

and releasing the mouse button. This causes the system (after the user selects the command) to pull in all documents into the pile controlled by this script which contain "design competition update" in the regarding line/field of the document. In this manner, the system may automatically classify information for the user and the user may instruct each pile separately according to the needs and criteria of the user or the basis of a user defined specification.

Region 611 of the script window 601 contains a list provided by the system which indicates the most unique used words in the particular document from Richard Mander (i.e. the document which was added to the pile and then caused the system to display the script window 601 on the display screen 22). As will be described below, the system can provide a ranked list of the most frequently used words in a document which also best characterize the document relative to all other documents in the file system of the computer. This is done by the use of, in one implementation, the vector described below for the document. In this implementation, each word listed in region 611 is one component of an n dimensional vector which describes the document and provides a representation of the contents of the document. As shown in script window 601, the word "design" is the most characteristic word followed by "competition" and "interface". The system provides this listing in region 611 and the user may scroll through the various words in this listing by selecting the scroll arrows 627 and 628 or by using the scroll bar 629. The scrolling of the contents of region 611 is accomplished in the normal manner of the prior art. The user may select any one or all or some of the words in region 611. When these words are selected, they appear in region 612 indicating to the system that the pile's script should be modified such that the representation of the pile (e.g. the pile's vector), as described below, is modified to take into account the

47

words selected by the user. Alternatively, the user may type words into region 612 to thereby modify the script of the pile. In this manner, the script of the pile is modified to take into account the new document from Richard Mander which was added to the pile by the user before the script window 601 was displayed.

5 For example, the user could select "competition" and the word "children" and these words would then appear in region 612; when the user selects "add to script" button 620 this causes the system to update the script for the pile causing the internal representation of the pile to be modified in order to allow the user to control the way the script behaves in the computer system. The selection of the  
10 words in region 611 may be accomplished in any of a number ways which are familiar to user's of computer systems having graphical user interfaces (e.g. positioning the cursor over a word and selecting the word or "shift-clicking" to select multiple words).

15 In a similar manner, the script of the pile may be modified to take into account labels which specify certain criteria such as label 614 shown in region 630. These labels contain certain "canned" criteria specified by the user and the user may select these labels by positioning the cursor over the particular label and selecting the label causing the icon of that label to appear in region

20 617. These labels may be the programmable user interface elements  
*now U.S. Patent No. 5,120,828,*  
described in the application having serial number 07700,729 which was filed  
May 15, 1991 and is entitled "A User Interface System Having Programmable  
User Interface Elements" and is assigned to the assignee of this application.  
These labels provide a programmable means for causing the execution of a  
2.5 series of instructions and may include, for example, a requirement that all documents in a pile be dated after a certain date.

48

Once the user has determined and selected the various criteria which are to be used to modify the internal description of the pile (pile's script) then the user may indicate to the system that the script should be modified by selecting the "add to script" button 620 and then by selecting the okay button 623 which  
5 causes the execution of the script wherein the system checks for documents managed by the filing system which match the modified script. If the user does not like certain criteria which have been previously selected, the user may cancel the selection by selecting the cancel button 625. If the user selects the "script..." button 621 the system provides a window for the user to type in words  
1 0 or other criteria which will then be used to modify the script of the pile. This modified script may then be selected by selecting the add to script button 620 and okay button 623. If the user has requested a script window similar to window 601 while not adding to a pile, then the user desires to modify the script of the pile; this is accomplished by entering commands into a script window in a  
1 5 manner similar to that described for window 601.

The manner in which the filing system of the computer takes into account the representation of each document and each pile and performs the various operations with piles which have been described above, will now be described  
2 0 by referring to the various flow charts of Figures 15 through 21.

Figure 15 shows one method according to the present invention for creating the internal representation used by the filing system in organizing piles, which internal representation is for each document in the filing system (or for a  
2 5 certain selected group of documents as specified by the user) and for every pile maintained in the filing system (or for piles which have been selected by the user to have their internal representation determined). Typically, the filing

system will perform this operation for every document (usually user documents, not system documents such as operating system files) and for every pile in the system. If the filing system is using the method of the invention for the first time, then the method shown in Figure 15 must be applied to every document in the system; that is, every document in the system must be treated as a new document. The method 700 for indexing documents begins at step 701 which is merely a node point in the process. The first operational step of the system involves a test for whether the document is a new document at step 702. A modified document is considered a new document as well as a new document which the user creates or which comes into the user's computer system via electronic mail. It is also noted that a document that is deleted affects the internal representation of all other documents in the system and the designer of systems using the present invention may want to eventually (e.g. in background processing on the computer system) recompute each document's vector and each pile's vector in this circumstance also. If there is a new document in the computer system processing proceeds to step 703 in which the computer system counts the number of times certain words are used in a document. Only certain words are counted because many words (e.g. the, it, a, and, etc.) are used too commonly and would not provide a good internal representation of the contents of any document. Therefore, these words (referred to as "stop words") are discarded in analyzing and indexing documents. This may be accomplished by the system establishing a "stop word" table which is examined for each word in a document which is being indexed. The system should also "stem" each word being indexed in order to prevent unnecessary redundancies between the singular and plural of a word (e.g. "horse" and "horses") and different tenses of a verb and other techniques which are known in the art may be applied for stemming the words (e.g. child and children would be considered

