PragmatiX: An Interactive Tool for Visualizing the Creation Process Behind Collaboratively Engineered Ontologies

Simon Walk, Jan Pöschko and Markus Strohmaier

Knowledge Management Institute, Graz University of Technology, Inffeldgasse 13/VI, 8010 Graz

Keith Andrews

Institute for Information Systems and Computer Media, Graz University of Technology, Inffeldgasse 16c, 8010 Graz

Tania Tudorache, Natalya F. Noy, Csongor Nyulas and Mark A. Musen Stanford Center for Biomedical Informatics Research, Stanford University, 1265 Welch Road, Stanford, CA 94305-5479, USA

ABSTRACT

With the emergence of tools for collaborative ontology engineering, more and more data about the creation process behind collaborative construction of ontologies is becoming available. Today, collaborative ontology engineering tools such as Collaborative Protégé offer rich and structured logs of changes, thereby opening up new challenges and opportunities to study and analyze the creation of collaboratively constructed ontologies. While there exists a plethora of visualization tools for ontologies, they have primarily been built to visualize aspects of the final product (the ontology) and not the collaborative processes behind construction (e.g. the changes made by contributors over time). To the best of our knowledge, there exists no ontology visualization tool today that focuses primarily on visualizing the history behind collaboratively constructed ontologies. Since the ontology engineering processes can influence the quality of the final ontology, we believe that visualizing process data represents an important stepping-stone towards better understanding of managing the collaborative construction of ontologies in the future. In this *application paper*, we present a tool – *PragmatiX* – which taps into structured change logs provided by tools such as Collaborative Protégé to visualize various pragmatic aspects of collaborative ontology engineering. The tool is aimed at managers and leaders of collaborative ontology engineering projects to help them in monitoring progress, in exploring issues and problems, and in tracking quality-related issues such as overrides and coordination among contributors. The paper makes the following contributions: (i) we present *PragmatiX*, a tool for visualizing the creation process behind collaboratively constructed ontologies (ii) we illustrate the functionality and generality of the tool by applying it to structured logs of changes of two large collaborative ontology-engineering projects and (iii) we conduct a heuristic evaluation of the tool with domain experts to uncover early design challenges and opportunities for improvement. Finally, we hope that this work sparks a new line of research on visualization tools for collaborative ontology engineering projects.

Keywords: Collaborative Ontology Engineering, pragmatic analysis, ontology monitoring, ontology engineering visualization, ontology evaluation, ontology tool.

INTRODUCTION

While collaboration, negotiation, and consensus represent an integral part of ontology engineering processes, it is only recently that disciplined tools and infrastructure for *collaborative* ontology engineering have emerged. Tools such as Collaborative Protégé (Tudorache, Noy, Tu, & Musen, 2008) not only provide an infrastructure for collaboration and coordination, but also provide a structured log of all ontological changes, which users have made via the tool. These logs can, for example, include records of concepts added, properties changed, or relationships qualified. In aggregation, such logs can essentially capture the entire evolution of an ontology from its inception to its final stages on a very fine-grained level. At the same time, the availability of fine-grained logs poses new challenges and opportunities for studying and analyzing the history of collaborative ontology engineering projects. While there exists a plethora of visualization tools for ontologies, they have primarily been built to visualize aspects of the final product (the ontology) and not the collaborative processes behind construction (e.g. the changes made by contributors over time). To the best of our knowledge, there exists no ontology visualization tool today that focuses primarily on visualizing the creation processes behind collaboratively constructed ontologies.

This *application* paper sets out to present a visualization tool that primarily focuses on visualizing pragmatic aspects of collaborative ontology engineering, i.e. the social processes that yield collaboratively constructed ontologies. We present a tool – PragmatiX – that taps into structured log of changes provided by tools such as Collaborative Protégé and visualizes them via network-based and other kinds of visualizations. The tool is aimed at managers and leaders of collaborative ontology engineering projects to help them in monitoring progress, exploring issues and problems, and tracking quality-related issues such as overrides and coordination among contributors.

Our initial prototype demonstrates its capabilities by tapping into change-logs produced by variants of Collaborative Protégé, where changes and notes as well as comments on changes are represented in the Change and Annotation Ontology (ChAO) (Noy, Chugh, Liu, & Musen, 2006). Because several large collaborative ontology-engineering projects in the bio-medical domain use Collaborative Protégé (and its derivatives) for tool support, we have access to change-log data from a series of different projects. For example, the International Classification of Diseases (ICD-11) project uses WebProtégé, a Web version of Protégé that is built on the collaborative framework of Collaborative Protégé, to collaboratively engineer a bio-medical ontology consisting of more than 30,000 concepts (Tudorache, Falconer, Nyulas, Noy, & Musen, 2010). Almost all changes to this ontology have been captured and are available for further analysis. The International Classification of Traditional Medicine (ICTM) ontology represents another example, where a sufficiently large record of changes is available. In this paper, we will use data from these two projects to demonstrate the general applicability of our tool for visualizing pragmatic aspects of collaborative (ontology-) engineering projects. While the illustrations in this paper are limited to these two projects, there is nothing in our implementation, which prevents other collaborative ontology engineering projects (e.g. outside the bio-medical domain) being visualized with our tool in a similar manner, given that data about the creation process is available in a structured form (see section *Implementation*).

Our application paper makes the following contributions: (i) we present the new and extended version of iCAT Analytics, called *PragmatiX*, which allows for visualizing the creation process behind collaboratively constructed ontologies (ii) we illustrate the utility and generality of the tool by applying it to structured change-logs of two large collaborative ontology-engineering projects and (iii) we conduct a heuristic evaluation of the tool with domain experts to uncover early design challenges and opportunities for improvement. Our research is relevant for managers of collaborative ontology engineering efforts aiming to analyze and visualize the social dynamics of the development process.

This paper is structured as follows: The *Related Work* section provides a detailed overview of existing and relevant published work, which has influenced this paper. In the *Materials and Methods* section we describe the resources used by *PragmatiX*, such as various data sets and their structured logs of changes as well as the algorithms used to calculate the features displayed in the different views. In section *The PragmatiX Visualization Tool*, we describe *PragmatiX* itself in greater detail, including all the visualizations and views provided by the tool. The results of a formative evaluation of *PragmatiX* are presented in the *Evaluation* section; the benefits and limitations of *PragmatiX* are outlined in the *Discussion* section. The paper closes with concluding remarks in section *Conclusions*, where we additionally address possible future work.

RELATED WORK

The following areas of research are relevant to our work: Collaborative Ontology Engineering, Collaborative Ontology Engineering Tools, and Ontology Visualization Tools.

Collaborative Ontology-Engineering

The field of ontology engineering covers many different topics ranging from best practices for creating ontologies (Cristani & Cuel, 2005; Noy & McGuinness, 2001; Spyns, Meersman & Jarrar, 2002), identifying and implementing semi-automatic processes to create ontologies from different resources such as plain text (Maedche & Staab, 2000) to the task of ontology-evaluation (Brank, Grobelnik & Mladenic, 2005) in order to determine and quantify the quality of an ontology, for example with respect to its intended use-case.

In contrast to traditional ontology engineering, the task of collaboratively developing and engineering an ontology represents an emergent field of research with new problems, risks and challenges.

For example, Noy & Tudorache (2008) and Falconer, Tudorache, & Noy (2011) focus on identifying, defining and surveying requirements for collaborative ontology-engineering applications. Their work demonstrates that an analysis of change-logs of collaboratively engineered ontologies allows users to be grouped according to their change behavior. Pöschko et al. (2012) have shown that analyzing the structured log of changes in collaborative ontology engineering projects using *iCAT Analytics*, the predecessor to *PragmatiX*, yields interesting results, such as how work is distributed among authors or which areas of the ontology already received a large amount of contributions, which can be used to enhance the collaborative engineering process and to help encourage users to contribute. In contrast to iCAT Analytics, PragmatiX allows the import and visualization of multiple data sets in one instance. It also provides additional functionalities such as various statistical overview pages (e.g. the heat-map as described in Section *Concept Network Visualization*, or the dashboards in Section *Dashboard*). Additionally, a heuristic evaluation has been performed on *PragmatiX*, providing interesting results for future work.

Collaborative Ontology-Engineering Tools

Many collaboratively engineered ontologies, such as the Gene Ontology (GO) (Harris et al., 2004), the National Cancer Institute Thesaurus (NCIt) (Golbeck et al., 2003), the International Classification of Diseases revision 11 (ICD-11) and the International Classification of Traditional Medicine (ICTM) (Tudorache et al., 2010), are created using tools that provide special methods and functionality to help users collaborate. This special functionality often includes mechanisms to comment single concepts, to engage in discussion, and to justify changes and design decisions, all of which support collaboration among ontology editors.



Figure 1. Exemplary Screenshot of the ICD Collaborative Authoring Tool (iCAT) that is used to create and maintain ICD-11.

