SERC (the Science Education Resource Center at Carleton College) builds educational portals on an infrastructure suffused with digital library technologies and sensibilities. Yet our users—largely geoscience faculty looking for resources and information that will help them improve their teaching—don’t think of us as a digital library (if indeed they have any notion of what that term means). Our portals identify focused topics of importance to educators and present information on these topics through a variety of structures—there is no presupposition that a collection of metadata and a search engine will be the best solution. Yet as we built the tools to support the creation of these portals it was clear that collection-building (at least on a small scale) and more generally, a systematic approach to collecting and organizing the information within our sites would be a recurring theme. These are traditional strengths of a digital library. So we built portal tools that integrate and interoperate with digital library elements at a number of levels.

This approach, building bridges between the educational digital library world and the broader set of technologies through which educational information is funneled, is not unique to our efforts. Many educational digital libraries are reaching across this divide in an effort to connect more directly with those they are trying to serve. The National Science Digital Library is incorporating Pathways projects that may build educational portals on the foundations of the extant central metadata repository and search services[1]. Merlot has announced a partnership with course management system provider Blackboard—tying their collection into the tools faculty use to connect with their students[2]. A recent Digital Library Federation report “Digital Library Content and Course Management Systems: Issues of Interoperation”[3] highlighted the importance of these sorts of efforts. The potential benefits of connecting digital libraries more directly to the educational environment are clear. The details of the mechanisms through which this might be best achieved are not. The aforementioned report concludes that “we are at the point where some convincing demonstration projects are badly needed”.

The portals we are building and the tools upon which they are built offer one such demonstration. As we built the technical infrastructure that supports our sites we identified areas where digital library technologies or approaches could be profitably integrated. Our tools reflect a judicious
choice of those digital library elements we felt offered the biggest payoff in making our portals easier to manage and more effective for our users. This article presents an overview of our portal technologies with a particular focus on our integration of digital library elements. It highlights one possible path connecting ‘traditional’ digital libraries to educational portals and begins to illuminate the benefits and challenges of such an endeavor.

About SERC Portals
SERC education portals follow a number of different models for pooling and disseminating community expertise on focused topics. For example the “On the Cutting Edge” project [4] is built around a series of face-to-face professional development workshops for geoscience educators. Each workshop pools the expertise and energy of the participants to catalyze the development of collections supporting themes and topics explored in the workshop. At the conclusion of each workshop its portal evolves from a workshop support website into a website that highlights these collections. For example a workshop on teaching petrology was the launching point for a collection of exemplary teaching activities, a database of analytic instruments available in the community, and an extensive bibliography of petrology teaching resources from across the internet.

The “Starting Point: Teaching Entry-Level Geoscience” project [5] takes a different approach: engaging editors on pedagogic topics relevant to educators teaching entry-level geoscience college courses (e.g. teaching with models, interactive lectures, field labs). Each editor pulls together community expertise to build guided introductions to the topic. These overviews serve as wrappers for collections of classroom ready activities that exemplify each topic.

The web sites that support these projects (as well as others at SERC) present their content in a variety of ways. A portal may include a mix of a simple linear set of pages providing a directed introduction to a topic, more loosely structured sets of pages each providing a jump off point to external resources, and search interfaces that enable exploration of larger collections. To support this diversity of information structures we’ve created an authoring environment with a great deal of flexibility. The tool makes it easy for our authors to choose the most effective method in communicating their message—including drawing on digital library tools when appropriate.

The SERC CMS
SERC web sites are built using a tool we call the SERC content management system, or simply the CMS. [6] Our CMS is an integrated set of tools which allow site authors to create and edit our web
sites from within a straightforward web browser-based interface. Authors visit our administrative web site, authenticate and then navigate to the editing interface for a given web page. There they can edit the page text within a standard web form text box. They format their text with a limited set of HTML tags supplemented with some custom tags, such as the tag to refer directly to a cataloged resource. The results are stored in a central relational database along with other editable properties of the page (navigation menu, page template, title, URL, etc…) The final site viewed by the public is generated dynamically from these elements.

This approach has benefits both in terms of the technical quality of the actual HTML, as well as in the operational details of coordinating work among science educators from across the country. It provides an environment where the expression of content expertise is explicitly separated from the technical issues of constructing a high-quality website. The CMS ensures the resulting pages follow current best practices as far as web standards and accessibility.[7] It also allows for the intertwined larger scale problems of design, usability and information architecture to be addressed uniformly across the site without requiring the each site author to master these disciplines on their own.

