



Faculty Publications

---

2009-09-15

## The Role of Web Services in Portal Design: Approaches for an Algerian University Library

Mohamed Ridda Laouar

University of Tébessa, Algeria, ridda\_laouar@yahoo.fr

Richard Hacken

hacken@byu.edu

Mathew Miles

Brigham Young University - Idaho, mmiles@byui.edu

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>



Part of the [Library and Information Science Commons](#)

### Original Publication Citation

Library Hi Tech 27 (3): 460-479

---

### BYU ScholarsArchive Citation

Laouar, Mohamed Ridda; Hacken, Richard; and Miles, Mathew, "The Role of Web Services in Portal Design: Approaches for an Algerian University Library" (2009). *Faculty Publications*. 1577.

<https://scholarsarchive.byu.edu/facpub/1577>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

## **The Role of Web Services in Portal Design: Approaches for an Algerian University Library**

Whoever first thought of using a networked computer as a scholarly workstation to which (and from which) information might flow freely may well have been a librarian, at least in the honorary sense. But the dream of such a data nexus – along with all the tools of connectivity, discovery and documentation a scholar is likely to need – has been developing in steps, piecemeal, and is still not yet fully realized.

The need to establish an academic information portal for our university library led us to study the literature of documentation and design principles for existing portals as a means of choosing best practices that would fit our situation. The growing number of web applications to manage – the delicious dilemma of too much information – was a problem we hoped to solve by setting up a portal. We also felt that such a portal could provide answers to problems of data transfer and access (Ignat et al., 2006). In sum, we wished to design a methodology whereby each of our users could draw together multiple information tools and sources from a single access point.

In something of a balancing act between technology and economy, we wanted to find how extant computer systems might help us aggregate as much information into easy-to-use terminals as possible while perforce implementing such a system with low entry costs. In the following article, we suggest provisional definitions for portal phenomena, present a series of evolutionary steps we have witnessed in the development of information portals – looking at specific portals along with the desirability of their design properties – and then we analyze the lessons learned by discussing some web technology elements that seem promising for the architectural planning and infrastructure creation of a library portal at the University of Tébessa in Algeria.

### **1. DEFINING THE “PORTAL” AND ITS VARIANTS**

Architecturally and therefore literally, a “portal” is some sort of gateway or even grand entrance – especially into an important edifice such as a fortress, mosque or cathedral – whose features make it possible to control entry and exit and are often highly ornamented. The figurative meaning of “portal” in the virtual realm may benefit more from the borrowed concepts of entrance into an important space and entry control than from ornamentation

(though the latter could be tied to attractiveness or user friendliness). Thus in the broadest sense, while keeping in mind the architectural parallels, we could say that a “portal” represents a unique electronic port of entry to a wide range of resources, services or both that are needed within a particular information community.

The closest general term to “portal” is perhaps “information portal,” though a given portal may contain ancillary tools, often as part of a complementary “application portal” (see section 1.1), that are not directly associated with information retrieval and dissemination. Since our university library in Algeria operates as a bilingual Arabic and Francophone institution, we note that the general term in French “*portail documentaire*” is a false cognate for English speakers: the word “*documentaire*” in the technological sense means “information” and goes beyond the narrower meaning of “documentary.” A “library portal” would be one type of “information portal” and would necessarily consist, at least in the current understanding, of at least metadata and data that can be locally or remotely hosted.

The “scholarly portal” or “scholarly information portal” distinguishes the academic from the commercial portal. The “enterprise portal” or “corporate portal” is congruent with the design of a scholarly portal but with a very different bottom line and user community. Both scholarly and enterprise portals increasingly rely on de-centralized provision and management of content, which ideally keeps the content current and minimizes local labor costs at any one institution.

We are tempted nowadays to use the phrase “web portal” as a synonym for “information portal,” especially since web access is ubiquitous and forms the central engine of today’s portals. As we shall see in the following historical section, however, early types of portals were local and mostly self-contained – though some of them could be accessed in awkward ways remotely. Early library catalogs are one example of such early non-web portals. At the present time, though, a web portal can grant the information community (library users, company employees, etc.) access to the panoply of services involving electronic mail, discussion forum, digital publication space, proprietary databases, search engines, and more.

One of the strengths of a current web portal is that it typically allows the user to authenticate a single time to access all the resources offered. This saves users the time and frustration of having to authenticate each time they access a different resource. Authenticated users are allowed to register and connect to the portal for using a set of proposed services, notable among them the customization of their own workspace, which is

organized with the assistance of portlets, forming a basic man-machine interface (IHM : *“interface homme machine”*). For staff and, if desired, partners and affiliates, the web portal is thus the entry point for data retrieval within an organization. The aim of the portal is to situate the user centrally within the information system. In fact, the tools bundled in the portal allow the user:

- to have a single access point for multiple sources of information,
- to acquire information that is organized and structured,
- to shape and customize information to the user’s interests and proposed tasks.

A “digital library” is one specific type of portal that can involve a combination of searchable metadata and retrievable digital objects, though it may or may not limit itself to locally held objects and may include some of the peripheral services listed under “web portals” above. The term “digital archive” is sometimes used for an analogous online collection. As one answer to the scholarly communications crisis (aggravated in a time of financial hardship), an “institutional repository” is a type of local digital archive of a home institution’s professional publications – to the degree copyright clearance is possible.

A “federated search,” involving the retrieval and combination of results from multiple information providers, may be a variety of search portal in its own right, but it is often used as one cornerstone in a more variegated online library services approach. That is to say, one type of “portal” may be contained within another “portal.” Perhaps the term “P2P” that now refers to a “peer-to-peer” network might come to include the nested connotation of “portal-to-portal” access as well.

A “metadata repository” is a broad term that could include catalogs, digital libraries or portals that allow for searching of metadata. The item thus identified might be either digital or analog, either local or remote.

“Information systems” is perhaps the broadest term of all, which would subsume all types of “portals,” “catalogs,” “digital libraries” and “metadata repositories” along with other cyber-tools for information query and/or storage.

Smith (2004) describes the term “portal” as problematic, since it has been applied to widely different systems – from static and simple web pages to dynamic and sophisticated virtual conglomerates. He argues for adoption of “a definition that distinguishes portals from all other types of information systems and a General Portal Model (GPM).” (Smith, 2004, p. 93). His suggested definition is that of “an infrastructure providing secure, customizable, personalizable, integrated access to dynamic content from a variety of sources, in a variety of source formats, wherever it is needed.” (Smith, 2004, p. 94).