A large variety of ontology-engineering tools, such as OntoEdit (Sure et al., 2002), semantically extended Wikis, such as Wiki@nt (Bao & Honavar, 2004) and OntoWiki (Auer, Dietzold, & Riechert, 2006), or Collaborative Protégé and WebProtégé (Noy & Tudorache, 2008; Tudorache, Noy, et. al., 2008) provide special functionality supporting users in reaching consent and avoiding conflicting changes by actively encouraging collaboration.

Both of the collaborative ontology engineering projects that we use in this paper were developed using the web-based tools iCAT (see Figure 1) and iCAT TM, two very similar, customized versions of WebProtégé. The most important feature of WebProtégé and its derivatives for this paper is the fact that it provides a very detailed and fine-grained structured log of changes of the ontology, which can be used to analyze the creation processes in addition to the collaboratively constructed ontologies.

iCAT and iCAT TM both offer extensive collaborative features, providing authors not only with the ability to conduct collaborative work but also to engage in threaded discussions, to facilitate collaborative decision making.

Ontology Visualization Tools

The domain of ontology visualization covers a large set of applications providing various graphical representations for ontologies, which range from simple indented lists or trees, 2-dimensional graph representations to very sophisticated 3-dimensional layouts.

For example, Jambalaya (Storey et al., 2002b) was developed as a plug-in for an earlier version of Protégé. It uses a visualization technique called SHriMP (Simple Hierarchical Multi-Perspective) (Storey et al., 2002a), which supports the concept of interchangeable nested views representing an ontology in 2-dimensional space.

OntoViz (Singh et al., 2006), another Protégé visualization plug-in, on the other hand represents an ontology as a 2-dimensional graph, using the Graphviz (Ellson, Gansner, Koutsofios, North, & Woodhull, 2001) library. In OntoViz, every node represents a class or an instance, which in turn can display its name and some or all of its (inherited) properties and roles. Every edge represents a relationship between classes, instances, or both.

OWLViz (Horridge, 2012), TGViz (Alani, 2003), and OntoGraf (Falconer, 2010) are ontology visualization plug-ins for Protégé, which are similar to OntoViz. They represent an ontology as a 2-dimensional graph, where each node represents either a class or an instance and every edge represents a relationship between two entities. However, instead of visualizing all property and role information, OWLViz and TGViz reduce visual clutter by providing detailed information for each entity only when it is selected. Additionally, they allow filtering for specific parts of the ontology.

AlViz (Lanzenberger & Sampson, 2006) is a tool that was specifically designed to visualize and augment the task of ontology-alignment, i.e. mapping the classes and instances of one ontology to the classes and instances of another ontology. Classes and instances are represented as nodes, which are colored according to the result of the alignment process.

In contrast, OntoRama (Eklund, Roberts, & Green, 2002) uses a hyperbolic-type layout to visualize an ontology. This approach emphasizes classes or instances in the center of the visualization, as they are assigned more space and display a higher level of detail, while nodes near the edges of the visualization are minimized and only display a low level of detail. OntoSphere 3D (Bosca, Bonino, & Pellegrino, 2005) uses different types of 3-dimensional visualization to support users in browsing and exploring the structure and complexity of an ontology.

However, all of these ontology visualization tools have their focus on visualizing an already created ontology or (parts of) a static snapshot of an ontology. The idea and overall objective for *PragmatiX* is to create a tool, which not only visualizes an ontology, but also visualizes and analyzes the creation process.

MATERIALS AND METHODS

In this section we will first describe all identified requirements including the target audience of *PragmatiX* as well as the data sets and the structured logs of changes, which were used in our analysis and evaluation.

User Research

PragmatiX was specifically designed to augment; support and help to enrich and enhance the work performed by different user-types and classes (Schreiber et al., 2000) of knowledge-based system development processes. We have grouped these roles into the following three groups: The Administrative Personnel, which is composed of *Knowledge Managers* and *Project Managers*. The Engineering Staff, which in turn is composed of *Knowledge Engineers* and *Analysts, Knowledge System Developers* and *Knowledge Providers*. The final group of users is called Ontology Viewers. They consist of *Knowledge Users* and *System Visitors*. These user-

types differ with regard to their informational needs and overall goals and objectives, for example Knowledge Users and System Visitors are mainly interested in gathering a quick characterization of the data set or the tool. The Engineering Staff is concerned about the correctness of the underlying data, while Administrative Personnel wants to track the progress of the project.

For the purpose of identifying informational needs as well as the goals and objectives of different user types, meetings with members of the Protégé-Team at Stanford Center for Biomedical Informatics Research and with Domain Experts from the World Health Organization (WHO) were held. In those meetings, requirements for the tool were elicited and discussed with the stakeholders (consisting of the team that develops Protégé and the team that is in charge of ICD-11 development) in an iterative manner.

Implementation

The majority of *PragmatiX* was written in Python using the web-framework Django. We use NetworkX (Hagberg, Schult, & Swart, 2008) for all network calculations and use Graphviz (Ellson et al., 2001) to pre-calculate the different network layouts (visualizations). The data sets were exported from iCAT and iCAT TM using their Java API and stored in a MySQL database.

To visualize the different network views we make use of a combination of JavaScript, AJAX calls, and JSON. The asynchronous JavaScript and XML calls are necessary to update the graph after user interactions. Most of the visualizations and analyses available in *PragmatiX* are pre-calculated, resulting in reasonable response times and relatively low server load. This is especially useful for the network visualizations, where all positions are pre-calculated using Graphviz, and stored in the database. This approach minimizes calculation and loading times, since all required information can be extracted directly from the database without having to invoke additional computational tasks on either the client or the server. Additionally, most values displayed in a pie or line chart are pre-calculated.

PragmatiX additionally provides mechanisms to import data in a specific input format (basic txt-files), where each line corresponds to one concept followed by a very limited set of attributes extracted directly from the ontology, separated by tab-stops. This set of attributes consists of (i) a unique concept id, (ii) a concept title or username, (iii) a concept definition or change message, (iv) the assigned communities of interest (user groups; only if available) of the concept and (v) a concept's assigned display status color code, depending on the type of concept.

The ChAO change-log provides detailed change-information such as the user who performed a change, the concept it is performed on and a detailed change-description such as "*Moved class: R75.2b Niemann-Pick disease. Old parent: E75.2 Other sphingolipidosis, New parent: Sphingolipidosis*", allowing to omit the storage of additional attributes (other than the ones mentioned), as they can be automatically generated for each point in time by processing the change-descriptions of the structured logs of changes. Not having to know which properties are available for which concept additionally increases the generality of *PragmatiX*. All additional attributes (such as the features displayed in Table 2) are calculated after the initial import.

The used input format for *PragmatiX* can easily be reproduced independent from the original source, given a structured log of changes is provided which can be mapped onto the ontology. Using this convention, we were able to import and visualize all articles from the official (and freely available) Wikipedia change-data dumps, which are marked with an ICD-10 code, into *PragmatiX*. The extracted articles have been mapped to concepts while contributors in Wikipedia

represent users in *PragmatiX*. Relationships (edges) have been extracted from the original ICD-10 ontology.

Once the data is available as textual files in the required format, PragmatiX provides SQL scripts that import the content into the database. In a next step, the pre-calculations, a python script provided by PragmatiX, have to be initiated, which are then automatically performed. Once the pre-calculations are done, the new data set has to be added to the configuration-file and is afterwards ready to be browsed. We are currently working on refining the import process to automate all necessary steps and provide a step-by-step guide for all steps that cannot be automated.

Once PragmatiX reaches a stable version, we will consider releasing it as Open Source Software.

Application

PragmatiX focuses on visualizing the creation process behind collaborative ontologyengineering projects that provide (i) structural and contextual information about the ontology and (ii) a structured log of changes (and notes) that allows mapping every logged action to a specific user and the affected concept(s).

We have applied *PragmatiX* to five different collaboratively engineered ontologies from the bio-medical domain. Due to limitations in space, in this paper we will demonstrate the application of *PragmatiX* to only two of the five available bio-medical ontologies, which were both constructed using variations of WebProtégé. The two projects are:

- *ICD-11*: The structured log of changes comprises *152,955* changes and *31,197* notes over an observation period of 24 months. The ontology itself consists of *33,714* concepts and *76* users that performed all the changes.
- *ICTM*: This data set is of a smaller dimension and only consists of *1,311* concepts with a total of *21* users that actively work on the ontology. The change log consists of *39,495* changes and *1,449* notes over an observation period of 10 months.

Even though both data sets are maintained by WHO and have been created using either Protégé or one of its derivatives it is important to note that *PragmatiX* can be adapted to support every collaborative ontology engineering project which exhibits a structured log of changes.

Evaluation

Formative usability evaluation is usually performed during interface development, in order to identify potential problems to be fixed in future releases. Two classic methods of formative evaluation are widely used in software development: Heuristic Evaluation (HE) (Nielsen & Mack, 1994) and Thinking Aloud (TA) (Barnum, 2010) testing. The former involves a small group of specialist evaluators who inspect an interface and use a list of heuristics, combined with their knowledge and experience to identify and classify potential problems.