50

the same word). During indexing of each document, the system adds a new (non-stop) word to a table of unique words for the document or if the word is already in the document's table the system increments the count component for that word in that table. It will understood that this table will be maintained for a document during indexing and the table contains an index to each non-stop word in the document along with a count of the number of times the word appears in the document. The index typically refers to (indexes) a global document frequency table which contains a list of all non-stop words used in all documents managed by the filing system of the present invention. Also during indexing a document in step 703, the global document frequency table is updated in the following manner. That is, step 705 is typically performed at the same time as step 703 although it need not be. For each unique word in a document the system adds the word or increments the count for the word in the global document frequency table but only for the first time the unique word is found in the document. This global document frequency table is maintained for all documents in the system which are to be indexed and shows the number of documents in the system which use the particular word. If during indexing a unique (non-stop word) is found which is not in the global document frequency table then the word is added to that table but only for the first time the word is found in the document. That is, subsequent uses of that word in the same document will not cause a modification of the global document frequency table. If that same word is found in another document then, during indexing of that other document, the system increments the count for that word in the global document frequency table, thereby indicating that the word is contained in two documents.

51

After indexing a new document, step 703, processing proceeds to determining the number of documents in the system (documents which are previously indexed) which use each of the certain words in the new document. This is determined by examining the global document frequency table, which  
5 may be a hash table, for each word in the document which has been indexed. If all documents are new in the system, a special initialization process which is accomplished in two phases must be performed in order to properly index all documents in the system. In the first phase, word frequencies are computed for all the documents; in other words, step 703 is performed for each document in a  
10 system and then the global document frequency count table will then be completed. In the second phase the document vectors are calculated, which is shown as step 707, for each document in a system. Step 704 performs, in effect, this two phase initialization process by forcing step 703 to be repeated when system initialization occurs; step 703 is repeated until all documents are  
15 indexed and then the global document frequency table is completed. System initialization may be determined by searching for the global document frequency table; if it exists when a new document is indexed in step 703 then the system has been initialized and otherwise the system has not been initialized and initialization proceeds until all documents are indexed in step  
20 703.

In step 707 the internal non-normalized representation of a document is determined by calculating the document's vector. This is done by determining, for each of the non-stop words in the document the ratio of the number of times  
25 the particular word in the document is used ("tf"--obtained from the document's table) divided by the number of documents in the system with the same particular word ("df"--obtained from the global document frequency table). This

52

ratio is sometimes referred to as the term frequency (tf) of the word divided by the document frequency df of the word (tf/df). This ratio provides a weight for the particular word for this document as shown in step 709. Typically these individual weights are normalized in the manner shown below. Then the

5 system will typically sort, in step 711, the normalized weights of a particular document to select a predetermined number of strongest normalized weights. The vector for the document is the collection of the predetermined number of strongest normalized weights. In a typical embodiment, 50 to perhaps an

10 unlimited number of the strongest weights (which represent 50 to an unlimited number of the most characteristics words of the document) would be used for the document's vector. If an unlimited number of weights is used, the system will tend to operate slower. At this point, the system has the information necessary to display the proxy shown in Figure 4f (used in the case of mail documents) because the most strongest weights identify the most characteristic

15 words in the document, and it is these words which are displayed in this proxy (together with, if desired, words in preselected fields of a document, such as "To", etc.). At step 717, the system determines whether the new document is part of a pile. If it is not processing loops back to node A (701) which is a waiting node in which the system waits for a new document in the filing system

20 of the computer. If the new document is part of a pile then processing proceeds to step 719 in which the internal representation (e.g. a vector) for the pile is determined. This representation may be any one of several possible types, including the average of all document vectors in the pile (which is representative of the internal context of the pile) or the pile's representation may

25 be designated by a script as described above or designated by a sample document (i.e. the pile's vector is the same as a sample document's vector) or the pile's vector may be designated by certain key words having user

established weights. It will be appreciated that if the method 700 for indexing a document is being performed during initialization of the system then the determination of a pile vector will occur after all documents have been indexed and the global document frequency table has been completed. After step 719, processing loops back to node A wherein the system waits for an interrupt or other messages indicating a new document has been placed into the filing system of the computer.

