Just For Me: Topic Maps and Ontologies

Jack Park
SRI International
Menlo Park, California 94025
jack.park@sri.com

Adam Cheyer
SRI International
Menlo Park, California 94025
adam.cheyer@sri.com

ABSTRACT

The development of the IRIS semantic desktop platform has provided illumination of some important issues associated with the collection and manipulation of knowledge assets that are organized by an ontology. We explore those issues related to the personalization of the workspace and of the knowledge assets manipulated by IRIS users. We show that a topic map can provide a necessary mediation between the formal organization provided by an ontology to serve the needs of semantic interoperability between workstations and the individual's need to personalize the workspace in a *just for me* fashion.

Keywords

Topic map, ontology, semantic desktop, semantic interoperability, knowledge assets, personalization

1. INTRODUCTION

> *Experience of life has taught me that the only thing that is really desirable without a reason for being so is to render ideas and things reasonable.*
>
> –C.S. Peirce, *Science* 20 April 1900

Semantic desktop applications [1] exist to facilitate productivity and creativity in knowledge work. Two key use cases facilitated by such applications are *finding* and *reminding*. Both finding and reminding services derive from means by which information resources are organized and turned into knowledge assets. Ontologies are created to provide organizational guidance for local and for networked knowledge work. A new semantic desktop system called IRIS[1] [14] we are building at SRI has provided an opportunity to observe the onset and evolution of an interesting human computer interaction (HCI) issue. In this paper, we wish to share an interesting finding, a kind of tension that grows out of two distinct requirements for tools applied to knowledge work. Those requirements call for semantic interoperability between knowledge applications, and for user personalization of the workspace, something we label "*just for me"*. IRIS can be cast as a kind of topic map for personal knowledge assets, assets which must exist in a networked community. It is in that topic maps context that we find an opportunity to realize a candidate solution to this HCI issue.

To anticipate, the term "just for me" refers to the notion that an individual's workstation must satisfy the user's need to work in a *familiar* environment, describing (naming and relating) things in ways which are familiar to the individual, possibly less familiar to the networked community. This tension arises from the nature of a continuum which separates two concerns. On one end of that continuum lie representations of the objective universe defined by a consensus and empirical ontology, while at the other end lies representations of the subjective universe of individual users. The concerns are those of semantic interoperability and of usability. If the poles of that continuum are orthogonal, they are more so for some users than others. The closer a user is to the objective end, as the author of the ontology at the objective end would be, the lower the tension; *just for me* might not be an issue at all some users.

With this paper, we hope to open a discussion that centers on HCI, particularly in the context of knowledge workstations such as IRIS. During the ongoing development of the IRIS platform, we continue to bump up against a kind of barrier, one associated with user experiences with IRIS. The barrier we encounter is a simple one: users bring a lifetime of personally learned *ontology*, an instrument of knowledge organization, to their day-to-day activities with IRIS, while IRIS brings a different ontology, one created with the specific intent to facilitate semantic interoperability across a network, to the very same activities. As we shall show, the personal ontology of the user is often not sufficiently similar to the IRIS-supplied ontology; cognitive dissonance and unsatisfying user experience ensues.

We will argue that the presence of a *personal topic map* can serve as a mediator between the needs of a satisfying user experience with those of semantic interoperability. We suspect that HCI will eventually rise to be at least as important to the success of semantic desktops as is semantic interoperability among platforms. We believe that it is the specific relationships that topic maps forge between

[1]IRIS: http://www.ai.sri.com/software/IRIS

subject identity and *names for things* which facilitate the mediation process. Let us look closer at the issue.

2. DISCUSSION

> *The end goal of all of this research, design, testing and rumination is not just a software system that is easy for people to get. (If that were the case, let's just give people 1s and 0s, it doesn't get much simpler or more generic than that.) The end goal is a data structure that sits firmly upon the deep-seeded, some might say, hard-wired, natural structures of the human information architecture. The stuff of linguistics and grammar.*
>
> –Mimi Yin [12]

> *Clarifying the lens is more primordial than any particular perceiving or acting*
>
> *–Mark Szpakowski*[2]

In order to frame a discussion about the tension we observe to exist between a user's needs and those of semantic interoperability, let us imagine three conceptual spaces, one which is associated with all the information resources directly or indirectly available to a user, one which is associated with a *model* of those information resources, rendering an otherwise heterogeneous information space into an organized, classified body of information, and the last space, which is the user's *lens* or *view* into the other two spaces. We sketch those three spaces in Figure 1[3].

