An Investigation of Specifications for Migrating to a Web Portal Framework for the Dissemination of Health Information within a Public Health Network.

Meg Murray, Ph.D.
Kennesaw State University
mcmurray@kennesaw.edu

Abstract

Web-based technologies provide a framework that facilitates the dissemination of healthcare related information. A portal strategy comprised of six layers has been defined for healthcare. This layered architecture is well suited for employing an iterative approach to development, implementation and deployment of an emerging technology. Planning for Web portal implementation should be seen as a process of building an infrastructure and foundation for the future, not as the development of a single all-encompassing solution. Mapping requirements and specifications for the portal framework to a conceptual target architecture is part of this process. This research paper investigates a model for a healthcare portal framework and the mapping of user specifications of a major public health organization for organizing and structuring information to facilitate a plan of action for moving to an Internet portal strategy.

1. Introduction

Information technology has the potential to transform how an organization achieves its mission. Healthcare is reassessing the technology models used to provide better healthcare services via the exchange and dissemination of healthcare related information. Healthcare is a complex industry and, consequently, healthcare information technology encompasses a broad array of applications. These applications include clinical information systems focused at patient care, laboratory systems, radiology systems, pharmacy systems, administrative systems, financial systems, resource management and claims processing. Further, these systems span different settings of care such as physician offices, public health agencies and acute care facilities. Until recently, decisions related to healthcare information technologies were driven from within the industry and often from within an individual organization or department within the organization. As technology evolved so did the investment in healthcare IT solutions. Much of this occurred as purchases of ‘best-of-breed’ software, often developed by different information systems vendors resulting in a complicated architecture of technology platforms and the collection of data in a multitude of formats. Healthcare reform initiatives and other market forces have driven the need for communication between these systems. Healthcare services are not provided in isolation and there exists a need to share data between and among the various entities within the healthcare community. The proliferation of healthcare information via the Internet has also created a need to provide accurate and timely information to healthcare consumers. Searching for healthcare related information is one of the main uses of the Web. Reports indicate that of the 80 million unique users of the World Wide Web in 2000, one out of every four visited at least one health care oriented Web site and visitors to these sites grew faster that visitors to the Web in general [12].

The Internet and Web-based technologies provide a framework that facilitates the dissemination of healthcare related information. The inherent benefits in these technologies are many. Among these are easier maintenance, a single point of access to viewing information and launching applications, simplified remote access, easy deployment of end user applications and a reduced end user training curve. Most people are familiar with Internet browsers, Web sites and even Internet portals. The development of portals to provide access to healthcare information is a natural progression for a healthcare organization.

Another benefit to portal technology is flexibility. A portal is built upon layers of services and component modules. Various areas can be emphasized or deemphasized based on the needs of the organization. This is important because different healthcare organizations have different missions or employ varying levels of technology use. Portal technologies, though, are still in their infancy and the deployment of portals into enterprises that run heterogeneous systems is very complex [4]. On the other hand, the openness of portal-based technologies creates flexibility allowing for an evolutionary approach to the migration of a total solution. An evolutionary approach leverages existing software and
IS resources and provides a mechanism for building greater understanding of the emerging technology.

This paper outlines the investigation of employing portal technologies for the dissemination of information in a national public healthcare organization. The primary mission of the Center for Disease Control (CDC) is to “promote health and quality of life by preventing and controlling disease, injury and disability” [5]. The fundamental building block to carrying out that mission is the analysis and dissemination of data and information that enable public health action. The CDC views information technology as “critically important for the attainment, management, use and communication of public health and related data and information” [7]. Selecting a technology in these times of great change is challenging and identifying which direction or even which applications will yield the greatest benefit has not matured to the point where measurable evaluation can occur. The greatest challenge for CDC information technology decision-makers is to identify and develop strategies that result in the most effective use of emerging technologies. A portal framework provides a means for achieving their goal. An overview of a healthcare portal framework is presented followed by an investigation of CDC information dissemination requirements mapped to the healthcare portal framework.

