

Maintaining Ontology Implementations: The Value of Listening

Duane Degler and Renee Lewis
Extreme Markup 2004

Abstract

It's hard to argue against the concepts of self-describing data, contextual interfaces, and richer metadata for content that eventually will make up the Semantic Web. The need is just so great, and becomes greater by the day with the huge increase in unstructured content and non-integrated data repositories. However, it is easy to imagine semantic environments suffering from the same challenges that many content management system implementations and the Web itself suffer from: the preoccupation with publishing and storing metadata could easily leave us drowning in it. If there's one thing that "six degrees of separation" shows us, it is that everything can be related to everything else. Will we evolve meaning in a mass of loosely assembled associations? This paper focuses on the authors' maintenance design challenges, and the approaches being explored to promote feedback and user involvement for the maintenance of semantic representations, to ensure they remain useful and current.

Introduction

It's hard to argue against the concepts of self-describing data, contextual interfaces, and richer metadata for content that eventually will make up the Semantic Web. The promise is compelling – it's about computer agents talking to other computer agents – that is, computers having conversations with other computers on your behalf.

However, it is easy to imagine semantic environments suffering from the same challenges that many content management system implementations and the Web itself suffer from: the preoccupation with publishing and storing metadata could easily leave us drowning in it. If there's one thing that "six degrees of separation" shows us, it is that everything can be related to everything else. Will we evolve meaning in a mass of loosely assembled associations? Will *too much* structure allow meaning to decay over time? What constitutes a change in meaning?

Meaning changes organically and is subjective to the viewpoint of the consumer of the information. It changes with the surrounding conditions. It changes as the environment changes. It changes with the participant's activities. We – information management and content people – have been used to making everyone see the world through the lens of a particular application supported by an explicit taxonomy – the user's experience was limited to what we coded the machine to understand. Now we expect an application to listen, to recognize what it needs, and to interpret through discovery the knowledge needed to carry out useful actions. This is a monumental shift in how we think about, implement, and expand knowledge.

Any ontology must be current and represent the "here and now" to be properly understood, interpreted and acted upon. Ontology *maintenance* is the set of processes – both manual and automatic – that focus on keeping ontology representations current within the environment(s) where used. It is, for us, one of the most important aspects of ontology development.

As we design using ontologies and context-based navigation to support large-scale organizational resources, we run into the common challenges of how to optimize for maintainability while making our applications more responsive, adaptive, and scalable at the same time. We focus on user feedback and how it plays a role in creating maintainable ontology, content, and applications. We reflect on the basics of what we want ultimately to achieve to find the most practical approaches to implementation. To be successful, we first must understand the systemic communication issues and goals, then enlist the support of users through their interaction with the information application.

A Communications Perspective

“A complex adaptive system acquires information about its environment and its own interaction with that environment, identifying regularities in that information, condensing those regularities into a kind of ‘schema’, or model, and acting in the real world on the basis of that schema.” (Gell-Mann 1994, p.17)

Why start to explore ontology maintenance with a perspective on communication? Because the heart of the future of computing *is* communication: facilitating human-to-human communication, human-to-machine communication, and machine-to-machine communication. Helping machines find ways to support human communication is difficult and complex. It is valuable to understand where some of the complexities arise. As reflected in the above quote, any ontology is a schema or model of regularities, requiring that the system “acquire information about its environment” (context). Understanding communication has helped us understand the balance between creating systems and data, and maintaining/evolving them so that they remain useful.

"All of us are exposed to huge amounts of material, consisting of data, ideas, and conclusions - much of it is wrong or misunderstood or just plain confused... Humanity will be much better off when the reward structure is altered so that selection pressures on careers favor the sorting out of information as well as its acquisition." (Gell-Mann 1994, p.342)

Human Communication

The most basic model of human communication has a sender producing a set of symbols that are consumed by a receiver (Cummings, et al, 1983).

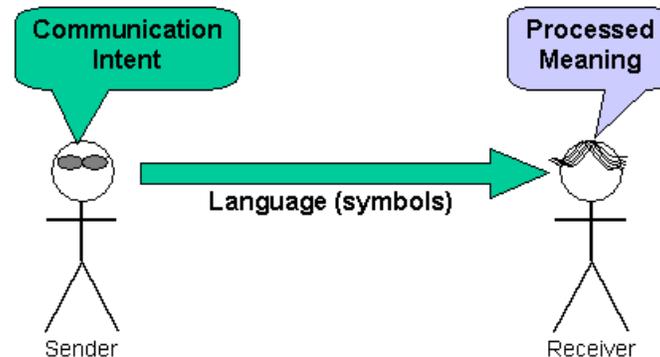


Figure 1: Direct communication between two people

The reason direct human communication works (well, usually works) is because there is a feedback loop that allows a process of refining the symbols and their interpretation to increase alignment of the meaning of the symbols between the sender and the receiver. The nuances involved in creating shared meaning also rely on the sharing of context, experience, and secondary information that help to round out the concepts that are being shared.

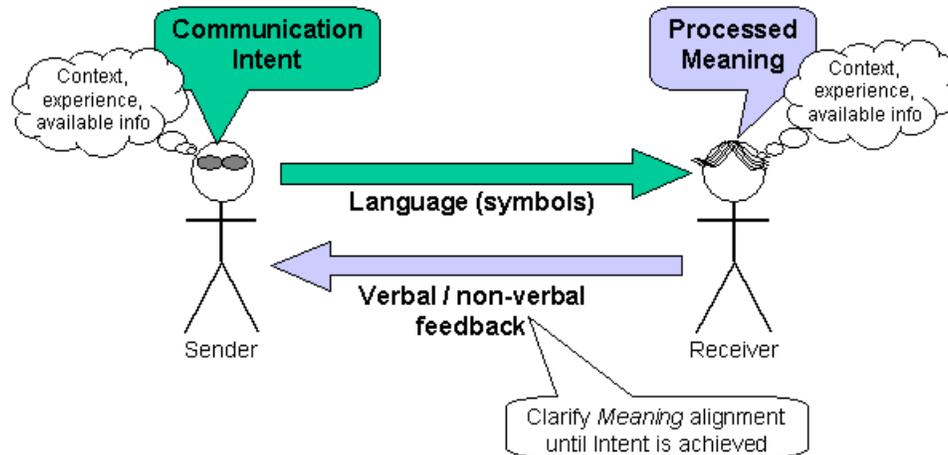


Figure 2: Direct communication relies on feedback to achieve the intent

Shannon and Weaver (1949) extended the simple, direct communication model to represent the role of technology devices (in their case, the telephone).

