NIH ENTERPRISE INFORMATION PORTAL

Implementation Plan: Summary

NIH Enterprise Information Portal

Institutes, Centers, and Offices (ICO) at the NIH create, receive and collect a vast amount of data and documents that reside in "stovepipe" systems and repositories – data systems and document repositories that perform or support one function, functional area or NIH organization but are not integrated with other systems/repositories. As a result, information sharing is often limited to top-down distribution through a functional or organizational hierarchy or to those who have a need to know because of their direct involvement in a functional area or other NIH activity. Outside of these well-established patterns of information exchange, employee access to information across organizational and functional boundaries is not always easy or quick. To address this and related issues, the NIH Chief Information Officer is leading an effort to develop the NIH Enterprise Information Portal. Portal technology can provide an environment where data, documents, tools and business processes are more easily and widely accessible outside the boundaries of organizations, systems and functions – without significant impact on underlying structures.

Vision

The NIH Portal will be a central point of access to data, documents, processes, and services available at the National Institutes of Health. NIH information, along with a wide variety of information from external sources, will be indexed and/or categorized to provide multiple mechanisms for finding and viewing data and documents within the Portal. The Portal will allow groups of users to set up community areas within the Portal where they can share information and interact collaboratively through whiteboards, threaded discussions and instant messaging. Initially, the NIH Portal will be deployed for NIH employees, followed by access for NIH grantees, other NIH partners and suppliers, and finally for access by the public.

Objectives

- To provide access from the NIH Portal to *all* the important data, documents, processes, and services that NIH employees need to do their job.

- To the maximum extent possible, provide direct access to the content of database systems and document repositories without necessarily requiring users to know or navigate those underlying applications.

- To facilitate person-to-person collaboration by supporting electronic meeting rooms, threaded discussions, instant messaging, and other utilities that allow groups with common interests to set up and manage their own community areas within the Portal.

- To develop mechanisms that connect authorized individuals to NIH business processes where they can initiate and participate in business transactions from their Portal interface.

- To create a central location where information and knowledge from those who are leaving the NIH can be captured, archived, organized and re-packaged in multiple ways for current and future NIH staff.

- To offer a continually updated array of portal tools that support personalization of displayed information and customization of user "MyPages".

- To support the NIH enterprise infrastructure.

Architecture

Network Architecture

The diagram below is a high level view of the NIH Portal network architecture. The pilot version of the Portal currently in place at http://my.nih.gov is using Plumtree Corporate Portal 4.0 software and a 3-server configuration for development of the NIH Portal. As the Portal goes to production the architecture will likely move to a 4-server configuration running Plumtree release 4.5. Plumtree is one of the founders, and the current market leader, of corporate portal software solutions and they have a proven track record of over 250 customer implementations to date. Their robust portal solution has the scalability and flexibility to grow and change with the changing needs and resources of the NIH. The extremely scalable architecture will allow the addition of multiple servers at key architecture points to support additional NIH communities and to provide redundancy and optimal performance for the NIH Portal.
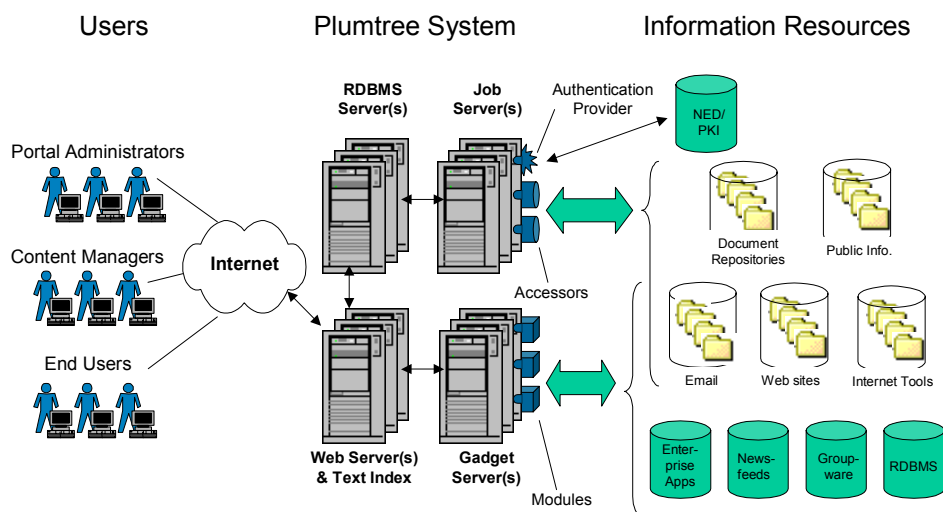
Functional Architecture

The NIH Portal functional architecture will include all of the functional elements that portal industry analysts associate with an enterprise information portal:

- **Integration.** Provides the means to access multiple disparate information sources such as enterprise applications, legacy data, email, records management repositories, network file systems, internal and external documents, web servers, newsfeeds, etc.

- **Categorization.** Provides the ability to create and maintain a taxonomy of subject categories, to capture and tag documents with meta-data, and to use these categories and meta-data to organize portal information into a hierarchical directory that users can search or browse.

- **Search.** Creates a full text index of all documents, and associated meta-data, as they are added to the portal and supports the query and retrieval of those documents using both full text terms and meta-data keywords with Boolean operators. Multiple search engines are envisioned for the NIH Portal.

