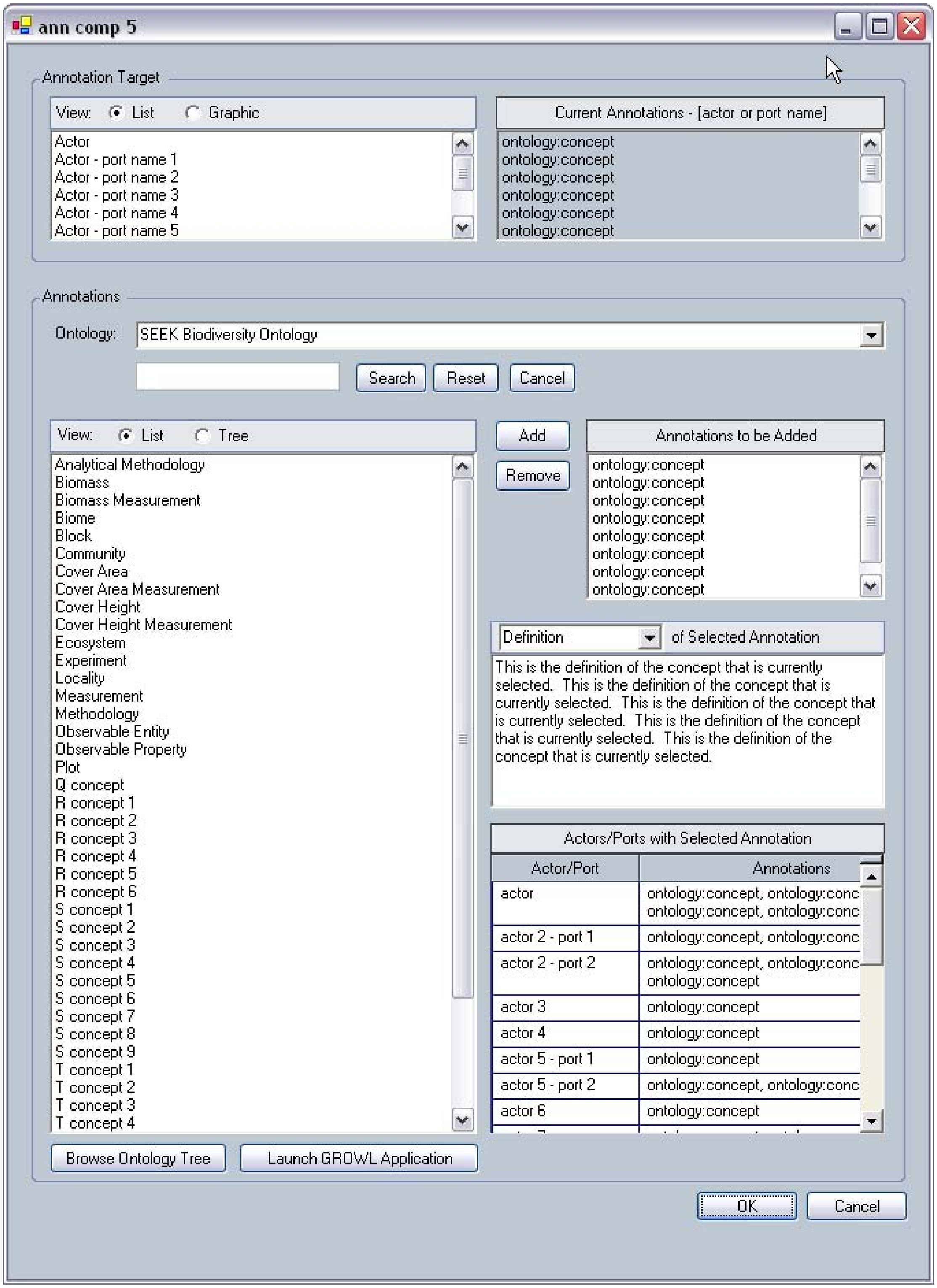



**Figure 1. Design idea for second iteration of screen design for annotation**

# 2 The Paper Prototyping Activity

The goal of the paper prototyping activity was to:

- Understand user expectations when performing annotation activities,
- Determine the steps users would take during annotation activities,
- Understand the information they needed to support decision making about annotation.