Because the authors work entirely through the web-based interface all the stages of site production can be coordinated centrally. The CMS provides versioning of site content, separate environments for development work and the live site, and central management of files and images used throughout the site. It eliminates the “who has the latest version of the file” problem that grows as multiple authors combine their efforts on a single site using traditional ‘static’ HTML pages. The CMS has proven to be an effective tool in enabling the collaborative authoring of our portals. Our CMS currently holds over 2000 pages of original content from a dozen projects. This content was pulled together by some 70 editors working at 40 institutions across the country. These editors, many with little or no previous web authoring experience, drew upon direct contributions from hundreds of their peers in the science education community in creating our web portals.

Beyond the benefits of coordination and consistency that are available in many web content management tools our CMS integrates tools and practices from the digital library community. At a most basic level the CMS allows for the expansion of page metadata beyond the obvious title, URL, author to include information like keywords and controlled vocabulary categorization. This additional metadata is exposed in <META> tags within the HTML and is used to drive our internal search mechanisms. Similarly, the CMS will make it straightforward to expose the sort of metadata
called for by frameworks like the Semantic Web [8] which hold promise for enabling more structured information use throughout the web.

The next level of digital library integration is evident in how we handle links to outside resources. Each link is backed with a complete set of metadata describing the external resource. This metadata provides concise descriptive information aiding our users in evaluating prospective resources. It allows us to provide more effective resource discovery, and it serves as an integration point for a number of services. Ideally this metadata could be drawn from existing digital library collections. Our sites would act much as a traditional library special exhibit; where choice items, exemplifying a focused topic, are drawn from the general collection and placed in prominent view in a display case by the front door. The existing digital libraries would act as the general collection, the ‘stacks’ from which we could cherry-pick. In practice there isn’t appropriate metadata available from outside collections for many of the online resources we’d like to reference. And so in the near term we’ve set up our own cataloging process, creating metadata for these resources and collecting them in small behind-the-scenes collections.

**Cataloging**

External references, identified by site authors or submitted by the community via forms throughout our sites, are moved into a cataloging workflow system whose end product is a complete metadata record for the resource. The first step in this system is to check existing digital library catalogs to see if appropriate metadata already exists. Many of our projects relate to Earth system education and so we coordinate metadata sharing with the Digital Library for Earth System Education (DLESE). We use the DLESE OAI software to harvest all 14 collections currently aggregated by DLESE. This nightly harvest uses the OAI-PMH [9] which we also use to share our locally cataloged collections back to DLESE for dissemination throughout the National Science Digital Library [10].

Currently this effort pays off about 25% of the time. The metadata has already been created and harvested from DLESE. Our local catalogers do a simple one-click import (making the record available in the CMS) and their work is done. The balance of the time we have to rely on our own efforts—building a digital library behind the scenes. The resource continues through our workflow system which provides online tools to coordinate the metadata creation. The workflow tools track cataloging progress, aid in distributing the work from different projects across our pool of catalogers, and record our quality assurance process. Most of our resources are cataloged to the
ADN metadata standard [11] using copies of the DLESE Cataloging System[12] which we run locally. We also have cataloging tools for creating qualified Dublin Core in line with the NSDL requirements [13] as well as for capturing basic bibliographic information for print resources. The collections we build in this way are strongly focused on the subject of the particular portal and are of exceptionally high quality. Since our goal is to build an effective portal (rather than a comprehensive collection) we focus on identifying only the best resources (which in many cases fall under close scrutiny by authors and reviewers) rather than canvassing freely for resources to bulk up a collection. As these collections flow back into DLESE and NSDL they represent a particularly strong set of resources within these larger repositories.

![Diagram of relationships between internal and external catalogs and the CMS.](image)

**Figure 1: Relationships between internal and external catalogs and the CMS.**

**In-Line Referencing and the User s’ view of Catalog Records**

Once web resources have been cataloged they can be incorporated into the portals in a number of ways. The simplest mechanism is to insert a link to a resource within the normal flow of the text of a web page. Rather than coding a traditional url, the author inserts an explicit link to the catalog record making use of its unique internal id: [resource 123]. When the page is displayed by the CMS a link to the resource appears along with a small ‘more info’ link which points to an enhanced
catalog record.

This ‘more info’ link provides a view of the metadata record reduced to those fields of most interest to users (title, description, url, author, terms from subject vocabularies). This concise synopsis of a resource is valuable for users in evaluating whether the resource might meet their needs. These pages also serve, unexpectedly, as a common entry point into our portals as their simple, information-rich format makes them high-ranking targets for web search engines such as Google. [14]