Kalyanaraman and Sundar (2008) seek to add some conceptual clarity to the portal concept by making functional differentiations. They uncover five related metaphorical notions for the portal – gateway, billboard, network, niche and brand – that in turn suggest five leading aspects of portal utility – respectively, control, content, community, customization and commerce.

For purposes of clarity, we will distinguish in the following article between the types of portals discussed in this section. Unless otherwise stated, we will generally be referring to an information portal tied to web services.

### **1.1 IDENTIFYING VARIANT FUNCTIONS OF PORTALS**

We generally distinguish functionally between three broad portal families:

- the information portal, granting access to multiple information sources (documents, reports, articles, internal or external databases, etc.) aggregated in a single point;
- the application portal, granting access to the various applications of the organization with corresponding data relevant to the user profile;
- the appraisal portal, sometimes called the standby portal, enabling analysis of portal usage as a means of enhancing the organization's knowledge base.

### **1.2 NOTIONS OF PORTAL USER PROFILES AND CUSTOMIZATION**

The web portal concept is closely linked to the notion of a user profile. Ideally, each user has access to resources of the information system that fit the user's unique profile, all in accordance with security policies defined by the organization. At the same time, the user profile can guide the way to customization (Chevalier et al., 2007), and so we speak of an "online work environment" or "virtual office." Following this terminology, the environment is constructed of modular bricks (generally called "portlets" or "webparts") from which the user can choose to organize relevant scholarly structures within his or her workspace. Popular customizations being built into portals include personalized search history lists, RSS feeds and the ability to store bibliographic entries for later printing, e-mailing or saving. The latter illustrates how a popular commercial idea (the virtual "shopping basket") has been adapted to scholarly uses.

## **2. HISTORY OF PORTALS**

The emergence and spread of new Information and Communication Technologies (ICT) – notably of the Internet, of research networks and enterprise intranets – have abetted the growth of a phenomenon now justly referred to as the "information explosion." Scholars would be casualties of the information explosion were it not for support systems that can help to pinpoint and extract relevant info particles from the vast oceans of data (Mothe et al., 2003). The sorcerer's apprentice of technology that has overloaded our neural networks, ironically, has also provided cyber-networks to help control the flood it enabled. Over the past few years, the use of web technologies within organizations (companies, universities, libraries, and so forth) has become an essential element in managing different sectors of information. The evolution of portals of all kinds shows an increasingly strong trend toward more loosely coupled modular systems.

### **2.1 LIBRARY CATALOG MODELS**

#### **2.1.1 EARLY LIBRARY CATALOG MODELS**

In the commercial online realm, some of the first portals, such as Yahoo!, were web directories; others were search engines like AltaVista. Just as these were essentially searchable catalogs of "holdings," i.e., of documents inside the World Wide Web, likewise early library catalogs (many predating Yahoo! and AltaVista) contained metadata for individual library holdings. As such, the holdings were limited to locally owned content. Search strategies were often command driven, requiring the user to memorize a command structure [e.g., "((A=Tolstoy or A=Tolstoi) and (A=Leo or A=Lev)) and (T="voyna i mir" or T="war and peace")"]. In contrast to researchers in house, remote users – if even authorized – required a complex combination of protocols to connect to the catalog. Relating to esthetics as well as utility, early catalogs were not yet amenable to a graphic user interface (GUI). Yet these catalogs could be considered proto-portals, since they combined in one place the query potential for finding millions of analog items.

The overwhelming emphasis in early library catalogs was on metadata rather than data. The metadata and the materials to which they pointed (mostly analog) had to be maintained locally.

### **2.1.2 CURRENT LIBRARY CATALOG MODELS**

The website for a university or research library today invariably offers much more than a catalog. Advancements from the early catalogs to the current models are in the breadth of ancillary tools and guides (at times leading to an inadvertent de-emphasis on the catalog itself) as well as increasing use of remote metadata retrieval and burgeoning linkable presentation of digital objects. In this section we limit our observations to the catalog itself; more broadly conceived websites with additional portal elements will be discussed under “hybrid portals” (See 2.5 below).

The current state-of-the-art catalog can offer attractive and intuitive interfaces that include images and external links while minimizing user aggravation with search jargon. Just as importantly, the scope of MARC records has expanded to locate holdings beyond the walls of the library and beyond tangible analog materials: e-books, web pages, electronic journals, streaming audio and video, and other local or remote documents. You are not only guided to material locations onsite but you can receive (view, see, hear, read) virtual copies of many items at your computer. The search interface yields to the study surface of that which was sought.

From the maintenance point of view, at least some of the metadata and holdings (analog or digital) are hosted elsewhere.

## **2.2 METADATA REPOSITORY MODELS**

Metadata repositories have evolved in parallel fashion as library catalogs as far as design goes. They differ from library catalogs in not indexing a single local library or collection but an aggregate of holdings at numerous other sites. The main progress has been from metadata-only searches to the current search-and-link connections to remote digital holdings.

### **2.2.1 EARLY METADATA REPOSITORY MODELS**

OCLC and RLIN are examples of early metadata-only repositories. These bibliographic utilities host metadata from a variety of sources, but do not own any content of their own. Typically the metadata does not link you out to content. It simply provides information about where the content is located.

### **2.2.2 CURRENT METADATA REPOSITORY MODELS**

The National Science Digital Library is a repository that houses and hosts metadata and indexes but no content. It gathers the metadata in a variety of ways, including OAI harvesting, ftp, email, direct entry and web crawling. This model works fine for those that have the resources to host a very large database. It begins to break down as the amount of metadata increases past a critical point. The advantage of not having to host data may be offset by a lack of control over data and the necessity of relying on other institutions for accuracy. (For more information on the NSDL see “hybrid portals” below.)

Within Algeria, the “Réseau Inter Bibliothèques Universitaires” (RIBU) is a metadata repository under development for a consortium of ten Central Algerian academic institutions. Sponsored by two European institutions, the Free University of Brussels and the Université d’Aix Marseille I, it was launched in 2005 with a standardized system for library management (SYNGEB). The objective is to establish an electronic union catalog for this part of the country, and the holdings of the Tébessa university library will become a part of this collaborative effort.

### **2.3 FEDERATED SEARCH MODELS**

Returning to the architectural metaphor, federated searching developed from the desire to enter and profit from a number of important spaces through one entryway. (“Information mall” doesn’t have the right feel.) Figuratively, the user can query a range of relevant databases simultaneously. The model has evolved and improved with developments in technology and the Internet. Current federated searching holds out promise as a model for the future, though it leaves something to be desired at present.