We have given those spaces the labels

- Documents—the space of all information resources
- Knowledge Structures—structured representations of the information resources
- Topic Maps—a user's *lens* into the other two spaces

The illustration does, indeed, have some of its ancestry in those marvelous illustrations in Steve Pepper's "The TAO of Topic Maps" [6]. But, Figure 1 is different in the sense that it injects an ontology layer between information resources and the topic map. Such a separation, by no means, implies that the topic map does not point into those information resources. Rather, it suggests that there is a marriage of ontologies and topic maps as suggested in Bernard Vatant's paper "Ontology-driven Topic Maps" [7]. Such a marriage contrasts with the case where the topic map *is* the ontology, as described by Eric Freese in Chapter 13 in [2] and by H. Holger Rath in Chapter 14 [2].

[2] Mark Szpakowski: http://collab.blueoxen.net/forums/yak/2005-08/msg00040.html#nid07

[3] This diagram was first conceived during discussions by the first author with Mary Keeler and Howard Liu. In some sense, it represents an interpretation of a Peircian view of inquiry space.

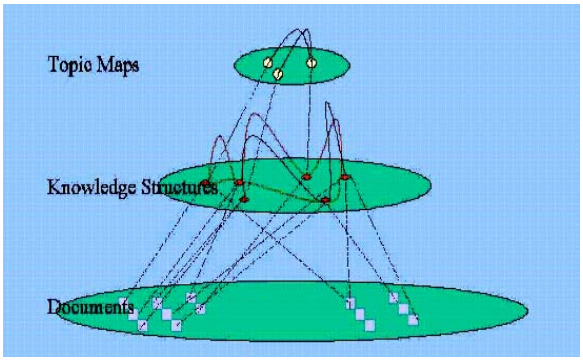

Figure 1: Conceptual Spaces for Information Architecture

The spaces just sketched relate to the issues we develop here in the following sense. There exist the dual needs of *user personalization* of the workspace, and *semantic interoperability* between the databases maintained by individual installations when applied in group settings. In some sense, the *user interface*, labeled a Topic Map in the illustration, can present the universe of information resources available to the user in a *just for me* fashion. The desktop application needs to present those information resources to external applications in a *consensus reality*, we say, semantically interoperable fashion. In both cases, *subject identity* must remain based on consensus reality.

2.1 Just for me

Just for me is perhaps best viewed as a fractal concept, one that starts at the individual level. Owing to the many relationships between individuals and their work groups, "just for me" morphs into "just for us" (locally), which then morphs (onion skin fashion) to "just for us" (globally). The Topic Maps layer of Figure 1 delineates the context for what follows. It is precisely the topic map, perhaps, the entire user experience associated with interacting with IRIS that constitutes the "just for me" discussion that follows. To anticipate the arguments, "just for me" suggests that the topic map should be user-constructed. The ontology (Knowledge Structures of Figure 1), in the case of IRIS, is constructed by teams of researchers; it could have just as well been constructed by a committee created by an enterprise that uses installations of CALO; indeed, it could be constructed by an individual user in the first place, but that's neither the case for CALO, nor the context of this discussion. The issue is this: entities other than individual users craft the ontology, and users craft their own topic maps. In some cases, the ontology can provide all that is necessary to populate such a topic map; in other cases, the user might see things differently.

It is probably useful to digress for just a moment. Consider the physician, or the scientist, or the lawyer. For those individuals, the ontology is probably already *just for me*. Such users operate closer to the objective end of the continuum mentioned above. They invented it in the first place, and they live in it. Certainly, they might personalize aspects of it, but, for the most part, their ontology is their normal universe of discourse. For the rest of us, for the

office users of IRIS, the ontology might or might not represent the individual's way of knowing. Office workers don't generally invent ontologies; for them, knowledge engineers provide the ontologies.

The story behind "just for me" is the story behind constructivist epistemology, which suggests[4] that constructivism is about focusing on *personally constructed* reality as opposed to *ontological* reality, where ontological reality might reflect either fiat or consensus reality. The central notion is that people construct their own reality through social interaction; they construct their own names for things that are identified in social settings, and they construct relationships between those things. It follows that users of semantic desktop workstations are going to have their own way of organizing what they know and their own names for things with which they interact. The tension, as we see it, lies in the fact that the user's constructed reality must co-exist with the group's consensus (ontological) reality. In some cases it co-exists, and in others, tension remains.

