Julianne Beall  
Library of Congress

Dewey for Windows

Abstract: This paper describes how the features of Dewey for Windows (DFW) facilitate use of the Dewey Decimal Classification in CD-ROM form and suggests future enhancements to make DFW a more efficient tool. The strength of Dewey for Windows lies in the many different approaches that it provides to the classifier.

1. Introduction

Dewey for Windows provides a Windows interface for the Dewey Decimal Classification. The version described here is the first field test version. This version uses the same CD-ROM database as the earlier DOS version of Electronic Dewey, a CD-ROM that holds the 20th edition of the Dewey Decimal Classification (DDC 20). When Dewey for Windows goes into production, its CD-ROM will hold DDC 21. While the user interface of DFW is very different from that of the DOS version, the search engine (database access program) is similar. See Carpenter (1995) and Trotter (1995) for descriptions of the DOS version. This paper does not attempt to cover all the features of Dewey for Windows; it focuses on those features that seem most useful to the working classifier. See Vizine-Goetz and Bendig (1996) for a more comprehensive description of DFW.

2. Searching

Dewey for Windows offers multiple ways to search for a particular topic, but they all involve use of the Index window or the Search window. The most useful approach with the Index window is the Relative Index (Phrase) scan. With this approach, one scans and selects from a display of the Dewey Relative Index that is similar to the printed index. This approach is particularly useful for terms that appear in many different places in DDC, when the classifier wants to be guided to the main numbers for the topic, e.g., business, literature, programming, taxes. The built-in Scan view has a half-screen Index window (tiled vertically), a quarter-screen Search window, and a quarter-screen DDC Number window (see figure 1).

In DDC 21 additional nonprint Relative Index entries are being added to the CD-ROM version of DDC, entries that cannot fit into the crowded print index volume. The Dewey editors plan to add many more nonprint index entries over time.

The Search window has an input box used primarily for initiating keyword searches for terms or Dewey numbers, and it lists in DDC order the Dewey numbers and captions of the hits that result from searches. The Search view has a half-screen Search window (tiled vertically); that window is useful for viewing long hit lists. With the Search window one can do keyword searches on every word and number in the DDC database (except a very short list of stopwords), using Boolean combinations and truncation and mask symbols (asterisk [*] for zero or an unspecified number of characters at the end of a term, question mark [?] for zero or one character anywhere in a term). One can use index labels to limit a search. The most useful index label is dd: as in a search for dd:616.1-616.9; it

limits retrieval to the Dewey number as the entry number (or one of the entry numbers) in the record. A search for 616.1-616.9 without the index label retrieves every record with that span anywhere in any note, including the frequently appearing add note "Add as instructed under 616.1-616.9."

Keyword searches are very efficient for distinctive terms that do not appear in many places in the DDC, e.g., batik. One can of course also scan for such terms in the Relative Index (Phrases), but unless one already has an Index window on the screen for another purpose, it is not efficient to use the Index window when a keyword search will suffice. Keyword searches are useful if one does not know the Dewey Relative Index well enough to predict the order in which terms will be presented. For example, there is only one Relative index entry under World War II, for the main number 940.53. To index all the relevant records under World War II would violate the Dewey editorial rule against recapitulating the schedule in the index; however, there are many Relative Index entries elsewhere in the alphabet, with World War II as a subfield, e.g., Refugees—World War II. A keyword search on world war ii and refugee? will retrieve the relevant record 940.53159. Keyword searches are useful as shortcuts for particular combinations that are typically not given in the Relative Index. For example, if one is interested in compensation by industry, one could search for compensation and industr* to retrieve 331.28. If one used the Relative Index approach, one would scan for compensation, look at the full record for 331.21, and follow the reference to 331.28. Keyword searches are also useful if one recalls a particular example in a Manual note or an add note that would enable one to retrieve the desired record, e.g., Seth Thomas clocks given as an example in the Manual note 300 vs. 600 Social sciences vs.
Restrictors are special indexes that allow a search to be limited to a specific main class (one significant digit, e.g., sl:300), division (two significant digits, e.g., s2:330), or section (three significant digits, e.g., s3:331) or to a specific auxiliary table (e.g., tn:tl). For example, a search for waste? and s2:340 will retrieve waste(s) only in the 340 Law schedule, a search for germany and tn:t2 will retrieve Germany only in Table 2 Geographic Areas. Classifiers who are familiar with a particular part of the DDC schedule often remember the first two or three digits of the schedule number they need or which table they need and are well prepared to make good use of restrictors. Even without that familiarity, classifiers can use the restrictors when they are following add instructions that specify a particular part of the schedule or a particular table. For example, a classifier at 155.413 Conscious mental processes and intelligence [in child psychology], following the instruction to add "the numbers following 153 in 153.1-153.9", can search for, e.g., learning and s3:153; a classifier at 808.819 Poetry displaying specific features, following the instruction to add “notation T3C—1-T3C—3 from Table 3-C”, can search for, e.g., marriage and tn:t3c.