2. The Healthcare Portal Framework

Web portals began in the consumer market with the larger search engine sites. Portals offer end users fast, centralized access to Internet type services and information found on Web sites. In an effort to insure customers return to their sites, these portals began providing end users with the ability to personalize their interaction with the Web. Corporate data portals or Enterprise Portals, built on the same principles, began to emerge in the workplace a few years ago as an efficient way to organize information that employees routinely need to access. Consumer healthcare portals such as WebMD and HealthCentral.com appeared on the Internet as a means to provide searchable health information and health assessment tools [9]. Local healthcare organizations saw the benefit in this technology not only as a means to provide an access point to consumers for general health related information but also as a way to draw consumers to the services they provide. Further, they recognized the potential of the portal platform as a way to build an enterprise infrastructure that promotes the exchange of data among internal users. Consumer portals provide access to external users while the corporate or enterprise portals are built to address internal business needs within a secure environment. Access to data and applications contained within the enterprise portal is restricted. Permission is granted based on the user’s role or need of access.

Portals offer a central aggregation point for data and links to many different applications and systems as utilized by diverse participants located in many different settings. Portals are a fairly new technology but their use has risen sharply in a short period of time. A portal is analogous to an internal World Wide Web. Portal software is server based and employs Web technologies such as HTML, XML and Java to display information in a Web browser.

Portals have the ability to consolidate and deliver information and applications in an organized way. The primary objectives of an enterprise portal are twofold: to provide a single point of access to resources and a common presentation of content regardless of where the content resides. Portals must facilitate the integration of a wide range of data, provide efficient access to relevant structured and unstructured content, incorporate the ability to analyze information and provide personalized user interfaces. To accomplish this, the Web portal framework is built upon five layers, each of which addresses a primary functional requirement of the portal platform [13]. Within healthcare, a sixth layer for patient identification services is also required [13]. The framework is depicted in Figure 4.

![Web Portal Framework](image)

2.1 Security Layer

Security is included in the healthcare portal framework due to the sensitivity and confidentiality of patient information. A healthcare portal should encompass security measures such as encryption, authentication of users, single sign on and auditing and logging services [17]. A single sign on environment for all users is a basic component of providing a single point of access. Single
sign on supports granting of permission to data, content and applications to individuals based on user profiles.

The primary means of authentication is the user name and password. These names can be stored in the portal database or linked to an LDAP accessible directory. This provides a centralized means for managing user accounts. Many healthcare organizations are requiring advanced methods of security that go beyond user names and passwords. The open nature of the portal technology allows for the use of many different alternative authentication methods such as Public Key Infrastructures (PKI) that use digital certificates, tokens, smart cards and biometric devices. Alternative authentication techniques are most easily addressed by designating security as a separate services layer within the portal framework.

### 2.2 Presentation Layer

Portals offer a secure, single point of access to enterprise resources and the presentation layer provides the interface between the user and those resources. In essence, presentation services manage the look and feel of the portal interface. This is the layer where text, images, audio, video and data requested by the user are assembled into a coherent picture and then rendered for viewing. Presentation of these resources is generally via a Web-browser. Browser access allows for the spawning of a new window for each new application or the partitioning of the screen into frames allocated to individual applications. Within the presentation layer are also communication services that provide a means for converting output for display in alternative devices such as the Personal Digital Assistants or other wireless devices. The footprint for the display on these devices is extremely small and requires a different presentation of information. The portal must be able to provide customized presentation display regardless of whether access is from a PC or a handheld wireless or other mobile device.

### 2.3 Business Application Logic Layer

The application layer ties the user interface capabilities to the enterprise's business processes and existing systems [15]. This may be achieved by Application Programming Interfaces (APIs), but Web-enablement ultimately purports a migration to applications developed using Web-based architectures and programming languages. Web-based applications are designed to be modular and are built on a distributed n-tier architecture; not dissimilar to that employed in client/server systems. The typical architecture incorporates a standard Web-based browser interface on the client desktop, a middle-tier that hosts application logic and connectivity and a tier that hosts backend servers such as database servers. Also incorporated in the middle-tier are business logic and rules processing. Overtime new business applications and capabilities can simply be added to existing resources within the business application layer.