Consider a short story. Jared Spool[5] recently spoke to a web designers' meeting, speaking about the conclusions he was able to draw by observing web users purchase cameras online. The problem was framed in the context of a sum of money given to the purchasers, comparing the sum given to the amount of money actually spent on a purchase. He observed that, at most websites—for example, amazon.com—something like 110% of the money given for the purchase was spent on the purchase. One particular website stood out by capturing far more money in the purchase than was originally budgeted by the buyer—the consumer spent more than allocated for the purchase. The analysis is revealing. It goes like this. In spite of all of the many faceted attributes of the cameras displayed, consumers were mostly interested in the pictures they would get from the camera, not with number of pixels, lens speed, and so forth. To most consumers, cameras were a vehicle to get pictures. The site that achieved large sales was the site displaying photos taken by the camera. That site had discovered a *just for me* "sweet spot". The faceted attributes were still there, but those attributes were not the ones those consumers were using to make the purchase. A moral drawn from that story is this: ontologists can design elegant taxonomies, and those may, or may not, satisfy the needs of the consumer (user). It may be that the ontologists simply cannot be expected to "think like a user" or anticipate all of the many ways in which users categorize their personal universe.

The *name for things* issue animates discussions of some existing and emerging techniques. The distinction between name and identity comes together through the emerging notion of *folksonomies*,[6] and the well-known *faceted classification*[7] [12]. While there are some institutional faceted classification schemes in existence,[8] as for example DMOZ,[9] other such schemes are crafted as needed by individuals and group. Those individual activities are indicative of the human need to individualize, or *personalize*, the ways in which information resources are named and identified.

Consider the *delicious*[10] website, where users are encouraged to "tag" various websites. Jon Udell explains[11] this tagging process as a means to associate names for things, websites in this case. He goes further and casts this as a *language evolution* process. Constructivism, indeed! Tagging, as a process, is not without its critics. For instance, L. Jeffrey Zeldman argues[12] that tagging replaces user-centered taxonomies with usage frequency. Tag *clouds* form which separate related topics from each other. Another way to look at tagging is that it is social constructivism at work, where individuals are practicing their "just for me" behaviors, and evolving those behaviors during social interactions. Indeed, Mimi Yin, who works on the Chandler[13] project, has much to say about these issues. In her recent paper "Hierarchies versus Facets versus Tags" [11], she argues, among other things, that hierarchies can become "messy" and unmaintainable, and that facetted systems don't offer some of the benefits of hierarchies and become too flexible to be useful.

Yin's writing constitutes anecdotal evidence that open questions remain. Adam Mathese, in his paper "Folksonomies—Cooperative Classification and Communication through Shared Metadata," [13] concludes his discussion with the notion that a folksonomy is fundamentally uncontrolled in nature, suffering imprecision and ambiguity, but at the same time, free-form tagging allows for self-organization of information resources.

In the context of IRIS and "just for me", we take the view that tagging really is just for the individual user, and not necessarily for the group, unless the group dynamic chooses to encourage it. The role of a topic map with groups of IRIS users would be to permit personal topic maps to reflect personal naming conventions locally while linking to a group topic map, say, on a server, which reflects consensus naming conventions. Mappings between the two are mediated by subject identity properties reflected in the consensus ontology. At the desktop level, we are certainly not alone in this notion. Murray Altheim has implemented faceted classification in Ceryle [18], a semantic desktop application created to organize those knowledge assets necessary for authorship.

---

[4] Statement adapted from http://en.wikipedia.org/wiki/Constructivist_epistemology

[5] Jared Spool: http://www.uie.com/

[6] Folksonomy: http://en.wikipedia.org/wiki/Folksonomy

[7] Faceted Classification: http://en.wikipedia.org/wiki/Faceted_classification

[8] Just about every online catalog uses facets, e.g., price, shipping weight, lens speed, color, etc.

[9] DMOZ: http://www.dmoz.org/

[10] delicious: http://del.icio.us/

[11] Jon Udell: http://weblog.infoworld.com/udell/gems/delicious.html

[12] Zeldman: http://zeldman.com/daily/0505a.shtml

[13] Chandler: http://www.osafoundation.org/

2.2 Tensions between "ontology-driven" and personal topic map

> *…when thinking about ontologies and semantic web it is easy to focus on the requirements of precision and data integration to the exclusion of the requirements for end user navigation*
>
> –Dave Reynolds, et al. [17]

Semantic interoperability demands consensus identity, and some consensus names; personal views, *just for me*, call for personal naming conventions alongside consensus subject identity. Let us contrast these views:

- Separate ontology and topic map
  - Topic map, itself, doesn't necessarily have to support consensus names for things
- Topic map as ontology
  - User's *lens* topic map directly supports user names for things along with consensus names for things

If IRIS is to have a topic map,[14] then, in the near term, it will be separate from the ontology. It is not inconceivable that some future version of IRIS could migrate to a topic map as ontology architecture. For now, the tension exists because IRIS, indeed, all semantic desktop workstations, are, essentially, *just for me* platforms and the intuition is that a topic map can mediate between the interoperability need for a consensus ontology and the user's need for a personal lens into that workspace.