Using restrictors has the advantage that the search goes much faster than if one used only truncation (waste? and dd:34* or germany and dd:12*), but using restrictors has the disadvantage that relevant Manual entries are not included. If one really needs to find Manual notes as well as table or schedule information, then one may wait for the results of searches like biograph* and dd:36* or biograph* and dd:338*. Both searches retrieve useful advice in the Manual about classifying biographies. Or one may do a simple search for biograph* and scan the 172-record hit list (arranged in DDC order) looking for relevant records. One hopes, however, that the restrictor indexes will be expanded in a future version so that they can be used to retrieve Manual entries. On the other hand, if one wants to exclude Manual notes, which are sometimes intrusive, the current disadvantage of restrictors may be seen as an advantage; there is no other way to exclude Manual notes as a category from a hit list. Ideally, a better way to exclude (or limit to) Manual notes as a category will be provided in a future version of DFW, possibly by means of a Manual index and index label.

Library of Congress Subject Headings have been statistically linked to DDC numbers to provide additional search terms for the Dewey numbers—the five headings most frequently appearing with that Dewey number in bibliographic records produced by the Library of Congress. Special efforts are being made to provide LCSH for completely or extensively revised DDC 21 schedules, where statistical matching with DDC 20 numbers in bibliographic records would not be helpful. Using the Index window, one can scan and select from the LCSH (Phrase) index. The list that one scans is not a list of LCSH authority records, rather just a list of headings stripped from bibliographic records. One can also do a keyword search for the LC Subject Headings associated with a DDC number, using the index label su:. Since the statistical matches for the most part have not been editorially reviewed, there are accidental bad matches; one must always check the full information in a record to be sure that one has found the correct Dewey number. The LCSH are useful as a searching aid when the editors have not indexed a particular sought term or synonym that fits a Dewey number, e.g., Parental leave in DDC 20. These terms will later be added to the Dewey Relative Index, though in some cases only to the nonprint index. The LCSH are also useful for terms that are excluded by policy from the Relative Index. For example, specific battles in a war are not named in the Relative Index unless they are named in the schedule. Thus one can find World War I's Verdun (940.4272) and World War II's Pearl Harbor (940.5426) by name in Dewey for Windows only by searching on the Library of Congress Subject Headings. Of the three hits in the LCSH index for Pearl Harbor, two are obviously bad matches (European theater; East and southeast Asian theater), but one (Pacific Ocean theater) is good. For a classifier
weak in geography, there is another clue to which match is best: a click on the LCSH button for each record reveals that the bad matches were based on one bibliographic record each, while the good match was based on 27 bibliographic records.

One common reason for people to get no hits in the DDC database is that they are searching for a concept that would be represented by a built number. The pieces needed to build the number are in the database, but the built number is not. Educating classifiers is one answer to that problem, but adding more built numbers to the database is another. There is plenty of room on the CD-ROM, and there are plans for a continuing project of adding additional highly posted built numbers to Dewey for Windows, with their LCSH. Meanwhile, with the local annotation feature of Dewey for Windows, classifiers can add built numbers that they frequently use, with one numerical and one verbal index term searchable in a separate Notes index.

3. Selecting and Identifying Disciplinary and Interdisciplinary Numbers

Dewey for Windows offers several methods of finding a number in the correct discipline for a work and identifying the interdisciplinary number for a topic. With a display of the Relative Index (Phrases) in the Index window, the discipline of a number is shown by the subfields of the index entries, and the interdisciplinary number for a topic can be identified because it appears opposite the term without any disciplinary subfield. For example, the following Relative Index entries show that the social welfare number for pollution is also the interdisciplinary number, 363.73, and offer other numbers for the topic in the disciplines of law and toxicology.

