From Static Web Site to Portal

Villanova University provides a case study for integrating a dynamic and individualized Web system

by **Christopher G. Connolly**

When the concept of portal Web sites emerged, Pennsylvania's Villanova University realized this concept accurately described many of its objectives. Portals—sites that serve as a starting point for people when they connect to the Web or that they tend to visit as an anchor site—offer the potential to organize and customize information. Villanova wanted to create a Web system that informed people of the latest news and events, could be customized, and took advantage of information stored in the university's records system. Given these objectives, we have spent more than 18 months overhauling our static Web site to a dynamic and individualized Web system.

Before we look at preliminary steps for developing and integrating portal technology and enabling Web-based solutions, it is important to understand the differences between Internet Web sites, intranets, and portals.

An Internet Web site is the most basic manifestation of Web technology, providing information through hypertext markup language (HTML) that allows for cross-referencing with hyperlinks. Consumers worldwide can access the Internet; thus the information presented on an Internet Web site is intended for public consumption without restriction.

An intranet Web site is one that is contained within an organization. Its purpose is to provide more relevant, useful Web-based solutions to an internal community. People have access to an intranet through an authentication process, usually involving usernames and passwords that they must remember. Intranets afford a way to present information to a restricted audience.

For Web-based application developers, the dichotomy between Internet Web sites for presenting public information and intranet Web sites for presenting restricted information creates a development burden. Several Web sites must be maintained to provide security and confidentiality for certain kinds of information, possibly even duplicating some information on more than one site.

A portal, on the other hand, is a gateway to the Web that allows the plethora of information available on Internet and intranet Web sites to be organized and customized through a single entry point. A good portal provides seamless access for nonauthenticated users until sensitive information is requested, when it then prompts for a username and password. Authenticated visitors or those known to the site by cookies (textual information passed to the client to be stored on the client's system) are presented with a more individualized view of the organization's Web site.

There are, however, many products on the market that are simply advanced intranet Web sites that take better advantage of information known about the user. These portals can fail developers by requiring them to maintain an Internet Web site for nonauthenticated visitors and a portal for those with authentication. Storing a university's student campus activities schedule in a password-protected portal prevents prospective students from viewing the university's activities, but placing the schedule on the public Internet site prevents campus Web designers from personalizing the information. For example, designers could place sophomore activities at the top of the schedule when those students view the content.

Unlike Internet and intranet sites, most portals are proprietary and thus often do not conform to any standards. Although they may provide application programming interfaces (APIs), each provides a different set of APIs. Portals are meant to be a solution for multiple intranet username/password systems, which are based on many open standards; however, more vendors are offering portal solutions for different functional needs such as athletics or the student record system. What is the solution once an organization has several different vendor-based portals?

In-House Versus Commercial

One of the challenges an organization faces when considering a portal approach is determining whether to develop an in-house software solution, outsource such development, or purchase a commercial solution. The organization must assess the need for a competitive advantage in the technology arena, the ability of an off-the-shelf product to meet those unique needs, and the resources necessary to develop an in-house solution. The Netscape browser is a good example of a technology that meets almost any organization's unique needs. Most organizations, with the exception of competing browser vendors, gain little to no competitive edge by developing their own custom-built browser. Advanced, underlying technologies such as Web browsers, Java, and e-mail are examples of open standard solutions that generally should not be developed by organizations that are not

Glossary

API—Application Program Interface; the way in which software communicates with other software

ASP—Active Server Pages; Microsoft's Web server-scripting language that provides dynamic content

Cookies—Textual information passed to the client to be stored on the client's system. Cookies allow Web servers to identify clients.

HTML—Hypertext Markup Language; a formatting language that allows programmers to communicate with Web servers and browsers

Java—A high-level, object-oriented programming language developed by Sun Microsystems

Java servlet—A Web server application, usually intended to provide dynamic content, that is written in Java and conforms to Sun Microsystems' specifications

Java taglet—An HTML tag that can be placed in a Web page to access a Java Servlet before being served to the client
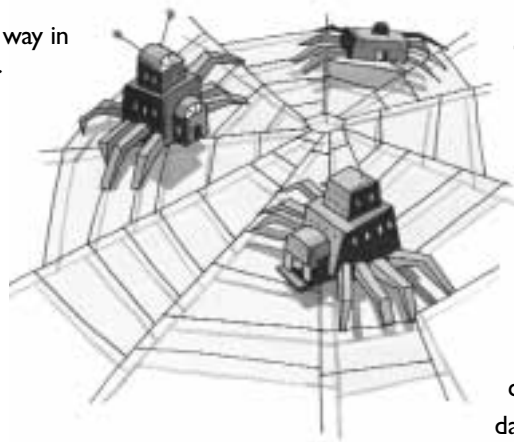
JavaScript—A scripting language developed by Netscape based on Java that can be embedded in Web pages to add dynamic functions to the page

LDAP—Lightweight Directory Access Protocol; Netscape's strategic directory protocol. It defines a simple mechanism for Internet clients to query and manage a database. LDAP is compatible with many commercial and freeware products.