#### **2.3.1 EARLY FEDERATED SEARCH MODELS**

Dialog Corporation’s DialIndex was a very early example of federated searching. It allowed trained information seekers to sift through years of multiple databases via Telnet with a single search interface. The drawbacks were its very high cost and its lack of scalability in such a way that it could be distributed to the public. The user would be required to come to the library and have a librarian trained in Dialog intricacies – as an intermediary for the actual information seeker – perform the DialIndex search for a fee.

### **2.3.2 CURRENT FEDERATED SEARCH MODELS**

The current principle of federating searching involves software “translators,” usually provided by a vendor like WebFeat Knowledge Prism or ExLibris Metalib, that search the indexes of other content providers who host the indexes, metadata and content. This is the most modular, scalable and loosely coupled system to date; it is made possible by the use of standard web protocols that have been widely adopted. The federated search does not require the housing of content, indexes or metadata and thus shares similar advantages and disadvantages as current metadata repositories – with the added caveat that at present the result sets reflect the “lowest common denominator” among individual database features from which the metadata arrives.

At the turn of the millennium, the Grainger Engineering Library at the University of Illinois at Urbana-Champaign wished to provide “web-based asynchronous simultaneous searching of multiple secondary information sources and integrated reference linking between bibliographic resources.” (Mischo et al., 2002, p. 119). Today we might call that “federated searching.” The designers decided to create a custom federated search, for which they combined two different search technologies. One employed lightweight scripting against off-the-shelf web-based services (see section 4 below for more about web services), while the second was a more sophisticated, distributed approach requiring advanced programming techniques. The advantages of the first were simpler implementation and greater scalability. The disadvantage was the lack of flexibility in allowing the search code to interact effectively with each individual database. The second, more advanced programming method was easier to adapt to the idiosyncrasies of each individual database. The Grainger Library also implemented OpenURL resolving with these federated databases, which allowed greater access to full-text articles without having to perform multiple searches. Our resource limitations in Tébessa preclude such a model for us.

### **2.4 DIGITAL LIBRARY MODELS**

Strictly speaking, a “digital library” would be the fully automated flip-side version of an analog library, i.e., it would house both online metadata and the full range of targeted digital documents held locally, with links between the two.

“Institutional repositories” of locally produced digital publications, found at increasing numbers of universities or corporations, could be regarded as a special subset of digital libraries.

BAMBI (Better Access to Manuscripts and Browsing of Images), a European project for transcribing and annotating scanned ancient Greek and Latin texts, is a digital library model that includes collaborative functionality. It is aimed at both general users of a library who want to examine manuscript sources and, more importantly, at professional students of texts: philologists or critical editors of classical or medieval handwritten works : papyrologists, epigraphists, palaeographers, and codicologists (Calabretto et al., 1998). BAMBI has true Web 2.0 functionality, meaning the ability to tag or annotate using a web interface; it represents asynchronous work.

DEBORA Project (Digital Access to Books of the Renaissance) is another European project involving collaborative functionality. Its goal is to allow different users and scholars to participate in the digitization, metadata creation and posting on the Internet of rare sixteenth century books. The collection being created within DEBOPRA consists of digitized images of books from libraries in Lyon, Rome and Coimbra. (Le Bourgeois et al., 2001). With remote access, collaborators can use system tools for image processing, for metadata creation (tagging, annotations, rating of materials) and error correction. Since a special Java-based client – which is not widely available – is used for viewing and annotating, the use is mostly limited to scholars from the participating libraries.

The University of North Texas Libraries have assembled what they call a Texas History Portal, yet its architecture and functionality seem to brand it a digital library. Digital assets and metadata expressing the history of the state of Texas are self-contained on the UNT server. They flow into three different interfaces for distinct classes of potential users: Young Scholars, Researchers and Educators. Each of these areas provides auxiliary materials such as lesson plans and PowerPoint presentations. The most difficult part of creating this project was to design and build connections between the archival system of holdings and the tripartite front-end presentation. Metadata creation presented similar challenges: labor-intensive human input is very expensive (Nordstrom et al., 2004).

Gallica, meanwhile, the online digital library of the Bibliothèque de France, remains the gold standard. It has been in operation since 1997, has easy access to magnificent and rare texts within its own collection, and enjoys the budget support of the French state. It has expanded to include tens of thousands of books, journals, newspapers, maps, scores, engravings, photographs and sound recordings. Gallica with its high-end costs does not represent a model we can follow in Algeria.

## 2.5 THE MOVE TOWARDS HYBRID PORTALS

As web portals evolved, companies began adding more services like email, chat and news to attract and retain users. The academic web portal concept received more attention with the growing success of other types of portals on the web. Design concepts led to a type of one-stop research environment characterized by distributed access over multiple platforms with ease of use and desirable levels of authentication and data security. The hybrid academic portal has become a tool that integrates all the work applications that the user needs on a single site; it can be targeted to on-campus users as well as to external members of the academic community. It offers consolidated access to relevant information sources and, with its bias toward finding rather than searching, it saves research time.

As we mentioned above, some so-called “digital libraries” tilt heavily towards a model that combines metadata-to-data links with ancillary services. The National Science Digital Library (NSDL), for example, defines itself as a “metadata repository, search and discovery services, rights management services and user interface portal facilities.” (Lagoze et al., 2002, p. 201). Though all of its data is offsite, NSDL defines the entirety of its project as a digital library with the various user interfaces being called portals. The success of the \$24 million project was based on the idea that the ability to partner with, and aggregate, a variety of information sources would lower entry costs. Another aspect of lowering costs was the use, within the search phase, of simpler web-based protocols such as http and html rather than “more complex and rich mechanisms such as SGML and Z39.50.” (Lagoze, p. 202).

There are at the present time innumerable examples of a library catalog expanded to hybrid portal proportions, or, as it is often informally referred to, the library’s “home page.” Most academic or research libraries are now taking advantage of its modular features. One such library site, which we offer as typical of many, serves the library of the London School of Economics. Besides linking to the online catalog, this portal links to other electronic resources (many of them password protected for LSE users); to the archival collections; to news and calendar events; to service and local facilities information; to interactive Web 2.0 “ask a librarian” features, subject guides, online exhibitions and more. In true hypertext format, each secondary level, such as the archival collection, then offers a new range of choices within the new context: archives catalog, library catalog, pamphlet collection, rare books, and focused ancillary information for the archival user. An overview of the various trunks and branches of the web tree comes in the form of an A-Z site index, and a search box

allows for pinpoint accuracy. An alternate means of reaching the user community is through “Services for you,” a frame that supplies targeted and contextualized support to staff, students, alumni, visitors, or users with disabilities.