### 2.4 Portal Services Layer

Portal service modules provide the capability to manage the user interface and administrative tools to edit and modify the user environment. They include such features as a search engine with taxonomic capabilities for searching and indexing structured and unstructured data, content management facilities, integration capabilities including XML specifications and open APIs allowing custom solutions to extend access to most any existing information source, and provisions for personalization and security. Providing personalized views for individual users is one of the major advantages to the deployment of a Web portal. Personalization is achieved through the creation of personal profiles that contain user information and preferences. Portal access may be tailored by user preference, need, security level and access authorization allowing the IT administrator to expose (or not expose) applications, programs and proprietary information based on the roles of the user. The ability to personalize portal access is critical to portal success for it provides the means to protect the integrity and confidentiality of information and other resources [3].

Content management tools provide a means for creating and managing dynamic, evolving sources of information. The term content extends the idea of information beyond traditional data and text formats to include various document formats such as email, PDF files, and HTML files as well as hypertext links and other media types such as animation, audio and video. Content management includes provisions for creating and modifying content, organizing and storing content and mechanisms for accessing existing repositories and other data stores [18].

The ability to identify and extract pertinent information in an efficient manner is paramount. According to Tripp [18], the challenge is not so much in the storage and retrieval of various media formats as in the retrieval of content relevant to the user Content delivery engines should be programmable and flexible so that information can be dynamically selected from various sources (databases, Web pages, etc.). Retrieval of content can occur through open-ended searches or taxonomies. Metadata based information may also be employed. Metadata is used to annotate information sources and can be stored separately from the documents allowing for the use of agents to perform basic searching tasks.

### 2.5 Enterprise Application Integration Layer

A fully functional portal must bring together unstructured data such as emails, Web pages and text
documents, structured data such as relational data and integrated application content. The portal must provide access to data from multiple sources regardless of origination or storage location. Data can originate in databases, spreadsheets, word-processing documents, Web-pages, email or even presentations. Applications made available through the portal can reside on a wide range of hardware platforms including mainframes, AS400 systems, Unix servers, NT servers and PCs. Most portal software available on the market offer content management services, a search engine and a library of Application Programming Interfaces (APIs) to commonly used applications. However, it requires additional development efforts to incorporate proprietary applications and data. These middleware technologies and integration capabilities are seen as key to a successful portal implementation.

The ability to exchange data between applications and provide application integration is a fundamental component of a successful Web portal. The Enterprise Application Integration (EAI) layer provides the integration tools necessary for data translation or other data wrap-arounds and provides the mapping from a host system to that required by the browser-based object-oriented technology. This model is highly extensible allowing host databases, regardless of their architectural platform, to be incorporated into browser-based access without the need to recreate interface code. Different methods are employed for providing access to data stored in various formats and generated from various software applications. Some of these methods include proprietary middleware for point-to-point integration, integration software engines that translate data from one format to another, flat file and ASCII translation, screen-scraping and direct copy utilities. Perhaps the most talked about method for providing data interchange is the eXtensible Markup Language (XML). XML is an open standard meaning that XML output from one application can be read and manipulated by other XML-enabled applications without the need for data translation [19].

XML is a markup language for presenting documents on the Web. The premise behind XML is that is separates data from presentation. XML has been called a meta language for defining other markup languages. An XML style sheet tells a Web browser exactly how to display the data in an XML document and an XML Document Type Definition or schema stores the definitions of tags related to industry specific data or fields of knowledge. Document Type Definitions (DTDs) or schemas, often referred to as dictionaries or vocabularies, serve as a uniform source of data definitions to promote the exchange of information between organizations within the same industry. There is a risk that different groups might produce different dictionaries or DTDs. Within healthcare, the Health Level Seven (HL7) standards organization and the wide acceptance of the HL7 messaging standard minimize this risk. HL7 specifications include a standard for electronic data exchange. The HL7 standard defines the format and protocol of messages containing key sets of data that are exchanged between healthcare applications. The most recent version of HL7, Version 3 currently in draft form, will enable systems to create XML documents that incorporate HL7 message content [10].