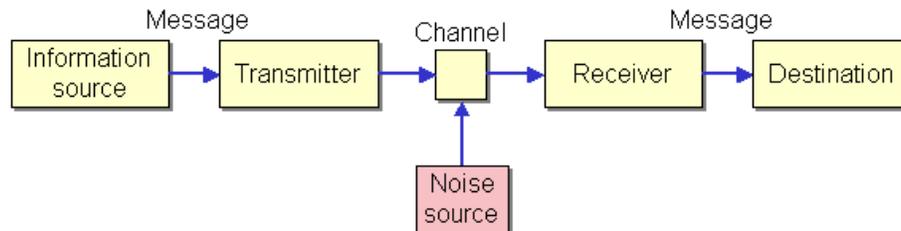


Figure 3: Shannon & Weaver's Communication Model for Telephone Systems

Their focus was on the direct signal processing of audio via telephone equipment, and the need for technical quality (noise reduction) to limit the potential for miscommunication. They did not go into the limitations of audio-only communication on the quality of the communication experience more generally (lack of visual non-verbal cues, dependence on synchronous communication, etc.). If we applied that same perspective to computing, it would be like focusing on the screen's pixel resolution and number of colors displayed – i.e. we are still left to explore what messages and symbols are being facilitated or filtered by the characteristics of the specific technology involved. Exploring these latter issues is increasingly important the more we mediate human communication through computers, the more we use terms like *meaning*, *intelligence*, and *knowledge*, and the more we load our interfaces with inconsistent, culturally-bound metaphors and representations.

Chandler (1994) identified gaps in the Shannon and Weaver model as it has been more generally applied to organizational communication. Of particular interest to us is our understanding of the goals and practical implementation challenges of ontologies in computer/information environments, and specifically the lack of *feedback* and *context*. Here, feedback refers to the active role of the “destination” in successful communication, and context refers to situational information (either explicitly or implicitly provided). Chandler goes on to say:

“Each medium has technological features which make it easier to use for some purposes than for others. Some media lend themselves to direct feedback more than others. The medium can affect both the form and the content of a message. The medium is therefore not simply 'neutral' in the process of communication.”

The Computer's Non-Neutral Role

There is a relentless drive toward using computers and the Web as the primary point of interaction between people and organizations. So many areas of our lives are now online experiences, whereas once we used to go to an office to conduct a transaction (buy goods, renew a drivers license, apply for benefits, manage money, find reference books in a library) or make a telephone call to receive/give information (product support, customer services, making travel arrangements, or complaining to our local Congressional representative). We used to interact with other people, not with e-forms and search engines and site maps as a proxy for people. Once communication was more synchronous and direct, now it is more often asynchronous, indirect, and impersonal.

What is the fundamental implication of this? *We, the users, must fit our communication into the pre-determined constructs of the software we are interacting with.* A huge burden is thus placed on designers (of software, interactions, and information) to *predict and support* widely divergent user needs. Computers are extremely weak in their ability to tune services or information based on recognizing and interpreting the context of our particular situation. What has been lost? Listening!

We have been exploring the impact of computer-mediated communication as part of our practical project work with both transactional systems (people carrying out tasks, often with more than one party involved) and content systems (people seeking and using formal/semi-formal written communication). Next, we'd like to explore the computer's role and potential impact areas on communication. It has been by understanding where the impact lies that we begin to understand where to focus our design efforts and technologies.