To accomplish these goals, the decision to conduct a formative exploratory activity with a small group of scientists was made. Formative usability techniques are used in iterative development and are useful in quickly acquiring valuable user input with a small number of participants. Paper prototyping [9] was chosen because:

- It is an exploratory evaluation technique [10]
- It does not require extensive preparation (we were taking advantage of an existing meeting where users would be present)
- A lot can be learned from a small number of users
- It does not require a working software prototype.

Three scientists from the SEEK Biodiversity and Ecological Analysis and Modeling (BEAM) group were brought in to participate in a one hour paper prototyping activity that centered around annotation of data and components in a scientific workflow. Annotation in this sense means users choosing categories (classes) from formal ontologies as the annotations rather than free form notes and comments in the more general sense [11].

## *2.1 Procedure*

Participants were given:

- A brief introduction on paper prototyping
- A brief introduction on what we mean by annotation and how the system could exploit workflow components that are annotated. (See Appendix A – Script).
- A set of instructions describing their task and explaining the data set and associated workflow they would be using in the annotation exercise. (See Appendix B – Instructions).
- A set of paper prototyping materials. (See Appendix C).
- A list of terms (in tree form) to use during the annotation (See Appendix D).

Once the exercise began, participants were asked to:

- Verbalize and explain the steps they took to accomplish the annotation exercises
- State what kinds of information they needed to make their annotation decisions
- State their expectations of labels
- State where and how they expected to find and/or access functions or information
- Draw, or direct the facilitator to draw or make notes, on various paper materials.

