

# Ebsco Discovery Services

## Abstract

*Ebsco Discovery Services (EDS) represents the first of the trio of web scale discovery services debuting in 2010. Built off the established EBSCOhost platform, EDS extends the platform into the web scale discovery space through a preharvested, centralized index encompassing content sourced from Ebsco databases and beyond. This chapter provides a brief history, overview, and a few insights into the future development path of EDS, describes the local and remote content associated with the EDS index, and highlights some of the features, functionality, and flexibility associated with the EDS interface.*

## Overview

Ebsco began development of Ebsco Discovery Service (EDS) in 2008. Public announcement occurred in spring 2009, and after a beta period concluding later that year, public release occurred in early 2010. At the time of this writing, late summer 2010, approximately twenty-five customers have gone live with an EDS implementation. Ebsco indicates that new customers can generally be set up and ready to go live within eight to ten weeks. EDS is based in large part on the infrastructure and interface associated with the popular EBSCOhost platform, which debuted around 1994. Early EDS development partners were generally academic customers, though other library types, including at least one public library, have more recently begun trials of the service. EDS is offered as a hosted platform; no local installation options are available. When initially released, EDS required user authentication prior to conducting a search. In mid-2010, Ebsco released a Guest Mode option providing

unauthenticated users with some limited search capabilities.

The annual subscription pricing model relies primarily on the institution's full-time equivalent (FTE) count and level of service desired. Level of service can include factors such as the number and types of local library resources harvested and indexed (such as local digital collections and institutional repositories). Multiyear and consortial discounts are available. Ebsco provides underlying application or interface updates to portions of EDS approximately every three months. Ebsco telephone customer support is available 24/7 Monday through Friday and for reduced hours on the weekend; in addition, customers can report issues through Ebsco's website or via e-mail.

## Content and Scope

### Publisher Content

At time of writing, the base index underlying the EDS service includes content from nearly 20,000 providers, in addition to metadata drawn from tens of thousands of book publishers. This base index presently includes metadata for more than 45,000 journals, more than 800,000 CDs/DVDs, nearly six million books, and more than one hundred million newspaper articles; this base index is searchable by all EDS customers. EDS includes items from several open-access repositories, such as materials from the DOAJ (Directory of Open Access Journals), OAISTER, and arXiv.org e-Prints. Ebsco creates a unique index for each EDS customer, which includes local harvested content (ILS catalog records, digital collections, etc.), the base index content, and additional content pulled from Ebsco-sourced databases for which the library has a

subscription. For example, an EDS customer whose library maintains a strong EBSCOhost-based collection may find the number of included journals in the central index could rise to over 80,000. Ebsco indicates the central index encompasses metadata from hundreds of thousands of sources from more than 20,000 providers. Ebsco has agreements with a large array of providers. Examples include Alexander Street Press, Association for Computing Machinery, Cambridge University Press, Emerald, IEEE, Ingenta, LexisNexis, NewsBank, Oxford University Press, Readex, Sage, Springer Science + Business Media, Taylor and Francis, Wiley, and Elsevier. In mid-2010, Ebsco announced a partnership with Thomson Reuters, providing access via EDS to Web of Science data.

Given that Ebsco has been an established aggregator of content for years (nearly 300 full-text and secondary research databases) and that it leverages this established content and existing detailed indexing within EDS, the company describes its metadata as deep and detailed. This rich metadata often includes elements such as author-supplied abstracts and keywords for many major publishers, such as those referenced above. Ebsco notes that controlled vocabulary subject headings can be searched for content found in many EBSCOhost databases and that in general, an EDS search includes a search of the indexing, abstracts, and full text of EBSCOhost full-text databases for which the library maintains a subscription (including resources such as the popular Academic Search database). Ebsco indicates an overall broad level of full-text searching within content, dependent in part on publisher agreements in place. With Ebsco's recent acquisition of NetLibrary, full-text searching of these e-books is provided from the EDS interface (presuming the library has this e-book subscription).