In addition to the core metadata we augmented these display pages with various sorts of related information. Since each reference to a resource within our CMS is tracked we are able to provide ‘this resource is also referenced here’ links so users can explore related areas within our sites by following these citations. We’ve also integrated 3rd party services that provide information related to a given resource. For instance the DLESE community review system [15] collects users’ reviews of the teaching materials in the DLESE collections. We harvest the review status for DLESE resources (provided in a prototype annotation metadata format via the OAI-PMH) and where appropriate provide direct links to the reviews from our resource display pages. Additionally we take advantage of the archiving service provided through NSDL, linking to archived versions of the resources we reference so that users can investigate a resource even when it is temporarily off-line [16]. The integration of these 3rd party services is completely automated: users benefit from the direct linking of this information within our site without effort by our authors or catalogers. As similar services, connecting additional information to websites through catalog records, grow up around the digital library community we’ll be able to integrate them into this display page.
Linking to outside sites through a reference to a catalog record in this way also enables a number of efficiencies as we manage these outside links. For instance link-checking can be centralized and automated. If the url of a referenced resource changes we need only change the single central catalog record and all references to this site are automatically updated throughout the site. These sorts of efficiencies are magnified as record sharing is more widespread. Updates made to metadata (e.g. a change in url) by any of the DLESE collection providers automatically propagate with no human intervention throughout our sites.

Annotated Bibliographies
Another way we make use of our cataloged resources is by building annotated bibliographies. Our authors frequently find the traditional annotated bibliography format-a citation list of core resources on a topic annotated with brief, critical descriptions of each resource-a strong approach to presenting key references. The CMS provides a simple tool for aggregating and annotating groups of resources based on their catalog records. The resulting lists can be easily inserted into existing web pages. Authors may annotate the list with descriptions specific to that list or let the system fall back on using the existing generic description field from the metadata. As with stand-alone links
these annotated lists link to the resource metadata and the value-added services that can be tied to the catalog record. [17] The ability to identify existing resources in peer digital libraries, and with a few clicks bring the resource into an online annotated bibliography, complete with all the benefits of its backing metadata, exemplifies the value of working within a website authoring environment that incorporates digital library tools.

Building Collections of Community Expertise

Although identifying existing websites and print resources is a common way to build a collection of materials around a topic, in many cases valuable community expertise is not available in an easily reachable form. The great teaching activity may exist as only a set of hand-outs in a filing cabinet combined with the time-tested wisdom of the faculty member who uses it. There is often little motivation for individual community members to commit time to building web pages (or writing articles) to share their expertise on topics that fall outside the normal channels of scholarly exchange.

In an effort to tap this sort of knowledge we’ve developed a process and set of tools for gathering this sort of expertise. Educators are solicited to fill out a relatively simple web form carefully structured to tease out the information needed to make that bit of knowledge more easily reused. For instance we collect educational activities related to petrology, asking for not only a description of the activity but also the course context in which it’s taught, learning goals, assessment, general teaching tips and allowing for uploads of relevant files (e.g. the Word file they print and hand-out to students, the Excel spreadsheet that contains the base data) [18]. The form is built around the elements we’ve identified as necessary for other educators to adapt and use activities successfully in their own environment.

The CMS provides a simple online tool for building the collection form (often just a matter of building a variation on an existing form) and embedding the form within an existing page. Submissions automatically collect in an online queue where they can be vetted (e.g. for duplicate or incomplete submissions and appropriate use of controlled vocabularies) and then with a single click become a new web page within the portal. These pages aggregate into ‘mini-collections’ which, by virtue of the topical categorization the submitter has provided, can be automatically exposed in search/browse interfaces[19].
We can follow this same model to aggregate community knowledge beyond teaching activities. For example we currently use these tools to build searchable mini-collections of geoscience course syllabi (with associated information about the design of the course), and a registry of geochemical analytic instruments. While the tools themselves don’t eliminate all the challenges in bringing together this sort of community expertise, they do reduce several critical hurdles. The tools minimize the effort needed on the part of the contributor to add their expertise to a community pool; they need only fill out a web form and upload any existing supporting files they have. The tools also drastically reduce the editorial effort needed to transform the contributions into a coherent collection; a single click turns a contribution into new page in a searchable collection. By minimizing these logistical barriers we hope to empower communities of educators to collect expertise and insights previously unshared.

**Faceted Search**

When a site author sees the need to create a collection of resources (either cataloged references to external websites or content gathered through the site into a mini-collection) the CMS provides an easy mechanism to set up search interfaces specific to the collection. Our current interfaces are based on ideas explored in the Flamenco projects’ work on faceted search [20] and which are starting to show up in other digital libraries[21]. The central theme is to simultaneously provide both a full text search interface and a browse interface over multiple controlled vocabularies. Users can freely combine text search with browsing using either technique to further refine a search.