The methods employed in steps 703, 705, 707, 709 and 711 are known in the art and various ways of determining an internal (or content) representation of a document are available in the information retrieval art. See, e.g. "Term-Weighting Approaches in Automatic Text Retrieval," G. Salton and C. Buckley, Information Processing & Management, Vol. 24, No. 5, pp. 513-523, 1988. However, these methods have not been used in the context of piles as described in this invention. In one implementation of the invention, the normalized weight for a particular word ("tw<sub>1</sub>") in a particular document is determined according to the formula:

T<sub>10540</sub>

$$tw_1 = \frac{tf_1 \left( \log \frac{N}{n_1} \right)}{\sqrt{\sum_{i=1}^a \left( tf_i \left( \log \frac{N}{n_i} \right) \right)^2}}$$

20

where:

N = total number of documents in the filing system which are selected for indexing (default selection is all user documents);

25

54

54



- 5  $n_i =$  df for the word "i" (df is document frequency for word "i" from the global document frequency table)
- $tf_i =$  term frequency for the "ith" non-stop word in the particular document
- 10  $a =$  number of unique non-stop words in the particular document (i.e. the number of components in the vector for the particular document)
- 15  $n_1 =$  the df for the particular word having its normalized weight being determined
- 20  $tf_1 =$  tf (term frequency) of the particular word in the particular document;  $tw_1$  is the normalized weight for this particular word in this particular document

Other implementations may be used according to the needs of the system designer.

- 25 Figure 16 shows a method for performing the operation of browsing of a pile. From node E, the system begins at step 751, in which the system examines the location of the cursor and determines whether or not it is pointing to a pile. It will be appreciated that the system will typically monitor the position of the cursor in any one of the many ways known to those in the art. If the cursor
- 30 is not pointing to the pile processing loops back to node E. If the cursor is

pointing to a pile then processing proceeds to step 752 in which the computer determines whether the cursor has continued to point to a document on the pile for a predetermined period of time. If not (e.g. cursor moved away from pile) processing loops back to node E in which the system continues to monitor the position of the cursor. If the cursor has pointed to a document on a pile for a predetermined period of time then processing proceeds to step 753 in which the system arranges the pile for viewing. This step may involve nothing or it may involve shifting a portion of the pile as shown in Figure 4e. Processing then proceeds to step 755 in which the view cone is shown and a proxy of the selected document in the pile is displayed. Then the system continues to monitor the position of the cursor to determine whether the user has pointed to another item in the pile (step 757). If this is true, processing loops back to step 753 to arrange the pile for viewing (e.g. pile is re-arranged to reflect a new shifted portion) and processing continues to step 755 in which the view cone continues to be displayed and a new proxy of the new selected document is also displayed in the view cone. If the same document is still being displayed because the pointer has not moved away from the document processing proceeds from step 757 to step 758 in which the computer determines whether or not the user has signalled to the computer to page through the selected proxy to the extent the proxy has multiple pages. If not, then processing goes to step 760; otherwise processing proceeds to step 759 in which the system shows the desired page of the multiple page proxy in the view cone on the display screen. Since the computer continues to monitor the position of the cursor and the status of the mouse's button, the computer system continues to determine whether the cursor is pointing to the pile (step 760); if not processing loops back to node E and otherwise processing loops back to step 755.

54

As noted above, the system provides various automatic or user invoked actions on piles in order to assist the user in organizing the documents stored in the user's filing system on the computer. The various ways in which the system of the present invention performs these functions will now be described with reference to Figure 17. Figure 17 shows three distinct processing flows from three nodes (G, G1 and G2). It will be appreciated by those in the art that any of these nodes may be entered into on the basis of an interrupt provided to the system or some other user action so that entry may occur into the flow chart of Figure 17 at any node point, such as node point G, G1, or G2 depending on the action by the user. Processing begins from node G at step 801 which determines whether or not the user has given a document to the system to file in existing piles. This may occur by the user providing a "put away" instruction to the filing system after selecting a document to be put away. If this occurs, then processing proceeds to step 803 in which the system attempts to find a pile with the closest pile vector to the document's vector and places that document in that pile. This is typically done in such a manner that the similarity between the content of the document and content of each pile is determined. In a typical embodiment this similarity is determined by computing the unit dot product of the pile's vector and the document's vector. A system or user supplied predetermined threshold may then be used to compare against each dot product and the highest dot product above the threshold will indicate which pile to place the document into. Alternatively, the document (copies thereof) or aliases (pointers) thereof may be placed into all piles which produce dot products with the document's vector which exceed the preselected threshold. It should be noted that step 803 could be performed automatically by the system for any new or modified document if the user so instructs the computer system to perform this function automatically. That is, the user may

instruct the filing system on the computer to always be active in putting away documents into piles rather than doing this only when specifically invoked by the user (e.g. a "put away" command). An example of a method for determining a measure of similarity will now be provided. Consider a document vector and

