The Human Interface: an Ongoing Study of OPAC Usage at Adelphi University

Terry Ballard
Gill Library, College of New Rochelle, tballard@cnr.edu

Follow this and additional works at: http://digitalcommons.cnr.edu/gill-publications

Part of the Library and Information Science Commons

Recommended Citation

This Article is brought to you for free and open access by the Gill Library at Digital Commons @ CNR. It has been accepted for inclusion in Gill Library Publications by an authorized administrator of Digital Commons @ CNR. For more information, please contact lfazzino@cnr.edu.
The Human Interface: An Ongoing Study of OPAC Usage at Adelphi University

Terry Ballard and Jim Smith
Adelphi University

History of the System

In September 1990, the libraries at Adelphi University migrated from a CLSI circulation system to an Innopac online catalog and circulation package. The database contained just over 300,000 bibliographic records, and approximately 320,000 item records. Although there had been public access to the collection on CLSI, searches were done by means of the circulation commands—i.e., "&" would precede a title search. The command structure was augmented by function keys on the Wyse 60 terminals. However, the system, known as ALICAT (Adelphi Library Catalog) was not really designed for end-users.

In general, the database migration was smooth, with most of the problems occurring in the area of location ambiguities for items that were held in branch collections. There was also a substantial problem with the migration of diacritic characters, but since most of the holdings were English language materials, this problem was not immediately apparent.

The Systems Librarian, who had joined Adelphi the month before the cutover to the new system, was working with an OPAC that was already quite effective and user friendly. He saw his task as one of taking a system that was good and improving it as time went on.

A quote from literature or the arts was added to the bottom of the Main search menu (see figure 1), and was changed on a weekly basis. This helped emphasize the philosophy that the computer was a tool for human information needs.

Another early modification was a rewrite of the help screens. Search commands are initiated by typing in a mnemonic letter—"A" for author, etc. There is then a brief help screen that describes how to formulate that type of search. It includes a few examples of correctly composed searches. Since the original screens were generic options designed by the company, the search examples were not always represented in Adelphi's collection. For example, if a user typed "T" they would see the message in figure 2, below.
Welcome to Adelphi University Libraries
For Assistance, Please See the Reference Librarian

You may search for library materials by any of the following:

A > AUTHOR
T > TITLE
S > SUBJECT
K > KEYWORD
L > LC CALL #
N > NON-LC CALL NO
G > GOVT DOC #
I > ISBN
R > RESERVE Lists
B > Library INFORMATION


"Knowing others is to be clever. Knowing yourself is to be enlightened."
- Lao Tzu

Figure 1. Main Menu.

TITLE:
Type as much or as little of the title as you want
for example --> Zen and the Art of Motorcycle Maintenance

or just --> Zen and the

... then press the RETURN key

Figure 2. Generic title entry template.

Studying Usage in General

One of the most striking features of the Innopac system is the wide array of data that is available concerning searches. It is possible to monitor all searches in the system as they are being keyed in. A file that logs searches by time of day, terminal, search type, and unsuccessful searches is available. It is large enough to store a week's searches during peak usage periods. However, it was noticed that when the file was more than 50% full, it had the effect of slowing down access to the system's file information screen.

The librarians were quite interested in knowing how well the system was performing, and this log information was regarded as an important tool in measuring the OPAC's overall effectiveness.

Normally, the log is saved to an ASCII file each weekday morning,
and kept in a hard disk for eventual retrieval in WordPerfect as the need arises. There is an ongoing project studying keyword searches, so the actual searches are saved each morning. Otherwise, only the percentages are recorded.

Although the system reports the total search count and the number of unsuccessful searches, there is no automatic list of the percentage of searches that are unsuccessful. After keeping a separate count of the ratio of unsuccessful searches to total searches on a Quattro database, we have found that the failure rate of approximately 25% has remained constant.

The Literature of Other OPAC Usage Studies

In a presentation given before the RTSD/RASD/LITA Pre-conference in 1982, Pauline A. Cochrane looked at the historic background of catalog use studies through a review of the literature. She projected the future of online use studies, specifically related to transactional log analysis. And she made suggestions for improvements in OPACs, drawn from the analysis of transactional logs.