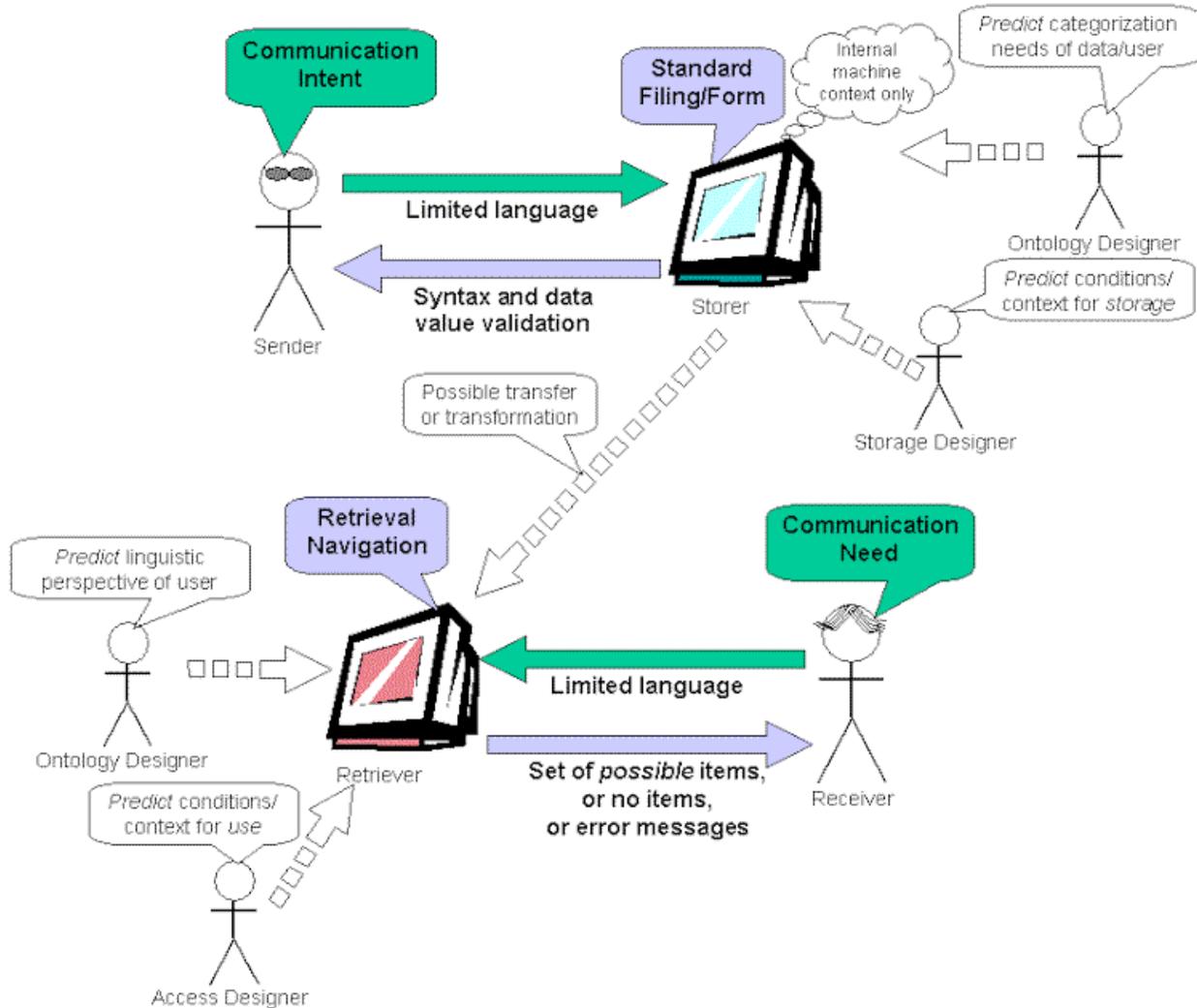


Figure 4: The impact of computer involvement in human communication

As you see in Figure 4, the communication situation becomes much more complex. The Sender transmits the content/data into the computer, which has some form of standardized filing system. The structure of the filing system and the sophistication of the application that manages the activity dictate the level of the “conversation” and feedback that the Sender experiences. The conversation often consists of confirming whether the data provided conforms (or not) to a limited, standardized schema. Data conformance requirements are based on decisions (which are in many ways *predictions*) made by the designer of the data capture application months or years before the Sender decides to provide the information.

It is possible (in fact, likely) that the designer responsible for predicting how the content/data will be used is different from the storage designer who designs a retrieval application to support that intended use. Data may undergo transfer or translation, where again some aspects that are valuable for shared meaning may be altered in some way. This is increasingly the case when the idea of reusable content becomes a priority among technologists and user organizations.

Finally, the Receiver expresses a need in the limited language of the computer retrieval application. The feedback available to the Receiver is limited, in many cases today, to a list of available data items – in the case of no items being available there *may* be a message, which *may* describe error conditions. That is the extent of the feedback to the user, and often it is completely ignorant of the person’s context, or why they are asking to retrieve information.

Does an Ontology Framework Help?

We believe that there can be an improvement – and simplifying – of the model with the introduction of semantic technologies and an ontology framework for applications and content. We believe the model could begin to look more like this:

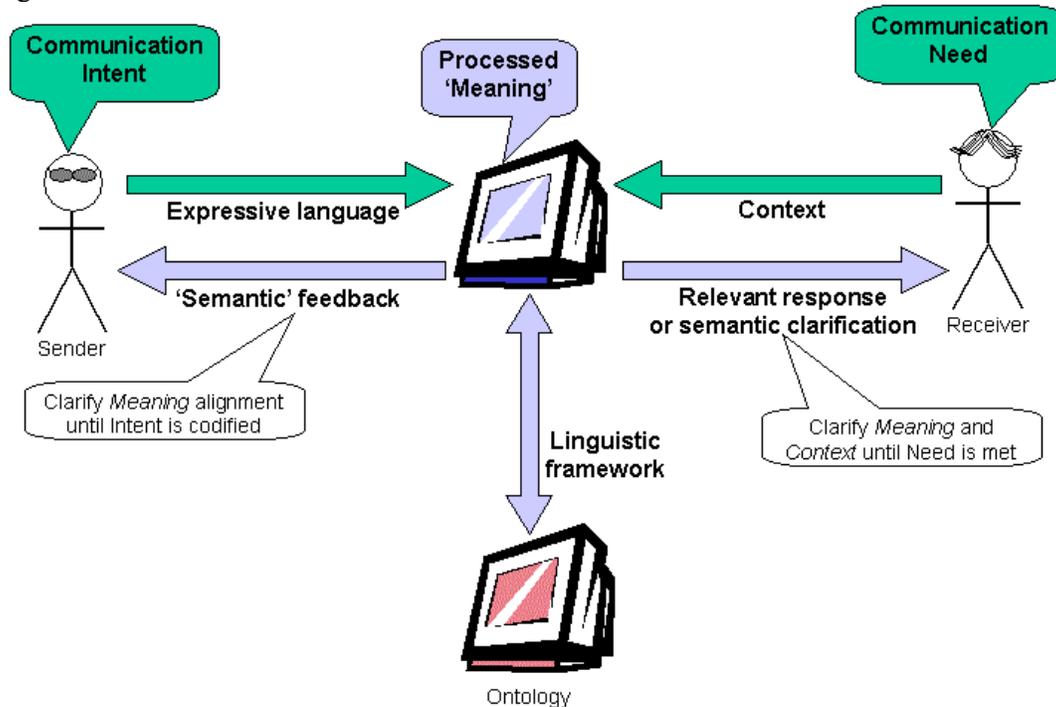


Figure 5: Ontology-based communication, using semantic technologies

The key improvement here is a framework for harmonizing the language and definition bases used by the computer during the course of interactions with both Sender and Receiver. The ontology in this case participates in two key activities:

- Providing a map to standardized language that represents the contexts and domains served by the computer application
- Providing a framework for managing language evolution – and potentially meaning – in order to maintain the mapping between the parties (human and computer) involved

It is important to understand that the simplification occurs regardless of whether the communication system is defined through a more informal localized system where agreements are reached through a handful of entities to perform a specific function, or through a global system pushed through industry standards and alignment, for example.