From a design and implementation perspective, the library portal of the Université Laval in Québec, Canada, is centered on a modular series of “portails thématiques” (subject portals). Each of these allows the addition of content with a simple and intuitive content management system (CMS). Similar to the editing interface of LibGuides by Springshare, this allows for a distributed model of labor among librarians trained in a subject but untrained in computer programming. A disadvantage might be an irregularity of contents subject by subject, since they are not the product of an integrated system but are created in traditional ways that may be idiosyncratic or uneven (Chicoine, 2008). From the user’s perspective, there is an opportunity to enter a customizable environment and seek information and guidance within a niche topic of interest (Teasdale, 2008).

For glimpses into the future of the hybrid scholarly portal, the Société Jouve (Jouve Company) in France has developed a number of academic information portals, each often referred to as a “Service Commun de Documentation” (SCD, roughly translated as “Collaborative Information Service”), at universities such as Nantes and Valenciennes. These sites were designed around the use of “open source” software.

### **3. PORTAL DESIGN: LESSONS LEARNED FOR THE TÉBESSA PORTAL**

In portal projects that we analyzed, we saw the importance of virtual communities clustered around interactive tools and expert facilitators. The main functions we seek are authentication, federated searching and content management. Beyond that, we wish to create relevant new services and applications. Features we found intriguing for us to emulate are:

- intuitive searchability and resource management with a simple interface;
- federation of multiple tools around a central communications system;
- value-added services that integrate teaching and research;
- creation of a virtual workplace with collaborative functionality;
- adaptation to individual user profiles while allowing full access to all authorized users;
- integration into scalable systems.

We also identified certain elements missing in some portals that we would nevertheless like to design into our portal – to the extent resources allow:

- Many portal projects do not track users' actions. Storing these keystrokes assists their work, since the tracks allow the research, annotation or search work to continue from the point where the user previously left off. In addition, these tracks can be used for customized help requests.
- Many portals have no personal storage space for regular users. This space, or access to a bibliographic management tool such as RefWorks or EndNote, retains personal bibliographic information for later use.
- The majority of portal projects have no area for collaborative work on documents, space that would allow more extensive projects to be worked at in tandem.

Our aim is to design a hybrid portal – at the center of users' workflow – that can make a synthesis between the classical supply of library resources (catalog, metadata search, electronic resources) and additional enrichment tools and services based on a dynamically changing collaborative world (Tamine et al., 2007). Web services and widgets (see section 3.1) seem well suited to help us achieve our goals for the University Library of Tébessa without prohibitive costs.

We live in an age when access to enormous quantities of information on all topics is facilitated by the web and the boundaries between source documents acquired remotely and those produced locally are vanishing. At the same time, the scholarly holdings of teaching and research institutions in Algeria – judged to be quite insufficient when contrasted with the needs – are extensively underutilized due to such constraints as a lack of tools and lack of human expertise. We see this as an opportunity for growth rather than a reason for despair, building on the idea that social production through a networked environment can bring distributional consequences for the benefit of less affluent nations (Benkler, 2006). It is imperative for us to put technologies into play that are within the means of Algerian institutions.

### **3.1 CONTRIBUTION OF WEB SERVICES TO PORTAL ARCHITECTURE**

Web services are technologies based on standards that allow heterogeneous applications to interact remotely between themselves through the universal protocols (HTTP, XML, and so on). The architecture of web services is based on three elements: the

supplier of services (who writes a web service), the web service itself (written to an open standard) and the client (who takes advantage of the standard as a *lingua franca* to enable communication). The development of a portal based on web services requires the preparation and harvesting of content while recycling (or readapting) existing applications (Alonso et al, 2004).

Interoperability between the portal and the various applications is the key element. These services provide a frame for finding, describing and executing the applications. Widgets and gadgets are importable modular tools that enhance use without adding cost or local human resources. Today, web services and portals have come together to integrate such systems as electronic trade, e-services and e-learning platforms (Jimenez and Espadas, 2006). Mashups are examples of inventive utilities built around the concept of combining web services.

Indeed, web services are of interest to libraries from several points of view:

- integration of partnered applications: web services represent simple and reliable means of interoperability (Pierce et al., 2002), since they are written to an http standard that is rapidly attaining a de facto standard. Web services can interpose common query protocols between the distributed databases in ways that are transparent to the user and that do not require any changes or additions to the applications themselves.

- document access: besides the possibility of integrating metadata as with the OAI (Open Archives Initiative) protocol and with SRU/SRW (Search/Retrieve via URL ; Search/Retrieve via the Web), we can also integrate the applications themselves. Thus we can share data thanks to exchanges taking place in real time, but we can also take direct advantage of the various parts that come with the developed applications: search engine, inquiry interface... and why not the contents themselves and the ontologies to describe them !

- public services: web services open the door to integrating the scholarly work station with a certain number of online functions, especially Web 2.0 social functions, whose importance is increasing. They permit us to combine simple tools that our users commonly access on the web: retrieval of news feeds with RSS, mingling of catalog data directly with those of Google or Amazon, that is, pointing in the increasingly popular direction of content integration regardless of format.

### **3.1..1 WEB SERVICE FEATURES**

The main features of a web service are:

1. Web based: web services are based on protocols and languages of the Web, in particular HTTP (Hypertext Transfer Protocol) and XML (Extensible Markup Language).
2. Self-described, self-contained: the web services framework contains all the information necessary to utilize the applications, essentially by combining three functions: finding, describing and executing.
3. Modular: web services operate in a modular and discrete manner. This means that in place of integrating all functionalities into one global application, we create (or port) several specific applications into an interoperating system in which each performs one of the functionalities (Chaoying et al., 2006).

### **3.1.2. FORMATS AND PROTOCOLS OF THE WEB SERVICES' ARCHITECTURE**

Formats and protocol standards do exist for web services, specifically as the symbiotic trio: UDDI - WSDL - SOAP. These three technologies constitute the architectural spectrum of web services. They are XML-based technologies, in theory allowing use by web service applications in a stand-alone manner (without human intervention) from beginning to end. This architecture functions like mailing a letter via the post office – respectively, creating an address, writing the message and placing it in an envelope (Chauvet, 2002):

- i. UDDI (Universal Description, Discovery and Integration) permits creation of a directory of web services to be consulted (analogous to creating an address).
- ii. WSDL (Web Services Description Language) is a language based on XML used to describe the services (analogous to a message). It is indispensable to the UDDI as a means of allowing clients – via tools such as JBuilder, Delphi or Office – to find methods that can help them evoke the web services (informing them which information sources to search, which language, data types, processes and commands to use).
- iii. SOAP (Simple Object Access Protocol) is a form of web service, an XML-based protocol (analogous to an envelope) that defines the mechanism of information exchange between the clients and the web service suppliers. So it is in SOAP that we place the message we wish to have arrive at the target application. The application in turn uses a SOAP protocol to return the answer.