- **Publishing & Distribution.** Provides a mechanism for posting documents directly into the Portal, bypassing the more common 'crawl and post' process. Also provides support for push/pull methods of document distribution, including both portal and email notification to users when new documents are posted.

- **Process.** Provides a means for defining, automating, viewing, monitoring, and managing business processes using workflow and business intelligence tools.

- **Collaboration.**   Supports community areas within the portal where there will be synchronous and asynchronous mechanisms for groups of NIH Portal users to communicate, interact and share information, using electronic meeting rooms, instant messaging, whiteboards, and threaded discussions.

- **Personalization.**   Supports user preferences and personalization on many levels within each user's personal view (their "MyPage") of the Portal, including the documents that are retrieved, the applications that are selected, and other choices that vary with the content source.

- **Presentation.** Provides a customizable web page interface from which a user can access information in the NIH Portal – a "MyPage" whose content, layout, and color can all be changed to accommodate the user's personal preferences.

In addition to the core functionality described above, the NIH Portal will provide a security structure that preserves the access restrictions that exist in the native repository of each content source. When any data, document or object in the portal is restricted to a specific user group, only that group of users will see it and know that it even exists within the Portal. Security on portal content will be determined by NIH content managers, system administrators, and other managers/owners of information included in the NIH Portal. When all necessary components are operational, the NIH Portal will use the NIH Enterprise Directory (NED) along with Microsoft's Active Directory (AD) to authenticate users who access the portal. Until the portal is able to utilize these directories, most users will access the portal using their current NT network account and password.

Component Architecture

The NIH Portal will have the three primary components of a Plumtree portal: MyPages, Communities, and the Directory.

- **MyPages.**   The first thing a user sees when they sign into the portal. An NIH Portal user without an account will see a Guest MyPage, a user who signs in for the first time will see a Default MyPage, and a user who selects and arranges content based on their preferences will have created their own personal MyPage(s). MyPages are composed of NIH Portal 'modules', which are mini-applications sometimes referred to in other portals as gadgets, portlets, applets, or OM-lets.

- **Communities.** A shared resource center or workspace set up and managed by a portal user. A Community Page is an area within the portal where NIH Portal users who share a common interest or practice can communicate and interact in a number of ways. Any NIH Portal user can set up and manage a Community after submitting a request and following a set of guidelines. Communities are composed of MyPages containing community-oriented 'modules' and collaboration tools such as an electronic meeting room, whiteboards and threaded discussions.

- **Directory.**   A hierarchical list of subject categories sometimes referred to as a taxonomy. The Document Directory is used to organize information for browsing and retrieval by clicking through increasingly narrower categories until reaching the desired collection of documents. Browsable directories are an alternative to text and keyword searching. They are often favored by users who aren't familiar with the text terms and keywords that are needed when using search technology.

Deployment

As of December 1, 2001

A pilot version of the NIH Portal has been operational for about a year. It is using Plumtree Corporate Portal 4.0 software and a 3-server configuration. Functional elements already implemented in the Portal include Integration, Search, Personalization, and Presentation. Two of the three portal components, MyPages and Communities, are operational although collaboration tools are still in test mode. Over 120 modules are available for selection and use on MyPages. Access is limited to NIH staff who have a Microsoft NT account and password.

Near-term (within 6-12 months)

Within this timeframe, content will be added to the portal in a phased-release by functional area. An effort will be made to coordinate simultaneous inclusion into the portal of data, documents, 'modules' and communities related to a specific functional area. Functional areas that will be covered include:

- Extramural Research
- Human Resources
- NIH Library
- National Library of Medicine
- Acquisitions & Supply
- Financial Management

- Property Management
- Travel Administration
- R&D Contracts
- Information Technology
- Office, Conference and Food Services
- Research Support Services

- Safety, Security and Facilities Services
- WorkLife Services
- Clinical Center
- Intramural Research
- Public Information
- Technology Transfer/Development

IC-specific data, documents, 'modules' and communities will also be added during this time period, as time and resources permit. Near-term plans include a software upgrade to release 4.5 of Plumtree's Corporate Portal software and the server configuration is expected to go from 3 servers to 4 or 5. The functional elements Categorization, Publishing & Distribution, and Collaboration should be fully operational – and the third of the 3 portal components, an enterprise-wide Directory, will be operational with links to documents relevant to most of the NIH functional areas. Additional modules supporting each of the functional areas will have been added, and single sign-on should be operational as should authentication using NED and AD. Access will continue to be limited to NIH staff.

Long-term (12-24 months)

Long-term development will focus on the addition of workflow and business process intelligence, knowledge management content and capabilities including knowledge capture and expertise locators, and possibly on electronic learning initiatives. Within this timeframe, development efforts will also focus on extending the NIH Portal to external audiences such as NIH grantees and other partners, suppliers, and the general public. Modules will continue to be added and the ongoing effort to increase the functionality of existing modules will continue.

Critical Project Dependencies

As an enterprise initiative of significant scale, the NIH Portal project requires the full support and cooperation of NIH Institutes, Centers and Offices to be successful. The NIH Portal project team will work closely with all ICO's to identify their information needs and to determine how the portal can effectively deliver the content and applications that NIH employees need.

As content from each area is added, Portal Content Managers will need to be designated within ICO's to manage access to their documents and applications and to identify and recommend new information and information resources for the Portal. Only by ensuring that portal documents and data are current - and new documents, applications and tools are made available - can the Portal successfully serve as the central point of access to NIH data, documents, processes and services.