Ontology-driven applications can be a very good thing, and make the conversation with the computer better – if the ontology is relevant and up-to-date! If the ontology begins to ossify, it loses relevance to the human participants in the dialog, and would increasingly become a barrier to communication. This particular problem is a bigger risk to global standards where there will be more resistance to change, but that resistance could easily make the standard not useable and possibly replaced by improved versions. The key issue is that *it is impossible to predict all possible language requirements and changes in advance*, so the ontology must be able to learn and evolve. But how?

The Role of Feedback

Feedback Paths in Computer Applications

What's missing in the above models is feedback between the Sender and the Receiver. It is feedback that allows the content/data (and the descriptive mapping of that content to the Receiver's situation – the ontology) to be refined and improved. What types of feedback mechanisms are available currently, between Sender and Receiver? Figure 6 illustrates the current, limited feedback in many applications.

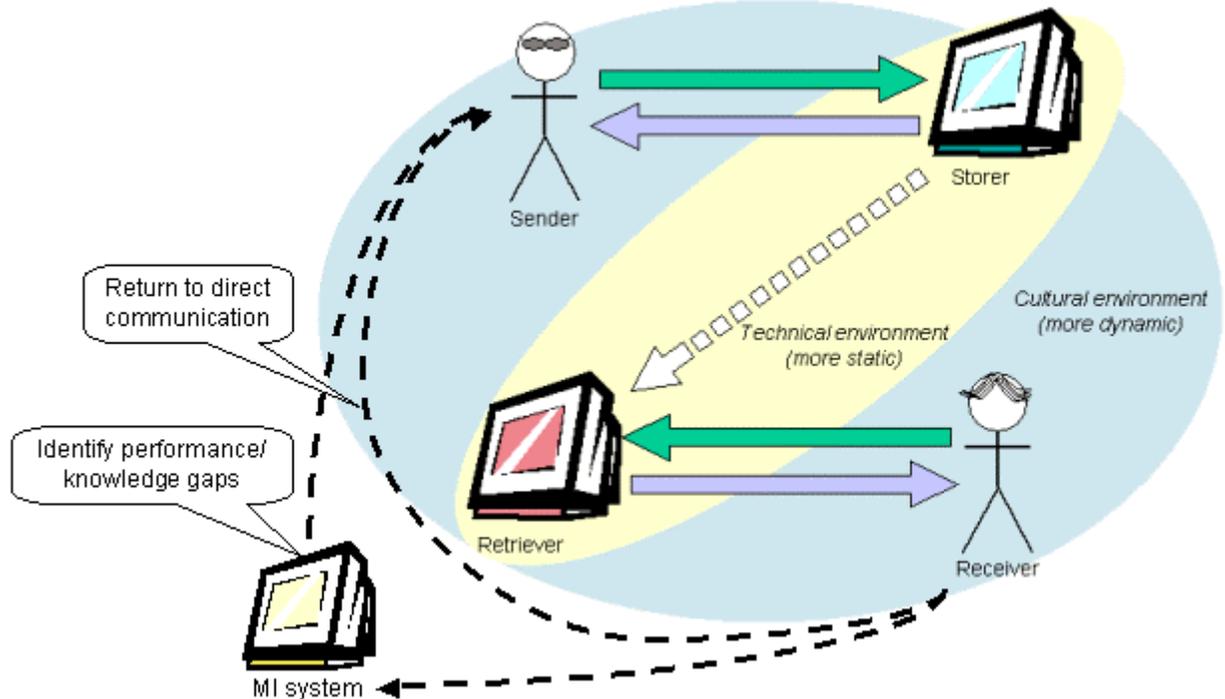


Figure 6: Two types of feedback, both outside the communication channel

Feedback channels fall into two broad categories. One (increasingly of interest to the KM community, through a focus on collaboration and Communities of Practice, e.g. Wenger 2002, Preece 2000) is to return to the simpler model of direct communication – sending an e-mail to the author to start a conversation, calling a help desk, or asking your neighbor.

The other feedback mechanism is slower because it requires identifying meaning mismatch through the symptoms of poor task performance. Most MI (management information) systems today are not set up to accommodate the integration of performance data from transactional systems and information access systems to allow people to track and correlate findings. They also do not necessarily distinguish between types of interactions that could help distinguish "conversational" problems from other forms of performance issues.

Neither of those two types of feedback lends itself to broad-based, systemic, ongoing improvement of the overall application, content or ontology. There is clearly a need for improved conversations between computers and users, and richer mechanisms for feedback to support refining the design (hence, dialog) of applications in a way that evolves naturally with continual improvement.

Ontology Feedback

So, if the ontology used by semantic technologies within/supporting applications, enhances the ability of the computer to *listen* in the course of a conversation with a user, then it has to become self-aware to be effective. The big question is: how well does it *understand* the language of the user?