The latter involves a small number of representative test users from the target user population, who talk out loud whilst performing representative tasks, thus providing insight into their thought process when problems occur. Summative evaluations (Rubin, Chisnell, & Spool, 2008) involve the objective measurement of performance metrics and statistical analysis and are often used to compare alternative designs or competing products.

Because of the early stage of development, we have limited our evaluation efforts, and concentrated on conducting a Heuristic Evaluation on *PragmatiX* with three ontology-

engineering experts who investigated and explored our tool in sessions of 60-120 minutes. We have gathered feedback about the utility and problems assigned to *PragmatiX*, which we will further discuss in our section *Evaluation*.

THE PragmatiX VISUALIZATION TOOL

PragmatiX represents an evolution of the iCAT Analytics tool (Pöschko et al., 2012) and goes beyond iCAT Analytics by (i) being applicable to collaboratively engineered ontologies in general (and not specific to a particular ontology) and (ii) by adding several new views and visualizations to its repertoire. PragmatiX provides several different ways to interact with the analyzed data sets, which will be described in this section in greater detail. Users can perform exploratory analyses using different kinds of visualizations including **three network visualizations, ranked overviews** and **detailed statistics views** for all concepts and users. To further accommodate the needs of *Administrative Personnel*, we extended the tool by implementing a *dashboard*, which lists rather general statistics that can be used to interpret and monitor the progress of the underlying ontology engineering process.

Tool Overview

PragmatiX provides several different views that allow for different types of interaction with the imported data sets (see Figure 2). These different views and network visualizations are listed in Figure 2 and consist of:

- The **concept network visualization** hierarchically visualizes the concepts of an ontology via *is-a relations* and simultaneously allows to further visually inspect/explore conceptual features, such as the number of changes performed on every concept of the ontology.
- The **author network visualization** visualizes the relationships across users by identifying and quantifying commonly edited concepts (or collaboration).
- The **property network visualization** displays properties of concepts, and their *pragmatic* relationships with each other (e.g. what property was edited after what other property?)
- The **dashboard & community views** are used to visualize and list general statistics which support *Administrative Personnel* in monitoring the progress of the engineering process.
- The statistics overviews feature rankings of all concepts, authors and properties according to several different pre-calculated features.
- The **detailed statistics views** provide detailed information about the change-history of a single concept or a single author.



Figure 2. An overview of all statistics views and visualizations provided by PragmatiX. The tool features three network visualizations (Author, Concept and Property Network), different types of dashboards and overviews as well as very detailed statistics views. The data displayed in this figure is from the ICD-11 data set. The nodes in the network views either represent an author, a concept or a property. While edges in the concept network visualization depict is-a relationships, they vary in the author and properties network visualization according to the currently selected features.

Concept Network Visualization

This network is used to visualize change data specific to hierarchical structures, relations and to the complexity of and between the concepts of an ontology. Every node represents a single concept in the ontology. Every edge represents *is-a* relationships between concepts. The color assigned to each concept represents the *display status* of the concept – a property of the ICD-11 ontology, which can be used to represent the current development state or progress of a concept. We have adopted the color codes from *iCAT* and visualize them if they are supported by the

underlying imported collaboratively engineered data set. In the case of ICD-11 the colors have the following meanings:

- Gray: no display status assigned
- Red: the concept requires extensive work
- Yellow: the concept is worked on, but is not finished
- Blue: the concept is ready for subsequent phases

These color codes are usually assigned by managers or leaders of collaborative ontology engineering projects, but could also be assigned by other roles (e.g. editors) through other mechanisms (e.g. voting). The color codes provide a quick overview of the current state of the ontology, which is especially important for collaboratively engineered ontologies as they can help to minimize the difficulties of identifying concepts or areas of an ontology that still need work, without having to allocate additional resources for that task.

Additionally the concept network visualization allows users to decide what determines the diameter of each node by selecting from a set of conceptual features (see Table 1) that are used as weights. This allows *PragmatiX* to help answer a series of questions about the creation processes behind a particular collaboratively engineered ontology.

Table 1. A list of features, used to weight nodes in the concept network. Per feature, we highlight corresponding questions that could be tackled with this kind of visualization. Additionally, a list of author specific concept network features, used to weight nodes in author specific network visualizations, and their corresponding research questions are provided.

Feature	Question addressed
Concept Network	
Changes and Notes History	
Number of changes and/or notes	Which are the highly edited/discussed areas in the ontology?
Changes and notes	Which are the highly active areas in the ontology?
Distinct authors of changes/notes	Which concepts attract many different authors?
Authors Gini coefficient	Which concepts are edited more "democratically", i.e.,
	in a more evenly distributed manner? Contrarily, which
	are the areas/concepts that are dominated by many
	changes of a single author?
Overrides	Which concepts cause most disputes (i.e. have the
	highest amount of changes performed on the same properties of a concept)?
Edit sessions	Which are the highly active areas (with aggregated
	consecutive changes of the same property by the same
	author being 1 edit session)?
Distinct authors by property	Which concepts have many properties that are edited by many different authors?

Network Features

Number of parents/children	Which concepts have many parents? (This is particularly interesting in the case of ICD-11, as multiple parents were not possible in ICD-10 and are therefore introduced gradually.) Which concepts have many children? (i.e. Number of parents/children in the
	ontological structure)
Depth in network	Which concepts are at what levels in the ontological structure (i.e. what is the shortest path of each concept to the root concept)?
Betweenness centrality	What are central/popular concepts in the ontology, when
(directed/undirected), Pagerank,	looking at different attributes of the network structure of
Closeness centrality	the ontology?
Number of changes by community	How many changes of a concept did each community performed?
Number of titles/definitions and	How many different titles, definitions or language codes
language codes	are available for each concept?
Author Specific Concept Network	
Changes and Notes History	
Number of changes and/or notes	What concepts or areas of the ontology did a user either edit or comment frequently? What concepts or areas of the ontology did a user both , edit and annotate frequently?

In addition, *PragmatiX* allows limiting the visualization of the concept network to only display the set of concepts a specific user has edited, weighted according to the features listed in Table 1. The network itself is created analogously to all other network visualizations, resulting in an empty network if the selected feature renders an empty set of concepts (e.g., a user has not made any edits yet).

The author specific concept networks can be used to analyze a variety of different aspects related to user behavior, such as the role of a specific user (i.e. generalists vs. specialists, see Figure 3) during the engineering process or concepts, topics and areas of interest of specific users.



Figure 3. Two author specific concept network visualizations taken from the ICD-11 data set with the selected feature "Number of changes". The left author-specific concept network indicates a user who seems to be more of a generalist (editing a vast variety of concepts in the ontology) while the right concept network indicates a user who seems to be more of a specialist (editing a more narrow area of the ontology).

In addition to these features and visualizations, we have also implemented a heat-map (see Figure 4), which allows users to visually monitor and track activity within the ontology. The heat-map can be used in all concept networks (including all user specific concept networks) combined with every feature listed in Table 1.



Figure 4. PragmatiX showing the concept network visualization of ICD-11 as a heat-map and the selected feature "Number of changes". The warmer the node colors the more recent the last edit has been performed.

The *Administrative Personnel* can use the concept network visualization to track activity and progress as well as to identify domain specialists. The *Engineering Staff* can use the tool to identify parts of the ontology that are (or are not) very active to adapt the engineering process or the underlying knowledge representation. *Ontology Viewers* can use the visual representation to explore the complexity of the ontology and to identify areas of community interest according to different pragmatic features (e.g. number of edits).

Author Network Visualization

In addition to the user specific concept network, *PragmatiX* provides a visualization of all authors, displaying the extent of collaboration (see Figure 2) they engaged in during the engineering process. The following two features can be selected:

- 1. *Commonly edited categories (collaboration)* shows a network of authors, connected by weighted edges according to the number of commonly changed or commented on concepts. The node size represents the total number of changes performed by each author.
- 2. *Overrides* shows a network of authors, connected by weighted edges according to the number of changes by one author that were overridden by another author. The node size represents the fraction of all changes performed by the author that were overridden by other authors.

The author network supports *Administrative Personnel* to measure if and to what extent authors engage in collaboration and perform overrides. Additionally it can help members of the *Engineering Staff* to explicitly identify the "*importance*" of a user (e.g. for the collaboration graph, according to the connectivity and weights of the edges, indicating very active and collaborative users). *Ontology Viewers* will use the author network to explore the complexity of social interactions and might be interested to compare the extent of collaboration with the amount of performed overrides to see whether the project is led in a more or less democratic way and who is responsible for keeping the order and can be contacted in case problems arise.

Property Network Visualization

In collaborative ontology engineering projects, it is interesting to study the pragmatic relationships between different properties, to show – for example - what properties are edited first, or in what sequence properties are edited (where a property refers to a property of the ontology, i.e. a data type, object or annotation property). This could allow for identifying patterns of property editing behavior, which could have implications for the design of more effective user interfaces. To that end, we calculated and visualized the property network. Each node corresponds to a property and every edge represents the number of changes on a property followed by a change on a different property.