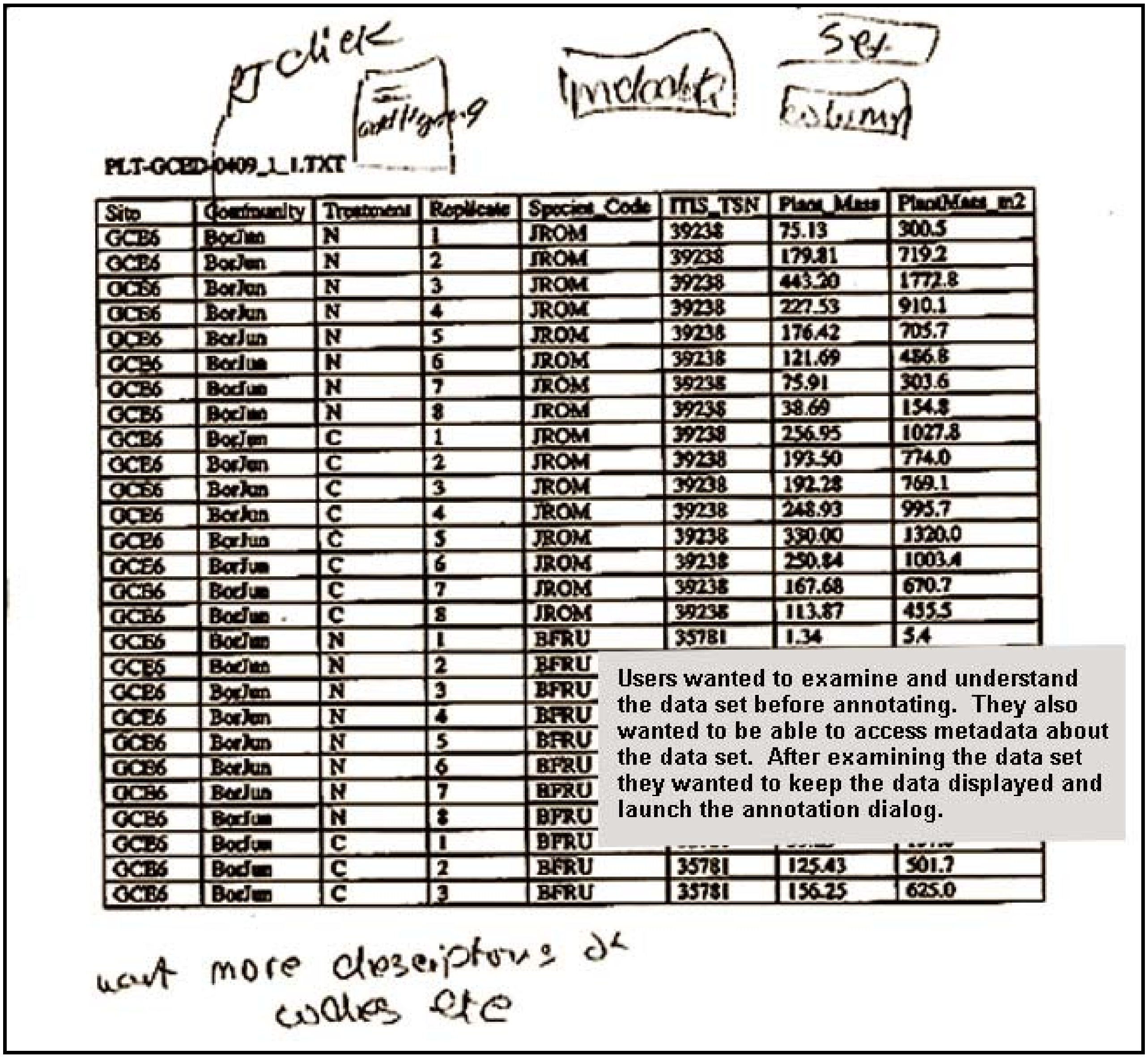
Participants spent one hour on the paper prototyping activity. After one user completed the hour, it was clear that there was only time to do the data set annotation and one component annotation. Column annotations were not done. Also users didn't spend much time on the component annotation, just a few minutes at the end of the hour. The majority of the time was spent annotating the data set.

## 2.2 Results

There was a significant amount of information obtained during the paper prototyping activity. Results are listed in four general categories of: screen designs, annotation, tasks analysis and user interface characteristics.

### Screen Designs

The screen designs produced during the design activity revealed a trend towards simplicity. They verified the assumptions about what users minimally expected on an annotation dialog: list of terms, access to term definition, and running display of terms selected. Figure 3 gives an example of one participant's annotation dialog. Figure 2 is an example of the data set screen and Figure 4 is an example of a dialog for suggesting additional terms to the ontology.