2.3 A Solution: Topic Maps as Mediators

Consider a patient-doctor scenario, the context of which is a vision problem where the patient says "…the world is going *Picasso* on me." The doctor replys with "Well, we call that syndrome *scintillating scotoma*; it's one of the many kinds of migraine headaches people can experience." The doctor is working from a medical ontology, and, based on years of experience, has a pretty good idea of how patients are liable to describe their signs and symptoms. In the simple *office world* of IRIS, let us imagine some office worker relating the proper name *Joseph P. Sixpack* to a favored way to recall that person, say, *Bubba*. The rest of the world doesn't need to know about "Bubba", but, if association of a specific person to a favored name renders the user's life simpler, less hectic, and more productive, then it's useful to provide for such mediation.

[14] A small, focused topic map existed in the original prototype, but was set aside to facilitate evolution of other functionality. A focused topic map means that a small map of specific aspects of the user's world was created as the user entered that world. For instance, a general topic map of all people and of all projects existed in the background, and fragments of those maps appeared, for instance, in the tasks view.

Topic maps offer a simple *paradigm* from which a candidate solution to the *just for me* issue emerges. The topic maps paradigm simply suggests that, for each *subject*, there can exist in a given topic map, at most, one *topic*, where a topic is best-imagined as a container, nexus or proxy for the subject, at which all presently-knowable information related to the subject can be found. The nexus concept evokes the image of a topic as a *hub* around which all information resources related to the subject radiate.

While topic mapping, the paradigm, itself, continues to evolve, there remains a core set of notions, a core ontology which guides application developers. Core to topic mapping are the elements: *topic*, and *association*. With those two core elements, a slightly broader ontology grows. We draw on the concept of *scope* to construct a solution to the *just for me* issue.

A scope represents a *context*, and, as such, is, itself, a subject represented by a topic. Names for topics can take scopes. That is, a particular name for some subject might be a string written and scoped in the English, a string written and scoped in the German, and name strings written in other languages, all representations of a name for the subject. Scopes can also be used to provide context other than language. Scoping *Bubba* as a *private* name string in a topic map might mean that this name string will not appear in the public transmission of data between IRIS installations. Using *scoped names* thus permits a user to inject favored names for things.

Using *scoped associations*, a topic map permits a user to inject favored relationships into the knowledgebase without risk of those relationships altering the semantic interoperability of the larger ontology. Topic maps thus provide a useful means by which users of ontology-driven workstations can personalize their working environment.

Implementing a topic map in a semantic desktop application as a means of satisfying a *just for me* requirement avails other potential benefits. For instance, the inherent indexical capabilities of topic maps are suited to many of the needs already satisfied by the ontology. That opens the door to two larger questions: *could a topic map satisfy the need for an ontology*, and *could the ontology satisfy the need for a topic map*? Both, great questions, and each suggests avenues for future research. While other workers are already exploring those questions, we believe the opportunity to implement an ontology as a topic map remains an important opportunity since that one structure can satisfy both semantic interoperability and just for me requirements. For the present work, it is shown that a topic map can mediate between two important needs, those related to the user and HCI, and those related to semantic interoperability between workstations.

We have discussed personalized names for things and personalized relationships between things. Another issue is, and will remain for a long time in the future, that of subject identity. For instance, IRIS exists in an email-rich universe, where the names of unknown (to IRIS) persons appear frequently. CALO provides a framework wherein new

persons are isolated and studied by a variety of means, the intent being to disambiguate identities. For instance, one email might come in from, say, `jpark@foo.org` and another from `jackpark@bar.com`. A question is this: do both emails refer to the same individual? That is a nontrivial question, and it mirrors the subject identity issues facing topic maps during merging processes. IRIS includes a harvesting framework which includes some tools for name resolution. More powerful tools are included in CALO.