The relationships between properties could be of great interest for the *Engineering Staff* and explicitly for *Knowledge Engineers* to enhance and adapt their knowledge representation and tools to better fit the natural working process of its users by closely grouping properties that exhibit highly weighted edges.

Network Visualization Implementation Details

The nodes, which represent either concepts or users, and edges, which represent either ontological relations or collaboration and overrides, in the network visualizations are weighted according to a set of independent features (see Table 1 for more details) and are visualized by adjusting the size and/or color of the nodes and edges. Currently, *PragmatiX* features the following network visualization layouts, both generated using Graphviz (Ellson et al., 2001):

- 1. *twopi* (radial)
- 2. *sfdp* (multi-scale force-directed "spring model")

The radial layout allows for a clear visualization of ontological or hierarchical structures similar to that of taxonomies or trees. The force-directed layout, on the other hand, is better suited to visualize highly interlinked ontological structures and networks. Due to the fact that all layouts are pre-calculated (i.e. the x- and y-positions for each concept are stored in the database for each layout), *PragmatiX* can be easily extended to support any given 2-dimensional layout algorithm and does not necessarily depend on Graphviz.

To navigate the graphical representations users can either use the arrow-keys on their keyboard combined with the graphical user interface that allows for easy zooming and jumping back to the center of the network, or they can explore the network by common drag-n-drop principles and adjust the zoom level by using the mouse wheel.

For reasons of usability and to avoid visual cluttering, *PragmatiX* only displays a specific fraction of nodes for large data sets, rather than the whole network at once. This is aimed to enable users in identifying and exploring top (i.e. the most interesting) concepts regarding specific features and attributes rather than analyzing the layout of an imported data set in general.

To that end, *PragmatiX* knows the coordinates of the user view's bounding box and selects the corresponding part of the network to display. To determine which nodes are displayed to avoid visual cluttering in large data sets, we have implemented an intelligent filtering algorithm, which divides the bounding box (or field of view) into 10x10 raster boxes where each box displays the node with the highest weight of the currently selected feature within its boundaries. To avoid disconnecting components that are physically connected, all nodes from any selected node and all edges from any selected node to the root node are displayed as well if available, thus forcing the network to stay connected.

Dashboard

The *dashboard* (see Figure 5) was created to provide something similar to an "overview" page that provides overall and generalized statistics about the whole collaborative ontology-engineering project.

Figure 5a shows the changes and notes distribution over time as a line chart and visualizes the aggregated amount of performed notes and changes in the ontology over time. The distribution of changes across users (Figure 5b) is represented as a pie chart. The basic statistics table (Figure 5c) is a textual representation to quantify the size of the ontology, its users and their performed changes and annotations. Additionally, the category display status statistics (Figure 5d), provides additional information about the amount of concepts (and their average number of changes) with the corresponding assigned display states.

The Community Statistics pie charts (Figure 5d) provide information about the percentage of changes performed by each user group on their corresponding concepts. For example, as can be seen in Figure 5, the corresponding assigned primary community performed 25.99% of all changes across all concepts.



Figure 5. A screenshot of the dashboard, providing analytical information for the complete ICD-11 engineering process, divided into four elements. Figure 5a shows the total amount of notes and changes performed over time. Figure 5b shows the amount of changes performed by the Top 15 contributors. The Basic Statistics table (Figure 5c) can be used to quantify the size of the ontology, its users and their performed changes and annotations. The Concept and Community Statistics (Figure 5d) provide additional details about the distribution of changes across assigned communities and the overall progress of the project.

The *dashboard* was specifically designed to fit the requirements of the *Administrative Personnel* to provide a quick overview of the current progress of the ontology. Additionally it supports the *Engineering Staff* to identify the distribution of edits across time and users. *Ontology Viewers* might be interested in parts of the *dashboard*, depending on their personal motivation towards *PragmatiX* and the imported data set.

Community Views

If an ontology exhibits different Communities of Interest (or user groups), *PragmatiX* provides "smaller *dashboards*" for each community, called "*community views*". They are analogously designed to the dashboard (see Figure 5) and provide the same graphical visualizations, displaying only relevant data for each community.

In addition to the dashboard, each community view displays textual information of the amount of changes and notes performed on the concepts, which are assigned to the community as well as the total number of changes and notes performed on these concepts by all authors in the ontology.

Community views are specifically designed to meet the requirements of *Project Managers* and their assigned areas of the ontology. Similar to the dashboard, both *Administrative Personnel* and *Ontology Viewers* might be interested in specific community views, depending on their current motivation or tasks. In the case of the ICD-11, a community of interest is usually referred to as a Topic Advisory Group (TAG).

Category and Author Views

In addition to the network visualizations, *PragmatiX* provides overviews and very detailed statistics views for all authors and categories in general and every single author and category in detail. In these overviews, we rank all concepts and authors according to our implemented features (see Table 1 for a list of features). This allows users to quickly identify the top (and worst) concepts (see Figure 6) or authors for every feature without having to browse the graphical network visualizations.

The detailed concept and author statistics views can be reached by either clicking on a node in the corresponding network visualizations or by following the links (as displayed in Figure 6) on the ranked overviews. These links are represented by the title of each concept and in the case of ICTM, which provides multiple languages for each concept; all available title-translations for the corresponding concepts are displayed in the listings of Figure 6.

PragmatiX 🔀 Search 🔤 🖬 🛄	Dashboard	Categories Authors	Properties
Number of changes			
3			
Category	Value	2	
'37.1.1.3' {'Soeum-type Exterior pattern NEC ', '少陰人 表病證, 他處未分類 ', '달리 분류되지 않 음인 표명증'}	¦은 소 257	7	
ICTM Categories'	183	3	
'37.3.2.3' {'Taeeum-type Interior Pattern NEC ', '太陰人 裏病證, 他處未分類 ', '달리 분류되지 8 태음인 리병증'}	¥은 141	1	
'37.2.1.1' { Soyang-type Lesser-Yang Wind-Damage pattern ', '少陽人 少陽傷風證 ', '소양인 소 풍증'}	양상 133 -	3	
'37.1.2.3' {'Soeum-type Interior pattern NEC ', '少陰人 裏病證, 他處未分類 ', '달리 분류되지 않 음인 리병증'}	은 소 124	4	
'01.01.100' {'Palpitation disorders (TM)', 心悸, 心悸}	121	1	
'37.2.2.3' {'Soyang-type Interior Pattern NEC ', '少陽人 裏病證, 他處未分類 ', '달리 분류되지 않 양인 리병증'}	:은 소 112	2	
- <mark>'37.2.2.1' {'Soyang-type Chest Heat-congested pattern ',</mark> '少陽人 胸隔熱證 ', '소양인 흉격열증']	} 111	1	
'22.04.420' {태양축혈증, 太阳蓄血证, 'Greater yang blood accumulation pattern (TM)'}	108	5	
'20.01.510' {'血虚动风证 [血虚生风证]', 철허생풍증, 'Blood deficiency associated wind pattern '血虛動風証 [血虛生風証]'}	(TM)', 100	0	
'27 01 002' {'between excess and deficient abdomen' 腹力中等度}	(0	
'27.01.016' {seichushin. 正中芯}	(0	
'27.01.014' {'shofukufujin OR lower abdominal numbness', 小腹不仁}	(0	
Show all			
Number of notes			
Category	Value	2	
Z 'Discussion about iCAT'	8	8	
'Traditional medicine conditions originating from Chinese medicine'	6	6	
{Ecthyma, 賺疮}	3	3	
'Blood deficiency Pattern, 血虚証'	3	3	
'1-29' {'Traditional Medicine Disorders (TM)', 传统医学疾病, 傳統醫學疾病}	3	3	
'90-99' 'Other Concepts in Traditional Medicine'	3	3	
{Scrofula, 漯疬}	3	3	
'00' 'Constructs of the Nervous System'	3	3	
15' Mental and Behavioral Constructs TM'		3	
'U1.U1.32U {'Fluid collapse disorder (TM)', 液脱, 液肌}	3	3	
'04.21' {'Allergic rhinitis disorder (TM)', 鼻劍, 鼻鼽}	(D	
1021 (17anafu Luna System) 肺系 肺系)	(n	

Figure 6. An excerpt of the concept overview for ICTM showing the Top 10 (and worst 3) concepts according to the features "Number of Changes" and "Number of notes" as clickable links. As ICTM provides multiple languages for each concept title and definition, all available titles for the corresponding concepts are displayed in the listings of Figure 6.

0

'32' {'Zangfu and Meridians or Channels Patterns (TM)', 脏腑和经络证, 장부병증, 臟腑経絡証(鍼灸) }

The detailed concept statistics views (see Figure 7) provide further information about the parents and children of a concept, the change and note history of a concept, the group dynamics (e.g. Who contributed what amount of edits or notes when?) and a table that lists all feature-values used in the concept network visualization. Due to reasons of space, the table containing the pre-calculated feature-values has been omitted in Figure 7.