Cochrane stated that many searches retrieve nothing, that many searches retrieve too much, that users do not enter their search as required by the system, and that users often misuse keyword searching by using common words.

Cochrane wrote that transaction logs give information about the OPAC's operation that could not have been obtained in any other way. She said that this data should be used to make corrections to the OPAC that will help future searchers.

Christine Borgman performed a study of an old, established, and heavily used system at Ohio State University. Part of the systems maintenance included regular monitoring of use. In conjunction with the monitoring data, observation data from another study was used.

The OPAC studied had no log-on or start session procedure. That made for easier use, but resulted in difficulty when trying to determine user “sessions” from the transaction log. To establish sessions, observational data was collected and later combined with the transaction log. They claimed they were able to match 61.5% of the logged transactions with the observed sessions.

Some of the questions the study attempted to address were: How long did searchers spend at the terminal? How did session lengths vary? How were session commands distributed? Did user behavior vary by session? Did user behavior vary by the length of the session?

The Library Control System (LCS) began operation in 1970. It is a command-driven system, with author, title, author/title, subject and call num-
ber access. To begin a search one must type a three letter command to indicate the type of search; for example, "AUT/" followed by another command key, for an author search.

Most sessions were short, with some variation by library for the duration of the session. There was an observed difference in session length by academic quarter, by library, and by location of terminal. One third of all sessions included at least one subject search. Subject searching was used less during the summer term.

All error sessions were examined and compared to the location of the terminal where the session occurred and by the school term when the session took place. Most errors were found to occur in the winter term and at the location adjacent to the card catalog. The error sessions, sessions in which errors consistently occurred, were also found to be comparatively short sessions 43.3% of the time.

Borgman felt that a random sample from one library or one quarter or one terminal would not result in an accurate representation of user behavior. She did not dwell on searchers' errors; but found 13.3% of all commands to be either typographic or logical and that 12.2% of all sessions consisted entirely of errors.

Three stated goals of Thomas Peters' study were to determine failure rates, to study usage patterns, and to determine possible causes of failures. Peters notes that reference librarians could use the information obtained from transactions logs in bibliographic instruction, reference interviews and collection development.

The catalog studied was LUMIN, a union online catalog used by all four campuses at the University of Missouri. The system was originally developed by the Washington Library Network. Only 60% of the holdings were online. Exclusions were government documents, music scores, recordings, manuscripts, and periodical titles.

A brochure was placed by all the terminals to explain the operation of the OPAC. Online information is displayed on the screen during different phases of the operation. However, there was no help key for users to press to gain more in depth help related to their search type.

Four terminals were studied over five months in this project. The study was divided into two phases: first an attempt was made to establish use patterns and failure rates for each type of search. The second phase was an attempt to assign probable causes for the failures.

Peters found failure rates to be consistently high: just over 40%. Ten percent of failures were related to spelling errors; with more than a fifth of the errors related to spelling and typing.

Peters recommends transaction logs as an excellent low cost way to improve communication among librarians, patrons, and online catalogs. If li-
library use instructors are familiar with users' problems observed in the transaction log, they can guide the user past problem areas and make instruction more relevant to the user.

Collection Development librarians could also use information gathered from transaction logs to determine what types of material and areas of the collection need strengthening. Designers of OPACs could use the log information to make system enhancements which would anticipate user errors. Peters listed three factors which decide the success of OPACs:

1. The technical performance of the computer system and the communication system.
2. The design of the interface between machine and people.
3. The quality of the database being searched.

Peters finds it difficult to understand why some users will spend hours working with software on their personal computers to make it do what they want it to, but then refuse to spend more than five minutes learning how to use an OPAC. He finds that basic skills like typing and spelling, as well as more specific conceptual knowledge about online database systems and Boolean logic, are important for OPAC users' success.