2.6 Enterprise Master Patient Index Layer

An important component of a healthcare portal is that it must provide the capability to access patient data from multiple healthcare systems located at multiple sites. (While there has been discussion in the United States of assigning individuals a unique national health identifier, at the present time each healthcare provider assigns their own number to each patient.) This will require the ability to clearly and uniquely identify individuals across systems with a common patient identifier. In the healthcare industry, this has become known as an Enterprise Master Patient Index (EMPI). Individual healthcare systems maintain their own indexing systems for identifying patients. The EMPI is designed to uniquely identify patients and correlate their records at an enterprise level while maintaining relationships between all of the different source system patient identifiers. This allows navigation from one system to another via the portal while maintaining reference to the same patient in each application. Systems that employ the same person identification services work together without the need for custom interfacing. A person identification service standard specification has been developed by the Object Management Group’s CORBAmed Task Group. The OMG is a non-profit organization that develops vendor independent specifications for middleware and other component based technologies. The CORBAmed Task Group is charged with developing standards specific to the healthcare industry [14].

Providing a single point of access to data and applications to facilitate the dissemination of healthcare related information moves healthcare organizations closer to the vision of comprehensive access to healthcare information. Portals integrate numerous applications and data sources while allowing access to be customized to meet the needs of individual users. The defining factors for a successful portal include:

- Architecture built on layers of services and component modules.
- Providing the ability to inter-mingled data and content from multiple sources stored in multiple formats.
- A framework that is extensible by employing open standards in the development of portal services.
• Providing for gradual customization so that organizations can add and integrate content, applications and services at their own pace.

3. Case Study

Building and adapting Web technologies that address the needs of the enterprise are not simple ventures. There is a limited understanding and knowledge base of the underlying functional requirements and potential difficulties associated with the Web-based platform. These systems require the integration and consolidation of structured data, unstructured data and integrated application content. The technologies have not yet matured to the point where this type of integration is seamless. However, Web-based technologies and their associated standards, such as portals and Extensible Markup Language (XML), are evolving and the increasing success of this framework suggests that improvements are on the horizon. The Centers for Disease Control (CDC) are embarking upon defining an enterprise-wide information technology architecture that meets their mission and objectives. This is a major undertaking given the size and complexity of the CDC enterprise, their need to manage and disseminate a multitude of content (their current computer center maintains over a trillion characters of data) and the immaturity of Web and Internet-based technologies.

3.1 Description of the CDC

The Centers for Disease Control (CDC) is a $4 billion federal agency within the Department of Health and Human Service of the United States Government. The CDC is comprised of 11 Centers, Institutes, and Offices primarily structured around various diseases, health conditions, or health risks. These offices include the National Center on Birth Defects and Developmental Disabilities, National Center for Chronic Disease Prevention and Health Promotion, National Center for Environmental Health, Office of Genetics and Disease Prevention, National Center for Health Statistics, National Center for HIV, STD, and TB Prevention, National Center for Infectious Diseases, National Center for Injury Prevention and Control, National Immunization Program, National Institute for Occupational Safety and Health, Epidemiology Program Office, Public Health Practice Program Office, Office of the Director which includes the Information Resource Management Office who holds the primary responsibility for information technology and informatics.

The mission of the CDC is to “promote health and quality of life by preventing and controlling disease, injury and disability” [5]. To carry out this mission, the CDC engages in a variety of activities including public health leadership, research, surveillance, health statistics, policy and program development, emergency response, education and the dissemination of information and guidance. The CDC seeks to accomplish its mission by working with partners throughout the nation and world to monitor health events. The CDC has strategic partnerships with healthcare and public health agencies, municipal public health agencies, national and international public health organizations, university and institutional public health policy makers and research organizations. Scientific credibility and service are the characteristics for which the CDC is most widely known [5] and maintaining this edge in an era of information explosion is critical to maintaining their reputation. The CDC has identified its user community to comprise three categories: general public users/consumers, partners and CDC employees.

The CDC is an information intensive organization that identifies data and information as the fundamental building blocks that enable scientific discovery and public health actions [7]. Public health informatics deals with resources, devices and formalized methods for optimizing the storage, retrieval and management of information to facilitate public health problem solving and decision-making [5]. Consequently, information technology is critically important for the attainment, management, use, and communication of public health related data and information and the CDC’s dependence on information technologies continues to grow [6].

