The Intraspect Knowledge Management Solution: Technical Overview

Summary

The Intraspect Knowledge Management System provides a collaborative environment for knowledge work where people search for, collect, organize, share and collaborate around information. As they work in this environment, their work is captured and maintained as a *group memory*, making it available for knowledge sharing and reuse. Intraspect's group memory contains only the information that has been *put to use*, and it captures the *context of its use* — who collected it, when it was used, for what task, how it was combined with other information, and what people in the organization said about it. The Intraspect solution is a comprehensive integration of intranet technologies for creating, working in and harvesting a group memory: distributed object storage, full text search, desktop integration, web publishing, email processing and distribution, agent-based monitoring, and brokering to other intranet services. Information in the Intraspect repository includes documents on desktops and file systems, static and dynamic web pages on remote servers, email from any Internet-compatible client or server, and any other information source served by the Internet standards HTTP or SMTP

1. The Group Memory Approach to Knowledge Management

Intraspect KMS is a ready-to-use, integrated solution to enterprise knowledge management that taps the inherent power of the intranet for sharing and reusing information. Unlike hand-built libraries or custom applications, Intraspect captures and harvests the collective knowledge of an organization *as people work* in a networked computing environment. The result is a **group memory** — a repository of the information used in an organization, and meta-data about how it is used to solve business problems.

The key to an effective group memory is to provide an on-line environment that delivers productivity benefits to individuals and collaborative teams. People and teams can't be asked to do extra work to make knowledge available for use by others; they must find it personally useful to work in a shared environment. To achieve this effect, Intraspect offers an integrated family of services useful to knowledge workers: private and collaborative workspaces in which to organize and share information, integrated discussion lists, access control, document versioning and commenting on any object. A thin-client user interface makes it possible to collect almost any sort of information object from its natural source: web pages through browsers, email messages from email clients, and files and documents from the desktop.

To discover information in the group memory, Intraspect offers full-text and meta-data search over all object types across multiple information sources; one can search a personal workspace, a shared workspace or the entire enterprise. To keep informed of the information flows and collaborative activity in the group memory, users can assign

subscription agents to monitor searches, workspaces, discussions or any object in the group memory, and notify the user when something relevant happens in the group memory. Together, these services make Intraspect a compelling environment for individual knowledge workers and collaborating teams, and the resulting group memory becomes a self-sustaining intranet resource for the entire organization.

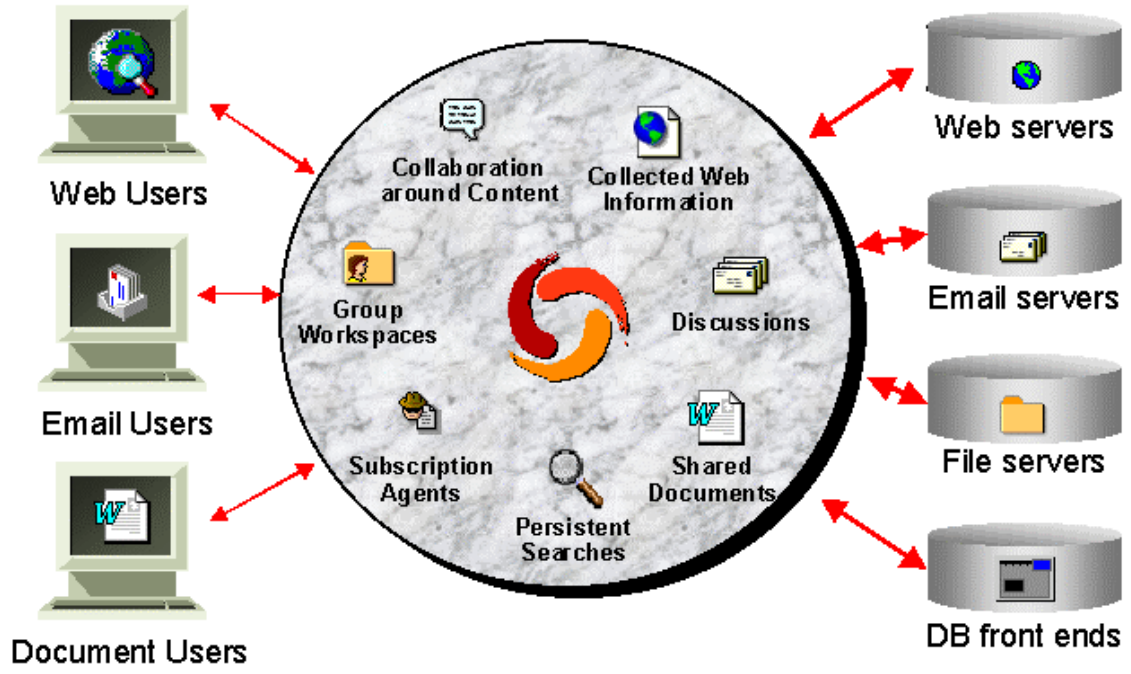

Figure 1. A Group Memory offers a context for working on on-line, mediating access to the major sources of information available in the intranet environment. The group memory is a clearinghouse and repository for web information, messages, files, databases reports, and any services that broker information, such as search engines. Users can interact with the group memory in a way that suits their preferred work style using a thin client, a web browser or an email application.