Mary Richards at the University of Wisconsin-Stout conducted a three stage study from 1982-86, during the library's conversion from card to online catalog. Richards found that a majority of searches were subject searches. Author, title, and subject were the predominant search forms. Use of Library of Congress Subject Heading (LCSH) increased for those using the OPAC, compared to the card catalog. Previous computer use had no effect on user's success rate with the OPAC. Users preferred the online catalog to the card catalog.

Subjects were asked if they would answer a few questions on their use of the catalog. The questionnaire was lengthened later in the study and this caused a decrease in the total respondents in each succeeding phase.

75% to 80% of all searches were subject searches. Subject searching took longer to do than known item searching. It was more important than ever for users to know an exact LCSH, and to be able to type it accurately. The percentage of searchers who reported finding no matches to their search ranged from 31% (in 1984) to 9% (in 1986). Part of this drop could be attributed to increased education of users in use of LCSH directories. Some search failure could be attributed to collection failure and not to catalog or user failure; the library just did not have the item. The 1984 phase of the study compared card catalog use to online catalog use side-by-side and found users preferred the online catalog. However, it should be noted that the card catalog had not been regularly maintained since the online project began. Also, in-
struction in use of the card catalog had been discontinued.

More than half of the users reported that they had no problem using the online catalog in 1984 and 45% reported no problems in 1986. Problems mentioned included software bugs and problems with subject searching.

As a result, the approach at the University of Wisconsin-Stout has been an attempt to emphasize subject searching in bibliographic instruction. The effectiveness of their online instructions was found difficult to measure and evaluate. They found that the use of LCSH had not been a part of search strategy in using the card catalog, but it grew significantly.

Rhonda N. Hunter described a transaction log study conducted at the BIS online catalog at North Carolina State University. Subject searches accounted for more than 50% of total searches in the system, and they produced no hits 62% of the time.

The overall failure rate in the study was 54%. Hunter believes that the high failure rate is due to the fact that the system is command-driven rather than menu-driven. She concludes that it would be more effective to improve the OPACs themselves rather than to educate the users to search more efficiently.

Beverly Barrett studied failed searches on OPACs at the Dixson Library, University of New England in Australia. She reported that the main reason for failure was that users did not match the controlled vocabulary of the database. The librarians wanted to plan improvements to the OPAC and increase user ability to access information. They used transaction log analysis in the study to determine a preponderance of subject searching and the high failure rate of the searches. Subject searches were popular but difficult to perform successfully. Their catalog provides subject access through use of LCSH and FLASH. Other search types included author, title, call number, ISSN, and ISBN. The catalog has keyword access potential that is not available to the public. The transaction logs from five days of OPAC use were examined.

To searches that failed, the OPAC responded in two ways, either: "No exact match—try again being less specific" or "No qualifying entries found—please enter new command," depending upon the system recognizing some or none of the search items.

The analysis of errors began with the sorting of spelling and keyboarding errors. Barrett checked them against the Macquarie Dictionary. Remaining errors were checked against LCSH or FLASH. Then, using keyword searching, errors, except spelling and keyboard, were again checked.

This study analyzed individual searches, not user sessions, because of the limits of the transaction log. Of 3215 subject searches, 32% failed. These failures fell into three categories:
• Search did not match a subject term (70.2%)
• Spelling error (15.8%)
• Incorrect or inappropriate search strategy (10.0%)
• The cause of failure could not be definitely determined (4.0%).

The most significant cause of failure was the use of non-LCSH terms, accounting for 36.5% of the failures.

Ray Larson's study covered six years and involved over 15.3 million searches in the MELVYL catalog at the University of California. He found that subject searches accounted for 59% of use; that subject searches were most difficult to perform for 43% of those reporting difficulty; and that upgrades of subject searching was the most desired OPAC improvement.

The study found a decline in use of subject searching and an increase in title keyword searching as its replacement, and the author places the blame for this on the difficulty users have found in using subject searching in the OPAC. Subject searching is still the most commonly used method of search in an OPAC, but it is also the most likely to fail by either providing no hits or too many hits for a user to manage. Misspelling and lack of understanding of LCSH account for subject searching's high failure rates.