Object-oriented—A style of design in which systems and information are viewed as objects that interact with each other

SQL—Structured Query Language; the international standard language for defining and accessing relational databases.

Tags—Short commands that are the "words" of the HTML programming language

---

specialized in these technologies. However, Java servlets (Web applications that run on a server) and active server pages are programming interfaces developed to allow organizations to create custom applications. Writing a Java servlet that displays the photographs of the students enrolled in a course by bringing together information from vendor A's student record system and vendor B's picture ID card system is a good example of an in-house application.

Are portals an advanced, underlying technology such as e-mail or an applied one based on existing technologies? The answer is that today's portals are an applied technology application based on existing technologies. Currently, most portals are simply several underlying technologies packaged together that are fairly simple to develop. The organization must assess its situation to determine whether to develop a portal that will provide a competitive edge or whether to purchase a commercial solution that will allow it to focus on other issues.

A number of commercial portals geared towards the higher education industry have incredibly low prices because the companies offering these products realize that university students represent a lucrative demographic. A vendor sells its product to a university at a low cost by displaying advertisements on the portal pages. Each university will need to assess carefully whether to use a low-cost ad-driven portal, a higher-cost non-ad-driven portal, or a portal developed in-house.

Steps for Integration

Once an organization decides whether to develop in-house or purchase a commercial portal, the following steps must be completed. Some of these steps come with a commercial product. For each step listed, I've described how Villanova approached and resolved that issue.

USERNAMES AND PASSWORDS

A system of centralized usernames and passwords is key. There is little point in providing Web tools if people do not know their usernames and passwords.

Villanova chose Netscape's directory or Lightweight Directory Access Protocol (LDAP) server for its central repository of usernames and passwords. LDAP is an open-standard protocol that is compatible with many commercial and freeware products, which reduced development and integration costs. Sun Microsystems provides a Java Naming and Directory Interface for LDAP, supplying an open-standard programming interface for our username and password solution. Currently, Villanova's remote access server (Merrick Radius Server), Netscape proxy server, Apache Web server, Internet protocol (IP) registration system, Netscape Usenet News server, Netscape messaging server (e-mail), Netscape calendar server, Microsoft FrontPage, and Novell servers are linked to the LDAP system. Most of these systems talk natively to LDAP, requiring no code development to integrate. Villanova plans to integrate its Windows NT/2000 environment into its centralized username and password system during the summer of 2000. Villanova required fewer than 12 months to build and integrate its centralized username and password system.

CUSTOMER RECORD SYSTEM

For a portal to communicate effectively and efficiently with students, faculty, staff, and alumni, it must know as much as possible about each individual. If a university operates different systems for payroll, students, and alumni, it becomes a difficult hurdle for a portal to overcome.

Villanova centralized its operations into a single database system. The university installed SCT's Banner system five years ago, and most departments are now integrated. The information technology (IT) department's helpdesk system, university card system, and telecommunications billing systems are examples of departments using databases that are not yet fully integrated into the central record system because the vendor did not offer the needed functions. The helpdesk and university card systems have been linked with the LDAP system so information can be retrieved with minimal effort.

CENTRALIZED INFORMATION

Although a university may have a centralized customer records system, it is often too complicated to develop Web content from these data. A university's record system might have a customer's information stored in more than 20 different locations (called tables in database terms). Filtering through these tables when attempting to customize the look and feel of a Web page can be time consuming both to develop and execute. In addition, it is likely that the

---

Universities will need to assess whether to use a low-cost ad-driven portal, a higher-cost non-ad-driven portal, or a portal developed in-house.

---

portal will need to store and manipulate its own information that the records system may not be able to handle.

There are two key issues to consider when compiling information into a central repository system: what information needs to be warehoused (meaning a compact and easily accessible formatted version of the customers' information) and how that information is kept up to date. For example, if the portal organizes information based on the student's major, then it is important that the system be made aware of students who have changed their major. A repository may also be necessary if the centralized records system must be brought down once a day for backups or once a semester for upgrades. Most customers want their Web sites to be up 24 hours a day, seven days a week, so it is important that the portal can access its information 24/7. Repositories may be stored in simple database tables, a directory server such as Novell directory services or LDAP, or object-oriented databases.

Villanova chose LDAP as its central repository for relevant customer information. College, major, class year, permanent address, campus address, phone number, department, and university card photo are just a few examples of the information stored in the LDAP server. LDAP is updated from information in the university records system in one of two ways: via a nightly synchronization process or a database event listener. The nightly synchronization process checks each LDAP account's information against what is stored in the university records system and updates as necessary. The database event listener checks relevant tables in the records system, such as the address table, and triggers an event when a change occurs. The triggered event is then processed and the information is checked against LDAP. This process provides real-time updating while the nightly synchronization process catches any changes missed during the day or as a result of downtime.