Representing the *Context* of Information is Key

Information has value as knowledge when it is put to use. A document discovered through search on the Internet is not as useful as the same information accompanied by its *context of use* in an organization, e.g., how it was used to solve a problem, answer a question, or make a decision. No software can *infer* the true purpose of a piece of information because to understand the purpose of information would require modeling every aspect of how people use information to solve problems, which is impossible today. However, Intraspect provides capabilities that help to *capture* contextual information without requiring people to maintain rigid vocabularies of keywords or worry about where to place a document in a large, centralized taxonomy. Intraspect provides users with convenient tools for organizing, and annotating their information, such as hierarchical folders, long titles, description strings, and threaded comments. When a document is viewed along with this contextual meta-data, an intelligent human can interpret the folder names, labels, descriptions, and annotations to make sense of the information.

Borrowing the well-known office metaphor from desktop file systems, Intraspect allows users to group objects into folders and cabinets and arrange them into hierarchies. If a document is in a folder we say that the folder is a ***context*** of the document. The notion of context is extended to objects other than folders — documents, other files, email messages, discussions and references to web pages hosted on a remote web server. All of these objects can appear in the same context, and can *be* contexts for other objects. Figure 2 shows an example.

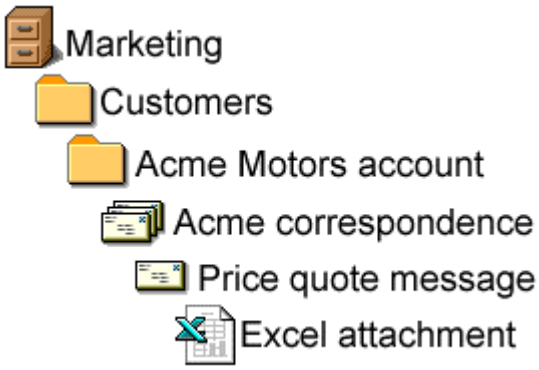

Figure 2: A hierarchy of contexts. The context of the Excel document is the email message to which it was attached. The context of the message is the ongoing discussion with the Acme customer. The context of the discussion is the folder on that customer. The context of the Acme folder is the folder containing all of the customer folders, and its context is the marketing department's file cabinet.

Furthermore, objects can have ***multiple contexts*** — the same object can appear in multiple places as it is re-used by different people to address various tasks. This is similar to the notion of "aliasing" in some file systems: an object can appear in several folders, although it is always the identical object. Allowing for generalized contexts and multiple contexts per object allows some interesting possibilities for the reuse of information. The same file that was once sent by email can now be put (linked) in a project folder. The key email message can be linked into a different folder as a document. A single web page can be "bookmarked" by many different individuals and workgroups, each applying the information to *their* respective contexts. (Figure 3 shows an example of two uses of the same web document by different groups.)

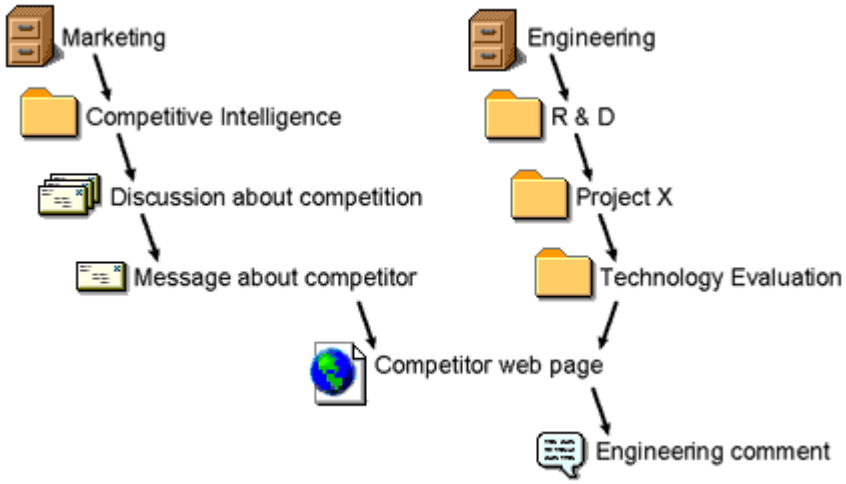

Figure 3: Multiple contexts of the same web page. The "Competitor web page" object has two contexts: an email message sent to a discussion run by Marketing, and a Technology Evaluation folder created by the Engineering team. The web page was originally collected by a marketing person, who sent it around to her colleagues in email. When an engineer collected the same URL, Intraspect linked the same object into a second context, the engineer's folder on "Technology Evaluation". Another engineer then read the document and made a comment about it. The comments made in the engineering context are kept separate from commentary in other contexts; marketing could hold their own threaded conversation about the same object. Sharing objects among multiple contexts encourages cross-organizational reuse while supporting group collaboration.

End users are part of the solution

Different work styles require different ways to organize information and capture context. Intraspect provides a variety of simple ways for *end users* to put information into contexts, without requiring them to fit their work into a centrally managed structure or rely on a programmer or webmaster to manage their information. ***Projects and virtual teams*** can use Intraspect public folders to share documents, plans, meeting notes, work in progress and deliverables. Intraspect discussions can be used to collaborate on project planning and execution. ***Departments*** can organize their official policy and process information into Intraspect cabinets with tidy folder hierarchies with descriptions about purpose and use. ***Customer relationships*** can be tracked using folders to capture customer data (including presentations made, meeting notes, and key contacts) and by using discussions to archive email correspondence with the customer. ***Supplier relationships*** can be tracked with folders to keep contracts, PO's, and web-based product descriptions. Comments made on these information objects can capture experience in working with the suppliers. ***Sales forces***, who "live in email," can receive timely information and share leads using Intraspect discussions and their personal email clients. The design philosophy underlying Intraspect assumes that empowering end users with simple tools for collecting and organizing their information will produce relevant, useful, and self-sustaining group memories.