In a typical configuration, EDS works with an institution's proxy server to provision off-site access. For resolution to the full text, the institution's link resolver is used, whether Ebsco's LinkSource link resolver or another link resolver typically used by libraries. Given that Ebsco is also a content provider, resolution to the full text often does not require a link resolver step, depending on the amount of Ebsco-sourced content the library subscribes to.

### Local Resources

EDS can harvest local collections, such as ILS catalog records, digital collections, and institutional repositories based on various underlying schema, such as MARC, Dublin Core, XML, and EAD. EDS utilizes various harvesting and delivery mechanisms, such as OAI-PMH and FTP. For local collections, there is no mandatory minimal set of field information required; relevancy ranking (discussed below) is, to a degree, dynamic, depending on the resource type and the level

of metadata present. Metadata associated with locally harvested collections is mapped and transformed to an underlying EBSCOhost schema. Depending on library need, updates can be harvested on a daily basis if necessary. As noted above, the price of the EDS service is partially dependent on the count of harvested local materials.

### Relevancy

By default, an EDS search is a keyword search against full text and fielded metadata, with results returned ranked by relevancy. Ebsco indicates EDS can load records having different levels of metadata and tunes relevancy appropriately based on the data source. For example, relevancy for records with more replete metadata would take into account items like subject, author-supplied keywords, and title, while a record with less fielded metadata may have full text weighted more heavily; relevancy is then normalized across these different data types. Factors such as term or document frequency calculations, which field a word appears in, and the uniqueness of the word in the overall index play a role in relevancy determination, as do other factors, such as currency, number of times cited, type of document, and so on. Subject headings from controlled vocabularies, titles, author keywords, abstract keywords, and full-text keywords also play a role. A detailed overview of EDS's relevancy ranking is freely accessible on the open Web at the Ebsco support website. The library does not have influence over the relevancy algorithms nor the ability to promote certain items in a given results list. Records are deduped in the index based on analysis of the various citation fields and matching algorithms developed by Ebsco.

*How does the EBSCOhost search engine determine relevancy ranking?*

[http://support.ebscohost.com/knowledge\\_base/detail.php?id=3971](http://support.ebscohost.com/knowledge_base/detail.php?id=3971)

In instances where content related to a single unique item is sourced from multiple providers, Ebsco creates a composite record, combining what's considered the strongest metadata and indexing from each source in the creation of this composite or super-record. Such records can include the subject headings from each content provider. For example, a particular article could be crossdisciplinary, and as such, may have different subject headings from different providers (each more closely aligned with users in that particular discipline), which are combined in this merged record. In 2011, a level of FRBRization for book records is expected, in the sense that available editions or manifestations of a unique item will be presented

in a single record view under the heading Available Formats/Editions.

## Interface Features: Overview, Results, and Navigation

### General

EDS offers a template that libraries can customize to their local environment. Within the EBSCOhost administrative interface, the library can customize various branding elements, such as colors and logos, and specify some layout details, such as the positions of logos. Libraries can choose to have a custom toolbar, which appears at the top of the interface, and can include and name library-specified hyperlinks appearing in the toolbar. Libraries can choose which elements appear in the toolbar, such as Sign In, Folder, Language, and can customize the names of the text labels. In addition, libraries can provide custom text to appear at the bottom of the interface, and choose which screens will include such custom bottom branding. Ebsco provides an EBSCOhost integrated toolkit, an API utilizing SOAP/REST Web services. At present, customers do not have direct access to the style sheets that control the layout, the effect of mouseovers, and so on. At the time of this writing, the interface is available in dozens of European languages, and by the end of 2010, it was expected to be available in dozens of Asian and other non-European languages.