| Site | Community | Treatment | Replicate | Species_Code | ITIS_TSN | Plant_Mass | PlantMass_m2 |
|---|---|---|---|---|---|---|---|
| GCE6 | BorJun | N | 1 | JROM | 39238 | 75.13 | 300.5 |
| GCE6 | BorJun | N | 2 | JROM | 39238 | 179.81 | 719.2 |
| GCE6 | BorJun | N | 3 | JROM | 39238 | 443.20 | 1772.8 |
| GCE6 | BorJun | N | 4 | JROM | 39238 | 227.53 | 910.1 |
| GCE6 | BorJun | N | 5 | JROM | 39238 | 176.42 | 705.7 |
| GCE6 | BorJun | N | 6 | JROM | 39238 | 121.69 | 486.8 |
| GCE6 | BorJun | N | 7 | JROM | 39238 | 75.91 | 303.6 |
| GCE6 | BorJun | N | 8 | JROM | 39238 | 38.69 | 154.8 |
| GCE6 | BorJun | C | 1 | JROM | 39238 | 256.95 | 1027.8 |
| GCE6 | BorJun | C | 2 | JROM | 39238 | 193.50 | 774.0 |
| GCE6 | BorJun | C | 3 | JROM | 39238 | 192.28 | 769.1 |
| GCE6 | BorJun | C | 4 | JROM | 39238 | 248.93 | 995.7 |
| GCE6 | BorJun | C | 5 | JROM | 39238 | 330.00 | 1320.0 |
| GCE6 | BorJun | C | 6 | JROM | 39238 | 250.84 | 1003.4 |
| GCE6 | BorJun | C | 7 | JROM | 39238 | 167.68 | 670.7 |
| GCE6 | BorJun | C | 8 | JROM | 39238 | 113.87 | 455.5 |
| GCE6 | BorJun | N | 1 | BFRU | 35781 | 1.34 | 5.4 |
| GCE6 | BorJun | N | 2 | BFRU | | | |
| GCE6 | BorJun | N | 3 | BFRU | | | |
| GCE6 | BorJun | N | 4 | BFRU | | | |
| GCE6 | BorJun | N | 5 | BFRU | | | |
| GCE6 | BorJun | N | 6 | BFRU | | | |
| GCE6 | BorJun | N | 7 | BFRU | | | |
| GCE6 | BorJun | N | 8 | BFRU | | | |
| GCE6 | BorJun | C | 1 | BFRU | | | |
| GCE6 | BorJun | C | 2 | BFRU | 35781 | 125.43 | 501.7 |
| GCE6 | BorJun | C | 3 | BFRU | 35781 | 156.25 | 625.0 |