5 a pile vector shown below:

<u>Document Vector</u>		<u>Pile Vector</u>	
<u>index</u>	<u>tw</u>	<u>index</u>	<u>tw</u>
5	0.1	5	0.2
27	0.2	16	0.17
503	0.15	257	0.17
600	-	-	-
-	-	-	-
-	-	-	-

T10580  
10

Each index may be a pointer/index to an entry to the global document

1 5 frequency table where the index "5" is for one word in the table and index "27" is for another word in that table. Note that in this example there is only one word (non-stop word) which is common between the document's vector and the pile's vector. The dot product, which is a way to measure the similarity between the content of the document and the content of the pile, for this example is

2 0 determined by searching for matches in the indexes (e.g. index 5) and multiplying the tw's for each matched index. In this case, the dot product is:  $0.1 \times 0.2 = 0.02$  since there is only one match. This dot product may then be compared to other dot products showing the similarity between the document and other piles.

2 5

Step 805 is performed any time after entry from node g1. In step 805, the system determines whether or not the user has given the system a sample

58

document which should be used in order to start a pile. If not, processing proceeds to node g2. If the system has been given a sample document and been instructed by the user to start a pile on the basis of that "seed" document, processing proceeds to step 807 in which the system uses the vector of the  
5 sample document to find all documents which are closer than the preselected threshold based on the internal representation of each document, such as the document's vector. Copies or aliases of such documents are stored in the pile which was created from the sample document.

1 0 Step 809 may occur from an interrupt to the system or a command to the system; this interrupt may occur from the action of adding a document to the pile when the user has instructed the system to monitor the addition of documents to a pile having a script where the internal representation of the document does not match the criteria of the script of the pile. As an alternate embodiment, the  
1 5 user may instruct the computer to observe the "cohesion" of each pile (whether or not the pile has a script) and indicate a warning to the user that a newly added document to a pile does not match in a close enough manner (determined by a similarity measurement being compared to a preselected threshold), which warning gives the user an opportunity to remove the  
2 0 document from the pile. It is noted that the user could actually instruct the computer to ignore any such steps which in effect results in an automatic "no" to the decision made at step 811. If the user has instructed the filing system to monitor the addition of documents to a pile having a script, then step 809 will be  
2 5 invoked when the user adds a document to such a pile which does not match the document. In other words, the script of the pile does not closely match the internal representation of the document. If this occurs processing proceeds to step 811 in which the system displays a dialog box or window to the user asking

the user whether the pile's script should be modified. This dialog box or window may be similar to the window 601 shown in Figure 14. If the user indicates that the pile's script is to be modified, processing proceeds to step 813 in which the system displays existing criteria and/or the script for the pile.

5 Alternatively, the script and criteria could be displayed in step 811 with a decision box requiring the user to decide whether or not to just add the document to the pile or modify the pile's script. At step 815, the system displays proposed script/criteria changes to the pile's criteria or script. This is shown in Figure 14 by the words found in the Richard Mander document which are

1 0 shown in the region 611 or by the items in region 605 such as the author of the document (from "Richard Mander"). In step 817, the user selects changes to the criteria and indicates to the computer that these changes are acceptable. This causes the system to respond in step 819 by pulling in copies or aliases of

1 5 vector. After step 819, the system proceeds back to monitoring any of the various actions/interrupts which cause entry into the processing shown in Figure 17. The use of a script has many possibilities, for example, it is noted that the script could include the use of the vector representation of the pile as well as other features which may be matched in a boolean "AND" fashion or

2 0 matched in a boolean "OR" fashion.

Figure 18a shows a method of the present invention for creating subpiles from a pile or other selected set of documents (e.g. documents within a folder). This method would be invoked in various ways by the user or perhaps under

2 5 system control automatically (e.g. as a result of a search where the user requests the search results pile to be processed into subpiles after the search). For example, the user could place a pile such as pile 550 into the visualization