By default, EDS offers a single search box (figure 22), with a Search Options hyperlink below the box providing more choices (figure 23). These choices include a Search Modes category with radio buttons including Boolean/Phrase, Find All My Search Terms, Find Any of My Search Terms, and so on. A SmartText Searching option allows users to enter long strings of text, which are summarized by EDS into search terms then applied against the search; Ebsco indicates that this search feature works with some database content incorporated into the EDS experience. In addition, users can indicate if they wish to display only results with linked full text, or scholarly or peer-reviewed journals. Users can search by title and author and indicate a publication date range. Several of these options are also available from within the search refinement pane, described shortly, once search results are returned. Users also have the option to enter advanced search or visual search modes. The advanced search mode provides the ability to conduct fielded searches and the use of Boolean operators via pull-down menus. The visual search mode offers another method for users to refine searches; as a user clicks on subheadings, a visual display maps the relationship between the original results set through to the final refined items. Visual search returns up to the top 250 query results.

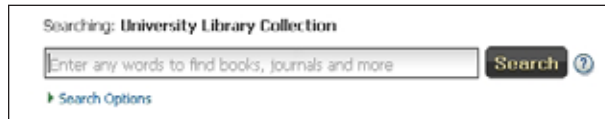


Figure 22  
EDS single search box

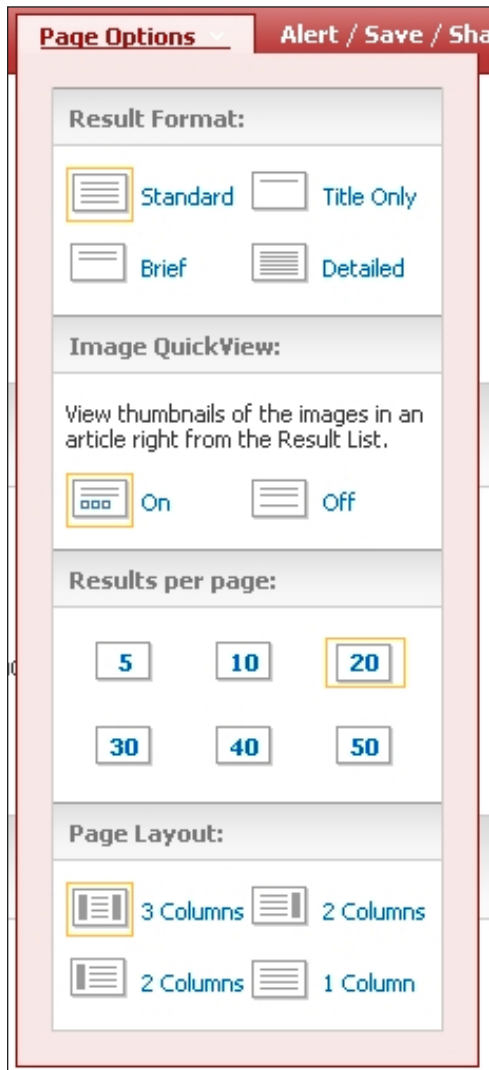


Figure 23  
EDS advanced search

### Returned Results—Brief View

Once a search is conducted, the full interface is displayed. In a typical installation, the full interface is divided into a large central section bounded by one or two vertical panes; functionality offered within these panes is described shortly. For all content, by default, items are returned ranked by relevancy; other sort options are chosen by a pull-down menu and include date descending, author, and title. The user can define various layout options and other parameters via a Page Options link in the header area (figure 24), and the local library can define which layout option is the default.

The majority of the EDS interface is dedicated to presenting results from a search. Each content type, such as journal articles, newspapers, and so on, has a unique icon. Book cover images, sourced from Baker and Taylor, are provided for all EDS customers. If a library subscribes to additional enrichment information (from Baker and Taylor or another provider), such content can be included in the EDS experience. For books (figure 25), typical information provided includes title, author, publication information, page count, size, and