Pollution 363.73
Pollution—international law 341.7623
Pollution—law 344.04632
Pollution—social welfare 363.73
Pollution—toxicology 615.902

If one of these numbers is found by some other approach, the index entries for the record can be seen in the full record display in the DDC Number window; that is another way to identify the discipline or subdiscipline, or to determine whether the number is the interdisciplinary number for the topic. One can also identify the interdisciplinary number for some topics by searching for notes in the schedules and tables, e.g., a search on death and interdisciplinary; however, many interdisciplinary numbers are identified as such only in the Relative Index.

Another way to find the numbers for a topic within a particular discipline or subdiscipline when the searcher knows the first one to three digits that specify the discipline is to use restrictors; for example, one could search for pollution and s2:340 to limit the search to the discipline of law, or pollution and s3:341 to limit the search to international law.

Another way to determine the discipline of a number is to view it in a Hierarchy window, and check its upward hierarchy. The Hierarchy window is described in detail below.

4. Moving Up and Down the Hierarchy

Dewey for Windows facilitates moving up and down the DDC hierarchy in several ways. The Hierarchy window presents the entire upward hierarchy for a number and the next step down in the downward hierarchy. For example, the hierarchy for 154.6 Sleep phenomena is
The hierarchy for 612.821 Sleep phenomena is

600 Technology (Applied sciences)
610 Medical sciences Medicine
612 Human physiology
612.8 Nervous functions Sensory functions
612.82 Brain
>612.821 Sleep phenomena

Within a Hierarchy window, the user can highlight any number in the hierarchical display, click on Hierarchy, and see the upward and downward hierarchy for that number. To see coordinate numbers, the user need only highlight the next higher number and click on Hierarchy. This makes it easy to start at any point in a schedule or table, then move quickly and easily up or down as necessary to identify the most appropriate number for a work.

To clarify the difference between two or more apparently similar numbers, a user may wish to compare their hierarchies. It is possible to put several Hierarchy windows displaying hierarchies of different numbers on the screen at the same time. One convenient way to do that is to customize one of the four user views with a quarter-screen Hierarchy window and a quarter-screen Search window both on the right side of the screen, where they will not quickly be covered by additional Hierarchy windows, which appear in a cascade pattern starting in the upper left corner of the screen but can subsequently be dragged to other positions.

A user exploring up and down in the DDC may need to see more information than is found in the concise summaries presented in the Hierarchy windows. To view nearby numbers a user can drag and drop a number into a DDC Pages window to get a display much like that of the printed page. This is a good way to see centered entries, which are excluded from Hierarchy windows except when the record for which hierarchy information is sought is a centered entry. The DDC Pages window does not, however, offer a good way to see numbers that are distant from the starting number, paging up and down through many numbers is slow.

A good way to get information about the downward hierarchy of a number in a Search window hit list is to drag the number to the input box of the Search window. The index label dd: will automatically appear to limit the search to instances of that number as the main number. The user can add an asterisk truncation symbol at the end of the number to get all the subdivisions of the number. To get information about coordinate numbers, the user can delete the last digit of a number before adding the asterisk. The resulting hit list will include all subdivisions of a number, not just the next step down; and it will also include centered entries, built numbers from the Relative Index, and Manual notes. Since built numbers from the Relative Index and Manual records are excluded from the DDC Page display, this approach gives the fullest listing of subordinate entries for a number. Since only the number and the caption are shown in the Search window hit list, many entries can be scanned quickly, especially if the Search window is maximized.
If the result of using the asterisk truncation symbol is too many hits for easy viewing even when the Search window has been maximized, a user can substitute for the asterisk two question-mark truncation symbols (each question mark specifying one character), or whatever number of question marks is appropriate. Thus the user can create tailor-made summaries of the downward hierarchy for a number.