2. The Intraspect Knowledge Management System

The goal of knowledge management is to make good use of the collective knowledge of an organization: to capture it, make it widely available, and make it easy to share and reuse. Organizational knowledge exists in many forms: in the heads of employees, in conversations flowing through the telephone circuits, in institutional practices and as written information. Intraspect deals with knowledge in the form of *on-line information put to use*. On-line information includes not only official company literature, but documents of all sorts, financial models, designs, demonstrations, email messages, and even web pages collected from the Internet. Information gains value as knowledge when it is put to use to solve problems, make decisions or create new intellectual property. It takes human expertise and labor — "knowledge work" — to find the relevant information and apply it to a task.

The purpose of knowledge management software is to make knowledge work more productive by improving the way information is acquired, distributed and used in an organization. People should be able to easily find colleagues, create a shared workspace, contribute to the project, and collaboratively discuss the evolving work. They should be able to make the results of their work available online, without spending extra effort to prepare it for publication. They should be able to easily discover and reuse the knowledge of others in the organization, without relying on others to anticipate their information needs and keep them informed. Intraspect is designed to meet *all* of these needs, supporting the work practice of knowledge workers while building a group memory.

Intraspect Functional Specification	
Knowledge Management Service	**Functionality**
Information Collection	End-user capabilities to import any object that is visible in a client web browser; upload any files visible to the client file system; send in any email message, with attachments, from any SMTP-compliant email application.
Information Sharing	On-line workspaces for collecting and organizing information, with versioning and access control. Personal and public spaces integrated into enterprise object repository to support individual usage and simplify wide-area collaboration. Integrated commenting to allow collaborative review of content, or for taking personal notes.
Information Organization	Hierarchical organization of information, with single copies of objects in multiple folders to support multiple contexts of information (re)use. Identity of objects is independent of their location in hierarchy, to support dynamic reorganization.
Integrated Discussion	Discussions held in the same workspaces where documents are shared. Messages automatically threaded and parsed for attachments. Existing discussion lists can be routed to Intraspect. Compatibility with all SMTP mail clients, and independent of mail server, to lower barriers to participation in on-line work.
Discovery and search	Automatic indexing for full text search of all collected documents including binary formats, email messages and remote web pages. Attribute search integrated with full-text search.
Universal Subscription	Subscription to any object, including workspaces, individual documents, and persistent searches. Notification of changes to subscribed objects delivered asynchronously to end-users.

Table 1: Intraspect Functional Requirements

2.1 Collection: Getting information into the group memory

When collected, an information object is sent to the server and becomes available for organizing, searching and all other knowledge management services. A user interface device called the *collection bin* makes it easy to collect from the user's desktop environment (see Figure 4).

Intraspect Knowledge Management is unique in supporting the collection of information from *all major intranet information sources* including web, email, and desktop applications.

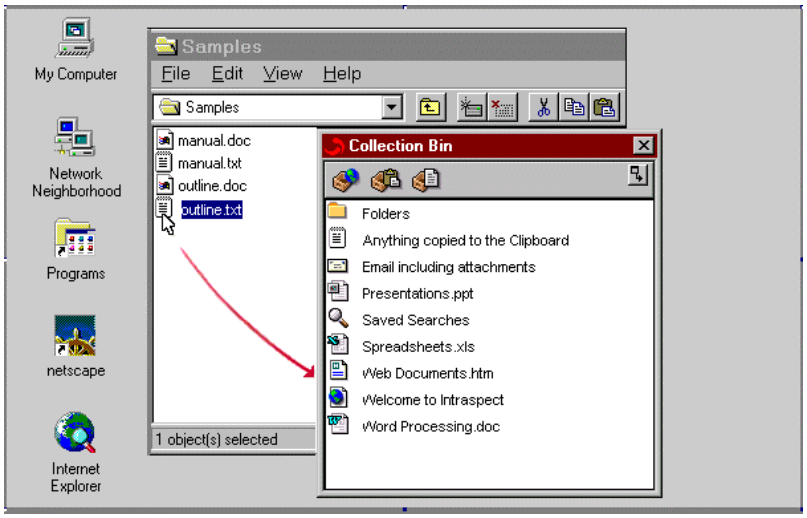

Figure 4. The Collection Bin is a general-purpose user interface for collecting information into the group memory. Any file or web page can be dragged into the collection bin. The leftmost button allows the user to collect any web page that is currently displayed in a web browser. The center button invokes an "open file" dialog, providing the ability to upload any files available to the user's machine including files from distributed file systems. The rightmost button is for collecting from the clipboard, which is a convenient way to collect snips of text or graphics (such as screen snapshots). Email messages and attachments can be sent directly to the collection bin from an email application. The collection bin is also used as a temporary holding place for moving objects around the folder hierarchy on the server.

Collecting from the web

The widespread acceptance of the World Wide Web has established HTTP as a standard for delivering all kinds of information to universal clients on every desktop. Intraspect allows users to browse and view the information in their favorite web client, and then collect pages with the click of a button. Any information object that is identified by an URL — including binary documents, multimedia presentations, and the results of interacting with dynamic web-based services — can be collected. When a web page is collected, Intraspect's server fetches the information, stores a copy in the object repository and indexes it for search. The original content remains hosted on the remote web server. If it changes, Intraspect updates its copy and notifies any subscribed user of the change

(subscription is described in Section 2.5). A snapshot feature (also described later) allows one to keep a series of copies of a web page that changes over time, such as a news service's front page.