5  
10  
15  
20  
25

window 551 shown in Figure 13a and then select the command pile by content causing the system to respond, as shown in Figure 13b with four subpiles created by the system. If the system determines, in step 851, that the user wants to create subpiles from a pile, then processing proceeds to step 853 in which the system determines whether there any documents which are left in the original pile. If no documents are left in the pile processing proceeds through node F2 to step 871. In the beginning of a subpiling process, typically documents will be remaining in the pile and thus processing proceeds to step 855 in which the first document is taken from the pile and is designated as the current document. Then processing proceeds to step 857 in which the system compares the similarity of the current document to each existing subpile. This can be done in many different ways as described above, including the calculation of the dot product between the vector of the document and each subpile's vector. The similarity measure, which may be the dot product described above, is compared to a threshold which may be established by the system or by the user. For example, the user may be presented with a graphical slider, the knob of which may be selected and moved along the slide to vary the threshold which will then vary the amount of subpiles which are created by the system during the process of subpiling. In the case of the first document from a pile, this test in step 859 will result automatically in the creation of a new subpile since there are no existing subpiles. After the first document in the pile, the test in step 859 will determine whether a new subpile is created, which occurs in step 861. In step 861, the new subpile is created with the current document as the seed document of the new subpile which thereby determines the internal representation of the subpile. If processing proceeded through step 861, then it loops back to step 853 as shown in Figure 18a.

41

If the result of the test in step 859 is a "yes" then processing proceeds to step 863, in which the system sorts the matching subpiles (i.e. subpiles producing a similarity measure with the document which exceed the threshold) by similarity measurement to the current document. Then in step 865, the system determines whether a document is allowed to be in more than one subpile. If not, processing proceeds to step 869 in which the current document is added to best matching (on the basis of a similarity measure) subpile. Also in step 869, the system updates the running similarity score by adding the similarity score, which may be the dot product between the document's vector and the best matching subpile's vector, to the running total of the similarity score for the entire subpiling system, which is initialized at zero prior to beginning the process shown in Figure 18a. If a current document is allowed to be in more than one subpile, then processing proceeds to step 867 in which the current document is added to all matching subpiles having a similarity measurement which exceeded the threshold. As in step 869, the running similarity score for the system is updated by adding the dot products or other similarity scores (one for each matching subpile to which the current document was added) to the running total for the system. Following step 867 or 869, processing loops back to node F1 and the various steps described above. It is noted that if the document is to be added to more than one subpile, either a copy of the document or an aliases/pointer to a copy of the document may be stored in the subpile, in a manner which is similar to the conventional prior art technique of storing a file or an aliases/pointer of a file in a subdirectory or folder.

If at any step of the process of creating a subpile, there are no documents left in the original pile, then the decision required in step 853 results in the system proceeding to step 871 in which the system determines whether there



are any subpiles which contain only a single document. If no such subpile exists then processing proceeds to node F3 which results in step 876 being performed, otherwise, processing proceeds from step 871 to step 873 in which a miscellaneous subpile is created. Then, in step 874, each document from a  
5 single document subpile is transferred to the miscellaneous pile (and the running similarity score is updated) and the single document subpiles are removed from the system in step 875. Then, in step 876, the system determines whether any improvement to the total system similarity score has occurred since the last total similarity score. This system similarity score is computed as  
10 described above in steps 867 and 869 each time a document is added to a matching subpile and after all documents have been removed from the original pile this test in step 867 is typically performed; obviously, in the first pass through this step, the system will place all documents back in the original pile by proceeding to step 877 and then back to node F1, repeating the process while  
15 maintaining in storage a record of the prior total similarity score. It will be appreciated that this total similarity score is only for documents in the original pile and only for similarity measurements between those documents and subpiles created during the subpiling process described here. The test in step 876 may include a comparison to a threshold where the difference, which  
20 should be an improvement, between the prior total system similarity score and the current total system similarity score is compared to a threshold and if that difference does not exceed the threshold, then no effective improvement has occurred and processing is completed. A designer of a system using the present invention may decide that it is efficient and reasonably accurate to  
25 merely repeat the process of subpiling only once or some fixed number of times rather than attempting to access an improvement at the end of each subpiling pass.

Figure 18b shows an alternative subpiling process according to the present invention where the user chooses the number of subpiles which are to be created from the original pile. This is typically done in order to avoid unnecessary clutter of subpiles. Without this feature, it is possible that the computer filing system of the present invention could create too many subpiles making it more difficult for the user to manage and organize the piles created in this manner. The process of Figure 18b beings at step 880 in which the system determines the number of subpiles which are allowed by the user when a subpiling process is invoked. Processing then proceeds to step 882 to determine whether the desired number of subpiles have been created so far in the subpiling process. If they have, processing proceeds through node J2 to step 891, and otherwise processing proceeds to step 884 in which the system computes the total similarity of each remaining document in the original pile to all existing subpiles. In the special case of the first document which is removed from the original pile, there is nothing to compute since no subpile exists. Then, in step 886 the remaining documents are sorted from least to most total similarity. Then in step 888, the document in the original pile with the least total similarity is removed from the original pile. In the case of the first document which is removed from the original pile, any document will do. Then, in step 890 a new subpile is created with the document which was removed in step 888, using that document as the seed for the similarity measurement comparisons (e.g. the document's vector is used for the subpile's vector). Then processing proceeds back to step 882 in which this process repeats through steps 884 through 890 unless processing proceeds through node J2 to step 891. In the second pass through steps 884 through 890, step 884 does involve the computation of the total similarity of each remaining document in the original