The CDC views information as their most important product. They define information resource management to consist of the collection, management analysis and dissemination of data and information for public health. Data, the raw material of information, is routinely collected from thousands of sources, analyzed and transformed into information resources. The CDC currently has over 110,000 Web pages in addition to an extensive document archive (electronic and hard copy formats). Providing the capability to effectively and efficiently create, store and retrieve relevant information is critical to the mission of the CDC. The CDC’s strategic planning efforts address the need to incorporate the best means available to support this mission while protecting the integrity and confidentiality of their information and data sources.

The CDC’s mission continues to be increasingly dependent on information technology, electronic communications and digital media. “Detecting health events and assessing health status trends in populations in a timely, comprehensive, reliable and cost-effective manner is only possible through IT” [7]. The CDC’s current strategic plan for information resource management is to develop and refine their Information Technology Architecture (ITA) to embrace a vision for next generation systems that incorporates existing technologies and integrates with the mission, business
directions and needs of the CDC. They envision an architecture that migrates towards Web-based technologies where possible and increases the integration of systems through standards and open system approaches. As specific needs exist for systems integration, standardization, distributed data/content management and communications with CDC public health partners and constituent groups, the ITA must facilitate internal and external communications and include standards to guide the design of new systems as well as the operation of existing systems.

3.2 Assessment of User Specifications

An initial attempt at solicitation for a vendor to carry out the CDC’s strategic information technology needs did not result in an acceptable response. It was quickly realized that a single vendor solution did not exist. This is in keeping with the current state of portal technologies. Building an enterprise portal is a major undertaking. Portals do not represent an out-of-the-box solution. Full deployment of an enterprise portal is evolutionary. Developing a portal with the level of sophistication organizations need is a long and hard development process [11]. Planning for a Web portal implementation should be seen as a process of building an infrastructure and foundation for the future, not as the development of a single all-encompassing solution. Some organizations have found it beneficial to start small and build their first portal focused on solving a pressing problem for a particular department or business function [15]. While the initial development efforts are focused, the portal platform is well suited for employing an iterative approach to development, implementation and deployment as illustrated through the layered architecture of the healthcare portal framework model. The development of the CDC’s information technology architecture parallels ITA in that both are viewed as continuous processes.

The framework for the ITA is designed to describe the relationship among work, information and information technology needs. The ITA is comprised of three components: Enterprise Architecture, Technical Reference Model and Standards Profiles and Implementation Plans and Maintenance Procedures. The enterprise architecture focuses on current and desired relationships among business and management processes and information technology. The technical reference model identifies the discrete set of conceptual layers and entities that will articulate specification standards. Implementation and maintenance plans will insure that the ITA is put into practice in a manner consistent with architectural specifications. Most of the work has been completed in relation to documenting the current technology architecture and identifying CDC functions and user requirements. The next step is to develop the conceptual target architecture from which vendor solutions can be solicited. Mapping requirements and specifications for the portal framework to the conceptual target architecture is part of this process. A discussion of the mapping of specifications to each layer of the portal framework follows.

3.2.1. Security

In 1998, the CDC launched a major security initiative that resulted in an expansion of their information security infrastructure. After a thorough investigation of available security technology, it was determined that no one product provided the needed security without imposing barriers to the mission of the CDC. The resulting plan, instead, included provisions for several categories of security including intrusion, computer incident response, security awareness and a secure data network. The secure data network which represents the portal security layer will employ digital certificate technology using the X.509 standard technology for the secure transmission of sensitive public health data over the Internet with their public health partners.

3.2.2. Presentation

In keeping with the dictate to migrate to Web-based technologies, presentation will be standardized on Web browser access. An organization-wide committee was formed to develop a common Web page template. The design was finalized and specifications were communicated to CDC entities. The template has been implemented by most centers and will be used as the template for future Web pages.

Moving to Web-based access to applications and data through a browser has an added benefit. It provides a mean to deploy a progressive strategy to Web-enablement of existing applications. A natural progression to an Internet platform is the introduction of a simple browser providing a launch point to existing applications and view access to existing data. The browser can be layered on top of the current infrastructure to provide access to existing legacy systems. This allows the CDC to standardize on the Web browser as the primary desktop interface to resources including content and applications.