**Figure 2. Sample paper prototype of data set dialog**

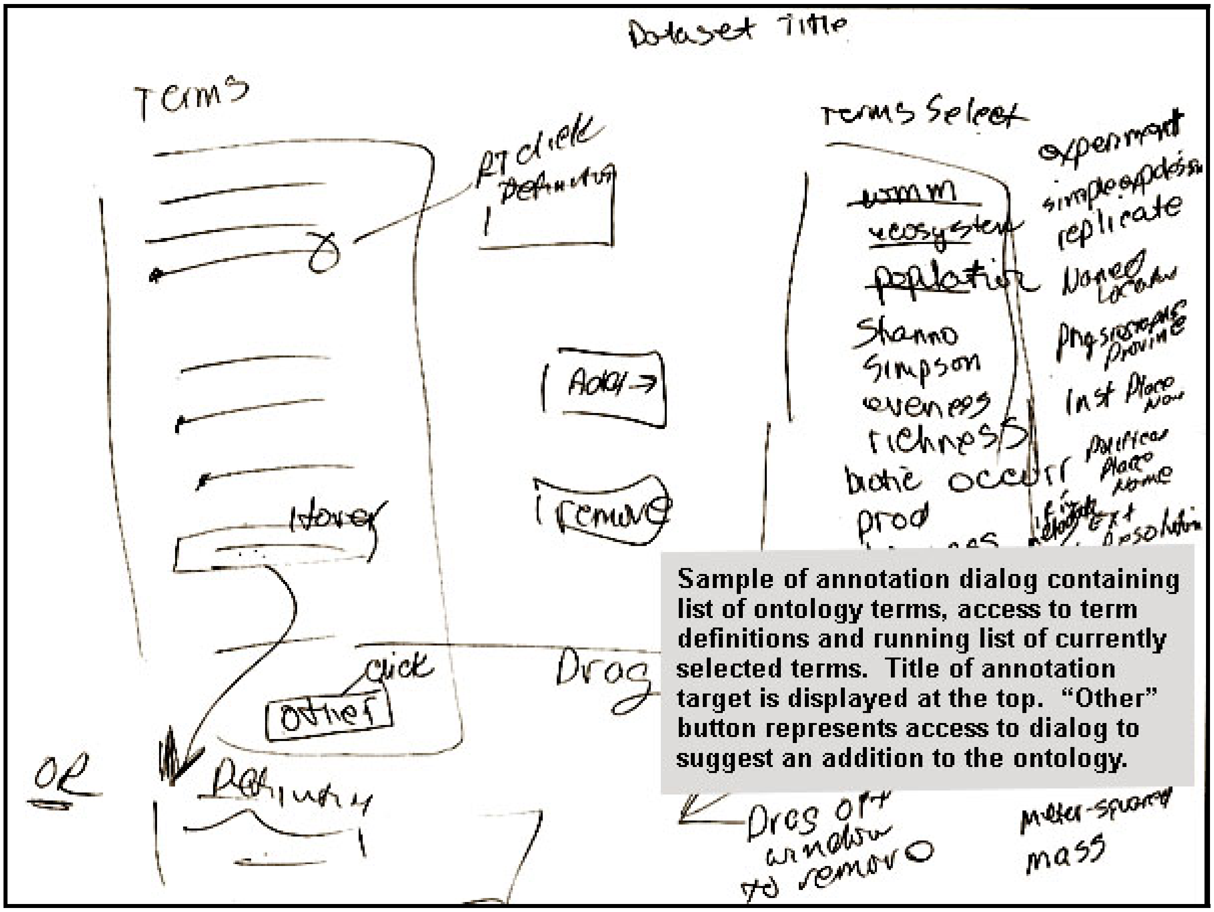



**Figure 3. Sample paper prototype of annotation dialog**

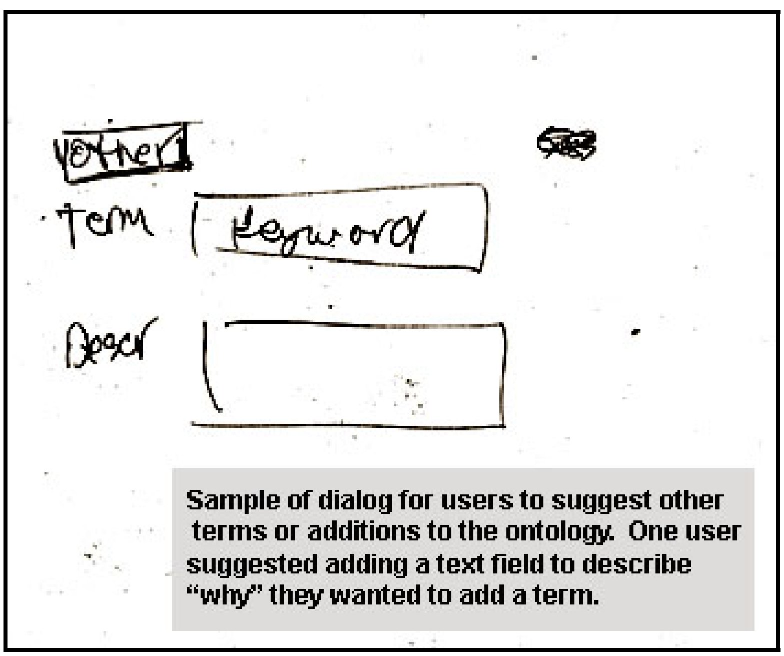



**Figure 4. Sample paper prototype of suggestion dialog**

## Annotation

The most interesting results of the paper prototyping exercise were:

- Participants' remarks on terms being presented to them that would never apply in a certain *context*
- The *variance* among the terms selected.

### *Context of Annotations Presented*

One user wanted to add two top level classifications to the ontologies: "data" and "analysis." He suggested this because he said none of the analysis terms would ever apply to the data. All users mentioned the fact that terms like independent and dependent variables would be common to every analysis so not something they would think would needed to be annotated. But when pointed out they might be useful to column level annotation, users agreed they might possibly be useful in that context.