One particular reason why classifiers need to move up in the DDC hierarchy is to find citation-order and precedence information. For example, a user is trying to classify a work on processing meat to be used for pet food and has identified two relevant numbers, 664.66 Food for animals and 664.9 Meats and allied foods. With the DDC Pages window, the user can easily browse from 664.66 to the centered entry 664.62-664.66 to learn that the number for meat should be given precedence. That approach often works; but sometimes the precedence information is too far away for simple paging upward to be efficient. For example, a user is trying to classify a work on wages of immigrant workers in the garment-making industry and has identified two relevant numbers: 331.28 Compensation in specific industries and occupations and 331.62 Immigrants and aliens [as workers]. The number in common is 331. The classifier can use the Hierarchy display to get to 331 or can simply search for \textit{dd:331} and display the record to find the table of precedence indicating that 331.62 is to be preferred. If neither of these approaches works, the citation-order or precedence information may be hiding in an elusive centered entry. Since every centered entry contains a note specifying where to class comprehensive works, one can search for \textit{comprehensive} and limit the search to the relevant portion of the schedule; for example, to get help in choosing between 331.52 Veterans [as workers] and 331.69 Indigenous ethnic groups [as workers] by consulting the centered entry 331.3-331.6, one can search for \textit{comprehensive and s3:331}. Only some of the hits will be centered entries, but those entries are easily spotted because they are marked with a “C” at the left in the hit list. Since centered entries are often crucial for various reasons, one would like to see the system provide a simpler and more obvious way to find all the relevant centered entries in the upward hierarchy for any given number. An alternative hierarchy display could be offered that would include centered entries. It would sometimes look awkward and cluttered, but it would be very useful.

5. Displaying Schedule, Table, and Manual Entries and Full Records for Built Numbers from the Relative Index

Dewey for Windows offers two basic ways to display schedule and table entries, and one way to display Manual entries and full information about built numbers in the Relative Index. Classifiers will select the display method depending upon their individual preferences and upon what information they need in a particular situation. The DDC Number window displays the fullest information about a single DDC number—the number and caption plus published notes, Relative Index entries for the number, and the LCSH linked to the number. The DDC Pages display shows multiple adjacent schedule and table records much as they appear in the printed DDC, with numbers, captions, and published notes. The user can page upward or downward in the schedule or table. The DDC Number window has a Hierarchy button to activate the Hierarchy window; the DDC Pages window gives no information about upward or downward hierarchy beyond what appears in adjacent entries displayed in the window. The DDC Number window shows slashes that indicate logical places to truncate DDC numbers; the DDC Pages window does not. The DDC Number window displays Manual records; the DDC Pages window does not. The DDC Number window displays built numbers from the Relative Index with their slashes, all the Relative Index terms and LCSH associated with them, plus the nearest upward nonbuilt number and its caption. The DDC Pages window does not display built numbers from the Relative Index;
but a quick way to find the add note used to build a number is to drag the number to a DDC Pages window and drop it.

The DDC Pages window displays local notes in blue before the published notes in black; the DDC Number window does not display local notes as part of the record, but has an icon that can be double-clicked to activate a separate window that displays local notes. Classifiers who make much use of local notes will favor the Pages display for notes that they want to see every time they view a record and that are short enough not to crowd much other information off the screen; they will favor the DDC Number window approach for long Manual-like notes that they need to consult only occasionally.

Both the DDC Pages and the DDC Number windows work well with short and medium-length records, neither works particularly well with long records, especially records with long add tables. The DDC Pages display is more compact than the DDC Number display, and thus offers some advantage for records with long add tables. One wishes a comparable compact display were available for long Manual notes. In the DDC Number display but not the DDC Pages display, the word or number used in the search query is displayed in bold; a classifier can then scan the long record looking for the bolded term. For example, a classifier at 670 Manufacturing, following the reference "See Manual at 338 vs. 060, 381, 382, 670.294, 910, T1—025, T1—0294, T1—0296," might search on 670.294 so that 670.294 would be bolded as an aid to finding relevant portions of the long note. What would really help, though, is a string-searching feature to allow the classifier who has retrieved a long record to search for words or numbers within that record.