3.2.3. Business Logic

The CDC supports several functional areas. Each of these areas, in turn, maintains various information systems and applications. Integrating these systems within a business logic layer of a Web portal will have many advantages in terms of customizing access to both internal and external users. Recent system initiatives that fall within the business logic domain include the Health Alert
Network designed to build IT capacity at the local public health office to facilitate more timely communications in regards to health emergencies; the National Electronic Disease Surveillance System designed to electronically integrate a wide variety of surveillance activities and facilitate accurate and timely reporting of disease information and the migration to next generation Enterprise Human Resource Planning systems.

3.2.4. Portal Services - Personalization

One feature of a Web portal is the ability to provide information that is personalized for each user. Web personalization allows for Web content to be tailored to user profile information as well as to user preferences. When the user accesses the Web site, the logic employed in the portal’s personalization services uses profile information to gather together information relevant to that person and display it in one place. This information might include generic publicly accessible information or confidential information specific to that individual. The three categories of end users identified by the CDC have different needs and specific users within these groups have unique needs as well. For instance, CDC partners maintain consistent contact with the CDC for the purpose of exchanging data, reviewing CDC published reports and receiving notification of health alerts. Internal users need personalized access to the data and applications necessary for them to perform their specific jobs. Personalization would allow the general public end users to customize their interaction with CDC resources. For instance, two areas often requested by the general public are health information for travelers and immunization information. Through personalization, these end users might design their interface to include frames for each of these areas. Each time they visit the CDC site, these frames would appear automatically.

Personalization services also provide means to log each user’s clickstream; an accounting of each page the user requests for display. Automated analysis of this data makes it possible to dynamically tailor user experiences. In reference to the traveler mentioned above, this means that the information about countries most recently searched could be displayed automatically. Personalization provides the methods to enhance user experiences and reducing the amount of time the user spends getting to the resources they need.

3.2.5. Content Management

A primary function and a core business process of the CDC is to manage information. This includes developing, reviewing, organizing, maintaining and disseminating public health knowledge and information to public health and medical professionals, policy makers and the general public. Content management tools are designed to maintain the content on a complex Web site. Content management includes four elements: content creation, content management, content access management, and content delivery. Content creation involves development and storage of elements in multiple formats. Content management enables the construction of Web content using standard templates that can be dynamically linked. Content access management allows for dynamic access to Web content through searching technologies. Content delivery involves the dynamic serving of content to the end user. Content delivery may be served through personalized services such as the use of profiles or clickstream analysis or customized through user preferences. Content management systems consist of data entry, web page development/presentation, site organization and interface capabilities to databases and other document repositories.

The retrieval of relevant information is a fundamental requirement of the CDC information technology strategy. The CDC is trying to enhance user access to the data and content by reducing search time and improving the retrieval rate of relevant information. Currently, information searching is done through the use of a third party search engine employing basic keyword searching using a domain specific lexicon for healthcare expanded with CDC specific terms. Keyword searching is based on matching words or synonyms. The match can occur within document titles, document text or against metadata if such information about the document has been extracted. Keyword searching can be effective but is often inefficient when the search results in the retrieval of large collections of documents. For instance, a search on the term “HIV” brings up a listing of 8665 documents; too many to peruse. Aggregating, filtering, ranking, visualizing, and other analytical techniques must be provided to turn content into actionable knowledge.

One way to achieve this is through the use of Information Extraction (IE) software. IE works in conjunction with information retrieval and organization tools to automate the finding of content. Information extraction pulls information from documents in various formats and converts them into a homogeneous form that can be searched and organized into taxonomies. In essence the information about these digital documents is transformed into a structured database format [2]. IE extends the traditional keyword searching used in information retrieval technologies. IE users define categories of information they want to capture. This makes IE fully customizable but requires much preplanning as structure for organizing information must be developed. Two processes are currently used for this. One involves Natural Language Processing (NLP) and the other wrapper induction. In NLP, documents are processed into a syntactic structure and tagged according to parts of speech. This is then matched to linguistic
patterns and compared against a predetermined answer key to determine relevance. Wrapper induction extracts content using predefined templates. Relevance is measured based on textual qualities that surround the data rather than on an examination of syntactic or linguistic qualities [2]. One area where wrapper induction is highly efficient is with dynamic Web pages where pages are assembled ‘on-the-fly.’ These pages are defined by form templates. Wrapper induction can be easily trained to learn these templates.