### *Variance of Selected Annotations*

Table 1 lists the annotation terms selected for the data set from the three study participants. It is easy to see there was a wide variation among the terms selected.

| Selected Annotation Terms for Data Set | | | |
|---|---|---|---|
| Participant 1 | Participant 2 | Participant 3 | |
| • Community<br>• Population<br>• Taxonomic biodiversity<br>• Richness<br>• Biotic abundance<br>• Biotic occurrence<br>• Species<br>• Productivity<br>• Biomass<br>• Internal code<br>• IT IS code<br>• Blocked experimental design<br>• Replicate<br>• Experimental Observation<br>• Regression<br>• Nitrogen addition<br>• Untreated control | • Community<br>• Taxonomic biodiversity<br>• Biomass<br>• Internal code<br>• Replicate<br>• Plot<br>• Mass | • Community<br>• Ecosystem<br>• Population<br>• Shannon Index<br>• Simpson Index<br>• Evenness<br>• Richness<br>• Biotic occurrence<br>• Productivity<br>• Biomass<br>• Species<br>• Common species name<br>• Genus species name<br>• Internal code<br>• ITIS code<br>• Experiment | • Simple experimental design<br>• Replicate<br>• Name location<br>• Physiographic provinces<br>• Institutional place names<br>• Political place names<br>• Extent (if in metadata)<br>• Resolution (if in metadata)<br>• Temporal (depends on metadata)<br>• Meter squared<br>• Mass |

**Table 1. Selected Annotation Terms for the Data Set from the Three Study Participants**

## Task Analysis

A task analysis revealed the basic sequence of steps [12], for annotating a data set and for annotating a workflow component. What emerged as a result of the task analysis was the need for supporting types of information to be concurrently displayed while the annotation dialog was displayed to aid users in deciding which categories to choose from the ontology.

Basic sequence of steps to complete annotating a data set was:

1. Examine data set (look at column headings, look at values, try to understand codes, note format of the data)
2. Examine any metadata about the data set
3. Keep this information visible while accessing list of annotation terms
4. Examine list of terms and display the definition of a term if needed
5. View any existing annotation terms (only one user mentioned this)
6. Select terms from list

Basic sequence of steps to complete annotating a component was:

1. Examine information about a component like name, description, variables, inputs, outputs.
2. Keep this information visible while accessing list of annotation terms
3. Examine list of terms and display the definition of a term if needed
4. Select terms from list

## User Interface Characteristics

Lastly, there were various user comments on user interface design and functionality.

***Access***

- Two of three participants commented they expected to find some kind of annotation function in a context menu.
- One user suggested accessing the annotation function from some button within a window.

***Interaction***

- Two of three users suggested "add/remove" buttons for navigation, one suggested drag and drop.

***Layout***

- All three users suggested two columns for the annotation window, with a list of terms to choose from and selected terms.

***Functionality***

- One user suggested a search function for the ontology but wasn't sure where the results would be displayed – he didn't want it to replace the ontology tree. He also wanted to select terms in the search results list and then those would get highlighted in the ontology tree so he could look at the terms around it.
- One user suggested an "Expand All" feature for the tree.

- Two users did note that there should be a way to suggest terms not in the ontology, perhaps a simple input screen with term and definition. One user also wanted another field that asked “why” someone was suggesting the term be added to the ontology.

***Navigation***

- One user suggested first letter navigation of the ontology, e.g., if you pressed the E key you would jump to the next term beginning with E.

***Terminology***

- One user remarked that the term “annotation” was very overloaded and he would rather see something like “Add Keywords” or “Add Descriptors” (however, “keyword” is also overloaded, and in this case is used to mean something specific in metadata).