Collecting from email

Many knowledge workers "live in email" — they use messaging to share knowledge and coordinate activity. Email messages are important documents in their own right, especially when combined with attached files and pointers to on-line sources. Unfortunately, most email messages end up in overloaded in-boxes or inaccessible archives and are not managed as a reusable knowledge resource. Furthermore, it is difficult for heterogeneous organizations to manage email content by standardizing on single-vendor email systems, especially when the participants include customers and suppliers outside the organization.

Intraspect provides a natural way to collect and organize important messages. Every object in the group memory has a unique email address including the predefined collection bin assigned to each user. Users collect email by mailing to these addresses. By default, the collection bin's address is simply the user's login name at the server's hostname. Typically, users assign this special address to a nickname in their email client so that collecting is simple. When an object is collected by email, the Intraspect server receives the message and parses it for MIME structure, indexing both the message and the content of attachments. Users collect an individual email message by forwarding it, collect a document by mailing it as an attachment, or make a new document with arbitrary content (including HTML) by simply composing a message and sending it to their collection bin. Collecting by email also provides a way to deposit information while off the network (assuming the user's email client can store-and-forward outgoing messages).

Collecting from desktop applications

Much of the valuable information in an organization is locked up on individual hard drives or lost in huge file systems used for storing everything. Intraspect lets people share any file they can see on their desktop by collecting it and making it public. When a file is collected from the desktop, a copy is sent from the client machine to the server where it is stored in the object repository and indexed for search. Once collected, a file can be put into multiple shared folders in the group memory (see multiple contexts). Files can also be versioned and updated, which supports collaborative review and revision.

Collecting from other repositories

Intraspect's server can import from any information source that presents its data over HTTP or SMTP. Intraspect can monitor web sources and receive email from distribution lists. It was designed to take advantage of web gateways to proprietary information systems, such as relational databases, groupware and document management systems. Gateways that provide "browser access" to these services make their information available to the Intraspect user for organizing, collaborating and searching.

2.2 Information Sharing: Knowledge Repository as On-line Work Environment

Intraspect provides a place for on-line work, a collaborative workspace where people can gather, share, discuss and deliver their work. Several metaphors have been proposed for collaborative workspaces, ranging from one-dimensional chat "rooms" to 3D virtual environments. Intraspect's first product focuses on *asynchronous, information organizing spaces* for collaborative work. This means that people can read and contribute without being logged on at the same time (like email), and that the user interface metaphor includes familiar elements such as folders, objects tagged with type icons and message-based discussions.

File types and information sources supported

Since Intraspect Knowledge Management isn't tied to specific tools for creating documents or composing messages, it can handle a large range of file types. For the purpose of sharing files, putting them in folders, making comments about them and making them available over web browsers, Intraspect supports any data file that can be collected from a file system, as a web page or attached to an email message. In addition, Intraspect can index for full-text search all of the common document formats including Microsoft and Lotus office suites, Adobe Acrobat, HTML, SGML and email messages.

Versioning

Any information object that can be collected can be versioned. If the object is a traditional document collected from a file system, new versions can be created when subsequent instances of the file are collected. (The user may also choose to replace the original.) A different mechanism is used to version information objects collected from remote sources, such as web servers or gateways. Since these objects may be updated frequently at the remote source (e.g., somebody else's web site), Intraspect does not automatically store their version history. Instead, Intraspect users make a *snapshot* of a remotely hosted object. The new snapshot is a copy of the original document and is a distinct object with a new URL that is thereafter hosted from the Intraspect server. All versions of an object are available to the same set of collaboration and organizational services. This supports the collaborative review of evolving documents and the history of contributions to important files such as designs.

Access control

Every object in Intraspect, including folders and discussions, is under access control restrictions. Access privileges are grouped into classes, including read-only, read-and-comment, contribute documents, create structure, and full ownership. Every user is automatically given a private workspace (folder hierarchy) where no one can see their work, and a personal shared workspace where they have control over content that is publicly readable. Group workspaces (file cabinets) can have arbitrary sets of users assigned any of the privileges. The access control model is designed so that users associate the location of a document in the workspace hierarchy with its access control protection. Each logical work group, for instance, may "own" a file cabinet for its private use. The

access control model minimizes users' need to think about access control; most of the time, access control for new objects is simply inherited from their surrounding workspace.

How Intraspect is different from web servers for information sharing

Frustrated with traditional file systems and messaging, many people are using web servers to share information. As web content creation tools get easier to use and web servers start to provide file sharing capabilities, more people can create their own "home pages" to distribute targeted information. Intraspect also makes it easy to distribute information to the web browser audience (e.g., by dragging it to a public folder in the server). However, there are important differences between Intraspect and a web site or catalog server. Intraspect is designed as a collaborative knowledge management system rather than a distribution center for published information. This difference is manifested in two properties worth describing in more detail:

- **Information sharing instead of publication.** People can share information objects, and make them available to targeted audiences, by simply dragging them into shared spaces. In the Intraspect group memory, the only difference between "published" and personal information is access control. There is no need for special content preparation or publication tools. Every object in Intraspect has an URL. People can search, browse, view and download shared documents through their web browser or the Intraspect thin client.
- **Organization independent of information source.** File systems, web servers, discussion systems and document management systems typically have primitives for organizing information (e.g., by directory, hyperlink, discussion list, and repository). In these systems, the identity of an object can be tied to its name and location (e.g., file path name, URL). In Intraspect, since every object is assigned a unique identifier independent of its name or location, the folders represent the *context* of the information, such as a project or customer, rather than the information source. As these contexts are changed and extended, the identity (and URLs) of Intraspect objects remain constant. This automatically ensures *link integrity* (maintaining link integrity in HTML file-based web sites requires editing the contents of the HTML documents when things move).