Regardless of the approach used, IE is based on pattern matching and patterns are established through machine learning [2]. Manually indexed documents are first created and used to ‘teach’ the IE software what attributes make up relevant content.

Language processing and information extraction are domain specific. Meanings of words or phrases in one context do not necessarily transform into another. For instance, the term ‘AIDS’ has one meaning within healthcare but can also be viewed as a verb which implies a different interpretation. The CDC is currently defining a structure for categorizing search capabilities and will need to extend the keyword search with natural language processing and wrapper induction.

3.2.6. Enterprise Application Integration

As Web portals transform raw data from diverse sources into useful information tailored to a single user, they require a high level of systems integration. The Web portal must be able to provide seamless communication between the enterprise, general users and partners and among legacy and other applications. This is achieved through Enterprise Application Integration (EAI). Point-to-point custom interface code as well as interface engines acting as a central gateway to translate and route data have been the primary technologies used to exchange data between applications. However, a more global and uniform method is needed.

To address this need, the CDC has defined a Public Health Data Model [8]. The purpose of the data model is to document information needs of public health to facilitate the exchange and dissemination of information with CDC partners. The premiere edition of the data model was published in July 2000 and was made available for comment by partners and health informatics standards setting bodies such as HL7. Compatibility with HL7 is deemed important as HL7 is the most commonly used standard for interoperability between healthcare information systems [10]. XML which is becoming the technology of choice for structuring data for exchange between systems has been incorporated within HL7 Version 3.

3.2.7. EMPI

The primary objective of the EMPI is to uniquely and accurately identify individual patients and their corresponding information across various healthcare systems. The CDC does not work directly with individual patients which greatly diminishes their need for an EMPI. Tracking of patients is done at local level. However, with the introduction of the electronic surveillance system, the CDC has a need to uniquely identify specimens across various CDC systems. An EMPI system would serve this need.

A summary of CDC specification mapped to portal framework layers is presented in Table 1.

Table 1 – CDC specifications mapped to the healthcare portal framework

<table>
<thead>
<tr>
<th>Framework Layer</th>
<th>CDC Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>Incorporate the existing security plan integrating the Secure Data Network within the portal security layer to facilitate continued use of digital certificate technology.</td>
</tr>
<tr>
<td>Presentation</td>
<td>Presentation will be standardized on Web browser access employing CDC Web page design templates.</td>
</tr>
<tr>
<td>Business Logic</td>
<td>Incorporation of recent system initiatives and existing business applications via portal access.</td>
</tr>
<tr>
<td>Portal Services</td>
<td>Provide personalization services to address the unique needs of the identified user communities working towards the ability to dynamically customize user experiences.</td>
</tr>
<tr>
<td>EAI</td>
<td>Incorporate the Public Health Data Model that adheres to HL7 standards including HL7 defined XML specifications.</td>
</tr>
<tr>
<td>EMPI</td>
<td>Secure an EMPI system to support tracking across CDC systems where necessary.</td>
</tr>
</tbody>
</table>
4. Conclusions

Planning for an information technology architecture that embraces emerging technologies is challenging for an organization with complex information needs. While the Internet and Web-based technologies have been recognized as the platform for the future, these technologies are still immature. Time-tested solutions do not exist. Therefore, an organization must be proactive in defining an architectural model that encompasses these technologies and meets the organization’s vision for the future. The healthcare portal framework provides such a model. The layered architecture of the portal framework supports an evolutionary approach to implementation thereby mitigating the risks inherent in the early adoption of new technologies.

The future challenges remain with being in close touch with CDC users, customers and partners in bringing the power and advantages of information technology to solve the scientific and business challenges of the CDC’s public health mission and vision of “Healthy people in a Healthy World Through Prevention”

5. References