The detailed author statistics views are similar to the detailed concept statistics views and provide information about the amount of changes and notes contributed by a specific user, the

communities a user is member of, concept-recommendations (Walk et al., 2012) the user might be interested to change as well as co-authors and overrides performed on a user.



Figure 7. Shows the detailed concept statistics view for the concept "22.04.420 'Greater yang blood accumulation pattern (TM)'" of the ICTM data set. Figure 7a shows the change and note distribution on this concept over time. The distribution of changes across users for the displayed concept can be seen in Figure 7b. A list of assigned communities is displayed in Figure 7c (empty). Parents of the currently browsed concept are listed in Figure 7d. As ICTM features a set of different languages, the various titles and definitions for each concept are depicted in Figures 7e and 7f.

The detailed statistics views for each concept and every author provide useful information for all members of the *Engineering Staff*. Project Managers and Knowledge Engineers can reconstruct the change history of a concept and identify the corresponding most influential/active users using the detailed statistics concept views. Knowledge Providers can, for example, use

their own detailed statistics author view to receive suggestions for concepts to edit. Both concept and author overviews are used by *Administrative Personnel* and the *Engineering Staff* to quickly identify users or concepts of interest, according to the implemented and ranked by features.

EVALUATION

PragmatiX was evaluated in May 2012 using the Heuristic Evaluation (HE) method. Three ontology engineering experts, who are all experienced in the task of engineering and gardening ontologies, thus represent a fraction of the actual target group of our tool, acted as evaluators and explored, tested and investigated the interface in sessions from 60 to 120 minutes. All three evaluators, without any direct request, assumed (multiple) different user roles during the evaluation task and included these perspectives in the provided feedback. The evaluation uncovered a total of 27 usability issues, ranging from rather simple problems, such as a misleading icon to display a legend while browsing a network visualization, to more serious usability issues, such as a confusing name for the concept of TAGs (Topic Advisory Groups, introduced by the WHO for ICD-11) which are better described as Communities of Interest. All 27 issues were classified according to a modified version of Nielsen's 10 Heuristics (Nielsen, 1994) called the "Andrews General Usability Heuristics", which are more concise and include small explanations or examples as clarification for each heuristic, which aid the evaluators in classifying identified problems during their evaluation task. We make the full HE Report available for download as Walk & Andrews (2012). A short excerpt of identified usability issues, manually filtered by significance and ranked according to their severity, is listed in Table 2.

Short Title	Description
Interpretation of results	It is very hard to interpret the currently browsed features (see Table 1) of the network visualizations as no information about the actual meaning (or interpretation) of chosen features is provided.
Details of initial concept network visualization unclear	Directly after login - when being confronted with the initial ontology concept network visualization - it is unclear what parts of the ontology are represented as nodes, edges, colors and diameter of the nodes.
Unreadable hover text	When the mouse is hovering over a node in the graphical representation, the hover text is unreadable if the hovered node has many children and parents.
Wrong vocabulary for Audience	Ontology experts might not be familiar with network analysis measures and vocabularies.
Can't move/drag nodes	The network interface should support "drag & drop" for nodes, as there are bigger nodes that sometimes conceal smaller underlying

Table 2. An excerpt of usability issues identified in the Heuristic Evaluation, manually filtered by significance. The issues are ranked according to the average severity assigned by the three evaluators

nodes.

All three evaluators, who have expertise and experience in developing and working with ontologies stated, that they were confused about the initial concept network visualization they were presented immediately after login. According to the feedback gathered during the HE, the confusion mainly arose as the evaluators never specified any features or relationships prior to the login and could not link the displayed information with the visualization. Additionally all information regarding the chosen data set, the selected features and the explanations of these are hidden within the interface of *PragmatiX*. The evaluators mentioned, that they are missing explicit information about the currently displayed visualization, which can help them to better understand what they are currently looking at. One evaluator specifically stated, that she is missing information on which nodes are currently displayed.

This leads to another very interesting observation. The evaluators had problems to really understand and interpret the different features. It was not immediately clear why it could be of interest to explore and visualize the number of changes performed on each concept or their number of distinct authors.

According to the evaluators this is mainly due to the lack of descriptions for the implemented features, their very unspecific presentation (i.e. describing the drop-down box to select the node-size features with "Feature", rather than "Feature to define node-size:") combined with an extensive usage of network theory vocabulary, which ontology experts might not be familiar with. One evaluator suggested, that we should use descriptive text-snippets rather than the actual name of the implemented measures.

This also correlates with the problems the evaluators faced, when trying to interpret the meaning of the visualizations after selecting different features. Not only was it unclear to the evaluators, what property of which element is influenced by selecting different features, but also how to make sense of the visualizations of these features.

In a few cases bigger nodes concealed smaller nodes or the displayed additional information when hovering a node was unreadable due to overlapping neighboring nodes, which actually amplified the problem of interpretation and is a direct result of pre-calculated graph layouts, which do not accommodate different node sizes.

As a possible solution to the problem of interpreting the visualizations, we were asked to provide additional textual information on the meaning of the currently selected feature and maybe some smaller examples to explain how to interpret the measure.

The majority of all uncovered problems are related to providing more and additional information about the implementation and meaning of various attributes of *PragmatiX*. All three evaluators mentioned the high utility of the implemented heat-map, the dashboard and community views.

It is noteworthy that, in addition to usability issues and positive findings, two evaluators also explicitly tried to locate specific features, which were not yet part of the system. This led to the introduction of a third category of findings called *"feature requests"*. One of these feature requests was, for example, the implementation of a timeline graph that would support browsing the state of the collaboratively engineered ontology and its engineering process at different points in time.

Additionally, it is important to note that *PragmatiX* is the first version of the tool, which was properly evaluated, thus we only have informal information about issues and complaints

regarding its predecessor *iCAT Analytics*, which are all related to missing features. Nonetheless, the biggest complaints mentioned by users of *iCAT Analytics* were the limitation of the tool to only support one data set per instance and the lack of group statistics, which were especially interesting to members of WHO.

Further improvements and formative evaluations are anticipated, and at some point in the future a summative study with end users from WHO is planned. However, the involvement of end users from WHO requires extensive planning and coordination, which is why further evaluations and the summative study are subject of future work.

In general, the visualization of very large networks is a very hard task, especially when it is desired to use the output for explorative analysis. The performed HE suggests that the currently implemented interface is useful for explorative analysis by browsing the visualizations, however new approaches and support have to be implemented for helping users in interpreting the visualized results.

In general, the feedback received during the HE was very valuable as it helped to uncover multiple design flaws that potentially confuse our target user-groups and aggravate dealing not only with *PragmatiX* but also collaboratively engineered ontologies in general. On the other hand, specific advice was gathered to derive potential approaches that can help to solve the identified usability flaws (i.e. to further explain the colors in all charts or to better describe the measures throughout the tool instead of only providing the name of the measure itself). However, it should be noted that an HE does not guarantee to identify all usability issues and is not designed to provide solutions to positively identified issues.

DISCUSSION

PragmatiX aims to visualize pragmatic aspects of the creation processes behind collaboratively engineered ontologies. We argue that the three implemented network visualization views are useful for a variety of exploratory tasks. For example, the concept network visualization can be used to monitor progress, identify generalists and specialists, to detect areas and concepts of high (and recent/past) activity. The author network visualization can be used to measure if and to what extent collaboration and overrides exist in the project and which authors collaborate with, or override other authors the most. The properties network visualization can provide insights into the creation process, which in turn can help to enhance the engineering tools for example by grouping properties that are frequently and successively changed. Our heuristic evaluation has demonstrated that – in principle – the *PragmatiX* tool can serve these purposes. It found that the majority of problems assigned to *PragmatiX* are related to insufficient descriptive textual information of implemented features. As a result, evaluators confused or misinterpreted vocabulary used in *PragmatiX*, which is mostly taken from the domain of network analysis. *PragmatiX* analyzes and visualizes the edit-, and contribution-behavior of all users that have contributed to a project and as a result are also named in the inspected change-logs, which can be a great benefit for management. Nonetheless, privacy poses an open issue and has to be addressed by the corresponding Ontology Administrators. One possible solution to this problem could be to obscure the change-logs by replacing names with consistent acronyms prior to importing them into PragmatiX. In the long run, identifying more profound, automatic and secure approaches to protect contributors privacy poses a very important subject of future work.

CONCLUSIONS

In this application paper, we have presented *PragmatiX* - a tool for visualizing the construction processes behind collaborative ontology engineering projects. Our main motivation for the development of this new tool was (i) an interest in making the otherwise hidden social processes and dynamics behind collaborative ontology engineering more visible and amenable to analysis and (ii) a lack of currently available visualization tools for that purpose. We have presented and preliminarily evaluated the main functionality and features of *PragmatiX*, and we have demonstrated its general applicability by using it for visualizing pragmatic aspects of five collaborative ontology-engineering projects. However due to limitations in space, only two projects are presented in this paper.