Where does IRIS presently stand in relation to resolving the *just for me* issue? We have cast IRIS as a kind of topic map for personal knowledge assets, assets which must reside in a networked community. We are not claiming that IRIS *is* a topic map in the sense understood in terms of XTM documents, merging tools and so forth. Rather, IRIS continuously orbits in the space of tensions between the necessity to use an ontology to organize information resources for purposes of interoperability between software agents involved in processing those resources and other semantic desktop installations, and between the user's need semantic desktop installations, and between the user's need to capture individual ways of knowing and doing.

3. BACKGROUND

There are several threads related to the background from which IRIS, topic maps, ontologies, and the issue we found. In the end, they all relate to the ever-increasing rate at which sound decisions must be made in the context of ever-increasing amounts of information to process in order to achieve those decisions. Out of the need to index a growing body of software documentation (one form of infoglut), topic maps were created. As Steven R. Newcomb said in his introduction to topic mapping in the book *XML Topic Maps* [2, page 32],

> *Information is both more and less real than the material universe. It's more real because it will survive any physical change; it will outlast any physical manifestation of itself. It's less real because it's ineffable. For example, you can touch a shoe, but you can't touch the notion of "shoe-ness" (that is, what it means to be a shoe). The notion of shoe-ness is probably eternal, but every shoe is ephemeral.*

As topic mapping technology matures and enters mainstream application in ever more complex indexical and organizational situations, user interface issues bubble to the foreground. Indexical and associative applications require attention to two core issues: *subject identity* and *names for things*. Recent innovation on the web,[15] the notion of *social bookmarking*, for instance, is pointing the way toward a web that satisfies a *just for me*[16] requirement. The combination of rising popularity and high level of innovation in this arena strongly suggests that the *just for me* requirement should be investigated at the desktop application level, along with the web.

We now live and work in a networked global village; the term *infoglut* has become the meme that reminds us of the information overload we experience in our daily lives, and about which Vannevar Bush eloquently wrote in his 1945 paper, "As We May Think" [10]. Indeed, it was that paper which inspired Ted Nelson, Douglas Engelbart, and many others to try to find solutions to the infoglut problem and augment human capabilities for solving complex, urgent problems.

IRIS has been developed as part of SRI's CALO[17] project, one of two projects funded under DARPA's "Perceptive Assistant that Learns" (PAL) program.[18] The goal of the PAL program is to develop an enduring personal assistant that "learns in the wild," evolving its abilities more and more through automated machine learning techniques rather than through code changes. In approaching the design and development of IRIS, we took much inspiration from the work of Douglas Engelbart, who performed much of his early work while employed at SRI. While Ted Nelson's Xanadu[19] [3] was arguably the first project to set the stage for modern hyperdocument processors, Engelbart's Augment[20] was the first system to find engagement in group document processing and sharing. In 1968, at the Fall Joint Computer Conference in San Francisco, Engelbart demonstrated Augment before a live audience.[21] Augment displayed many of the capabilities we now want to build into modern semantic desktop applications. Augment, the program, saw commercial application, and is still used today by Dr. Engelbart in his day-to-day activities.

Central to our work is the *augmentation program*, first proposed by J.C.R. Licklider (who funded Engelbart's work) in 1960 [5]. The emphasis of that program was to augment human capabilities with computers, as we see in the Engelbart work, as then compared to the *artificial intelligence program* (AI), in which human capabilities are mimicked or otherwise provided by computers. CALO represents a blending of the AI and the augmentation programs.

There is a clear and vibrant link between topic maps and the augmentation program. We see opportunities for that link in IRIS, because the program integrates several desktop office productivity tools, such as email, web browsing, calendar, instant messaging, and more. At the

[15] E.g. http://www.flickr.com/ and http://del.icio.us/

[16] The term *just for me* was first introduced to the first author by Nancy Glock-Grueneich in the context of pedagogy. The context is this: regular school learning is sometimes described as "just in case"; by contrast, constructivist learning is described as "just in time", and Nancy suggests that real learning is best described as "just for me."

[17] CALO is an acronym for "Cognitive Assistant that Learns and Organizes." CALO's name was also inspired by the Latin word *calonis*, which means "soldier's servant" and conjures an image of Radar O'Reilly from the M*A*S*H TV series.

[18] DARPA's PAL program: http://www.darpa.mil/ipto/programs/pal/

[19] Xanadu: http://xanadu.com/

[20] NLS/Augment at the Computer History Museum: http://community.computerhistory.org/scc/projects/nlsproject/

[21] Videos of the first online document editing project. Found on the web at http://sloan.stanford.edu/MouseSite/1968Demo.html

same time, IRIS provides a framework that supports aspects of artificial intelligence and machine learning, all in support of aiding the user in assembling, indexing, clustering and otherwise organizing a growing body of knowledge assets.