**Figure 24**  
EDS page layout options



**Figure 25**  
EDS brief results: book

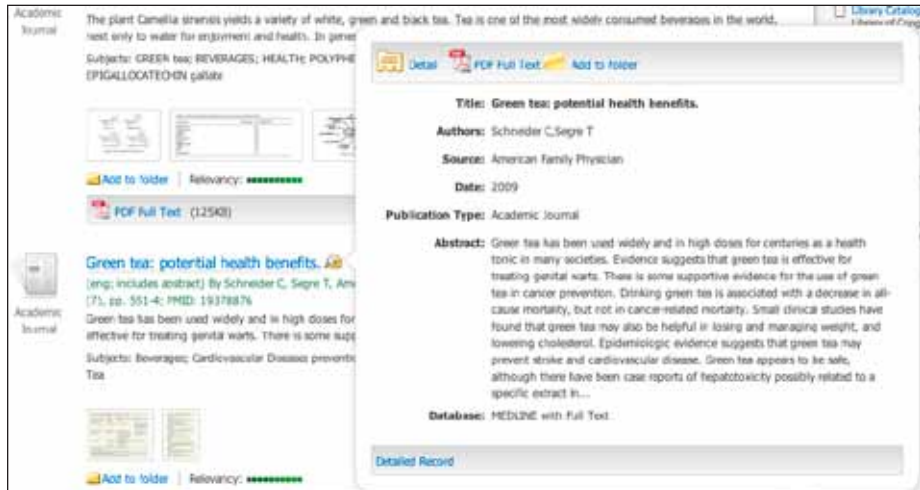


**Figure 26**  
EDS brief results: article

language. Subject headings are provided, and a real-time status call to the library catalog is made at time of search, which provides call number, status, and location information, as well as a link to the full catalog record. This link, when clicked, pulls up the record in the native ILS interface in another browser window or tab.

For journal articles and similar content (figure 26), typical information provided depends on the level of metadata information and may include common elements such as title, author, publication, volume and issue, and page numbers. For many articles, subject headings, the first few lines of abstracts, and thumbnail images of figures or tables within the article are also provided, depending on the source of content. Full-text availability for articles and similar content types is indicated through an icon and a link indicating the full text

in PDF or HTML; clicking on this link takes the user immediately to the full text, as described below. For other items with full-text availability, an icon and the link Check the Library Linkresolver for Full Text may be displayed, or, for some content providers, another message (e.g., for NewsBank-sourced content, Browse This Newspaper Title at NewsBank.). Which options exist depends on the source of content and the rights Ebsco has to display the full-text content natively within the EDS interface. For all item types, a magnifying glass icon appears to the right of the title. Mousing over this icon invokes a callout display (figure 27) providing much of the same information; in addition, it lists the database housing the record (such as the library catalog or a particular subject database) and, for particular item types, such as journal articles, a more complete abstract.



**Figure 27**  
EDS magnifying glass callout

### Faceted Navigation and Search Refinement

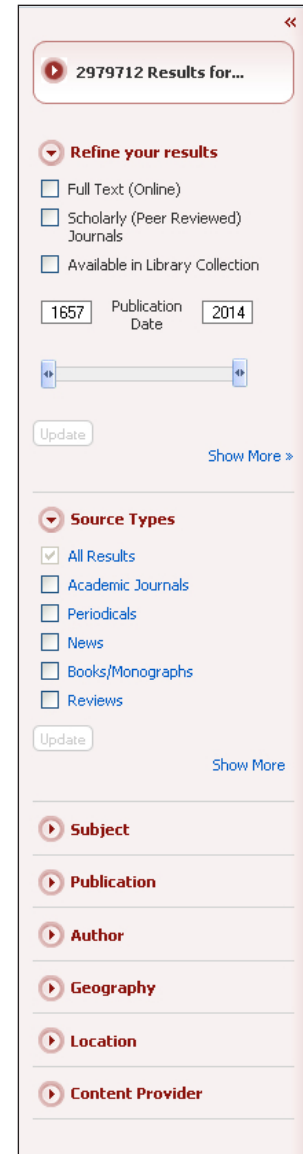
EDS offers various search refinement methods, including faceted navigation. By default, the left-screen pane is used to refine results (figure 28); the right-side pane is used to incorporate the optional EBSCOhost Integrated Search federated search component and widgets of additional functionality, described shortly. At the top of the refinement pane, users can choose to limit to full text, scholarly peer-reviewed journals, or the library catalog only. At the time of this writing, Ebsco is engaged in the classification of peer-review status for journals and expects that by the end of 2010, over 100,000 journals will have been classified. Users can also define a publication date range through a slider or by typing in boundary years. The majority of the refinement pane houses facet categories. Naturally, facets are dynamic according to the searches; not all searches will necessarily yield the same facet categories. For each facet category, several facet choices are displayed, and a Show More option invokes a callout that presents additional choices for that category. Facet choices have check boxes; thus, a user can limit to multiple facet refinements within each category. A breadcrumb refinement trail exists in that facet selections and limits applied by the user are indicated near the top of the refinement pane and can be x-ed out by users if they wish to easily remove the limits and expand their results. A Show More link underneath this top refinement section invokes the search options advanced search mode, described above.