We aimed to demonstrate that *PragmatiX* is a promising tool to visualize and analyze the pragmatic processes behind large collaborative ontology engineering projects. Future work on *PragmatiX* will likely focus on usability issues and feature requests gathered during the heuristic evaluation. We also anticipate including additional graph or network layouts that potentially provide additional insights into the social fabric, which would help to identify cliques or other groups of collaborators. Finally, we hope that this work sparks a new line of research on visualization tools for analyzing the processes behind collaborative ontology engineering projects.

REFERENCES

Alani, H. (2003). TGVizTab: An Ontology Visualisation Extension for Protégé. In *Knowledge capture 03 - workshop on visualizing information in knowledge engineering* (pp. 2-7). Sanibel Island, FL: ACM.

Auer, S., Dietzold, S., & Riechert, T. (2006). OntoWiki - A Tool for Social, Semantic Collaboration. In I. F. Cruz et al. (Eds.), *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006, Proceedings* (Vol. 4273, pp. 736-749). Berlin / Heidelberg: Springer.

Bao, J., & Honavar, V. (2004, October). Collaborative ontology building with wiki@nt - a multiagent based ontology building environment. In *Proceedings of the 3rd International Workshop on Evaluation of Ontology-Based Tools* (pp. 1-10).

Barnum, C. M. (2010). Usability *Testing Essentials: Ready, Set... Test!* Burlington, MA: Morgan Kaufmann Publishers.

Bosca, A., Bonino, D., & Pellegrino, P. (2005). Ontosphere: More than a 3D Ontology Visualization Tool. In P. Bouquet & G. Tummarello (Eds.), *Swap* (Vol. 166).

Brank, J., Grobelnik, M., & Mladenic, D. (2005). A survey of ontology evaluation techniques. In *Proceedings of the Conference on Data Mining and Data Warehouses* (pp. 166-170). Lubljana, Slovenia

Cristani, M., & Cuel, R. (2005). A survey on ontology creation methodologies. *International Journal on Semantic Web & Information Systems*, 1(2), 49-69.

Eklund, P. W., Roberts, N., & Green, S. P. (2002). Ontorama: Browsing an RDF ontology using a hyperbolic-like browser. In *The first international symposium on cyberworlds* (p. 405-411). IEEE press.

Ellson, J., Gansner, E. R., Koutsofios, E., North, S. C., & Woodhull, G. (2001). Graphviz - open source graph drawing tools. In *Graph drawing* (p. 483-484).

Falconer, S. (12.04.2010). OntoGraf. In *Protégé Wiki*. Retrieved October 17, 2012, from http://protegewiki.stanford.edu/wiki/OntoGraf.

Falconer, S. M., Tudorache, T., & Noy, N. F. (2011). An Analysis of Collaborative Patterns in Large-Scale Ontology Development Projects. In M. A. Musen & Oscar Corcho (Eds.), *Proceedings of the sixth international conference on knowledge capture* (pp. 25-32). New York, NY: ACM.

Golbeck, J., Fragoso, G., Hartel, F.W., Hendler, J. A., Oberthaler, J., & Parsia, B. (2003). The National Cancer Institute's Thesaurus and Ontology. *Journal of Web Semantics*, 1 (1), 75-80.

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th python in science conference* (p. 11 - 15). Pasadena, CA USA.

Harris, M. A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., & et al. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Research*, 32 (Database Issue), 258-261.

Horridge, M. (2012), OWLViz – A visualization plugin for the Protégé OWL Plugin, Retrieved June 02, 2012, from <u>http://www.co-ode.org/downloads/owlviz/OWLVizGuide.pdf</u>

Lanzenberger, M., & Sampson, J. (2006). Alviz - A Tool for Visual Ontology Alignment. In *Proceedings of the conference on Information Visualization* (p. 430-440). Washington, DC, USA:IEEE Computer Society.

Maedche, A., & Staab, S. (2000). Semi-automatic engineering of ontologies from text. In *Proceedings of the 12th Internal Conference on Software and Knowledge Engineering*. Chicago: KSI.

Nielsen, J. (1994). Enhancing the Explanatory Power of Usability Heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence* (pp. 152-158). New York, NY, USA:ACM.

Nielsen, J., & Mack, R. L. (1994). Usability Inspection Methods. New York, NY: John Wiley & Sons

Noy, N., Chugh, A., Liu, W., & Musen, M. E. (2006). A framework for ontology evolution in collaborative environments. *The Semantic Web - ISWC 2006*, 544-558.

Noy, N., McGuinness, D., et al. (2001). Ontology development 101: A guide to creating your first ontology. Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880. Retrieved from http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstrac

Noy, N. F., & Tudorache, T. (2008). Collaborative ontology development on the (semantic) web. *Nature Biotechnology*.

Pöschko J., Strohmaier M., Tudorache T., Noy N. F., & Musen M. A. (2012). Pragmatic analysis of crowd-based knowledge production systems with iCAT Analytics: Visualizing changes to the ICD-11 ontology. In *Proceedings of the AAAI Spring Symposium 2012: Wisdom of the Crowd*.

Rubin, J., Chisnell, D., & Spool, J. (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* (2. ed.). Indianapolis, IN: John Wiley & Sons, Inc.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W., & Wielinga, B. (2000). *Knowledge Engineering and Management: The CommonKADS Methodology*. Cambridge, MA: MIT Press.

Singh, G., Prabhakar, T., Chatterjee, J., Patil, V., Ninomiya, S., et al. (2006). Ontoviz: Visualizing Ontologies and Thesauri using Layout Algorithms. In *Afita 2006: The fifth international conference of the asian federation for information technology in agriculture, jn tata auditorium, indian institute of science campus, Bangalore, India, 9-11 November, 2006.* (pp. 709-719). Spyns, P., Meersman, R., & Jarrar, M. (2002, dec,). Data Modelling versus Ontology Engineering. *SIGMOD Rec.*, 31 (4), 12-17.

Storey, M.-A., Best, C., Michaud, J., Rayside, D., Litoiu, M., & Musen, M. (2002a). Shrimp Views: An Interactive Environment for Information Visualization and Navigation. In *Chi '02 Extended Abstracts on Human Factors in Computing Systems* (pp. 520-521). New York, NY, USA: ACM.

Storey, M.-A., Noy, N. F., Musen, M., Best, C., Fergerson, R., & Ernst, N. (2002b). Jambalaya: an interactive environment for exploring ontologies. In *Proceedings of the 7th International Conference on Intelligent User Interfaces* (pp. 239-239). New York, NY, USA: ACM.

Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R., & Wenke, D. (2002). OntoEdit: Collaborative Ontology Development for the Semantic Web. In I. Horrocks & J. A. Hendler (Eds.), *International Semantic Web Conference* (Vol. 2342, p. 221-235). Springer.

Tudorache, T., Falconer, S., Noy, N. F., Nyulas, C., Üstün, T. B., Storey, M.-A., & Musen, M. A. (2010). Ontology development for the masses: Creating ICD-11 in WebProtégé; In *Proceedings of the 17th International Conference on Knowledge Engineering and Management by the Masses* (pp. 74-89). Berlin, Heidelberg: Springer-Verlag.

Tudorache, T., Falconer, S. M., Nyulas, C., Noy, N. F., & Musen, M. A. (2010). Will Semantic Web Technologies Work for theDevelopment of ICD-11? In P. F. Patel-Schneider et al. (Eds.), *International Semantic Web Conference (2)* (Vol. 6497, p. 257-272). Springer.

Tudorache, T., Noy, N. F., Tu S., & Musen, M. A. (2008, October). Supporting Collaborative Ontology Development in Protégé. In *Proceedings of the 7th International Semantic Web Conference 2008* (Vol. 5318, pp. 17-32). Karlsruhe, Germany: Springer.

Tudorache, T., Nyulas, C., Noy, N. F., & Musen, M. A. (2011). WebProtégé: A Distributed Ontology Editor and Knowledge Acquisition Tool for the Web. *Semantic Web Journal*, 11-165.

Walk, S., Andrews, K. (2012). Heuristic Evaluation Report, Retrieved July 6, 2012 from <u>http://heuristiceval.simonwalk.at/he.html</u>

Walk, S., Strohmaier, M., Tudorache, T., Noy, N. F., Nyulas, C., & Musen, M. A. (2012). Recommending Concepts to Experts: An Exploration of Recommender Techniques for Collaborative Ontology Engineering Platforms in the Biomedical Domain. In *Proceedings of the 3rd International Conference on Biomedical Ontology (ICBO 2012)*, Graz, Austria, 2012.

Heuristic Evaluation of the Tool iCAT Analytics (publicly released as PragmatiX)

Report of 28th May 2012

Executive Summary

Three ontology specialists were assigned the task of evaluating the online ontology analysis tool *iCAT Analytics/PragmatiX*, the immediate predecessor (including the same functionality) to *PragmatiX*. All three evaluators have substantial knowledge of ontologies including detailed information about the way they are created, evaluated, and used. In addition to ontology specialists, the following groups of users were identified as potential users of (parts of) *iCAT Analytics/PragmatiX*: Administrative Personnel, Engineering Crew, and Ontology Viewers.