We can put what we like in this envelope as long as it is written in XML and is in conformity with what is described in the associated WSDL file. SOAP is object-oriented and manages compatibility status, so that the messages are capable of being transported in HTTP, SMTP, FTP, etc.

The structure of the SOAP “postal service” contains :

- an envelope that “deinitializes” the contents of the message for transport,
- a header with information such as authorizations or transactions (optional),
- a body containing information about the call and its response,
- an error management system to identify error conditions,
- attachments (optional).

#### **4. DESCRIPTION OF OUR METHODOLOGY**

One purpose in summarizing web services has been to introduce the methodological propositions involved in creating a scholarly web portal while expending a minimum of labor and financial resources.

If we take the example of a search engine, it is tightly integrated, in most cases, with the application that manages its data. But the search function can be externalized or outsourced (Google, for example, indexes data according to a proprietary, adapted periodicity that displays the results in its own environment). This represents usage of a resource external to the data, but the two systems are not integrated.

To answer these constraints, the three standards UDDI, WSDL and SOAP make possible the description, publication and invocation of web services (Kreger, 2001).

The foundation of web services interaction for a work station is based on a separation between the server, which manages the user interface, and the customer who uses a particular service. A SOAP server can be internal, in which case many services can be provided as SOAP intermediates via the intranet or local area network. If the SOAP server is external, the interaction is made via an extranet or the Internet.

#### **5. PHASES OF THE PROPOSED PORTAL DEVELOPMENT PROCESS**

With a web service, we integrate within our own environment a program that calls for external data sources according to criteria we designate. Before applying our methodology, we identified the following objectives:

- simplifying the users’ research,
- enhancing the visibility of the collections,
- promoting and developing the activities and services of the Library.

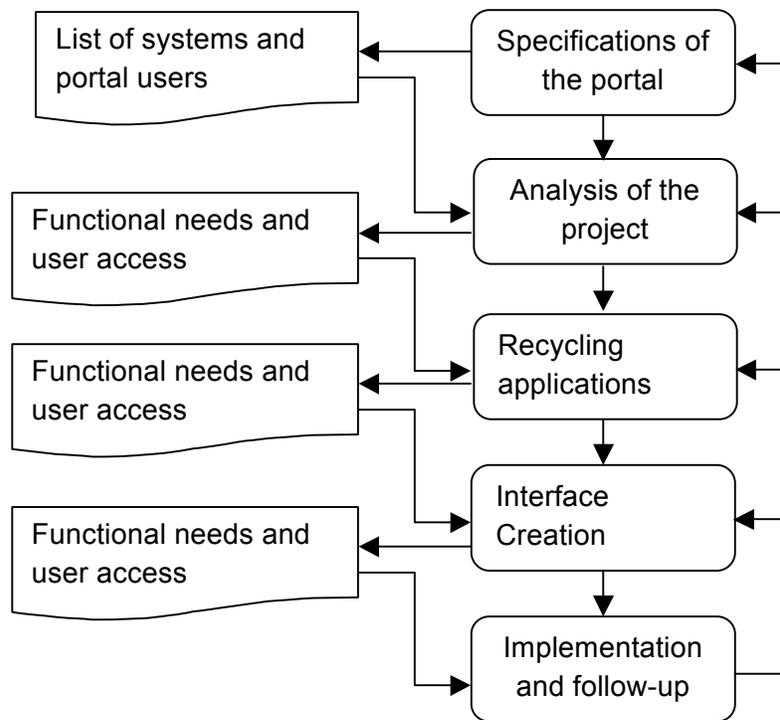
The portal possesses the following web-based advantages:

- data need only be present on a remote server;
- the remote server can possess computational power or storage capacities beyond those of the local user;

- the remote application can be used simultaneously by a large number of users, though the desired result set is only sent to each relevant location.

Thus, the portal development process we propose is generic and based on a global methodology. The main objective of the process is to combine and maximize existing applications into a set of modular web services. This architecture has the added benefit of being scalable so that future services not yet anticipated can be easily integrated into the environment as needed.

This process is iterative, and it follows the RUP (Rational Unified Process), that is, the phases are performed in a repeating manner until all requirements have been met. It is divided into five phases: (1) specifications of the portal, (2) project analysis, (3) recycling existing applications, (4) user interface creation, (5) implementation and follow-up (Figure 1). Phase four (interface creation) is analyzed on two levels, while the final phase is already divided into two stages, the first of which is itself divided into two further levels as described below.



*Figure 1: Different Stages in the Process of Developing the Web Portal*

## **5.1. PHASE 1: PORTAL SPECIFICATIONS**

The objective of this phase is to solidify the direction of the project, in other words, to give the big picture regarding the type of portal. Here we determine the list of internal and external systems and decide which users will be given access to which systems. Generally, the internal system (intranet) links to local area networks of the information organization (library). Each network or subsystem can contain one or more applications that require access to various data sources. External systems are generally provided by partner libraries or other subscription-based or open-access information providers. These systems are accessible through an extranet or the Internet, often involving public websites that can offer useful services and functionalities for the library (such as an interlibrary loan module at another institution). At the end of the specification phase, the designer can determine and evaluate the complexity of the project.

## **5.2. PHASE 2: PROJECT ANALYSIS**

This phase has two main objectives: the definition of functional needs and of user management.

### **5.2.1. DEFINITION OF FUNCTIONAL NEEDS**

During this stage, we define the global content of the portal. After drafting a preliminary list of services ratified at a meeting of library directors, we set up interviews and other types of dialogue with users. Various types of information are gathered during this stage, for example, a list of useful internal applications, technologies with which these applications have been developed, user categories and rules of management to be used.