2.3. Collaboration: communication and sharing in the same space

Since collaboration is an essential part of knowledge work, Intraspect tightly integrates collaboration services with its other services. Collaboration services include ***commenting*** (collaborative annotation) and ***threaded discussion***. In Intraspect, comments and discussions are managed in the same space as the other information objects (folders, documents, etc.). Instead of holding discussions in one tool, sharing documents in a second and working with web pages in a third, Intraspect users simply work in shared spaces and collaborate *around* their information. (Figure 4 shows a sample folder containing a mix of folders, office documents, web pages, and comments about these documents. It also shows a discussion object, which contains an entire conversation of threaded messages.)

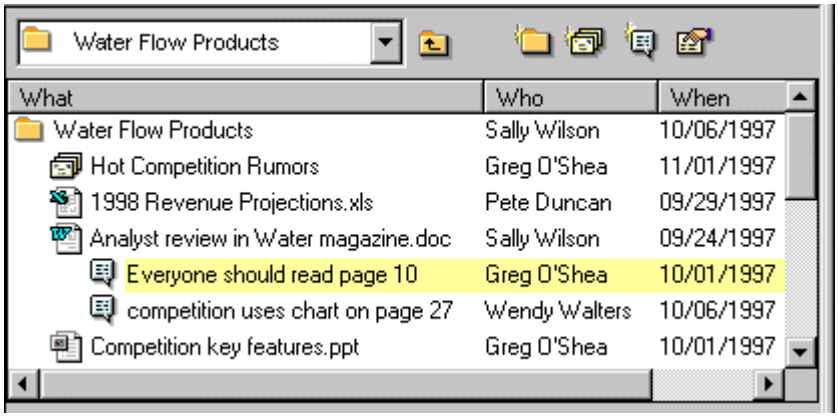

Figure 4. Comments about a Word document in the same space as other files and conversations. Comments are related to the objects they are about but are not part of them. For example, web pages that are hosted at remote servers can also be commented on, the comments appear under them in an Intraspect workspace. The example above shows a context for collecting information on a competitor. The analyst review, a Word document, has two comments about it. This folder also contains a discussion ("Hot Competition Rumors"), which is a different kind of a threaded conversation.

Commenting on information objects

Commenting about shared information is a common form of collaboration. We all get email messages that contain attachments or pointers to information with a few lines of commentary about why the information is relevant or worth investigating. Commenting is a structured way to talk about other information. Comments can be quick one-liners, akin to yellow sticky notes, or full-blown memos. Like email, comments can form conversational "threads" of replies. Unlike traditional email messages, Intraspect comments are displayed *with the referenced information* (e.g., document, web page). Only one copy of the referenced information is stored and its context of use is visible right from the comment. This is "collaboration around shared content" as opposed to collaboration around institutionalized conversations (mailing lists). Commenting is a powerful end-user technique for providing meta-data about shared information. People can say, in words and references, why a piece of information is relevant, good or tagged with a particular attribute (e.g., "highly recommended"). By using such conventions in the text of comments, end users can invent and maintain their own "collaborative review" schemes without programming databases, designing forms or buying into special purpose applications.

Managing multiple conversations about the same information

Comments have another property that is particularly useful for collaborative review of remotely hosted information (e.g., competitive analysis of web pages). Comments are situated in a context, which are represented by workspaces (folders) in Intraspect Knowledge Management. Since objects can have multiple contexts, there can be multiple streams of independent commentary for a single web page or document. Each folder in which a document appears has its own threaded commentary about that document. This allows, for example, engineering and marketing to add their separate perspectives to the same specification document. With access control, these multiple conversations may be

kept private while the document discussed remains public. Similarly, "feeds" of news articles or analysts reports can be automatically channeled into the group memory using email, and the attached documents can be distributed to and discussed by consumers of the information feed using Intraspect.

Hosting threaded discussions from Intraspect

Threaded discussion is the most familiar form of on-line collaboration. Some products use dedicated client/server protocols, such as NNTP, for managing group discussions. Others provide mailing-list distribution services in which a single email address represents a list of mail recipients and a server re-distributes messages to list members. Others use a general-purpose client/server system, such as HTTP/HTML or Lotus Notes, to build custom discussion databases (also known as "bulletin boards").

Intraspect provides threaded discussions that are independent of the client or server protocol. Intraspect discussions are simply objects that represent sequences of messages. Messages are treated as compound documents (using MIME as the embedding standard) that have special relationships to other messages (threading) and to objects being discussed. One can start a conversation from any folder, which puts the context of a discussion in the workspace where the collaboration is happening, rather than in its own namespace (cf. newsgroups). Discussions are started and managed by end users, not administrators. Once a discussion has been established, users can contribute to the discussion using the Intraspect thin client, their favorite email client or their favorite web browser. All messages in a discussion, like all Intraspect objects, are stored and indexed for full-text search. Users can choose to have copies of messages sent to them through email (via SMTP) or they can view them directly from the server through the Intraspect thin client or the web browser. Messages sent to discussions can contain complex MIME substructure, such as multiple parts in different formats (this is how the latest email composition tools achieve multimedia email and attachments using open standards). Of course, these attachments are stored and indexed for search as documents in their native formats (MS Word, PDF, HTML, etc.).