The initial intent of each of these user groups can, but does not necessarily have to, differ. Administrative Personnel are mainly interested in the progress of the ontology or the project itself. Compared to Engineering Crew, Administrative Personnel do not care about every single detail of the development process or the ontology. On the other hand, Engineering Crew are more concerned about investigating and identifying aspects of the engineering process which could have a positive or negative influence on the resulting ontology. Ontology Viewers are mainly interested in browsing and consulting the ontology, or parts of it, for their own purpose. They are not particularly interested in the analytical features of *iCAT Analytics/PragmatiX*, but do make use of the integrated visualizations.

To help identify and classify potential problems and positive findings, the evaluators used the "Andrews General Usability Heuristics", a slightly modified version of Nielsen's 10 Heuristics (Nielsen, 1992). In a first session of 60 to 120 minutes, each evaluator examined *iCAT Analytics/PragmatiX* to identify usability issues and classified these according to the previously mentioned heuristics. The evaluator assigned a severity rating to each of the problems in the combined list. The list was then sorted in decreasing order of average severity.

The evaluation identified a total of 27 usability issues across all three evaluators. The most positive findings involve the utility of the implemented heat-map feature and the Author Overview pages. The 5 most severe problems are:

- 1. *Interpretation of Results* It is very hard to interpret the currently browsed features of the network visualizations as no real info is given on what the chosen feature actually tells a user.
- 2. *Concept Recommender Titles* The headings for all concept recommender approaches, which are displayed in the detailed author views, are too complicated.
- 3. *Details of Initial Concept Network Visualization Unclear* Directly after the Log-In, when being confronted with the initial ontology concept network visualization, it is unclear what parts and properties of the ontology are represented as nodes, edges, colors and

diameter of the nodes.

- 4. *Legend Display Bug* If the legend is expanded and network visualization is switched from categories to authors or properties, the legend is still visible and cannot be deactivated.
- 5. *Unreadable Hover Text* The hover text if mouse is over a node in the graphical representation is unreadable if many nodes are displayed.

A complete list of all 27 usability issues is given in Section 6.

INTRODUCTION

Heuristic Evaluation (HE) is a usability inspection method, in which a small number of usability and/or domain specialists (the evaluators) analyzes and assesses the quality of a user interface. The evaluators use a list of *heuristics*, or general principles of good user interface design, to help identify potential problems. In this case, the evaluators used the "Andrews General Usability Heuristics", a slightly modified version of Nielsen's ten usability heuristics by Nielsen (1992). As each evaluator examines the interface, any potential problems or positive aspects, which are identified, are noted down and a corresponding screenshot is made.

Heuristic evaluation is usually performed with specialist evaluators who have extensive knowledge of either usability or of the domain in question or both (so-called double-specialists). For this study, it proved difficult to find double-specialist evaluators with extensive knowledge of both usability and ontology engineering. The three evaluators chosen were all single-specialists in ontology engineering.

Each evaluation was conducted independently to avoid any influence from the other evaluators. After the evaluation sessions, the findings from all three evaluators were aggregated into a single combined list, one for positive findings and one for negative findings (potential problems). The combined list of positive and negative findings was returned to each evaluator to assign severity (and positivity) ratings. The severity ratings ranged between 0 (no problem) and 4 (catastrophic problem). As a final step, the negative findings were sorted in decreasing order of average severity.

Note that a HE may or may not provide information on how to tackle the identified usability issues. In any case, HE represents a very cost effective method to identify design problems from minor to catastrophic. Additionally, a HE can be performed at an early stage of development, allowing problems to be recognized and fixed early on.

User Profiles

iCAT Analytics/PragmatiX is designed for users who are closely involved with the topic of collaborative ontology engineering, thus are related and can be mapped to the roles identified by Schreiber et al. (2000) for Knowledge Production Systems. These users might be part of the Engineering Crew, which needs to adapt the ontology to better fit the needs of the Ontology Workers, who provide content for the ontology. Additionally *iCAT Analytics/PragmatiX* also provides various network visualizations and statistics overview pages for Ontology Administrators and Project Supervisors to monitor progress. These overviews display additional information about the project, such as the distribution of changes across users, and allow domain experts for particular sub-branches of the ontology to be identified.

- 1. Administrative Personnel
 - a. Knowledge Managers
 - b. Project Managers
- 2. Engineering Crew
 - a. Knowledge Engineers
 - b. Knowledge Analysts
 - c. Knowledge System Developers
 - d. Knowledge Providers
- 3. Ontology Viewers
 - a. Knowledge Users
 - b. Visitors

Administrative Personnel are responsible for overseeing the progress of the project. They concentrate on using the statistical overview pages (dashboard, tags) and only examine more detailed statistical pages (concept, author) on special occasions.

The Engineering Crew is less interested in monitoring the overall progress of the project. Instead, they are particularly interested in very specific aspects of the engineering process, such as the distribution of changes across users or the identification of frequently changing areas of the ontology.

Ontology Viewers can exhibit various types of motivation, but generally consume rather than contribute to an ontology. They can exhibit either the behavior of Administrative Personnel or the Engineering Crew (or a combination of both). They can be interested in certain aspects of the ontology due to personal motivations, such as its use in their own projects or the application of iCAT Analytics/PragmatiX to their own ontologies.

Extent of the Evaluation

All three evaluators inspected nearly all parts of the iCAT Analytics/PragmatiX interface. There are no specific areas of the tool, which are only visible for specific users. The evaluators spent most of their attention (and time) analyzing and understanding the different network visualizations. The main focus set by all three evaluators (on their own initiative) was on the information that can be gained when assuming three particular roles: that of Project Managers, Knowledge Analysts, and Knowledge Providers. To that end, they first explored and investigated the different network visualizations and later used the overview pages (dashboard, authors etc.) to find and verify information they had previously extracted from the network visualizations.

EVALUATION ENVIRONMENT

The hardware and software used by each evaluator is shown in Table 1.

Evaluator	Expert 1	Expert 2	Expert 3
Age	35	28	31
Sex	Male	Female	Female
Web Browser	Firefox 12	Google Chrome	Firefox 12

Table 1: Hardware and software environment used by each evaluator.

		19.0.1084.52	
Operating System	Windows 7	Windows	7 Windows 7
	Professional	Ultimate	Ultimate
Internet	Graz University of	Localhost	Localhost
Connection	Technology		
Monitor Size	26" TFT	15,4" TFT	15,4" TFT
Monitor Colors	32-bits	32-bits	32-bits
Monitor	1920x1200	1920x1200	1920x1200
Resolution			
Browser	1920x1200	1920x1200	1920x1200
Resolution			
Date of Evaluation	2012-05-11	2012-05-22	2012-05-25
Time of Evaluation	10:00-12:00	15:00-16:30	09:00-10:00

POSITIVE IMPRESSIONS

Interesting and Useful Feature: Heatmap

The tool *iCAT Analytics/PragmatiX* provides a very interesting and useful feature called "heatmap". This feature allows tracking the activity of all users within the ontology and can be seen in Figure 1.



Figure 1: The concept network visualization with activated heatmap.

Also, evaluators remarked that this feature is very useful to tackle the problem of having to assign additional resources (opposed to non collaborative authoring systems) to monitor activity in collaborative authoring systems.

The dashboard

In addition to the heatmap, evaluators remarked that the dashboard is a very useful overview page that allows for easy monitoring of the progress of the project or the ontology. The dashboard for the International Classification of Diseases in its 11th revision can be seen in Figure 2.



2: The dashboard for the International Classification of Diseases in its 11th revision.

The author overview

The author overview page is a very useful page to quickly identify the different user types according to several different aspects (features). It allows users of *iCAT Analytics/PragmatiX* for example to identify the most (and least) active users within the collaborative ontology engineering process according to the amount of contributed changes or notes. The author overview page (an excerpt) for the International Classification of Diseases in its 11th revision can be seen in Figure 3.



3: The author overview page for the International Classification of Diseases in its 11th revision.

Clean Interface

Evaluators remarked that the interface of *iCAT Analytics/PragmatiX* is well thought through and does not feel overloaded or bloated. Especially the detailed concept statistic pages provide a variety of different information without overloading the interface. A detailed concept statistic page of the International Classification of Diseases in its 11th revision can be seen in Figure 4.