APIs AND DOCUMENTATION

Commercial portals or those developed in-house should provide an application program interface (API) and documentation for the information in its storage facility. The API allows a university to customize and tailor the information to its needs. In addition, a properly constructed, object-oriented API can provide a robust and resilient interface. For example, accessing a database table requires a raw structured query language (SQL) statement that is not very resilient to future upgrades and modifications. An API with simple "getters" and "setters," such as a getUserID

method, provides an interface to the information that is much easier to use and more robust. Although the APIs may use standard programming structures such as Java Naming and Directory Interface or Java DataBase Connectivity, the methods provided for each are not standard across portals. Case in point, vendor A's portal may have a getUserID method while vendor B's portal may use a getUsername method. These discrepancies can be a major hindrance and cause portals to become no more than high-end intranet products. For in-house developed portals, documentation for the API and the underlying code itself are required.

Because Villanova developed its own repository using LDAP, it had to develop its own API and documentation. The API, written for the Java programming language, and documentation are made available to any department or college within the university that wishes to use the information. Often the API is used for non-Web-related projects, such as configuring people's Windows machines to reflect their username, location, and workgroup. The API is currently in its second generation; it is easier to use and more object oriented, which makes it more resilient to change.

INFRASTRUCTURE INTEGRATION

The above steps, including developing or purchasing the portal software, require a significant investment of time and money. Existing and new systems integrate into and take advantage of the new infrastructure and that provides the true cost savings and benefits. There is more to most universities than their Web sites. Combining resources such as the Web, e-mail, e-commerce, and classroom instruction to take advantage of the same individ-

ualized repository of information reduces administrative and maintenance costs and presents a more seamless environment for the customer.

For commercial portals there are both specialized and generic portals for integration. A specialized portal is a one that is supplied for a specific target market or extends an existing product. For example, a vendor may enhance its athletic recruitment database system by offering a portal that immediately integrates with its existing database product. However, a registrar's office will have difficulty using an athletic office's portal for its own needs. A generic portal is one that is intended to be flexible enough for any organization's needs. A generic portal, however, usually exists outside

Existing and new systems integrate into and take advantage of the new infrastructure and that provides the true cost savings and benefits.

of an organization's present systems. This case requires an in-house or outsourced integration between the existing system and the portal.

Villanova has integrated many existing and new systems into its IT infrastructure. The integration with our e-mail system now allows our faculty and staff to target their specific student audiences using college, major, year, college-year, and major-year distribution lists. In addition, every course taught at Villanova is supplemented with its own e-mail distribution list and bulletin board.

Integrating our Apache Web server and Microsoft FrontPage extensions has provided several benefits. Every

faculty member, staff member, and student is automatically provided with a personalized Web site. Staff members are automatically granted or revoked rights to edit a department's Web site when they begin or end employment with that department.

Individualized Solutions

After the infrastructure is in place and a portal has been developed or purchased, it is important not to lose sight of the customers' needs. Many IT departments spend so much time and effort designing the underlying infrastructure that they believe the dynamic Web sites are just extensions of their infrastructure. However, the opposite is true. The dynamic Web sites provide the information and tools the customer needs while the infrastructure exists only as a means to support such a rich system. A common feature of a portal's MyHomepage is sports updates. From a university point of view, the idea of a sports section seems quite trivial and irrelevant. From a student's point of view, customized sports updates are a useful resource for the university Web to provide. Universities that make sports updates available will not only increase customer satisfaction and portal use, but will also enable the university to place its own news and events around the sports section, increasing student awareness of campus activities and opportunities.

Villanova's most successful portal tool has been its MyClassrooms feature. MyClassrooms provides a common link to classroom material and resources that are available on the Internet. All university courses include an e-mail class distribution list and a class bulletin board for faculty and students to collaborate throughout the semester. The key feature of MyClassrooms is the ability to view the photo ID of all the students enrolled in

a course. This has been extremely well received by students and faculty members.

Developing and implementing the customer-needed features of a portal are only the first steps of a properly deployed portal system. Typically an IT department represents only a small portion of a university's workforce. Therefore, it is essential that nontechnical users outside the IT department can take advantage of the portal system's advanced capabilities. Portal technology offers departments the ability to incorporate security and dynamic features easily using customized HTML tags. A good portal also should be able to work with existing Web servers. It is rare that an entire university operates on a single Web server. Therefore, the portal solution should be extensible enough for other departmental Web servers to

take advantage of the technology.

Villanova University has taken advantage of Java taglets, which provide a way to implement customized HTML tags to interact with Java servlets. This allows HTML developers to add security or dynamic content to their pages with

Organizations will need to advance their Web-based technologies to stay competitive.

simple HTML tags. It also provides a method to repeat dynamic information efficiently across different pages. For example, the university library system can display the same dynamic library

news module on the university homepage as well as on the library homepage.

Dynamic and individualized Web systems will become essential for organizations in the next several years as customer expectations grow and as organizations further develop their Web-based technologies to distinguish themselves from their competition. It is important not to lose sight of the original goals of the Web, which are to provide a seamless, heterogeneous environment to access information and avoid the bells and whistles of a proprietary portal solution. Villanova University has successfully integrated portal technology by constructing and using a solid IT infrastructure. *e*

Christopher G. Connolly (chris.connolly@villanova. edu) is assistant director of www and university IT at Villanova University.