Since discussions use the open standards SMTP and MIME, they can be used to involve "outside" participants in on-line collaborations, independent of whether they are registered Intraspect users are sitting inside the company firewall. This allows "outsiders" to receive copies of shared information and to contribute documents and commentary via their own email clients.

Mediating External Message Flows

Intraspect can also mediate existing message flows that are originated and distributed by outside mail servers, such as institutional mailing lists (distributions hosted on existing mail servers), mail-based information feeds (e.g., paid newsletter subscriptions delivered by email), and mail-based feedback channels (e.g., customer "suggestion boxes" or support requests). As a mediator, Intraspect receives email traffic addressed to a particular flow and assimilates the messages into a discussion object in the group memory. Intraspect users give symbolic email addresses to these discussion objects and then add these addresses to the external mailing lists. The Intraspect discussion objects subsequently receive mail sent to them as "mail agents." The messages are threaded and indexed for

full-text search. Intraspect can then redistribute the mail to users who request it or simply serve as an archive of the discussion in the group memory. Figure 5 shows the paths that email messages can take in the Intraspect server.

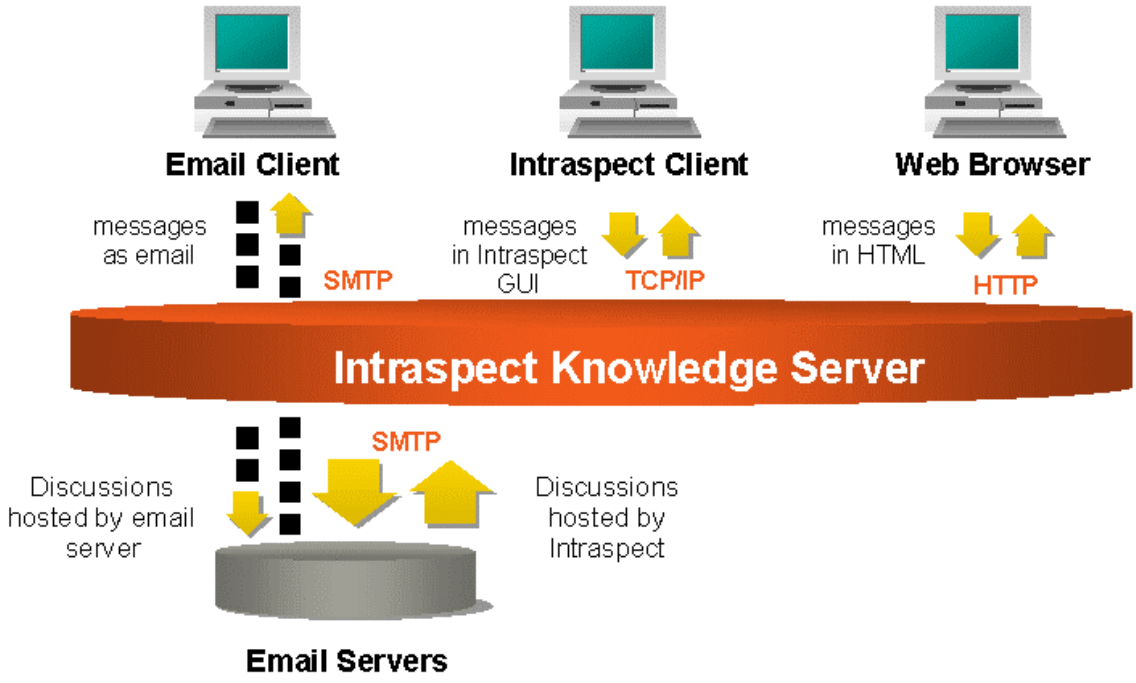

Figure 5: Message flows in Intraspect. To support existing collaborative work practice — involving heterogeneous mail clients, servers, and work styles — Intraspect supports multiple paths of information flow over SMTP. Users can read and compose messages using their mail client, the Intraspect Windows client, or the Intraspect Web interface. Email messages can be forwarded into the Intraspect server as documents themselves, as well as part of threaded discussions. Intraspect objects can be sent as email message attachments from any Intraspect interface. Users can even use email to comment about existing objects in the Group Memory.

Email mediation allows Intraspect to provide knowledge management services to existing message-based processes without changing the existing message flows. For example, a customer support team may receive requests through email. When a customer message is answered, a copy is kept in the Intraspect discussion. When the customer support representative replies to the message using an email client, a copy of the reply is also captured and threaded in the Intraspect discussion. The message thread can be linked into other Intraspect contexts, such as a folder used to collaborate around a problem that was escalated from customer support to engineering. Other customer support representatives can search this email correspondence to see whether a related problem has been solved, and Marketing can mine this information for feedback on product improvements. All of this is possible without changing the existing email addresses, or mail server, for the existing customer support process.

2.4 Discovery and retrieval through associative search

Discovery through focused search and navigation

In a large, dynamic group memory, the primary way to find information is through associative retrieval (searching for information objects by supplying identifying fragments and properties). Intraspect offers sophisticated full-text search integrated with attribute search. The same interface lets users find documents they've already seen (e.g., find word documents collected in the past month that mention "knowledge management") and discover what their colleagues have collected on a given topic (e.g., find web pages or email messages about "Lotus Notes administration"). Once they find a few documents that match a query, users can browse the local folder hierarchy to see related documents that their colleagues have grouped with them. They can also restrict a search to regions of the group memory, such as personal folders or specific file cabinets.