## *2.3 Discussion and Implications for Design of Annotation Mechanisms*

### Simplicity Trend

The results of the exercise clearly indicated a need for a simple design. Users didn’t suggest lots of functionality and various ways to explore and understand the ontology space. They were focused on quickly identifying annotation terms. It is possible that the expectation of simplicity is due to unfamiliarity with the annotation process, and/or with a lack of knowledge engineering expertise. As users become more sophisticated in performing annotation tasks, they may desire more functionality. In order to support both novice and advanced users, it may be more desirable to provide the simple design as the standard display with the option to view a more advanced design for more complex examination and tasks, similar to the simple and advanced search designs offered by many search engine applications. However, it remains to be seen whether more advanced functions and different types of information display will suit the task, and also be tolerated or desired by users.

### Information to Support Decision Making

During the task analysis**,** participants noted what information they needed access to before making an annotation decision. For data set annotation, they wanted to explore and understand the data, decode if necessary, look at the format, the associated metadata and any other explanatory information about the data. For components, users said they would examine characteristics such as name, description, variables, inputs, outputs as information important in making an annotation decision. Participants indicated they wanted this information to ‘stay displayed” at their discretion while also viewing the annotation terms. At the least, this means providing some non-modal dialogs. One idea is to organize this kind of supporting information so that it is easily accessible in a non-modal tabbed dialog that users can easily toggle with one click.

### Challenges/Issues in Understanding and Using Ontologies

***Presentation of Ontologies***

The three participants thought there were lots of terms to choose from and lots of noise, e.g., terms that would not apply depending on context. Participants stated this created more work for them because they had to sift through not only lots of terms but non-relevant terms to get to relevant terms. One participant commented that the ontologies provided were perhaps too extensive and that it was a lot to go through – which seemed interesting to note since according to the ontology creators, these were relatively simple ontologies and small, considering the size some could be. This raises issues on whether the system can automatically "prune" the ontology for presentation depending on the object to be annotated. Participants' desire for smaller ontologies and contextual presentation of ontologies also supports the simplicity trend mentioned earlier.

***Variance of Selected Annotation Terms***

As illustrated in Table 1, there was a wide variation among the annotation terms selected by the participants. This suggests that users had different ideas on the scope to which they needed to annotate. One participant commented that his annotations would not only take into account "what" the object to be annotated was, but would also take into account how the data or component might be used. So it is clear that "annotation" and the scope of it can be interpreted in many ways.

One participant asked if the purpose in choosing terms was to select the lowest level of granularity. It may be possible that by providing some annotation guidelines (like suggesting the user only pick the lowest level granularity), variance might be reduced. However, lower level granularity could imply richer more dense ontologies, which also means more terms for users to sift through which then becomes related to the presentation, filtering and context issues mentioned earlier.

Another participant comment was that making use of the annotations was only as good as the quality of the annotations themselves. This raises the issue not only of the quality of the ontologies affecting the system inferences, but also whether users are consistent in their choices of annotation and how that affects the system inferences. If two users annotate a data set or a component in very different ways, how will that affect future data or component discovery?

***Missing or Incorrect Terms in the Ontology***

Users did offer an idea for suggesting a desired annotation term to be added to the ontology. This could also be extended to allow users to question an incorrect term or the placement of a term within an ontology. This may be one mechanism that the system can offer to help bridge the gap between users and knowledge engineers that create ontologies.

# 3 Research Implications

Several research implications arose from the paper prototyping activity. They can be divided into two general areas: user interface design and annotation.

## User Interface Design

- As users become more sophisticated/experienced, will they want or need more features to support annotation tasks? How could this best be evaluated?
- Which type of information display (e.g., alphabetical list, tree view, graphical/visual view of ontology) is the most desirable? Do the different displays support different user

expectations and different user tactics when trying to learn, explore, understand and use formal ontologies?

- How many “levels“ of an ontology need to be displayed before a user forms a general idea of the structure of the ontology? Does limiting the detail of the list and keeping it simple conflict with the need for specificity (low-level granularity) in the tagging of data?
- Does pruning a graphical representation equate to filtering a list?
- How do we best represent relationships across classes and/or disjoint classes? Do users of ontologies need to understand these lower level details when performing annotation tasks?
- Is there a way for the system to contextually filter out ontologies or categories in ontologies that are not relevant depending on the selected annotation target?