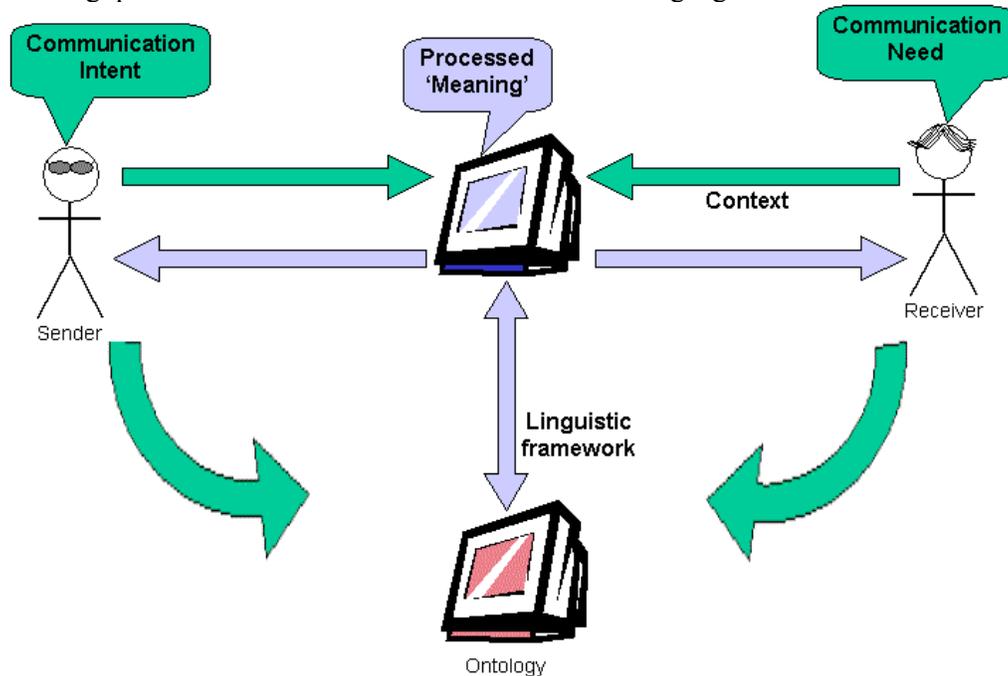


Figure 7: Feedback from users refines the relevance of the ontology

The feedback paths for the ontology are both explicit (suggestions and direct entry) and subtle (pattern analysis, dynamic listening, clarification-seeking, meta-learning, alerting for alternatives and drifts). Both are important and must be considered in system design. Some of the ways of doing this are explored in the next section. The interesting thing is that the design approaches that help us maintain an ontology can also gather feedback that allows us to use the ontology to help with more general feedback between users, content creators, and application developers.

Relevance = Context = More Useful

Keeping the ontology relevant to users means keeping it aligned with the user's context. Everything else is unimportant if we want computers to work more effectively on our behalf and communicate with us in a more meaningful way. Understanding, interpreting and maintaining context (rather than just representing it or codifying it in markup) is the truly difficult problem to be solved.

Many communications problems with face-to-face conversation are due to semantic differences – when one person does not fully understand what the other is saying. Context is not only situational, it is also experiential. Two people in the same place at the same time still come away with different perceptions and knowledge. Our understanding of context is informed by our knowledge relating to the situation, our prior experiences, our expectations, and our emotions. How well will software agents operating on our behalf know all the rich aspects of context, including our goals, expectations, and mood? How well will computers be able to represent that situation to other agents and people, so they can carry out their own actions on our behalf?

We predict it will take a long time to refine not only how to properly represent context so that it can be used in a meaningful way, but also to develop the required learning techniques that keep the context

relevant and current, and the social networks to refine them (Behrens and Kashyap 2002). This, too, will evolve through experience, but its importance cannot be lost or underestimated.

The Nature of Ontology Change

Changing ontology is something we inherently understand and comprehend as humans, but our systems struggle with this. Some changes are so profound that they can have rippling effects throughout an ontology, the frameworks that interpret that ontology, or the systems that take action on the outcomes of interpretation.

Ontologies need to change and evolve when we experience new understandings or arrive at new shared meaning. And, as one ontology changes its relationship to related ontologies, child ontologies and localized taxonomies can change or develop. In other words, each change can affect whether subject relationships exist and how subjects can be interpreted based on the meaning they represent.

A number of articles mention how semantic technologies enhance maintenance of the content that they describe (Pepper 1999, Garshol 2002, Euzenat 2001). Increasingly, we are seeing discussion of maintenance of the ontology itself (Broekstra and Kampman 2003, Biezunski and Newcomb 2000, Heflin 2004), and more literature is beginning to appear as these issues arise in early implementations. However, the historic calls for more systemic maintenance of our data and metadata (for example, Agre 1994) in all types of computing environments seem to go largely unheeded in software design and the establishment of standards.

Knowing that change is an important part of maintaining the value of ontologies, we must beg the question: what kinds of maintenance do we expect? What shifts in language and associations are we trying to accommodate? Here are a few key ones.

- **Expansion of scope for a particular subject.**

Subject definitions are not static, and their meaning changes based on legislative changes, colloquial uses or re-uses, cultural variations and general language adoption. For example, the introduction of ten new countries (and nine new languages) to the European Union will prompt more than just a huge demand for translation services. It also prompts a range of reinterpretations of existing laws and information, as the carefully crafted language that represents the intentions of legislators has to be aligned with phrases and cultural expectations of new member countries. This will include re-examinations when “we don’t have a word for that in our language” or “if we implement that directive, it will undermine these other tenets in our existing legislation.”

- **Changes in organizational structure.**

Mergers, acquisitions, reorganizations, selling off of business units happen every day for a variety of reasons. We often understand what the business rationale behind the decision was, yet we sometimes underestimate long-term language/meaning shifts when we bring entities together or split them apart. It is rare that separate entities have the same definitions and ontological views of even the same business. How do they come together? How do we align our old historic information? What constraints have been lifted where we can improve our definitions? Where are our relationships with new and old partners affected? How do we represent that? Where should we evolve and how?

The U.S. phone companies went through this entire process over the last 20 years since they first split up to be the “baby bells.” During the split, copies of the billing system were provided to each operating unit. Over time, the systems in these operating units evolved totally independently. Each established different descriptions to define the service being ordered and pricing models – contained in an ontology (that is, USOC/FID) – that triggered no less than 16 computing systems just to turn on one “plain old telephone” (that is, a POTs) line for a customer. Subsequently, the “baby bells” began to merge into new entities like Verizon, bringing together these systems that now contained totally different ontologies. In many cases the original staff who knew how the

codes were interpreted had retired, so the actual meaning of the codes was lost. New pricing models and operational consolidation has been nearly impossible. They still struggle with the issue of finding a common meaning encapsulated within a single, comprehensive ontology that describes the full range of pricing and services in a single product catalog that characterizes and is constrained by the switch capabilities of the network (a hardware constraint).