Integrated full-text and meta-data search capabilities

Unlike "keyword search" systems, Intraspect supports true information retrieval technology. The search system can match natural language phrases, factoring out small differences in word order and word endings. The system also ranks search hits by degree of match; documents that use the requested phrases more often appear higher in the ranking. Sometimes this feature is called "relevance ranking." This is misleading, since the relevance of a piece of information is more often determined by its context of use than the query used to retrieve it. Intraspect allows the user to specify true context-based relevance information in a search by allowing one to filter search hits by the contexts in which they are used.

Limiting searches by folder contexts is one way Intraspect includes "meta-data" in its arsenal of search parameters. Users can limit searches to their personal folders, to specific group workspaces or to entire servers. They can specify a range of dates during which the information was created or collected. They can specify restrictions based on who created or collected the information.

Brokered search

When users conduct searches in Intraspect, they are searching over the contents of the group memory. This is a powerful mechanism for information discovery and reuse. When knowledge workers are doing research, they also want to search outside sources, such as the Internet. Naturally, these outside sources have their own search engines. Intraspect allows users to specify their queries in the familiar interface used to search the group memory and *brokers* these requests to outside search services — automatically translating their queries into the appropriate syntax. This supports a natural work style called *cascaded search* in which users search in a series of sources, looking first in the places most likely to produce relevant information. For example, when looking for a document you've recently seen, you might search your local workspace first, then the local office, then the entire company and then the Internet.

Persistent search

Doing research using search engines on a large corpus of information can require domain expertise and time. The search query is formulated, the results analyzed and often the query is refined until the query retrieves the kind of information being sought. Searching is another form of knowledge work that Intraspect can help to make available for sharing and reuse. Once a user has gotten the desired results from a query, he or she can save the search query, thus making it *persistent*. This creates an Intraspect object that serves as a kind of virtual folder whose contents are those items that match the query. The persistent search object can be renamed ("how to find out what we know about customer X"), put in folders and even commented on.

2.5 Subscription and notification: Keeping informed

In any large organization, it is impossible to monitor every discussion list and keep track of every project. A group memory makes it easier to create discussions and projects, and encourages people to share their web collections and documents. Doesn't this create a burden on people to keep up with all this information?

The group memory architecture allows a creative solution to the traditional problem of keeping informed. A pure messaging approach to knowledge management, or a reliance on the web model of distributed publishing, puts the burden on humans to monitor message lists and publication sites for new information. The inherent problem with these approaches is they require that the people who post the message or publish the document know who would need to know, and the people who need to know something need to know where to find it. Search engines can help the latter case but they don't solve the problem of *where* and *when* to search. The group memory, while it contains messages and published information, doesn't depend on messaging or browsing to keep people informed. Instead, the group memory is a natural place to employ software agents to do the work. As people collect, create or discuss information in the group memory, agents can be triggered to inform others who are interested in the content. The originators need not anticipate the information needs of others and the information seekers don't have to know who to ask. Intraspect's first generation of agents, called subscription agents, can efficiently monitor the relevant activity in the group memory and notify interested parties. Let's consider the general notion of subscription first and then explore some of the uses of this powerful service.

Universal Subscription to information objects

Subscription is a general information-processing concept that has been applied to a variety of information types under various marketing concepts such as "personal web pages," "push," and "search agents." Subscription is essentially a service that delivers a stream of information to fulfill an information need stated by a subscriber. Stock quote services, news clipping services and subscription to public email distribution lists are familiar examples of subscription services.

Intraspect provides a Universal Subscription™ service, which delivers a subscription service for *all* types of information in the group memory. The unique power of Intraspect subscriptions is the precision by which one can indicate interest and completeness across

all information sources. For each kind of information object, there is an appropriate subscription service with special benefits:

- **Subscription to Workspaces.** To keep informed of the collaborative activities in a project, a user can subscribe to the project's folder or file cabinet. When any document in that folder is uploaded or collected, the user is informed of what happened, when and by whom. Folder subscription is also useful for notifying people when a document is ready for review, or a deliverable has been completed, or for other ad hoc notification tasks that people find useful.
- **Subscription to discussions, messages and comments.** To actively participate in a discussion, one can use subscription to turn on or off a flow of notifications of new messages and replies. This control over subscription is independent of subscription to the primary message source, such as a department distribution list or a customer feedback mail address. Users can safely delete email messages from their personal in-boxes because the conversations are safely stored, threaded and indexed for search in the group memory server. If a user goes on vacation, he or she can simply unsubscribe and catch up on missing messages using the Intraspect thin client or a web browser. In Intraspect Knowledge Management, one can also subscribe to *individual messages*. This is very useful for getting questions answered: a question is posted to a general discussion list and the user is notified only when *that* question is answered. The same service is available by subscribing to an Intraspect comment: the subscriber will be informed if anyone responds to the comment.
- **Subscription to documents.** A common collaborative task is to put up a document for review. By subscribing to a document, the Intraspect user can ask to be notified automatically when anyone makes a comment about that document. This allows free commentary without forcing the commentator to decide who should be sent a copy or when they want to be notified. Subscription to a document also informs the interested party when the document is revised.
- **Subscription to web-based information objects.** In Intraspect, web documents are like any other file or document in the system except that their source is a remote web server. Web information objects include not only documents, like HTML and Acrobat files, but also virtual documents such as the results generated by an Internet search engine. When a user subscribes to a web object, he or she will be informed when anyone in the group memory comments on the page. This is handy for collaborative review of web materials hosted remotely. (A common application is to review web-based examples when designing a web site or marketing message). The subscriber will also be informed when the web page changes at the remote web server. Intraspect efficiently polls remote web servers, using HTTP standard protocols, to determine whether a web page has changed.
- **Subscription to search.** Subscription is an especially powerful tool when applied to search. As described above, Intraspect allow users to focus their searches on specific types, by location within the group memory, by people, by time and by matching natural language text. The search parameters allow one to specify an