Adelphi University Study of Unsuccessful Searches

After one semester of using Innopac, the non-print librarian suggested making a study of unsuccessful searches. While there have been a number of user log-based studies in the literature, we were not aware of one that was performed on an Innopac system. It was decided that since the system did a significant part of the sorting, an in-depth study of searches producing no hits would be quite feasible.

Unsuccessful searches conducted from the period of February 18 to February 21, 1991 were saved in a printout. We examined 703 title searches, 357 author searches, and 880 subject searches. A numeric code was devised to cover the different kinds of problems that led to a hitless search. While these breakdowns are similar to Peters' codes, we had not yet seen his article at the time that we started our project. The following was the working code for all searches:

1. Misspelling
2. Garbage
3. Spurious entry. Usually some sort of scatological reference. (This was subsequently merged with #2).
4. Wrong form (such as a search for a title in Author mode).
5. Name reversal (first name first).
6. Material just not in the database.
7. Spacing error (words run together or split).
8. Misuse of initials or abbreviations (Newman M. A., for example, instead of Newman M., would block truncation of the first name).
9. We could not determine what the problem was. This category was dropped in the final count. Since there is some sort of an uncertainty factor in every search, we gave each one our best determination.
10. Two authors typed in at once (i.e., Masters and Johnson).

Two other codes were added just for Subject searches:

3. Initial article, which is overridden in title searches, but not in subject.

Entries that were coded “1” for spelling mistakes were given a further analysis. We also noted the word number that the mistake occurred in, the number of total words entered, and the letter position within the word. For instance, the title “Journal of nyursing education” was coded as “1 w3-4 p2.” In analyzing these positions, we could not find any inferences that were particularly enlightening. For instance, of the 144 spelling mistakes that were coded in title searches, 60 occurred in the first word, 37 in the second word, 22 in the third word, and 25 in a later position. The system gives “nearby’s” when it cannot make any hits on a search. Obviously, the most useful screens appear when the mistake is made after three words, so the high proportion of first word mistakes is somewhat discouraging.

In 1991, a study at Adelphi University concerning spelling errors in OPACs and other databases revealed a tendency for common words with three or more syllables and 8 or more letters to be quite commonly misspelled. In the list of 144 misspelled title searches, 54 fit this pattern—many variations of “education” or “introduction” or “psychology” were present with inverted or dropped letters.

Unsuccessful Author Searches

The chief source of no hits with author searches was the absence of the author in the Adelphi holdings. Out of 357 searches studied, more than 40% fit this category (see figure 3). Name inversions and misspellings accounted for 17.65% each. Another 7.84% were searches that should have been subject or title searches. The same percentage were searches that added two author’s names (such as Woodward and Bernstein). The system has to look for a sin-
Advances in Online Public Access Catalogs

Figure 3. Unsuccessful author searches by mistake type. Figure by R. A. Ferrick.

gle author per search. The remaining searches were either garbage or searches that used too many initials in an author's name.

Searches like “tional psychology” represent a special class of problem searches. A patron was looking for the title or subject “rational psychology,” but did not choose a category from the main menu. The initial “R” that was typed was ignored by the system, because there is no such option on the menu. The second letter activated an author search, and the patron completed the search without noticing what was happening. The same thing happened with the search “iting for godot.” Another type of search is represented by “poe edgar allen on criticisms.” This person was confused about the nature of an author search.

The misspellings in the Author searches were often due to the difficulty of certain names. There were many near and not-so-near misses such as “Nostradomous,” “Pulitzer, Joseph” and “Borge, Jose Luis.”

Unsuccessful Title Searches

Of the 703 title searches that were analyzed, 496 or 70.55 percent of them were phrases that seemed legitimate, so we classified them as valid searches for titles not owned by Adelphi (see figure 4). Just over 20 percent were recognizable spelling mistakes. Garbage entries were 4 percent, and there were only trace amounts of other problems.
We realized that it was very difficult to tell if the entry was legitimately a wrong form, *i.e.*, somebody who was doing a keyword or subject type search to see if there was a book with that title. Some of the longer and more convoluted titles looked like keyword searches.