From previous user polls we have found strong use of the library and a desire for the proposed resources. We have distinguished between different audiences (undergraduates, graduate students, researchers, etc.) with regard to varying services. Even the notions of ideal search interfaces differ, so we plan to target differing user types with custom interfaces as well. (See Texas History Portal in section 2.4). Most users consult the library catalog at least once a week in form of our card (yes, card!) catalog system. A good number of users avoid the online catalog

because it “lacks relevant tools and services.” Not all metadata for holdings have been digitally converted yet.

Asked about desired functionalities for the proposed local information portal, recurrently listed needs were:

- Improvement of metadata scope, interface design and capabilities for user interaction (collaboration, tagging, and so forth).

- Integration of bibliographic and electronic media systems, while drawing in remote resources and auxiliary services.

- Interlibrary loan capabilities.

We will continue to undertake user polls and studies (and other observations of actual daily usage) that can allow us to enrich the list of services.

### **5.2.2. USER MANAGEMENT**

This is a very important process linked to the operational needs of the portal: checking for permissible access, use and authentication. All these functions are essential for customization and the creation of user profiles. In addition, this phase allows us to specify which service corresponds to which user type.

### **5.3. PHASE 3: RECYCLING EXISTING APPLICATIONS**

The objective of this phase is to create a conceptual model for each service so that existing applications can be modularized into web services. This model is represented by a set of objects, each characterized by a name with attributes and operations. In order to determine the relationship between the internal services to be integrated into the portal and existing applications, it is necessary to measure present needs against the suite of currently used programs. If the anticipated needs can be met in existing applications, they will be ported across. For this, we match usage data to an application’s description using a conceptual model with UML (Unified Modeling Language) diagrams.

### **5.3 .1. OBTAINING USAGE STUDIES**

Usage studies are useful as a means of expressing the interaction between the system and its various actors (human users or other computer systems) and to determine different classifications within the conceptual model. Generally, these studies are made in a progressive manner, that is, the designer augments the study of varied aspects of usage one by one in order to recreate all the conditions and actions that the system will need to execute.

### **5.3 .2. DESCRIPTION OF THE USAGE STUDIES**

Before crafting the conceptual model, a scenario-based description of usage studies is absolutely necessary. This can be set forth textually or as interactive diagrams (with sequential or collaborative diagrams).

### **5.3 .3. DEFINITION OF THE CONCEPTUAL MODEL**

As the conceptual model is defined, the designer begins by using documents obtained during the step-by-step implementation of existing systems and in the evolution of present demands. The comparison is easier if the earlier conceptual planning was carried out with an object-oriented method that uses the UML notation. A second possible comparison can be performed between the conceptual model and the operations. Once the comparison is made, we can establish four possibilities:

1. Among the applications we have chosen to integrate in the portal, there are those we can port directly, that is, the functionalities of the applications that already exist are exposed, like a web service. This is feasible if the existing application has been implemented in a language that is compatible with the technology of web services as C#, vb.net, Java, etc.
2. Some applications are not directly portable. These applications require a transformation phase (migration) or “packing.” The transformation can be performed by hand and/or automatically from the source code in another language (Java for example). However, the packing operation consists in creating a layer of transition (a “wrapper”) that permits interaction with the source code. A good example is the creation of a Java envelope that uses Java Native Interface (JNI).

3. If a needed method or application doesn't exist, it will be created.
4. If a planned web service will be a composite of other web services, a “workflow engine” will have to be developed. If the workflow system already exists, we will need an overview survey for integrating a service into this system.

#### **5.4. PHASE 4: INTERFACE CREATION**

In the previous phase, we described the process of web services development from the supplier's point of view. In this phase we will explain how to configure the application for the user while exploiting available web services, both internal and external. Our goal is to offer information efficiently to users at precisely the moment they need it via an attractive interface that masks or erases any constraints caused by the multiplicity of data sources. To attain this objective, not only a mutually shared mechanism is necessary, i.e., web services, but also exchangeable data of interest to the users. Whenever possible, this interface will also provide simultaneous searching of multiple data sources. Such interoperability represents the dream of using one common language for the world of scholarly web portal users. In practice, this phase separates into two levels:

The first level involves the enrichment of the previous phases. Thus, we create the dynamic portion of the system – independent of platform – which will be implemented. To accomplish this task, we use diagrams of sequence, activity and collaboration. For external services, in general, some presentational elements of the interface are provided by the beneficiaries of services (in the form of user-selected portlets). Those presentation elements will be integrated directly into the portal if they are compatible with the portal frames. In other cases, we must develop customized presentations for external services.

The second level of this phase resolves the relationship between the findings of the first level and the software components used to implement the different interface elements. In order to facilitate the task and express the relationship with the implementation of the interface elements, a model of web applications programming should be used. The majority of designers propose using the MVC (Model-View-Controller) paradigm for the conception of Web applications, because it offers an important advantage: separation between the application model and the user interface.

### 5.5.1. PHASE 5.1: IMPLEMENTATION

The implementation phase is divided into two parts: development of the internal services (recycling of existing applications) and client implementation.

### 5.5.2. PHASE 5.2: FOLLOW-UP AND MAINTENANCE

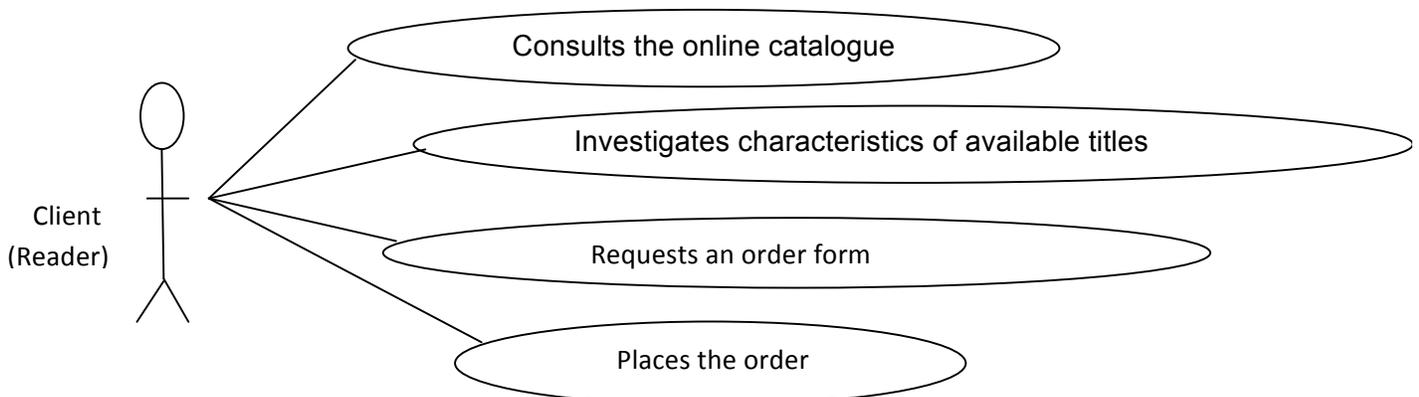
This phase consists of the management, administration and maintenance of the portal. This function is not bound to a given schedule, but it must be continuous to answer the evolving information demands of the organization, in this case the ongoing study and research of an academic library.