By default, the top facet group is the Source Type category, which includes choices such as All Results, Academic Journals, Periodicals, Dissertations/Theses, Books/Monographs, and so on. The Subject category appears next, allowing for limiting results with one

or more specific subject terms. Other examples of facet categories that may be presented for a particular search include Publication Title (such as newspaper or journal title), Author, Location (such as a library branch), and Content Provider. This last category provides a listing of databases that have content matching the query; this list also includes the library's catalog as a content provider source. The only facet category that indicates the number of matches next to each choice is Content Provider; the number of matches in each database is indicated in parenthesis. Ebsco plans to introduce hit counts for other facet categories in 2011.

### Returned Results—Detail View

In the brief results, clicking on an item title retrieves a detailed citation record view within the same browser window, though part of the interface changes (as there is no longer a need for facets, etc., at the single-item level). For physical books held within the library (figure 29), the detail view provides typical citation information (title, author, source, etc.) and, if available, an enlarged book cover image. Status and holdings information similar to that provided in the brief results view



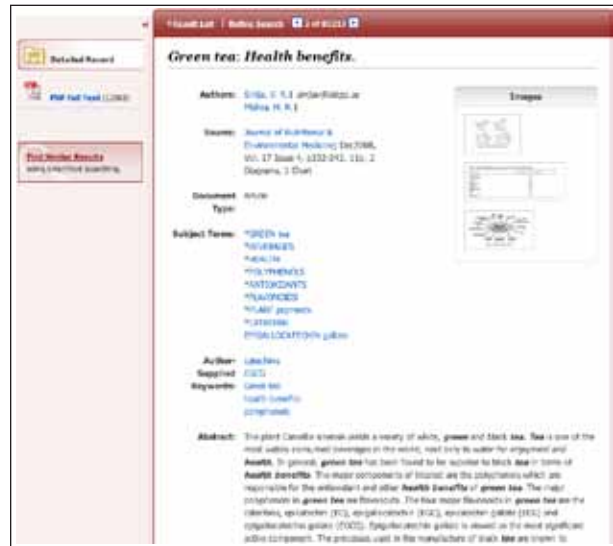
**Figure 28**  
EDS refinement pane



**Figure 29**  
EDS detail view: book

(call number, location, status) is provided in a pane to the left of the main screen content, as is a link to the native catalog record through the ILS interface (which, when clicked, opens up the native catalog in a new window or tab).

For articles and similar content, the detailed record view (figure 30) provides similar information to that displayed through the magnifying glass callout in the brief record view (i.e., title, author, source, subject terms, geographic terms, abstract, etc.). As described above, for full-text article content, an icon and link indicate full-text availability. As Ebsco is a provider of full-text content through many of its EBSCOhost databases, much of this EBSCOhost full-text content is displayed natively within the EDS interface, depending on local library Ebsco subscriptions. This display occurs when a user clicks on the PDF or HTML full-text indicator. In both cases, the user stays within the EDS interface wrapper. For PDF records, EDS uses the integrated Ebsco PDF viewer (figure 31). Within this view, the full-text content displays in the majority of the screen, while the left-hand pane allows users to see other articles within the same issue as the full-text item they are viewing and to navigate to other issues of the same journal. A large portion of the PDFs are in native PDF format; users can search the full-text item for a term. For items for which a PDF or HTML link does not exist, a link to browse the full-text content in the native interface may be provided; in such cases, a single click will return the full text. In other cases, a link to check the library's link resolver is provided, and, depending on how the local library has its resolver configured, the user may be immediately taken to the content or to a menu choice of content providers.