- **Changes in law, court interpretations, precedent, and legislative intent.**

Some changes in law have major impact on the way we see the world and where we get the information we need. One example of this is the introduction of “accessibility” legislation (such as “Section 508” in the U.S.) – such legislation both broadens definitions of who computer users are and what rights of access they have. At the same time, many software managers are narrowing the concepts of access to software functionality – where they focus on disabled people rather than a wide range of definitions for what constitutes “access.” Another recent change was the creation of the U.S. Department of Homeland Security. This newly created agency took on the aggregation of multiple agencies and parts of other agencies bringing them under a single umbrella. It created new departments and reorganized others. Whole new definitions for *homeland* and *security* emerged, and these concepts continue to evolve.

- **Human events changing our perception of what words mean.**

Some definitions evolve over time to become mainstream. Take for example the word *mouse*. It did not originally mean the thing you use to drive a cursor around the screen of a computer – but it does now. At first, that meaning was uncommon and rarely used. Chances are when the word mouse is used today, it is more likely to be in the context of a computer instead of a rodent. On the other hand, some definitions change or become more commonly used quickly based on events. A good example of this was how we all personally redefined the term *threat* after 9/11. Words like *terrorism* became commonly used words overnight, with both more clarity of meaning (in reference to that event) and at the same time much broader boundaries to its use.

- **Discovery that prompts new world views.**

Every day, science and technology advance. New capabilities are discovered, old capabilities are combined with others to create something new, and definitions are clarified, added or replaced. Current areas of discovery where our basic information of what we think we know is challenged includes genetics, space, technology, and evolution (for example, not that long ago, dinosaurs were not thought to be related to birds). New information in any one area can have dramatic impact on entire bodies of knowledge as well as related bodies of knowledge. This is particularly noticeable in medicine – we’ve seen how quickly *SARS* and *mad cow disease* have entered the language.

- **Natural evolution of, and differences in, language and culture.**

Knowledge is constantly growing. At the same time, our access to both historic and new knowledge is expanding. As a result, we increasingly interpret information beyond the reference points of where we live, who we live with, and any cultural biases we carry. The collective meaning of words like *retiree* or *family* have changed quite a lot over the years, and so references to content on these subjects may need to have different alignments depending on the age of the information being referenced. The same knowledge can have slight or significant variations in meaning when the context is taken into account.

Additionally, we have to consider that in some languages we have words that express very different meaning depending on context and other languages have different words to convey the different conditions. For example, *aloha* can mean both hello and goodbye. So, we may have word equivalents, but we may need context to determine which meaning to use.

The need to monitor and incorporate change is critical to maintain the value and quality of knowledge. Entities that make their ontologies available to others will need to incorporate methods of presenting information about both meaning and evolution so that others can judge whether the data is acceptable for

their use. To date, there has been little written into standards for creating the meta tags needed to present this rich – and subtle – information.

Going forward, value, quality and accuracy will be key elements in building trust associated with an ontology. More over, there is a need to continually validate the interpretations made and the use of any information source to maintain accuracy.

Supporting Ontology Change – the Art of Meta-learning

Meta-learning is really all about the context – the ability to define, interpret, and evolve meaning consistent with any given context. Context is a two-way street. On one side, we have the terminology, information or knowledge that must in some way nominate the allowable or reasonable contexts in which it can be used. On the other side, those who consume the knowledge need to reflect the context of use and the relationships with other knowledge that are fostered from use. This allows the contextual map to be expanded and refined.

We do not have control over the pace at which meaning and interpretation change. Nor do we have the luxury of coding metadata into a stable, slowly changing set of codes based on legacy repositories, as might be the case with more historic archives. Nor do we have the luxury of having information manager or archivist roles within the organizations we work with, where people are dedicated to the task of indexing and updating. In highly regulated, highly distributed environments with large user populations (up to 75,000 staff and over 50 million clients), we have had to find a range of ways to rapidly assess changes in organizational vocabulary that affects the topics used.

There are many considerations that go into creating a meta-learning environment that keep the ontology fresh and accurate. When incorporated into practical maintenance methods, an integrated set of techniques can begin to support the world of organically evolving metadata. We have started by using these methods:

- Capturing commonly used phrases
- Gathering simple feedback
- Listening to conversations
- Applying stealth knowledge management: learning from what people do

The following discussion provides some examples and priorities of where and how to monitor common information sources to improve the quality, accuracy and timeliness of an ontology. Collecting good information is the first consideration and knowing how, where and why to apply the new insights is the second most important consideration. By using these methods, not only is change managed but also the confidence of users is continually improved through their ability to perform their tasks.

Capturing Commonly Used Phrases

Listening to the language of users isn't as hard as it seems. We ask them to provide data all the time, and they liberally use both formal and informal language. What we fail to do is *use* that information to routinely refine our ontology.

We must continue to develop methods and frameworks to monitor, find patterns in, and reuse these common phrases. Today, we monitor Web usage logs to gain some insight in use patterns, but this doesn't go far enough to monitor language patterns. Some applications will log search strings as they are used and even optimize queries based on the commonly used patterns, but we do not often use this information to refine the ontology that defines the organization's view of the world.

The best information to tap into is the information that people create while simply doing their daily work assignments. It changes organically in response to the people who they interact with daily. We need to tap into the common language and identify the phrases, relationships, trends and shifts. Here are some ideas:

- **Gather new information during the course of daily work.**
 Prompt a user to provide information (via their topic selections and navigational choices) to understand how their current situation has prompted a need for information. If the data is found, then you have a clear match between commonly used terminology and available information, strengthening your awareness of the relevance of those terms. The most meaningful feedback from the user comes when there is no direct match. More information may then be solicited from the user to help locate information – and as a by-product to improve the ontology and/or the available content. Transporting the data and user suggestions using standard representation syntax (such as RDF or XTM) allows it to be more easily shared and further analyzed.
- **Provide capabilities that allow people to link documents to each other.**
 Today, the best we can do in most environment is ask designated subject experts or look at directory listings on computers of highly organized individuals to understand relationships between documents to improve how an ontology supports standard workflow. There is a growing need and opportunity to create frameworks and methods to improve how people share documents in common file systems. Unfortunately, most file sharing is accomplished through email attachments. With improved methods of sharing and linking documents (in ways that allow computers to read the resulting associations), an organization can gain powerful insights into what is really happening at a workflow level.
- **Select commonly referenced topics and associations, and then re-use them to query content resources and facilitate navigation to content.**
 There are relationships across topics that, when represented properly, can lead a user to more discrete topic areas or broader domains – whatever is most useful to them at the time. In this situation, each user can benefit by the actions of the collective. A secondary benefit from this is the ability to listen and detect a context different from the one most commonly used. In an ideal world, a user would be given the ability to have their own common patterns either as a separate source to draw upon or included in the retrieval methods. These do not have to be elaborate. For example, provide “myXYZ” and allow user’s to store their choices, or provide a search history to draw upon for similar situations.
- **Monitor changes in selections over time.**
 This begins to tell us where the language presented to the user may be in sync or out of sync with their situations. Shifts in selection patterns may refer to the whole (e.g., the affect of a legislative change or implementation of a new process) or to the individual (e.g., a job change). By aggregating from individual instances, we can begin to interpret what caused the shift. This is important to understand the nature of the change and the impact on the ontology.
- **Understand when users switch to alternative search methods.**
 This is a common behavior when user topics become misaligned with their situation. Reviewing usage patterns across various navigation methods helps identify when alternate search methods become preferred. Sustained changes in behavior are alarms that require further investigation.
- **Improve “bookmarking” capabilities.**
 In one of our applications, users are able to bookmark interesting sites and content items using a specialized bookmark application that maintains a repository of user-defined references to content (in the context of what they are doing at the time). At the time they mark the item, they can also describe it – thus providing us with further semantic information. Note that our approach to this has not used the standard bookmark capability of a browser, rather replaced it with a standards-based data process that provides the information we need to mine the information. There are also some sites like www.backflip.com that provide a shared service where favorites lists can be exported for daily use and future interpretation.

This area of information gathering, monitoring and interpretation is a major area of opportunity for individuals interested in improving the quality of ontologies that rely on common language in semi-changing environments. It is important to point out that sampling is an acceptable method of gathering insights.

Gathering Simple Feedback

For some reason, many application designers fear feedback from end users – either consciously or subconsciously. This is truly unfortunate, because the most relevant feedback available to support improvements is from end-user experiences. They are willing and able to tell you what you really need to know about the application – in context.

- **Make sure the user has a chance to provide open, unstructured feedback at every point.**
Always provide a method of communication. Always respond that their message was received. Invite them to identify themselves so that you can follow up for additional information where relevant.
- **Let users “speak their mind.”**
We have found that collectively a user community can clearly articulate priorities, shortcuts, and useful improvements that can define success. More generally, the success of discussion threads, blogs and open source development activities show this to be true, as well. Particularly in metadata-driven systems, users should be encouraged or even polled on areas of definition or description where designers are uncertain.
- **Capture the context in which the feedback was called.**
Do not just collect information about the page they are sitting on when they select the feedback button. There are tremendous insights in collecting context – the criteria that got them to that page, and preferably other session or user data that is able (and agreed with the user) to be captured.
- **Separate feedback submissions that are questions about working processes from feedback about the application.**
In customer-facing roles, user questions often reflect the type of question the individual staff member is being asked by the customer. The language in the question is often reflective of the language of the “outside world” and so should be considered carefully.
- **Allow users to nominate new terms.**
Synonyms become richer and the patterns inherent to different business units become apparent when users nominate new terms. Synonym language will reflect things that have meaning to groups within, and outside an organization.

Listening to Conversations

The best way to assess changes in language is really by listening to ordinary conversations, and then – most importantly – interpreting what you hear. Where are the human and organizational conversations taking place? Increasingly they are online, providing a unique opportunity to learn by monitoring (this is the same thing as listening).

- **Monitor every type of “conversation” that you are allowed to monitor.**
Monitor for the purpose of understanding language and priorities. Increasingly, users have collaboration tools available to them beyond e-mail. The use of collaboration and discussion areas is a fabulous opportunity, because the language used is more dynamic, and closer to the conversations that staff hear when they interact with people outside the organization (customers and the public). With each new collaboration environment that evolves within the organization, there is a new source of information to understand how users see their world. We are exploring the use of new technologies to mine these conversations, and then map that back to the organizational ontology. Profiling and discrimination are key capabilities for automated tool sets.

- **As new content is being created and reviewed, a collaborative conversation takes place around that new content.**

Capture the conversations surrounding changes in content, which helps to identify the potential impact on the metadata. Review and approval cycles present rich data about different interpretations and meanings, and also capture the reasoning behind *why* something is represented or phrased the way it is. Encourage the dialog where possible.

Applying Stealth Knowledge Management: Learning from What People Do

We actually know quite a bit about how people use applications and web sites. Stealth knowledge management is an approach where we learn about context and the decisions people make by observing the patterns of their transactional behavior when they use applications (Degler and Battle, 2000). As best practices are identified through actual performance, they can be reflected back to less experienced application users in the form of guidance and support. When gathering this type of feedback, we don't have to be concerned about each and every user, rather their overall behaviors. Discovering patterns, trends, shifts, and profiles really helps designers prioritize.

In particular, we need to monitor work tasks that may require multiple applications or information resources to understand user context at more than the individual software application level. Among other techniques, use periodic testing, job monitoring, and feedback loops. Optimize on the typical paths and monitor atypical patterns as this may be an early indication of knowledge shifts. When work patterns change, it is the time the new information is most needed.

We are beginning to see ontologies used to map between different management information systems and reports, but rarely see ontologies mapping across transactional systems and content management systems. Transactional systems and content management systems tend to be disassociated from each other semantically, yet both are key assets to most business. This has to change! It is important to monitor both task performance and information use by using the same ontology framework. This allows us to use the management information about each one to access insights and indicators from the other.