## 6. APPLICATION AT THE UNIVERSITY LIBRARY OF TÉBESSA

In this section, we present an example relating to an application for managing reader requests within the academic library of Tébessa (Algeria). Our objective is to port this application as a web service to be integrated into the scholarly portal.

### 6.1. USER SCENARIO

In Figure II, we present a case study diagram for a user transaction involving online discovery and request of a book, either from the library's closed stacks or for order through interlibrary loan. This diagram is very simplified because the actions taken by the reader have already been summarized in the transaction.



*Figure II: Web Service Scenario*

We will describe the user transaction of "placing a request" as follows:

**PRE-CONDITIONS**

- The reader has already consulted the online catalogue of available titles.
- The reader has already gathered together the features of the available titles.

**SCENARIO**

- The reader asks for an order form online;
- The system sends the order form to the reader/client;
- The reader fills out the order form;
- The reader makes the order;
- The system validates the information furnished;
- The system records the order;
- The system sends a confirmation to the client;
- The reader receives the confirmation.

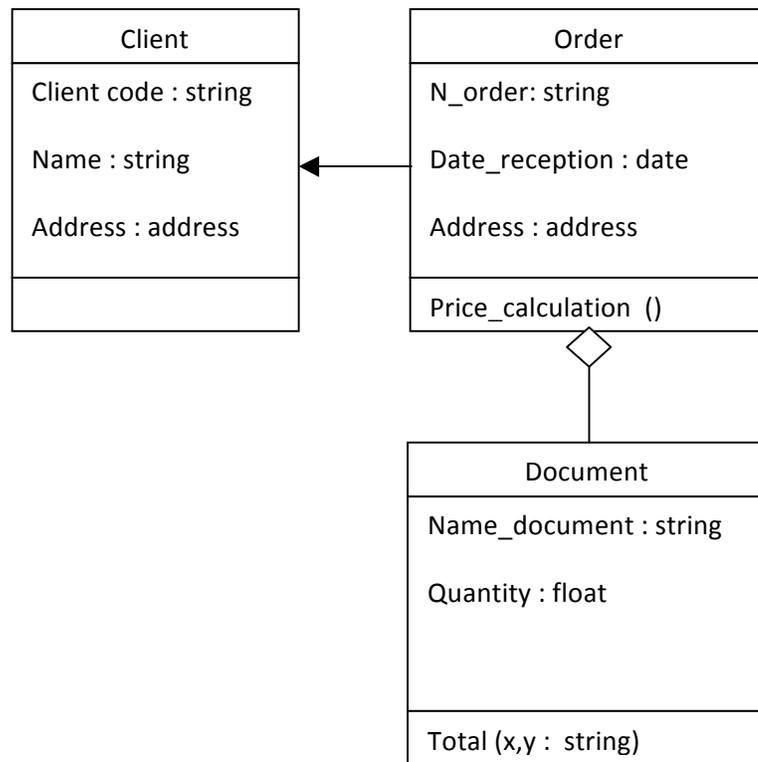
**POST-CONDITIONS**

- It is obligatory that the customer receive a confirmation.

**6 .2. THE CONCEPTUAL MODEL**

Through a syntactic analysis, the designer can derive classifications from a syntactic analysis of scenarios and from interaction of the diagrams. Other elements also play a role in guiding the conceptual model, such as rules of management that are very useful in describing the classes and in defining the cardinalities (minimum and maximum numbers of interactive associations).

After comparison between the conceptual model and the operations, we can decide on the names and parameters of the methods that will normally be implemented.



*Figure III: The Conceptual Model*

### **6.3. CHOICE OF API JAVA FOR THE DESIGN OF THE WEB PORTAL**

Java has numerous API (Application Programming Interfaces) useful for web services:

- API for the XML processing: JAXP
- API for RPC-XML: JAX-RPC
- API for the XML data exchange: JAXM
- API for the XML directories: JAXR

But above all, there do exist ready-to-use software “libraries” that are going to simplify the conception and consumption of web services greatly. These libraries present the advantage of being open source, and they benefit from great fame in the Java community. They seem therefore a judicious and economical choice for our web portal.

### **6.4. IMPLEMENTATION OF THE WEB SERVICE**

Initially, we let one function match up with one application developed in C++. In this application we port a procedure where we have conducted a migration automatically in the direction of Visual J#.NET, then we expose this function as a web service.

To develop the interface presentation (portlet), we can use an open source application such as Apache Jetspeed. Thus, for instance, three main stages exist for the client's role:

- Generation of a class proxy of the WSDL description of the service.
- Development of a portlet (presentation of the service) that uses the class proxy to evoke the methods exposed by the service.
- Integration of the portlet into a portal setting that conforms with administrative norms.

## **7. CONCLUSION**

Portals and web services are two new information technologies that have found frequent mention in recent research. In our work we have discussed history and evolution of portals as lesson plans for our own portal development. We have also discussed certain design aspects that have been studied in previous research and have presented a global methodology for the usage of web services in developing scholarly web portals.

The most important feature of our approach, we feel, it is to describe how to recycle (or port) existing applications in constructing new services. In our context, we consider a web portal to be an "integration project," that is, the portal is a combinatory internal and external platform of the web services.

For the call to external web services, we can use the WSRP standard (Web Services for Remote Portlets). Under that scenario, various developed interface elements (portlets) are published for direct use by IT professionals for their library clientele. In the same way, the institution can directly integrate into its portal the interface presentations published by other remote portals.

Moving this design discussion into the realm of practice, we are applying our methodology to a scholarly web portal project for the University Library of Tébessa. Current emphasis is on digitizing local resources (articles, theses, and so forth) for inclusion. Future plans include the expansion of user utility nationally by integrating the Tébessa portal into the regional Algerian network: *Réseau régional inter bibliothèques universitaires* (RIBU).