information need with great precision. Thus, search is a natural compliment to subscription. When a search is issued, the search hits are those objects in the group memory that *currently* match the query. When a search is subscribed to, the user is notified when any objects are created or collected that match the query *in the future.* Intraspect has licensed Verity's "agent server" technology to efficiently match a very large number of subscribed searches against a real-time flow of information events in the group memory.

Advanced uses of search subscription

By applying subscription to the power of full-text search, extended with parameters for meta-data such as person, workspace, date and object type, end users can create information flows that support many of the business information needs that formerly required programming. Some examples include:

- “Inform me when a particular news feed service mentions a topic of interest to me”
- “Inform me when anyone mentions my client's name or product in the group memory”
- “Inform me when anyone discovers a web page that mentions our product”
- “Inform me when the department manager says anything about my area of responsibility in a message, comment or report”
- “Inform me about *everything* the CEO says”

Notification Mechanisms

Subscription agents are a special case of information agents. Agents, in general, recognize some situations and take some actions on behalf of human clients. Subscription agents recognize when an information event occurs in the group memory, such as a when a document that matches a query is collected. The actions of a subscription agent are to notify the human client of the information event and provide access to the information. Intraspect's first release supports subscription agents with two notification mechanisms: an interactive GUI display of notification events and email reports. The GUI display shows up-to-the-minute status on all subscriptions and allows the user to sort and manipulate the display to make sense of the data. One can click directly from the screen to the objects mentioned in the display (the documents, messages, comments, etc.). The email reports list the latest notifications on a periodic basis, such as each morning. The email messages contain pointers to the information being described as URLs. Most modern email systems allow users to click directly from the URL in a message to launch the information in a web browser.

3. Intraspect Knowledge Server Architecture

Intraspect is a multi-tier, distributed application with a thin client for Windows, user interfaces for web browsers, and connections to email clients. The Intraspect server incorporates a high-performance search engine and scalable object database (both are hardened commercial products and are complete embedded in Intraspect). The Intraspect server is written in Java, with an multithreaded architecture. It includes a complete HTTP 1.1 compliant web server and an SMTP server for receiving and distributing email, plus

application logic for the Intraspect collaboration and knowledge management services. In intranet environments with existing web servers and Internet-compatible mail, the entire product can be installed in minutes. No special administration is required for the embedded database, search engine, web server, or email services. The entire product is designed to be run by end-users.

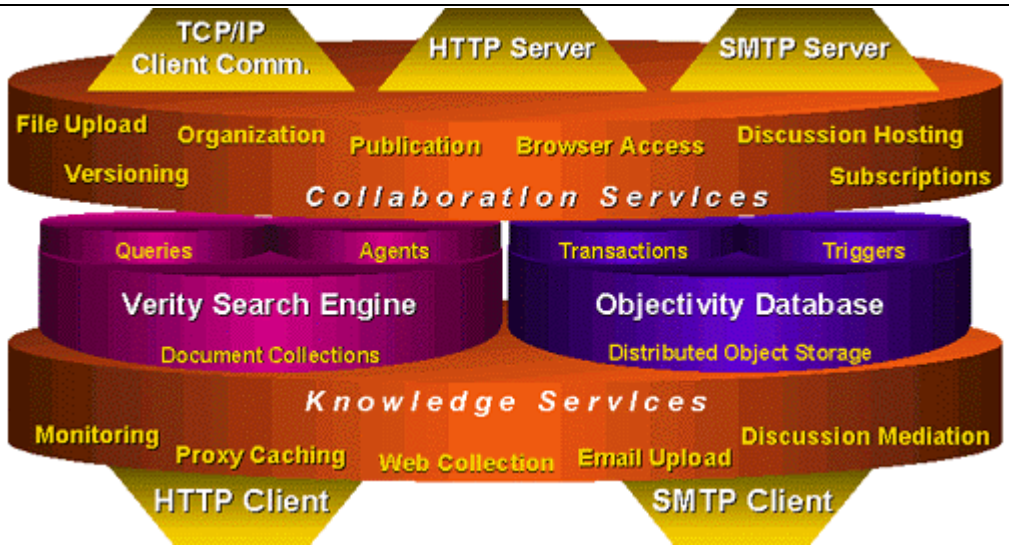

Figure 6. Intraspect Server Architecture. The Intraspect server is a full intranet peer for both HTTP and SMTP. It acts as an HTTP server for delivering the web browser view on data and for integration with other web-based information flows. It acts as an HTTP client when collecting and monitoring information from other web servers. It acts as an SMTP server so that it can receive mail and control its email address space. It acts as an SMTP client to send and redistribute mail to end users and other mail-based information flows. It also supports high speed, low latency, binary TCP/IP communications to Intraspect GUI clients. The database and search engines are completely embedded in the product.