In order to better understand how IRIS can be cast as a topic map for personal knowledge assets, we now briefly sketch those aspects of IRIS that make up the letters in the name. IRIS is first and foremost an *integration* framework. Whereas in today's packaged applications suites, where only loose data integration exists[22] (usually limited to the clipboard and common look-and-feel for menus and dialog boxes), IRIS strives to integrate data from disparate applications using reified semantic classes and typed relations. For instance, it should be possible to express that "*File F* was presented at *Meeting M* by *Tom Jones*, who is the *Project Manager* of *Project X*," even if the file manager, calendar program, contact database, and project management software are separately developed third-party applications. In a Topic Maps fashion, there should be a single instance that represents each concept, and all that is knowable about that concept should be directly accessible from that instance [2].

The IRIS framework offers integration services at three levels (Figure 2):

1. Information resources (e.g., an email message, a calendar appointment) and the applications that create and manipulate them must be made accessible to IRIS for instrumentation, automation, and query.
2. A knowledge base (KB) provides the unified data model, persistence store, and query mechanisms across the information resources and semantic relations among them.
3. The IRIS user interface framework allows plug-in applications to embed their own interfaces within IRIS, and to interoperate with global UI services.

The IRIS user interface provides the "shell" for hosting several embedded applications (Figure 3). Two side panels frame the main application window, one for selecting among available applications, the other for displaying and editing semantic links for the selected application object and presenting contextual suggestions from the learning framework. Applications can add toolbars to the IRIS frame, and when selected, an application's menu items are "merged" with IRIS menu functionality present for all applications. IRIS provides an extensible context-sensitive online help system and several methods for querying information resources within and across applications.

IRIS is used to semantically integrate the tools of knowledge work, to form *relationships* between knowledge assets. What do we mean by this? We use the term "semantic" in the sense used by the Semantic Web community, where markup technologies are being wedded to the tools of semantic representation (e.g., ontologies, OWL, RDF). This facilitates putting data on the web in such a way that machines can access it, make meaningful references to it, and perform manipulations on it, including reasoning and inference. In that sense, IRIS provides an OWL-based ontology and backside by which the artifacts of a user's experience such as email messages, calendar events, files on the disk or found on the web, can be stored and related to each other across applications and across users.

When defining the ontology to be used for IRIS, a design choice had to be made: Do we use a small, simple ontology or a complex, more-expressive ontology? We first implemented a fairly large, yet straightforward, ontology. However, the requirement that IRIS interoperate with CALO's reasoning and learning capabilities drove us to adopt CALO's preexisting ontology, which supports roles, events, and complex data structures.

Additionally, IRIS provides a framework for *harvesting* application data and *instrumenting* user actions in IRIS applications. The harvesting of data refers to importing external data into semantic (ontology-based) structures.

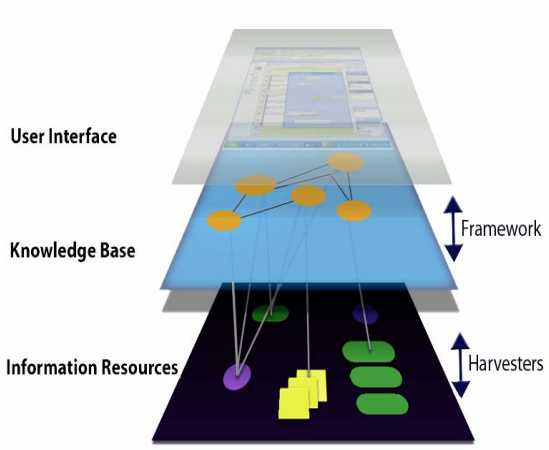

Figure 2: IRIS Architecture

One of the key differentiators of IRIS, compared to many semantic desktop systems, is the emphasis on machine learning and the implementation of a plug-and-play learning framework, providing the ability for IRIS to make *inferences*. We see machine learning as one of the solutions around a key issue limiting the Semantic Web's growth and mass adoption: Who is going to enter all of the required links and knowledge?

> *Prior to the Internet, the last technology that had any real effect on the way people sat down and talked together was the table.*
>
> –Clay Shirky[23]

[22] Even within a single application, deep data integration is usually pretty threadbare. Consider Microsoft Outlook: the email addresses displayed in a message are not linkable (or deeply related) to the people records in your contacts folder.