**Figure 30**  
EDS detail view: article



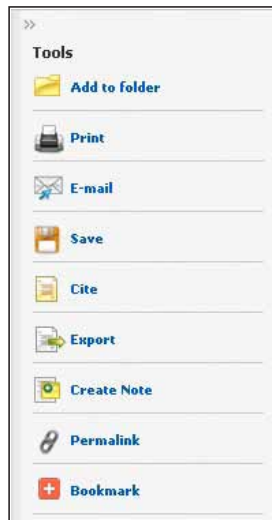
**Figure 31**  
EDS integrated PDF viewer

Note that the above description assumes a user is authenticated, usually behind the scenes via a campus IP address, or possibly through explicitly logging in through some other mechanism, such as a proxy server. As mentioned earlier, Ebsco released a guest access mode in mid-2010. This mode allows an unauthenticated user (such as someone from off campus who has not authenticated via a proxy server) to conduct basic searching of the library's EDS instance. In this unauthenticated state, citation information will be provided for a wealth of resources, but some citation information is not available or allowed for unauthenticated views, and the user will get a message: "This result from [content source] cannot be displayed to guests. Login for full access."

### Exporting Options, Shopping Carts, RSS Feeds

EDS offers a variety of export options (figure 32),

with printing, e-mailing, saving, and exporting to a citation management program all supported. A half dozen citation formats are provided, and the user can export to popular management programs, such as EndNote, ProCite, RefWorks, and BibTeX. From either the main interface (the screen with the initial list of returned results) or the detailed citation view of an individual item, users can place items into a folder (i.e., the item is marked for some future action by the researcher). On the main brief results page as well as at the detailed view, a folder icon



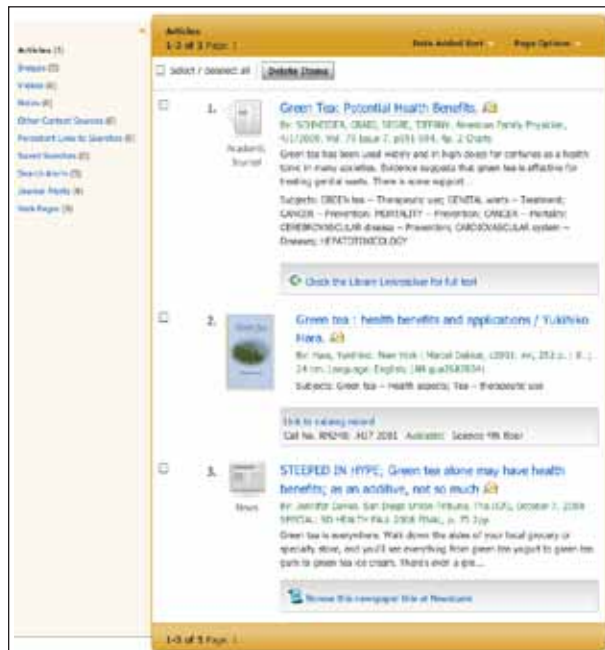
**Figure 32**  
EDS export options

appears at the top; a user can click on this icon to see a shopping cart-style view of all marked items (figure 33) and select which items to print, e-mail, save, or export into a citation management program (figure 34). In addition, in the detailed view, presented within a right-side pane is an inventory of current items the user has placed in the folder (marked). Carrying over functionality from the EBSCOhost platform, EDS allows users to create an Ebsco username and password account (or use an existing EBSCOhost account if they already have one), which allows them to store marked items, saved searches, and alerts. Users can set up alerts to automatically inform them via e-mail of new information; such alerts can be set up to periodically run a saved search or to alert users when a new issue for a selected journal title becomes available. Users can define some parameters, such as how frequently a particular search is run (e.g., daily, weekly, or monthly). Alternatively, users can set up alerts for a predefined search or new journal publication as an RSS feed. The personal account also allows the user to set preferences that provide partial control for some features, including how results are displayed (such as number of results per page) and export preferences. These preferences will be applied for subsequent EDS sessions, assuming the user logs in with his or her username and password at the next session.