The search screen provides a search text box at the top with search results below. On the right side of the screen the controlled vocabulary terms are displayed. Choosing a particular term limits the result set to the part of the collection that matches that term. The vocabulary terms are displayed along with the number of resources matching it. This number updates dynamically, reflecting the user’s current search text and browse choices. This provides immediate feedback about the scope of each search as well as providing direct access to deeper browsing through hierarchical vocabularies. The current search parameters are listed at the top of the page providing a simple mechanism for expanding as well as narrowing a search. The controlled vocabularies serve double-duty--a browse interface and a feedback tool for search scope.
The underlying technologies are those found in many digital libraries. Our full-text search uses the lucene library [22] implements Porter stemming[23], and has a search syntax designed to match our users most familiar touchstone: Google. The browse structures displayed are a natural expression of the controlled vocabularies used in our metadata. However, explicitly exposing the browse structure with it’s integrated search feedback alongside the full-text search provides users with a more direct sense of the information space they are moving through empowering them to explore more freely. Research results indicate that this sort of faceted search interface can lead to more successful search experiences than some of the more traditional options [24].

**Controlled Vocabularies**

The success of this sort of interface is highly dependant on the efficacy of the underlying controlled vocabularies. While we are making use of a number of existing vocabularies we’ve found it important to develop specialized controlled vocabularies to fit individual collections, or sets of collection. These specialized vocabularies have the right granularity and particular choice of facets that do the best possible job of exposing a particular collection. To support broad and flexible application of controlled vocabularies our CMS allows point and click creation of new vocabularies.
An integrated indexing system builds a continuously updated full text index of all web pages and catalog records, exploiting their structured nature to extract vocabulary information and using variable weighting of different page elements to improve search results. The CMS also supports various manipulations of web page specific metadata to enhance local search effectiveness including page keywords and arbitrary boosting of pages which site authors indicate are of especially high-quality on a given topic. The CMS provides a set of straight-forward yet powerful tools with which our authors build and fine-tune vocabulary driven interfaces for exploring their collections.

Conclusions
SERC educational portals leverage the techniques and tools of the digital library community to bring cohesion and clarity to their presentation of information. We make wide use of the rich descriptive metadata and services tied to our collections taking advantage, where possible, of overlapping efforts elsewhere in the community to share metadata and services. We build collections reflecting community wisdom in ways that are systematic, scalable and driven by the real needs of their final users. We exploit this order and intent through interfaces which implement current best practices enhancing users’ exploration of our sites and collections. We hope that our CMS serves as a demonstration of what is possible when the tools for building informational web sites are built on foundations that express the experience of the digital library world; organizing information in ways that help users better explore, evaluate and share the knowledge of their discipline.

For those looking toward similar sorts of integration between digital libraries and educational delivery tools, like Course Mangement Systems there are several key observations. First, it can be done. Our users are happily building bibliographic lists within a web authoring tool drawing the metadata from digital libraries. Second, while were able to exploit existing library content from closely aligned digital libraries there is still (at least in the areas that our users are working) a significant body of material not available in existing repositories. Having the ability both to create new metadata to describe resources from the larger web, as well as the tools to bring original materials online are crucial to our sites achieving its goals. The road is not one way from digital library into educational tool, but a loop perhaps tighter and with a different asymmetry than one
might expect. Finally, the tools and standards enabling interoperability are still in their infancy. Ideally a portal builder might be able to rely on search and display tools from existing digital libraries whose interfaces integrate with their portal in real time. But our digital library integration is largely depended on asynchronous data shuffling via OAI-PMH with search and display handled by locally built and installed tools. This was, for us, the most expedient way to reach our goals of building effective educational portals. We’ll know progress has been made when we can assemble portals of similar quality and functionality with a much simpler layer of digital library ‘machinery’ embedded in the portal.

Acknowledgements
This work is funded by the National Science Foundation under the grants: DUE 0127141, DUE 0226243, DUE 0226199, and EAR 0304762.

[6] CMS is a terribly overburdened acronym within the educational technology community even getting beyond the ambiguity over whether the ‘C’ stands for ‘Content’ or ‘Course’. We beg the reader’s indulgence in our use of the acronym as it has become standard parlance within our office and retrofitting a new name at this point seemed artificial.
[7] Web standards refers largely to adherence to w3c standards, http://www.w3.org/, as promoted by http://www.webstandards.org/. Adherence to these standards is intimately tied to creating sites that are as broadly accessible as possible. See for example http://accessnsdl.org/ for accessibility information related specifically to digital libraries
[13] NSDL metadata requirements are described in their metadata primer
http://metamanagement.comm.nsdlib.org/contributing.html and specifically in the xml schema
they provide http://ns.nsdl.org/schemas/nsdl_dc/nsdl_dc_v1.02.xsd
[14] ]In May 2004 10% of visitors arrived at our sites through a catalog record page.
[17] For one example see http://serc.carleton.edu/NAGTWorkshops/visualize04/recommend.html
[18] An example form through which community members can submit teaching activities,
http://serc.carleton.edu/1621
[19] An example search/browse interface across a teaching activity collection,
http://serc.carleton.edu/3717
[21] The Gateway to Education Materials (http://gateway.org) makes use of a similar faceted search
interface.
Swearingen, and Ping Yee, Communications of the ACM, 45 (9), September 2002, pp.42-49.