[23] Clay Shirky: at Emerging Technology Conference 2003 http://shirky.com/writings/group_enemy.html

Sharing information is one of the four key concepts that make up the IRIS vision. We feel that the ability to learn and leverage semantic structure in organizing one's work life will be greatly enhanced in a collaborative setting. Shared structures are essential for both end-user applications, such as team decision making and project management, and for infrastructural components such as machine learning algorithms, which improve when given larger data sets to work on.

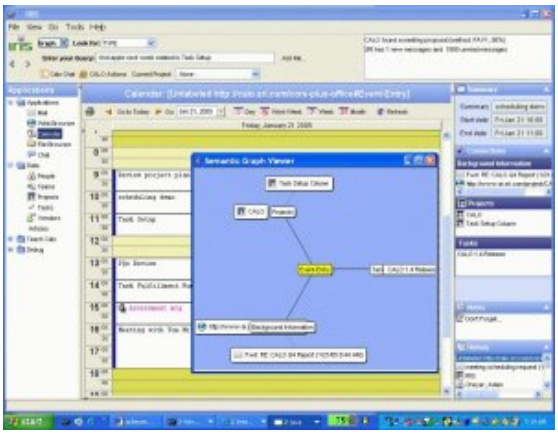

Figure 3: The IRIS Platform

4. RELATED WORK

Several projects exist and are similar in spirit and intent to IRIS. We sketch a few of them here,.Two projects of great significance to the personalization of information resources are Lifestreams[24] and WorldBoard.[25] Lifestreams is the vision of David Gelernter, and was developed as a dissertation project by Eric Freeman [15]. WorldBoard is the outgrowth of the vision of Jim Spohrer from his paper "Information in Places" [16]. Whereas Lifestreams speaks mostly to local individual needs, WorldBoard speaks to Global individual and group needs.

While developing IRIS, we explored Haystack[26] from MIT. When we discovered this project [8], we were amazed how well it fit our initial designs for IRIS, in terms of both architecture and user interface design, with the added benefit of being Java-based and open source. We learned much from a visit by Dennis Quan, one of the principal developers, and did, indeed, begin the task of adapting Haystack's significant code base to our framework. For a variety of reasons, we ended up moving in a different direction, but Haystack and Dr. Quan's deep knowledge of the subject gave us a solid start.

The next system we evaluated was the Radar Networks[27] Personal Radar, a very impressive semantic desktop that turned out to share many of the goals and requirements for IRIS: Java-based, ontology-driven, user-centric. We have combined elements of Personal Radar into the IRIS code base.

Well down the path of implementing IRIS, we discovered two additional projects. Gnowsis.[28] Gnowsis [9] appears to offer integration with many of the same third-party applications as IRIS, and to share many similar philosophies regarding application and data integration. MindRaider,[29] is a project arguably close to IRIS, Haystack, and Gnowsis in spirit and intent.

Ontologies have become the *lingua franca* of semantic interoperability. Programs such as Haystack, IRIS, and others are, essentially, driven by ontologies. This means that operations by users, data items such as email messages and calendar events, and communications between different platforms, are all performed in the context of a built-in ontology. Outside of desktops, SHOE [4] represents an early approach to providing for semantic interoperability on the web. SHOE is an ontology-based language. Today, the OWL web ontology language is rapidly becoming a standard for representing ontologies. The IRIS ontology is implemented in OWL.

5. CONCLUSIONS

IRIS is the platform that allowed its creators to experience the tensions we have discussed here. The system is now in daily operation as the primary office environment used by several members of the CALO community. In that daily use, and during trials by developers, we continue to encounter users who wish they could add tags or otherwise provide names for objects, or forge relationships between objects such as files on the hard disk and emails or calendar events. A growing awareness of this issue is propagating some rethinking regarding the design of the IRIS knowledgebase. This rethinking allows for the opportunity to implement a topic map linked to the ontology and driven by the user interface. With the addition of a topic map to IRIS, users will be able to provide their own names and relationships, and still remain semantically interoperable with the rest of the community.

We have argued that a *personal topic map*, one that mediates between a user and an ontology, is a candidate solution to the *just for me* issue. We believe that our work with IRIS and CALO supports our claim that semantic desktop applications which use ontologies for semantic interoperability can benefit from the application of personal topic maps. We offer a concluding conjecture that the topic map, itself, might server the role of the ontology, providing both semantic interoperability and *just for me* user support.