64

pile to all existing subpiles. This is done in the fashion described above wherein the internal representation of each document is used with the internal representation of each subpile to compute a similarity measurement between the document and each subpile. The total similarity for a document to all

5 existing subpiles is a summation of each similarity measurement between the document and all existing subpiles. Thus, for example, if four subpiles exist when the computation is made, four dot products may be computed as the similarity measurement, each dot product being between the document's vector and the seed vector of the four existing subpiles; these four dot products are

10 then added together to provide a total similarity measurement for the document. Then in step 886, these total similarity measurements for each document are sorted. The document with the least total similarity in this sorted list is removed from the pile in step 888 and a new subpile is created with this document as a seed as shown in step 890 and the processing continues back to step 882.

15 When the number of subpiles which exist by virtue of the subpiling process equals the desired number selected by the user, processing proceeds to step 891 which determines whether there are any documents left in the original pile. If the answer is no, the process is completed. However, if the answer from step 891 is yes, processing proceeds to step 893 in which the system takes the first

20 remaining document from the original pile and labels it the current document and then proceeds to step 895. In step 895, the system compares the similarity of the current document to each existing subpile and then in step 897, the system sorts the matching subpiles by similarity to the current document. In step 899, the system adds the current document to the best matching subpile and

25 proceeds back to step 891. The forgoing description of the process of subpiling from an original pile has been described with reference to a specific procedure; however, using known information retrieval technology, other implementations

65

will be available to those in the art upon reference to this description and the information retrieval art.

5 An alternative to the process of creating a subpile according to either of the methods shown in Figure 18a or 18b will be described. Prior to starting the subpiling process, the user may want to specify certain subpiles which must be created; this may be done by specifying a "seed" document which is representative of a desired subpile or a "virtual" seed document (specified by keywords which may be used to create an internal/content representation of the subpile). These desired subpiles are then used by the system as subpiles which exist immediately upon beginning the process (e.g. entering step 853) of creating subpiles from a pile. It should also be noted that the process of creating subpiles may be performed from several original piles merely by combining them into one original pile or by using one pile as an original pile and the other original piles as pre-existing subpiles.

10 The system and method of the present invention also provide a mechanism for searching documents in the file system of the computer whether those documents are in piles or conventional subdirectories/folders. This method is shown in Figure 19 and begins at step 901, in which the system determines whether the user wants to search for documents matching a set of terms which are specified by the user. Typically, the user will indicate that a search is desired and the system will respond with a question asking the user whether the user wishes to define a set of terms or to specify a sample document. It will be appreciated that, other possible means of requesting a search or formulating a search may be used. If the user wants to specify a set of terms, the system allows in step 903 the user to do so and then creates in the

step 905 a virtual document vector (or other representation of the terms) for the set of terms (and their user selected weights, unless by default the terms are equally weighted). That is, the user specifies terms and then those terms are used to specify a vector of a virtual document which is used later when

5 processing. Processing from step 905 proceeds directly to step 911. If the user wishes to search using a sample document (e.g. create a pile from a sample document manually by searching and obtaining a list of the documents or automatically create a pile containing the search results) then processing proceeds from step 901 to step 907 in which the system requests the user to

10 specify the sample document; if no sample document is specified processing loops back to node C waiting for an interrupt or other instruction indicating to begin the search process. After the sample document is specified in step 907, the system proceeds to step 911 in which it compares each document's vector (or other representation of the document) in the system with the document

15 vector (or other representation) of the sample document or the document vector (or other representation) of the virtual document created in step 905. This may occur by any of a number of ways of determining a similarity measure, including by determining the unit dot product of the sample document's vector against the vector of all documents in the system. It will be appreciated that alternative

20 similarity measures or search routines may be employed, such as measuring the similarity between the document and each pile or folder and then searching within those piles or folders having the best matches to the document. It is noted that this comparison can eliminate documents which are not close enough by testing each dot product against a preselected threshold.

25.

In step 915 the system sorts the list of documents based on the comparison with the document vector of the sample (or virtual) document. This

list may be limited by the comparison against a preselected threshold as described above. Then in step 917 the system creates and displays a pile of the sorted documents which are closest to the sample (or virtual) document. Following step 917 processing proceeds back node C which waits for a  
5 command to search from the user. Alternatively, the pile may be divided into subpiles according to the techniques of the invention. The user may also, after the search pile is created, manipulate that pile to perform other searches such as a new or revised search.