Because of that ambiguity, we made a special substudy of titles that seemed legitimate. Those containing four words or more were searched in OCLC to see if hits could be established in a database containing more than 20,000,000 records. The assumption was that most of these titles were made up. The results were just the opposite. Out of 117 type 6 searches (material not in database), 70 were found in OCLC and 46 were not. We concluded that this effect was caused by people searching from reading lists for books that Adelphi University simply does not own.

**Unsuccessful Subject Searches**

At first glance, the unsuccessful Subject searches have one basic problem—the searches do not match the LC Subject headings the library uses in its cataloging, and the entries were far enough from recognized analogous terms that they did not pull up a "see" reference (see figure 5). As in the author searches, a continuing problem is the practice of searching personal names with the first name first. A problem that is unique to Subject searching is related to initial articles. In a title search, initial articles are bypassed by the system. This is not done in a subject search—presumably because some of the headings
actually do begin with an initial article. However, except for those few cases, a subject search that begins with “the” will not only bring up no hits but call up a “nearby” screen that is entirely off track.

In looking at unsuccessful subject searches, one can see patterns where a user will try a number of different approaches for the same search series. These patterns can be more easily studied when we take advantage of the search log for the individual terminal report. One can sense the frustration in the following ten minute session where all but two of the search keys were non hits:

- puerto ricans in usa 11:47 AM, Feb 20
- puerto ricans in the united states 11:48 AM, Feb 20
- puerto ricans new york n y education 11:50 AM, Feb 20
- assimilation of puerto ricans 11:53 AM, Feb 20
- puerto ricans in american society 11:53 AM, Feb 20
- sociology of puerto ricans 11:55 AM, Feb 20
- puerto rican advancement 11:56 AM, Feb 20
- puerto ricans immigration status 11:57 AM, Feb 20

While we thought that some title searches were really keyword type requests, we were certain of it in the case of Subject searches. To that end, unsuccessful searches with multiple words were searched in the keyword file.
In one page of the printout, we identified 24 of the 56 searches that could be successfully searched in keyword (misspellings, garbage strings, and purely numeric searches were disqualified). From that group, 17 produced hits as a keyword search, while only 7 did not.

Most of the ones we checked were subject searches of two words, or two words linked with “and.” For instance, “african civilizations” got no hits as a subject search but we found 11 records when the same search was done as keyword. Normally, searches of 3 words will not produce hits in keyword because all three have to be in the same record, but the term “aged mental illness” found 1 record.

Results of the Study

Followup to the study has been made on a number of fronts. Using data from the most common kinds of errors, we made an extra screen in the OPAC in the Library Information section (see figure 6).

Information from the usage studies was made available to those librarians who perform library instruction. According to Valerie Feinman, Adelphi University’s Coordinator of Library Instruction, her classes warn students of the common mistakes that come up in searching the OPAC. She said that there is a particular tendency for students to type personal names in the wrong order in author and subject searching. Feinman also said that the problem of searching by more than one author is generated by students who are searching for article reprints that their professor has put on reserve. Since most of these classes are not exposed to bibliographic instruction, she felt that the problem had no easy solution.

Feinman said that most of the instructional sessions are performed...
SUBJECT:

ALICAT uses standard subject headings. If you are not sure of your heading you might try a KEYWORD search.

Type as much or as little of the subject as you want

for example ---> Sports medicine

or just ---> Sports

or ---> Shakespeare, William

or just ---> Shakespeare

... then press the RETURN key

Figure 7. Subject search tips screen.

for classes that are working on a specified project. She usually emphasizes subject and keyword searching. In spite of this approach, our daily usage reports show that keyword searching is inadequately used in the university’s searches.

Keyword Searching

In the 18 months that the Innopac ALICAT system has been operating, keyword searches have averaged only 5% of the total searches. There has never been a day where the percentage has topped 10%. Communication with other Innopac libraries on the Internet has confirmed that Adelphi’s reports are typical.