[24] Lifestreams: http://www.cs.yale.edu/homes/freeman/lifestreams.html
[25] WorldBoard: http://www.worldboard.org/
[26] Haystack: http://haystack.lcs.mit.edu/
[27] RadarNetworks: http://www.radarnetworks.com/
[28] Gnowsis: http://www.gnowsis.org/
[29] MindRaider: http://mindraider.sourceforge.net/

6. ACKNOWLEDGMENTS

This work was supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA or the Department of Interior-National Business Center (DOI-NBC).

We would like to thank those SRI and non-SRI CALO members, too many to name, who are working with us to integrate their cutting-edge technologies into the IRIS learning framework. We also would like to thank the anonymous reviewers, Glen B. Haydon, Patrick Durusau, Ray Perrault, and Josh Levy for comments on an early draft of this paper.

REFERENCES

1. Decker, Stefan; and Martin Frank, "The Social Semantic Desktop," 2004. On the web at http://www.deri.at/publications/techpapers/documents/DERI-TR-2004-05-02.pdf
2. Park, Jack, Editor, and Sam Hunting, Technical Editor, *XML Topic Maps: Creating and Using Topic Maps for the Web*, Boston, MA. Addison-Wesley, 2003
3. Nelson, Theodor von Holm, "Xanalogical Structure, Needed Now More than Ever: Parallel Documents, Deep Links to Content, Deep Versioning, and Deep Re-Use." ACM Computing Surveys 31(4), December 1999. On the web at http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/60.html
4. Heflin, Jeff and James Hendler, "Semantic Interoperability on the Web," 2000. On the web at http://www.cs.umd.edu/projects/plus/SHOE/pubs/extreme2000.pdf
5. Licklider, J.C.R., "Man-Computer Symbiosis", 1960. On the web at http://memex.org/licklider.pdf
6. Pepper, Steve, "The TAO of Topic Maps," Ontopia. On the web at http://www.ontopia.net/topicmaps/materials/tao.html
7. Vatant, Bernard, "Ontology-driven Topic Maps", XML Europe 2004, Amsterdam, The Netherlands. On the web at http://www.idealliance.org/europe/04/call/xmlpapers/03-03-03.91/.03-03-03.html
8. Karger, David R., Karun Bakshi, David Huynh, Dennis Quan, and Vineet Sinha, "Haystack: A Customizable General-Purpose Information Management Tool for End Users of Semistructured Data," in CIDR 2005, Asilomar, California. On the web at http://www-db.cs.wisc.edu/cidr/cidr2005/papers/P02.pdf
9. Sauermann, Leo, "The Gnowsis Semantic Desktop for Information Integration," in IOA 2005, Kaiserlauten, Germany. On the web at http://www.dfki.uni-kl.de/~sauermann/papers/Sauermann2005a.pdf
10. Bush, Vannevar, "As We May Think," The Atlantic Monthly, July 1945. On the web at http://www.theatlantic.com/doc/194507/bush
11. Yin, Mimi, "Hierarchy versus Facets versus Tags", July 11, 2005. On the web at http://wiki.osafoundation.org/bin/view/Journal/HierarchyVersusFacetsVersusTags
12. Yin, Mimi, "Preface to Hierarchy Papers", July 14, 2005. On the web at http://wiki.osafoundation.org/bin/view/Journal/PrefaceToHierarchyPapers
13. Mathes, Adam, "Folksonomies – Cooperative Classification and Communication Through Shared Metadata," December 2004. On the web at http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html
14. Cheyer, Adam, Jack Park, and Richard Giuli, "IRIS: Integrate. Relate. Infer. Share," submitted to the 1st International Workshop on The Semantic Desktop, Galway, Ireland, 6 November 2005.
15. Freeman, Eric Thomas, "The Lifestreams Software Architecture", Ph.D. Dissertation, Yale University, May 1997. On the web at http://www.cs.yale.edu/homes/freeman/dissertation/etf.pdf
16. Spohrer, James C., "Information In Places", IBM Systems Journal, Vol 38, No 4, 1999. On the web at http://www.research.ibm.com/journal/sj/384/spohrer.html
17. Reynolds, Dave, Paul Shabajee, Steve Cayzer, and Damian Steer, "SWAD-Europe deliverable 12.1.7: Semantic Portals Demonstrator- Lessons Learnt", Sept. 2004. On the web at http://www.w3.org/2001/sw/Europe/reports/demo_2_report/
18. Altheim, Murray, "Ceryle and "Grass Roots" Ontologies," Ceryle Blog, Feb. 10, 2005. On the web at http://www.altheim.com/ceryle/blog