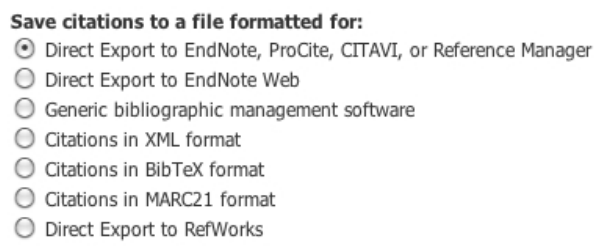
## Additional Features

### Did You Mean? Spelling Suggestions

EDS offers Did You Mean? functionality to address misspelled words. For a search term not matching any entry in the index, the user is returned to the initial



**Figure 33**  
EDS shopping cart



**Figure 34**  
EDS citation export

search box with the indication No Results Were Found and a Results May Be Available For . . . statement with one or more hyperlinked term suggestions; clicking on the linked term executes a search against that term.

### Embedding in Other Online Venues

EDS uses persistent URLs, so libraries wishing to provide a link for a “canned search” can do so and embed such links on different webpages. The persistent URL for a single item is displayed as part of the information in each detailed citation-view record. Persistent URLs to help define a canned search can be obtained by running the search and clicking on the Alert/Save/Share link near the top of the brief results screen. Bookmarks to EDS can be pushed to social networking/bookmarking sites, such as Digg, Facebook, and so on. In addition, a Search Box Builder tool helps libraries embed

## Vendor Website

Ebsco Discovery Services  
www.ebscohost.com/discovery

## Example Implementations

James Madison University  
www.lib.jmu.edu

Mississippi State University  
http://library.msstate.edu

Northeastern University  
www.lib.neu.edu

University of Oklahoma  
http://libraries.ou.edu

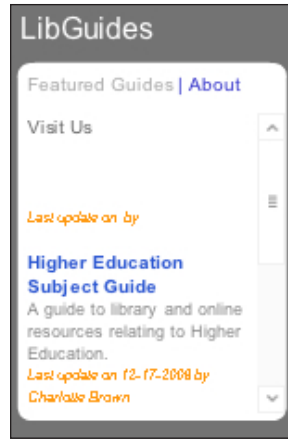


Figure 35  
EDS widgets



Figure 36  
EDS mobile view

the EDS simple search box on other webpages, such as the libraries' home page or institutional course management system. Should the library wish, it can use this tool to preconfigure limits for searches in this embedded search box; for example, a search box embedded within an engineering subject library guide could automatically be set to limit searches to materials available within an engineering collection, published after a certain date, in a certain format, and so on.

## Searching of Additional Remote Resources

As mentioned above, a right-screen pane is provided in the EDS interface, focusing on optional functionality the library may wish to incorporate into the EDS experience. Ebsco's optional federated search product, EBSCOhost Integrated Search, can be incorporated within this pane and allows for searching of additional sources whose content is not available or currently licensed to be preharvested and preindexed within EDS's central index. Traditional federated search products are not the focus of this issue of *Library Technology Reports*.

## Widgets

Widgets (figure 35), optional pieces of additional functionality, also appear in the right-side pane. Widgets are another avenue of pulling external content into the EDS experience. Examples of widgets could include integration of other library tools, such as LibGuides subject guides and blogs. Other examples include Flickr image, Google Books, and Wikipedia search functionality widgets and incorporation of an online chat client, such as Meebo. Some external sources, such as LibGuides, provide information helpful to widget creation; Ebsco has examples of widgets and their associated code at its support website. Widgets can be added to the interface using custom HTML code or as an iframe URL.

## Statistics

EBSCOadmin, the administrative module, provides, among other things, statistics related to EDS (and is the same module used for the EBSCOhost platform). For EBSCOhost databases within EDS, statistics available include number of sessions and searches (by month and database), number of abstract views, and number of full-text article requests (by month and journal). Statistics are customizable by a monthly date range. Statistical reports and graphs can be generated and displayed by date (month), by database, and by total searches. Statistical reports can be sorted and exported in a variety of formats.