We consider this a problem in a number of ways. Since the keyword index includes the contents notes, there is information about play titles, song titles, and short story titles in collections that can be obtained only from the keyword approach. Students who search the University’s wide collection of CD-ROMs use a keyword approach almost exclusively, but when they go to the OPAC, their search strategies center on the classic card catalog mode of Author/Title/Subject.

Also, as we have pointed out, a number of unsuccessful subject searches would have been successful keyword searches. To that end, we altered the Subject Help Screen to suggest a keyword strategy (see figure 7).

We also prepared a special handout explaining the advantages of Keyword searching, and some tips on how to do it (see figure 8).
The Key Is In the Word

One of the best kept secrets of our ALICAT system is that the 300,000 records in the online catalog are indexed by keyword. This means that you can search Adelphi's database in the same way that CD-ROMs are used. There are two major advantages to this method of searching. The first is that searching words in the title can be used as a substitute for a subject search. For example, one student typed in a subject search for "Persecution of Jews in the Crusades." That is not a subject heading that the library uses, so the student came away with nothing. However, a keyword search of "Jews" and "Crusades" brought up the title The Jews and the Crusaders, which contained the information that the student needed.

The second advantage of keyword searching is that it will bring up the name of a play, a story, or even a song title within a collection. For instance, to find the Beatles' song "I Saw Her Standing There," run a keyword search using the terms "Saw" and "Standing." You will be referred to the cassette "Rock 'N' Roll Music" which contains the song and which is available in the Nonprint Media collection. Similarly, a search of "Odd couple" will lead you to a collection of plays including Neil Simon's "Odd Couple."

To begin a keyword search, simply select K > KEYWORD from the main menu. At the prompt (which appears as WORD: on your screen), type in one or more words from the title of a work or one or more key concepts in your research topic. Don't use "and," "or," or other connecting words to link them.

Examples of Keyword Searches
WORD: Innocents Abroad
WORD: Jews Crusades

The ALICAT system has so many ways of getting to information that we are continually discovering new ways to search. If you don't find just what you are after, check with the librarian at the reference desk.

Other System Measures

Innopac has a number of capabilities for limiting a search. If a search brings up too many hits, a user can limit the results by date (i.e., after 1986), by location, by language, or by format. All of this is menu-driven and easy to work. There is even a very powerful feature for finding key words in the subject headings. From what we have observed, these limits are rarely used by our patrons. We have no way of quantifying this, because there is no log data on whether a search was limited or not.

Another feature that does leave data is the "Suggestions" option on the Library Information screen. Since the system went up, hundreds of users have typed in complaints, poems, obscenities, and even suggestions. These provide good input, and they are regularly reviewed by library administrators.
There was a one word suggestion left late one night on one of the terminals in the lobby that illustrates that there is still work to be done on the system. The suggestion was simply: "HELP."

A second option for suggestion is a format for requesting that the library purchase particular books. This also brings in a number of spurious entries, and numerous requests for books that do not actually support the university curriculum—chiefly titles by Stephen King.

Conclusions

After studying our user logs for 18 months, we are encouraged that our users find hits for their searches 75% of the time. Even the remaining 25% of searches do not necessarily mean that the user went away with nothing. The nearby screens provided by the system present the unsuccessful searcher with options that often contain the record that was intended. People who get an unsuccessful search always learn something about how to search better the next time.

In addition, the search data provides the university with valuable information about how to improve the system. As a direct result of this data, we have improved our help screens, our information screens, and our library instruction. We have begun the process of improving our subject heading authority file so that reasonable search headings like "Great Depression" will be linked to the correct heading.

Thomas Peters, Rhonda Hunter, and others have emphasized user log studies as a way to improve the connection between the user and the OPAC. At Adelphi University, we have also found that the data from the usage log is tremendously valuable; particularly information concerning unsuccessful searches. Every unsuccessful search can give clues that will lead to a more efficient system or more efficient users.

Notes

4. Richards, Mary, "Catalog Use at the University of Wisconsin-Stout: A