10 Figure 20 shows a method for using color to assist in organizing information in a computer and to assist a user in understanding the content and organization of piles within the following system. This method begins in step 951 wherein the user selects a pile and then in step 953 selects a method of visualization such as the visualization window 551 shown in Figure 13a. This  
15 causes the system, in step 955, and display the visualization window showing the pile within the window (or the user may have to move the pile into the window). Then in step 957, the user selects the command "order by date" and the system orders documents in the pile by the date of the document. It is appreciated that step 957 is a two-part process even though it is shown as one.  
20 This may be done in any of the well known ways of sorting and listing documents by the date of the document. This date may be either the date maintained by the filing system of the document or may be a date line (field) of the document which is contained in the contents of the document and was provided by the author of the document. After step 957, the user in step 959  
25 selects the command "color by date" and the system displays documents with different colors and organized according to the date. It is appreciated that step 959 is a two-part process even though it is shown as one. That is, the color of

68

the document is correlated with the age of the document in the manner described above. This is done typically by determining the range of dates for the entire pile by determining the oldest document and the most recent document (and thereby determining the range) and thereby determining the

5 range of the variable saturation for a particular color (e.g. 256 different saturation values for a given luminance and hue of a color). The oldest document would then be assigned the most saturated color (e.g. saturation level equal to 255 in an 8 bit color system) and the most recent document would have the least saturation (e.g. saturation level equal to zero for the same

10 luminance and hue of a color). A document with a date in the middle of the date range would have half of the saturation of the oldest document and other documents would be proportioned accordingly. Other well known techniques may be used for assigning colors on the basis of the dates of the various documents in the pile. The user may next select "color by content" in step 961

15 in the manner described above in conjunction with Figure 13a. This causes the system in step 963 to display the documents in the pile with different saturations of color according to the closeness (determined by a similarity measure) of the document to the internal representation of the pile. This may be done by comparing the document's vector to the pile's vector by the using the unit dot

20 product of the two vectors in the manner described above. The computer would determine all dot products for all documents in the pile and then sort those dot products from highest to smallest. The highest dot product indicates a document which most closely matches the pile's internal representation (e.g. pile's vector) and the smallest dot product would indicate a document which is

25 farthest away from the pile's internal representation. This sorted list of dot products corresponds to a sorted list of documents since each dot product is associated with a particular document and thus different saturations of a given

color at the same luminance and hue may be supplied to the different documents according to their dot product. In other words, the dot product for a particular document maps to a particular saturation of a color the same way that the date maps to a color as described above. The user may desire that this pile  
5 retain its colored form based on coloring by content by selecting the okay button 571 and then by closing the visualization tool window. This will cause the new colored pile to appear on the display means 22 in the filing system of the present invention. It is noted that the order of many of the steps shown in Figure 20 is not critical to the invention and these steps may be practiced in other  
10 sequences which are not shown.

In order to summarize many of the features of the present, a typical example will be given of how a user may interact with piles. Figure 21 represents a flow chart showing this typical example; the order of many of these  
15 steps is not critical to the invention and they may be practiced in other sequences which are not shown. Beginning in step 1001, the user creates a pile by moving one document on top of another. This has been described above in connection with Figure 3 and other figures. The user may also create a pile by selecting two different documents and selecting a command available  
20 in the computer system to create a pile. Next the user in step 1003 moves the newly created pile by positioning the cursor over the base of the pile and signaling the selection of the pile and then moving the cursor to move the pile. In step 1005 the user selects another pile by positioning the cursor over the base of the pile and signalling a selection of that pile. In step 1007, the user  
25 opens the pile by any one of the mechanisms for obtaining an alternate view as described above. For example, by rapidly depressing and releasing (double clicking) the mouse button while positioning the cursor over the base of the pile.



The system responds in step 1009 to the actions of the user in step 1007 by opening a pile window with the contents of the pile displayed in that window or alternatively by spreading out the documents in the form shown in Figure 8d. An example of a pile window containing the contents of the pile is shown Figure 5 8a.

In step 1011, the user may close the pile window by positioning the cursor over the close box and signalling a selection or by using any one of the other known techniques for closing a window in a graphical user interface.

1 0 Then in step 1012, the user selects a document which is not in a pile and drags it to the top of an existing pile to add it to the pile. This is typically done in one of many ways for moving a document in a filing system having a graphical user interface, such as positioning the cursor over the icon of the document and selecting the document and keeping the document selected while the cursor is

1 5 moved thereby moving the cursor and the document on the display screen. The system's response to step 1012 is step 1015 in which the system displays a script dialog box since the targeted pile to which the document was added is a scripted pile and since the new document has an internal representation which does not match the current criteria of the pile. The user selects in step 1015 to

2 0 modify the pile's script. This takes place in step 1017 in which the system modifies the pile's script on the basis of the user's input. For example, the user specifies particular criteria to use in modifying the pile's script such as selecting two unique words from the added document or by using a particular label such as the label 614 shown in Figure 14.