Firefox 🔪 🔣 Category XVII 'Developmental anomalies 🛨	1			
socialenalytics.tugraz.at/categories/http%3A//who.int/icd%23XVII/			☆ ▼ C	🛃 - Google 🖉 🔶 🖈 -
Meistbesucht Ki ICAT Analytics iCAT Analytics				Lesezeichen
		TU	About Logout	×
		Graz	Dashboard Categories Authors Properties	
	XVII 'Developmental an http://to.infic.dttVii 14.changes, 5 notes Show in network	omalies'	Change and Annotation authors	
	Jan Apr Jul Oct Jan Apr Jun Change Notes	Jul Oct 2011 2011		
	Category TAGs			
	Primary Tag http://who.int/icd#TAG_Rare_Disea Secondary Tag Involved Tag http://who.int/icd#TAG_H_Mortality Involved Tag http://who.int/icd#WHO_Assigned	ases		
	Parents			
	ICD Categories' Selected cause is Congenital malformations	s, deformations and chromosomal abn	ormalities'	
	Children			
	Anomalies of antenatal development' Dysplasia syndromes of antenatal origin' Former structure' Needing a decision to be made' To be retired - Developmental anomalies' Q00-Q09 Chromosomal anomalies'			
	Features			
	Feature	Value		
	Number of changes	14		
	Number of notes	5		
	Unanges + notes Distinct authors of changes and appointment	19		
	Distinct authors of changes and annotations Distinct authors of changes	4		
	Distinct authors of annotations	2		
	Authors Gini coefficient	0.543859649123		
	Number of parents	2		
	Number of children	6		
	Depth in network	1		

Figure 4: A detailed concept statistic page of the International Classification of Diseases in its 11th revision.

Quick response times

The response times of the network visualizations (the feedback after dragging or zooming the network) is very quick. This is due to the special filtering technique that is applied on the network visualizations, which results in displaying only a fraction of nodes and edges, thus avoiding visual cluttering. The concept network visualization for the International Classification of Diseases in its 11th revision can be seen in Figure 5



Figure 5: The concept network visualization for the International Classification of Diseases in its 11th revision.

ANALYSIS OF THE MAIN PROBLEMS

Interpretation of Results

The main problem according to all three evaluators is the circumstance, that it is very hard to interpret the currently browsed features of the network visualizations as no real info is given on what the chosen feature actually tells a user. A representative concept network visualization can be seen in Figure 6.



Figure 6: The initial concept network visualization directly after Log-In.

The circumstance that the legend is not expanded by default additionally amplifies the problem.

Concept Recommender Titles

The headings for all concept recommender approaches, which are implemented and brows-able at the detailed statistics authors views, are too complicated, as can be seen in Figure 7.



Figure 7: The implemented concept recommender systems at the detailed statistics views.

Details of Initial Concept Network Visualization Unclear

Directly after the login when being confronted with the initial ontology concept network visualization, it is unclear what parts of the ontology are represented as nodes, edges, colors and diameter of the nodes. Additionally it is unclear how *iCAT Analytics/PragmatiX* filters, which concepts are to be displayed as only a small fraction of all nodes are displayed at once, as depicted in Figure 8.



8: The initial concept network visualization of iCAT Analyitcs/PragmatiX with expanded legend.

Legend Display Bug

If the legend is expanded and network is switched from categories to authors or properties, the legend is still visible and cannot be deactivated as shown in Figure 9.



Figure 9: Legend Display Bug prevents users from collapsing the legend.

Unreadable Hover Text

The hover text, which is displayed if the mouse is hovering over a node in the graphical representation, is unreadable if one node has many parents or children, as all texts are displayed at once and can produce visual clutter as can be seen in Figure 10.



Figure 10: Visual clutter introduced by hovering a node that has many parents and children.

LIST OF PROBLEMS FOUND

An aggregated list of all problems identified by the evaluators can be found in this section. Problems are ranked in descending order of average severity.

Table 2: Identified Usability Issues

No	Short Title	Description	Heuristic	Found By				S	everity	r
				E1	E2	E3	E 1	E2	E3	Average
1	Concept recommender title	The headings for all concept recommender approaches are too complicated.	Aesthetic and Minimal Design	Y			3	4	2	3,00
2	Interpretation of results	It is very hard to interpret the currently browsed features of the network visualizations as no real info is given on what the chosen feature actually tells a user.	Recognition Rather Than Recall	Y	Y	Y	2	3	3	2,67
3	Details of initial concept network visualization unclear	Directly after the Log- In when beeing confronted with the initial ontology concept network visualization, it is unclear what parts of the ontology are represented as nodes, edges, colors and diameter of the nodes.	Recognition Rather Than Recall	Y	Y	Y	2	3	3	2,67
4	Legend display bug	If the legend is	-	Y	Y	Y	3	4	1	2,67

	expanded and network				
	is switched from				
	categories to authors				
	or properties, the				
	legend is still visible				
	and can not be				
	deactivated.				

5	Unreadable hover text	The hover text if mouse is over a node in the graphical representation is unreadable if many nodes are displayed.	Aesthetic and Minimal Design	Y			3	2	3	2,67
6	Usefulness of results	In addition to the interpretation of the features (No. 4), what can be achieved by calculating and interpreting these features?	Recognition Rather Than Recall		Y	Y	3	2	3	2,67
7	Categories list of features	The list of features when browsing "Categories" is confusing/misleading as it can be misinterpreted as category (of concepts).	Aesthetic and Minimal Design			Y	3	2	3	2,67
8	Description instead of color in charts	In all charts, instead of just listing for example "blue" concepts, the description for the blue display status should be used.	Recognition Rather Than Recall			Y	3	2	3	2,67
9	Wrong vocabulary for Audience	Ontology experts might not be familiar with network analysis measures and	Speak the Users' Language		Y		2	2	3	2,33

	vocabularies.				

10	Properties network "missing"	When selecting the properties network, there are only the top outliers visible and the actual connected graph is "missing" (hidden outside the current field of view).	Aesthetic and Minimal Design			Y	2	2	3	2,33
11	Show a concept in network	Very hard and confusing to identify the currently browsed concept when clicking "show in network" while browsing the concept.	Recognition Rather Than Recall		Y		3	2	2	2,33
12	Hide TAG options and features if not available	All features and statistics of TAGs should be hidden if TAGs are not supported by the imported data set.	Aesthetic and Minimal Design		Y		3	2	2	2,33
13	Selection process for visualization unclear	As only a small fraction of nodes is displayed it is unclear why and how this selection process is applied.	Recognition Rather Than Recall			Y	3	2	2	2,33
14	What type of relation do edges represent and which	What type of relationship is displayed by the edges	Recognition Rather Than Recall	Y	Y	Y	3	2	2	2,33

are shown?	in all network				
	which are shown due				
	to the selection				
	process?				

15	Missing colors in ICTM	There are no colored concepts displayed in ICTM. They are all grey!	Consistency	Y			1	3	2	2,00
16	Can't move/drag nodes	The nodes should be made "drag and droppable" as there are "overlaying" nodes.	Aesthetic and Minimal Design	Y			1	2	3	2,00
17	Action after click on node of properties network	Contrary to all other network visualizations, when clicking on a node of the properties network, the user will be redirected to the concept/categories network.	Feedback		Y		1	3	2	2,00
18	TAGs	Mistaken for TAGs (instead: Topic Advisory Groups)	Speak the Users' Language	Y		Y	2	2	2	2,00
19	Change context	The vocabulary "context of a change" is misleading as the context only discribes the performed action of a change.	Speak the Users' Language	Y	Y		2	3	1	2,00
20	Headings of statistic pages should be improved	Instead of having the names of the statistics as headings a more descriptive heading	Recognition Rather Than Recall		Y		2	2	2	2,00

	should be chosen.				

21	No "Features" in authors and properties network	Contrary to the concept network visualization the features are named "show".	Consistency		Y		2	2	2	2,00
22	Too many category features	There are too many features listed. Only the Top 10 most important features should be listed to prevent confusion.	Aesthetic and Minimal Design	Y			3	1	2	2,00
23	Improve Legend	Legend should include a more descriptive sentence of the currently shown feature.	Recognition Rather Than Recall		Y		1	2	2	1,67
24	Legend icon	The icon chosen for the legend is misleading and signals help rather than legend!	Recognition Rather Than Recall	Y	Y	Y	2	2	1	1,67
25	Changes by TAGs	The concept of TAGs in ICD-11 and ICTM was not understood/very unclear.	Speak the Users' Language	Y			2	1	2	1,67
26	Details of authors network and properties network	When browsing the authors and properties network it is unclear	Recognition Rather Than Recall		Y		2	2	1	1,67

	unclear	what attributes are responsible for each node size and edge weight.						
27	Missing "Home" Button	A dedicated home button is missing (instead or in addition to having the application logo- image linking back to the initial concept network visualization).	Recognition Rather Than Recall	Y	1	1	1	1,00

Severity	Meaning
4	Catastrophic problem
3	Serious problem
2	Minor problem
1	Cosmetic problem
0	Not a problem
Av	Average Severity

Table 3: Severity Ratings

REFERENCES

Andrews, K. (2011). Human-Computer Interaction: Lecture Notes, Retrieved Apr 05, 2011 from <u>http://courses.iicm.tugraz.at/hci/hci.pdf</u>.

Nielsen, J. (1992). Finding Usability Problems Through Heuristic Evaluation. *In Proceedings of the Conference on Human Factors in Computing Systems (pp.* 373-380). Monterey, CA, USA: ACM.

Nielsen, J. (1994). Heuristic evaluation. *In Proceedings of the Conference on Human Factors in Computing Systems* (pp. 152-158). Boston, MA, USA: ACM.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W. & Wielinga, B. (2000). *Knowledge Engineering and Management: The CommonKADS Methodology*. Cambridge, MA, USA: MIT Press.