## 8. REFERENCES

- Alonso, G., Casati, F., Kuno, H. and Machiraju, V. (2004). *Web Services: Concepts, Architectures and Applications*. Springer, Berlin.
- Arsanjani, A., Hailpern, B., Martin, J. and Tarr, P. (2003). "Web services: Promises and compromises." *ACM Queue*, vol. 1, no. 1, pp. 48-58.
- Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, New Haven.
- Calabretto, S., Pinon, J. M. and Bozzi, A. (1998). "BAMBI: système de manuscrits anciens pour historiens," *Les bibliothèques numériques*, vol. 2, no. 3-4, pp. 31-50.
- Chaoying M., Bacon, L., Petridis, M. and Windall, G. (2006). "Towards the design of a portal framework for web services integration," *Proceedings of Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services*, February 19-25, Guadeloupe, p. 163.
- Chauvet, J. (2002). *Services web avec SOAP, WDSL, UDDI, ebXML*. Eyrolles, Paris.
- Chevalier, M., Julien, C., Soulé-Dupuy, C. and Vallès-Parlangeau, N. (2007). *Personalized Information Access Through Flexible and Interoperable Profiles*. Berlin, Springer.
- Chicoine, P. (2008). "La bibliothèque 2.0 émerge à Québec," *Argus*, vol. 37, no. 2, pp. 11-12.
- Hailong S., Chunming, H., Jinpeng H. and Qin, L. (2007) "Design and implementation of an enhanced grid service container in the CROWN grid," *International Journal of Web and Grid Services*, vol. 3, no. 4, pp. 403-423.
- Hunt, L., Lundberg, M. and Zuckerman, B. (2005). "InscriptiFact: A virtual archive of ancient inscriptions from the Near East," *International Journal on Digital Libraries*, vol. 5, no. 3, pp. 153-166.
- Ignat, C., Steinberger, R., Pouliquen, B. and Erjavec, T. (2006). "A tool set for the quick and efficient exploration of large documents collections," *Proceedings of the Symposium on Safeguards and Nuclear Material Management. 27th Annual Meeting of the European Safeguards Research and Development Association (ESARDA-2005)*, London, 10-12 May 2005, available at [http://langtech.jrc.it/Documents/0505\\_ESARDA-JRC-LT\\_Explore-doc-collections.pdf](http://langtech.jrc.it/Documents/0505_ESARDA-JRC-LT_Explore-doc-collections.pdf) (accessed 30 March 2009).

Jimenez, G. and Espadas, J. (2006). "Implementation of an e-services hub for small and medium enterprises," *Proceedings of Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services*, February 19-25, Guadeloupe, p. 187.

Kalyanaraman, S. and Sundar, S. S. (2008). "Portrait of the portal as a metaphor: Explicating web portals for communication research," *Journalism and Mass Communication Quarterly*, vol. 85, no. 2, pp. 239-256.

Kolbitsch, J. and Maurer, H. (2006). "The transformation of the web: How emerging communities shape the information we consume," *Journal of Universal Computer Science*, vol. 12, no. 2, pp. 187-213.

Kreger H. (2001). "Web Services Conceptual Architecture (WSCA 1.0)," IBM Software Group, available at <http://www.cs.uoi.gr/~zarras/mdw-ws/WebServicesConceptualArchitectu2.pdf> (accessed 23 February 2009).

Lagoze, C., Arms, W., Gan, S., Hillman, D., Ingram, C., Krafft, D., Marisa, R., Phipps, J., Saylor, J., Terrizzi, C., Hoehn, W., Millman, D., Allan, J., Guzman-Lara, S. and Kalt, T. (2002). "Core services in the architecture of the National Science Digital Library (NSDL)," *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, Portland, Oregon, pp. 201-209.

Le Bourgeois, F., Emptoz, H., Trinh, E. and Duong, J. (2001). "Networking digital document images," *Sixth International Conference on Document Analysis and Recognition*, Seattle, Washington, pp. 379-383.

Melançon B. (2004). "Sommes-nous les premiers lecteurs de l'Encyclopédie?" in Salaün, J.-M., Vandendorpe, C. and Bachand, D. (Ed.) *Les défis de la publication sur le web: hyperlectures, cybertextes et méta-éditions*, Proceedings of the 15<sup>th</sup> ENSSIB Conference, Lyons, Presses de l'ENSSIB, Lyons, pp. 145-165.

Mischo, W. H., Habing, T. G. and Cole, T. W. (2002). "Integration of simultaneous searching and reference linking across bibliographic resources on the Web," *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, Portland, Oregon, pp. 119-125.

Mothe, J., Chrisment, C., Dousset, B. and Alaux, J. (2003). "DocCube: multi-dimensional visualisation and exploration of large document sets," *Journal of the American Society for Information Science and Technology*, vol. 54, no. 7, pp. 650-659, available at <http://www.irit.fr/recherches/IRI/SIG/personnes/mothe/pub/JASIST03.pdf>, accessed 30 March, 2009.

Nitecki, D. A., and Renfro, P. E. (2004). "Borrow direct: A case study of patron-initiated interlibrary borrowing service," *Journal of Academic Librarianship*, vol. 30, no. 2, pp. 132-135.

Nordstrom, K. E., Hartman, C. N. and Phillips, M. (2004). "The University of North Texas libraries' portal to Texas history: archival challenges and solutions," *JCDL '04: Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries*, Tucson, Arizona, p. 409.

OECD (2007). "Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking," available at <http://213.253.134.43/oecd/pdfs/browseit/9307031E.PDF> (accessed 23 February 2009).

Pierce M., Fox, G., Choonhan, Y., Mock, S., Mueller, K. and Balsoy, O. (2002). "Interoperable Web Services for Computational Portals," Supercomputing, *Proceedings of the Conference on High Performance Networking and Computing*, November 16-22, p. 39.

Smith, M.A. (2004). "Portals: toward an application framework for interoperability," *Communications of the ACM*, vol. 47, no. 10, pp. 93-97.

Strahilevitz, L. (2007). "Wealth without markets?" *Yale Law Journal*, vol. 116, no. 7, pp. 1472-1516 [response to Benkler].

Tamine, L., Boughanem, M. and Zemirli, N. (2006). "Inferring the user interests using the search history," *LWA*, vol. 1, pp. 108-110.

Teasdale, G. (2008). "La recherche intelligente se redécouvre à l'Université Laval: Un nouveau catalogue voit le jour," *Argus*, vol. 37, no. 2, pp. 7-10.

Tung-Ching L., Hsu-Ju, C. and Po-Hsun, C. (2008). Issue placement affecting factors for knowledge management systems initiation, *International Journal of Innovation and Learning*, vol. 5, no. 2, pp. 186-200.