### Annotation

- The mention by participants to include a way for users to suggest additional categories to the ontology may offer a mechanism to bridge the gap between the expectations of users that aren’t knowledge engineers and the knowledge engineers that construct formal ontologies. Studying the process of what terms get suggested and why, and if knowledge engineers accept the suggestions and how they insert them into a formal ontology, could reveal collaboration issues and suggest ways and means for supporting this collaboration.
- There is a need to study how users select terms and why, in their annotation tasks, and to study the variance among the annotations. Users who are very familiar with the annotation targets may select terms differently than users who are less familiar with an annotation target. Should only authors of objects annotate their own objects? How does the introduction of annotations from a user not familiar with the annotation target change how the system interprets things about an annotated target?
- Does a wide variance among selected terms affect the inferences the system might make?
- Can a set of annotation guidelines be formed to help reduce variance?

## 4 Conclusion

The paper prototyping activity was one small step in trying to understand if scientists with no formal ontology knowledge can exploit ontologies to successfully and adequately annotate scientific workflow components. The results suggest that some success is possible but most likely with a simple presentation. However, several issues were also raised that need to be studied further. There is clearly a great deal more to be learned about users’ annotation strategies, their understanding of ontologies and how they can best exploit the power of semantic technologies. Evaluation of user interfaces and user strategies with real users is key in designing and providing a usable interface in the semantic arena. Future plans include continuing with iterative and formative evaluation of the annotation mechanisms offered within the Kepler software. Plans also include the study of variance of annotations.

## References

1. Michener, W.K., Beach, J.H., Jones, M.B., Ludaescher,, B., Pennington, D.D., Pereira, R.S., Rajasekar, A., and Schildhauer, M., [in press], A knowledge environment for the biodiversity and ecological sciences, Journal of Intelligent Information Systems.

2. Altintas, I., C. Berkley, E. Jaeger, M. Jones, B. Ludäscher, S. Mock, (2004). Kepler: An Extensible System for Design and Execution of Scientific Workflows. 16th International Conference on Scientific and Statistical Database Management. IEEE publication number P2146.
3. Ludäscher, B., I. Altintas, C. Berkley, D. Higgins, E. Jaeger-Frank, M. Jones, E. Lee, J. Tao, and Y. Zhao. (2006). Scientific Workflow Management and the Kepler System, In Concurrency and Computation: Practice & Experience, Special Issue on Scientific Workflows, in press.
4. Bowers S, K. Lin, and B. Ludäscher, (2004). On integrating scientific resources through semantic registration. In: Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM), IEEE Computer Society.
5. Bowers, S and B. Ludäscher, (2003). Towards a generic framework for semantic registration of scientific data. In Proceedings of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, CEUR Workshop Proceedings, volume 83, ISSN 16130073.
6. Bowers S and B. Ludäscher, (2004). An ontology-driven framework for data transformation in scientific workflows. In Proceedings of the International Workshop on Data Integration in the Life Sciences (DILS), Springer, LNCS, volume 2994.
7. Passin, Thomas B., (2004). Explorer's Guide to the Semantic Web, Manning, pp 141-169.
8. Garcia, Elena and Sicilia, Miguel-Angel, (2003) User Interface Tactics in Ontology-Based Information Seeking, PsychNology Journal, Volume 1, Number 3, pp 242-255
9. Snyder, Carolyn, Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces, (2003), Elsevier
10. Stone, D., Jarrett, C., Woodroffe, M., Shailey, M., (2005). User Interface Design and Evaluation, Morgan Kaufmann, pp 552
11. Passin, Thomas B., (2004), Explorer's Guide to the Semantic Web, Manning, pp 90-105
12. Hackos, JoAnn T. and Redish, Janice C., (1998). User and Task Analysis for Interface Design, John Wiley & Sons, Inc., pp 71-72