Conclusion

While the main *technical* focus of the Semantic Web is machine interpretation of self-describing information and services, the ultimate purpose of that interpretation is to make the web **easier** and **more useful** for people. To fulfill the claim of increasing relevance and value, we need a way for every user to describe what they consider to be relevant and valuable to them – at any point in time, no matter how much that may change from one minute to the next. We have all been disappointed by the unfulfilled promise of software applications and utilities that claim to know what we are working on or what our preferences are evidenced by our continuing delivery of information and services that clearly illustrate an application's rudimentary and inadequate knowledge of users, circumstances and needs. Computers need to listen, and then reflect what they hear back into the data and metadata that drive applications.

“It takes a clever question to turn data into information, but it takes intelligence to use the result. Intelligence can create systems of enormous complexity, but it takes wisdom to determine which ones are worth the trouble.” (Wiener 1993, p.209)

As the real evidence of semantic systems emerges, we have to continually evaluate ourselves against this question: how do we manage the risks so that we end up with meaning, rather than a flood of meaningless complexity?

References

- Agre, Phil (1994). Living Data. *Wired*. Volume 2.11, November 1994. Online: <http://www.wired.com/wired/archive/2.11/agre.if.html>
- Behrens, C. and Kashyap, V. (2002). The 'emergent' Semantic Web: a consensus approach for deriving semantic knowledge on the Web. In: I.F. Cruz , S. Decker , J. Euzenat and D.L. McGuinness, (eds.) *The emerging Semantic Web: Selected papers from the first Semantic Web Working Symposium*, pp. 55-74. Amsterdam: IOS Press. (Volume 75: Frontiers in artificial intelligence and applications). December 2002. Online: <http://www.semanticweb.org/SWWS/program/full/paper29.pdf>.
- Biezunski, M. and Newcomb, S.: Eds., TopicMaps.Org Authoring Group (2000). XML Topic Maps (XTM) Processing Model 1.0. December 2000. Online: <http://www.topicmaps.org/xtm/1.0/xtmp1.html>.
- Broekstra, Jeen and Kampman, Arjohn (2003). Inferencing and Truth Maintenance in RDF Schema. In Volz, R., Decker, S., and Cruz, I. (Eds.) *Proceedings of the First International Workshop on Practical and Scalable Semantic Systems*. October 2003. Online: <http://CEUR-WS.org/Vol-89/>.
- Chandler, Daniel (1994). The Transmission Model of Communication. University of Wales, Aberystwyth. Online: <http://www.aber.ac.uk/media/Documents/short/trans.html>.
- Cummings, H.W., Long, L.W., Lewis, M.L. (1983). *Managing Communication in Organizations: an Introduction*. Dubuque, IA: Gorsuch Scarisbrick Publishers.
- Degler, D. and Battle, L. (2000). Knowledge Management in Pursuit of Performance: the Challenge of Context . *Performance Improvement Journal*, 39(6), July 2000, 25-31. Online: <http://www.ipgems.com/writing/kmcontext.htm>.
- Euzenat, Jérôme (Ed.) (2001). Report from the NSF-EU Workshop: *Research Challenges and Perspectives of the Semantic Web*, Sophia- Antipolis, France. October 2001. Online: <http://www.ercim.org/EU-NSF/semweb.html>.
- Garshol, Lars Marius (2002). Topic maps in content management: the Rise of the ITMS. October 2002. Online: <http://www.ontopia.net/topicmaps/materials/itms.html>.
- Gell-Mann, Murray (1994). *The Quark and the Jaguar - Adventures in the Simple and the Complex*. Great Britain: Little, Brown and Company.
- Heflin, Jeff (Ed.) (2004). OWL Web Ontology Language Use Cases and Requirements (section 3.2: Ontology Evolution). W3C Recommendation. February 2004. Online: <http://www.w3.org/TR/webont-req/#goal-evolution>.
- Pepper, Steve (1999). Euler, Topic Maps, and Revolution. March 1999. Online: <http://www.infoloom.com/tmsample/pep4.htm>.
- Preece, J. (2000). *Online Communities: Designing Usability, Supporting Sociability*. Chichester: John Wiley and Sons.
- Shannon, Claude E. & Weaver, Warren (1949). *A Mathematical Model of Communication*. Urbana, IL: University of Illinois Press.
- Wiener, Lauren Ruth (1993). *Digital Woes: why we should not depend on software*. Boston, MA: Addison-Wesley Longman.
- Wenger, E., McDermott, R., Snyder, W.M. (2002). *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Massachusetts: Harvard Business School.

About the Authors

Duane Degler, Pensare Group/IPGems

Email: ddegler@ipgems.com

Duane Degler has consulted on a broad range of organizational performance issues, with experience in user-centered design, information/knowledge strategy, process analysis, international system implementation, training, and multimedia. He is currently involved in the development of the Pensare Group, as well as supporting clients such as the Social Security Administration. His work and writings are increasingly focused on design and content/knowledge management related to the adoption of semantic technologies. He has worked with government agencies, legal firms, insurance companies, major auto manufacturers, and several health organizations. His interface/system design work has received three US awards. He has managed pioneering multimedia projects in the '80s and spent much of the '90s in the UK involved in knowledge management research and consulting, contributing to his multi-disciplinary approach to interaction design. Duane holds a BS in Broadcast Communications and an MS in Organizational Communications.

Renee Lewis, Pensare Group

Email: renee.lewis@pensaregroup.com

Renee Lewis has over 18 years of experience developing and delivering object and Internet-based solutions for government and commercial clients. She is currently President and Founder of the Pensare Group, LLC, a consulting company specializing in technology adoption and communication strategies. Today, she supports clients such as the Social Security Administration with a contextual retrieval engine, and three other advanced research companies focused on bayes net, analytic search engines, and collaboration portals. She has held leadership positions in technology companies focused on market areas such as drug safety, e-Learning, health care, and telecommunications. She also spent 10 years as a consultant for Booz, Allen and Hamilton, focused on implementing distributed database architectures for both commercial and government clients. She holds a BS from the Pennsylvania State University and an MS from George Washington University in software engineering.