## Mobile Interface

EDS can utilize the out-of-the-box EBSCOhost Web-based mobile interface optimized for various smartphone platforms (figure 36). For EDS customers, this interface searches the entire centralized index. Users can conduct a search, apply limiters, see results, select an item, and ultimately choose to access the full text or e-mail items. Some EDS features available in the full Web version are not currently supported in the mobile interface, such as automatic real-time status calls for ILS catalog holdings. At time of the writing, Ebsco was working on an EDS application.

## Upcoming Directions

While limited details can be shared to the broad public, Ebsco provided some directions that the company is currently exploring. Ebsco plans to continue expanding the multilanguage support of the EBSCOhost/EDS platform, including support for displaying Chinese, Japanese, and Korean characters. Ebsco is exploring



## Vendor Perspective: EBSCO Discovery Service

The goal of discovery is to allow libraries to maximize the visibility, value and usage of their overall collections. Through *EBSCO Discovery Service*™ (EDS) library collections are more accessible, allowing users to gain the most benefit out of the information the library makes available.

EBSCO Discovery Service creates a unified, customized index of an institution's information resources, and an easy, yet powerful means of accessing all of that content from a single search box. These custom solutions are created by harvesting metadata from both internal (library) and external (vendors) sources, and creating a pre-indexed service of unprecedented size and speed. In addition to the most comprehensive and robust collection of metadata from the best content sources, EDS also provides full indexing for *EBSCOhost*® databases and other partner databases including: Web of Science, Oxford University Press, Baker & Taylor, NewsBank, Readex, LexisNexis, Alexander Street Press and more. Not only does EDS search the most inclusive set of metadata, but superior relationships and licenses with academic publishers make EDS the most comprehensive service for searching the complete full text of journal articles and other sources—offering a truly integrated one-stop search experience for all of a library's journals, magazines, books, special collections, OPAC and more.

By leveraging the fast and familiar *EBSCOhost* platform, EDS offers a single interface for discovery and powerful features to heighten the research experience—everything the researcher needs in one place. When a library uses EDS on *EBSCOhost*, consistent

searching does not end with the result list. *EBSCO Discovery Service* offers a full-featured experience for users to remain within the structure of *EBSCOhost*—limiters and expanders, alerts, email, print, export, citations, bookmarkable URLs, persistent links, RSS, etc. In fact, *EBSCOhost* features are available for any applicable database, from any vendor including such features as: *EBSCOhost* basic and advanced search, screen functionality, subject clustering (facets), publication clustering (facets), sorting results by relevancy or date, date slider limiter, adding to folders and custom links. With EDS, EBSCO builds on the strength of the interface to serve the end user throughout their search, not just to find results based on a simple thin metadata search and then send searchers off on their own.

Another element that makes EDS different is the level of customizability that is available. The goal is to put the power and control in the hands of the customer, and have the “EBSCO” name take a backseat. *EBSCO Discovery Service* offers an unprecedented level of customization to the interface for prominent logo placement, interface colors, naming of the service itself, tool bar customization, etc.—all in such a way that can easily dovetail with university marketing/branding efforts. *EBSCO Discovery Service* allows sites to set up widgets (e.g., Library Guides, etc.) directly on the result page, as well as export bits of functionality from the EBSCO experience to be used in other sections of a university's website. With these options, the power of *EBSCOhost* remains but the look and feel of *discovery* becomes closely associated with the institution's website and identity.

a database advisor capability, suggesting databases the user may wish to search based on the query. Enhancements to the mobile interface will continue, and development of additional widgets that can be embedded in the EDS interface is expected. Ebsco is looking to expand information provided for monographs through incorporation of additional enrichment information (such as reviews) and through

provision of a level of FRBRization for book records in the sense that available editions, manifestations, and formats of a unique item will be presented in a single record view. In addition, Ebsco is exploring audiobook download capabilities. Ebsco expects to expand support for consortia-based catalogs or catalogs for libraries having close working relationships.