A classifier can easily take advantage of both the DDC Number and DDC Pages windows using the Browse view, which has a quarter-screen Search window, a quarter-screen DDC Number window, and a half-screen DDC Pages window (tiled vertically) (see figure 2). The quarter-screen DDC Number window rarely shows everything in the record, but it is easy to scroll down or to maximize the window for a quick check of information. For example, in the lower left comer of figure 2, the record for 341.754 has been scrolled down to show the DOC Index Terms, which identify the number as an international law number. A classifier who strongly favors the DDC Number window will prefer the Search view, which has two windows tiled vertically, Search and DDC Number. With a half-screen, vertically tiled DDC Number window, one rarely needs to scroll down or maximize.

For complicated number building, a classifier may customize one of the four user views to serve as a number-building view. An approach that works well is a tiled view with one Search window and either three DDC Number windows, three DDC Pages windows, or some combination of DDC Number and DDC Pages windows. The key is that the classifier should be able easily to display three records at once. For example (see figure 3), when building a number for a critical study of costume in fiction, a classifier might display the add note at 809.1-.7 in one DDC Pages window (upper right), the add note at 808.839 in another DDC Pages window (lower right), and the Table 3C record — 355 in a DDC Number window (lower left). Dewey for Windows provides a Work Area at the bottom of the screen where numbers can be built, then copied into a local annotation note or copied to the Windows clipboard and thence into a bibliographic record. After the first part of a number has been built, the relevant add notes technically no longer need to be on the screen; however, if they go off screen and then the classifier wants to review the number, checking to be sure that it has been built properly, then the Past feature must be used to bring them back.

A classifier can easily use different views in succession while classifying one work. For example, for a work that treats both the public finance and macroeconomic aspects of taxes, a classifier can begin with a scan of the Relative Index (Phrases) (see figure 1 again). A useful strategy is to select one entry from the Relative Index (Phrase) display (e.g., Taxes 336.2,
Fig. 2: Browse view

Fig. 3: Customized view being used to build number 809.39355
which the Relative Index presents as both the interdisciplinary and public finances number for taxes), drag the entry to the DDC Number window and drop it to see the full record; then select another entry (e.g., Taxes—macroeconomic policy 339.525) and either click on the Search button in the Index window or drag the entry to the Search window and drop it there. The classifier then switches to another of the views that contain a Search window and a DDC Number window (all the built-in views contain both windows). The information in the Search and DDC Number windows is carried over to the corresponding windows in the other view, but the Index window no longer fills half the screen. With the tax example, if the classifier tentatively plans to use 336.2 but would still like to check the full record for 339.525, one approach is to switch to the Browse view, drag 336.2 into the DDC Pages window to see it in the context of neighboring records, then drag 339.525 to the DDC Number window to see the full record.

6. Following References

In Dewey for Windows the primary device for following references is drag and drop. In an Index window with the Relative Index (Phrases) index, one highlights the line with SA (e.g., "Melanoma SA Cancer"), drags it to the Scan button, and drops it to go to the referred-to portion of the Relative Index. In a DDC Number or DDC Pages window, one highlights the referred-to Dewey number and drags the resulting suitcase icon to the place that will produce the appropriate response. If the reference is an add note referring to a record with a long add table (e.g., "Add as instructed under 930-990"), the classifier may want to drop the icon in a DDC Pages window to get the most compact display of the long add note. For other references to a schedule or table number, a classifier may display the record in a DDC Number or DDC Pages window, may drag the number first to a hierarchy window to see its hierarchical context, or may drag the number to the input box of the Search window and add an asterisk to create a list showing all the number's subdivisions. In short, the drag-and-drop approach to following references makes many choices available to the classifier.

The drag-and-drop approach, however, generally requires more effort than clicking on a hypertext link. When a classifier simply wants to look at the other number with the least effort, a hypertext link would be preferable. It is hoped that hypertext-like links can be added for those references where links are most likely to be desired, without obstructing any of the flexibility made possible by the drag-and-drop approach.

7. Conclusion

The strength of Dewey for Windows is that it offers many different ways to approach the problem of classifying works. Classifiers can select approaches depending on the nature of the topic, on their own familiarity or unfamiliarity with a particular part of the schedule, and on their own individual preferences.

References


Vizine-Goetz, Diane, and Mark Bendig. (1996). Dewey for Windows: accessing the Dewey decimal classification from the technical services workstation. To be included in a work tentatively titled *Planning and implementing technical service workstations* that will be published by the American Library Association.