2 5

Then in step 1019, the user instructs the system to add documents to the pile on the basis of the pile's modified script or the system does so automatically

71

after modifying the pile's script. Whether automatic or upon user instruction, the system adds to the pile on the basis of the new script in step 1019. The computer may display an animation of document icons moving from their old locations on the display screen into the pile.

5

In step 1021, the user instructs the system to create subpiles of a selected pile based on the contents of the documents. That is, the system is instructed to create subpiles on the basis of the contents (rather than for example the dates of the documents or other criteria upon which subpiling may be based). The system responds in step 1023 by examining the contents of each document in the selected pile and creating several subpiles from the original selected pile, each subpile containing similar content because of the operation of the present invention. The system as shown in step 1023 then assigns a name to each subpile on the basis of the words having the strongest weights for each subpile.

The user may then select, in step 1025, one of the subpiles by positioning the cursor over the base of the new pile and signalling a selection. The user then instructs the system to color the pile by date by selecting the command "color by date." This causes the system to respond in step 1027 by using the date of documents in the subpile to select a saturation of a color which is dependent on the date of the document. For example, the most saturated color represents the oldest document. The user in step 1029 then removes a document from this subpile by selecting the document and dragging it out of the pile. For example, the user may select the oldest document, based on the saturation color and remove it from the pile and examine the document. In step 1031, the user selects a document which is not in any pile and selects the

command "put away" which causes the system to put the document into the closest, in relevancy, pile or piles depending upon user preference.

Another example of a user's interaction with piles according to the present invention will now be described by referring to Figures 22a, 22b, 22c, 22d and 22e. As shown in these figures, the method and apparatus of the present invention may be applied to an object oriented programming environment where the use of piles is made for programming documents which are not the typical text/based office document, even though programming files contain text. As shown in Figure 22a, an object finder window 2201 includes three components which are the find box 2202, the results box 2204 and the discard box 2203. In this example, a search is specified by placing a sample document or a pile in the find box 2202 and the system will find objects that are similar to the objects in the find box 2202 provided they are not present in the discard box 2203. That is, the discard box 2203 acts as a "not" boolean logic search parameter. The results window or box 2204 shows the results of a search which have been automatically placed in three subpiles according to the present invention where the original search results were placed in a pile which was then divided into three subpiles 2210, 2211, and 2212.

2 0

Figure 22b shows a method and apparatus of the present invention where piles may be used to specify the search query rather than words or other query techniques. In this case, the find box 2202 contains the piles 2219 and 2220 which specify the search query (e.g. by using either internal representation of either pile as the search query). That is, the internal representation of pile 2219 is OR'ed together with the internal representation of the pile 2220 so that objects which are similar to either internal representation

2 5

of either pile will show up in the results window 2204. As shown in Figure 22b, the discard box 2203 contains the pile 2218 which is used to discard objects which are found in the search which match the internal representation of the pile 2218. It is important to realize that the descriptions for the search query are "fuzzy" in that they are specified by the pile's internal representation rather than a specific list of key words which are connected by boolean logical operators. The results of the search are shown in the box 2204.

Figure 22c shows another use of the object finder 2201 where attributes of objects 2227, 2228, and 2229 are used to search the computer's filing system for documents and other objects which match the search criteria or have those attributes. These objects may be piles or other objects which may be treated as a text document and from which an internal representation of the document may be derived.

Figures 22d and 22e show a use of the invention in connection with media objects which represent digitized movies. Each digitized movie ("clip") have all been indexed by key words when they were created or when they were edited. The user has placed two clips (digitized movies) into the find box 2202 in order to find similar clips on the basis of the text/keywords for each clip. The user instructs the computer to find such clips on the basis of either clip 2231 or clip 2232 within the find box 2202. The system responds displaying four piles 2233, 2234, 2235 and 2236 which match the internal description (or exceed a threshold set for comparison purposes against a similarity measure). In this case, the four piles shown in the results window 2204 of Figure 22d were created by subpiling the original search results pile created from the find operation. The user may the browse one of the piles as shown in Figure 22e to

74



and method of the present invention without the departing from the spirit and scope of the invention as claimed below.

We Claim:

Sub A2

5  
10  
15  
20  
25

1. A method for organizing information in a computer filing system having a display device and a first plurality of documents, said method comprising :
  - providing a graphical representation of a first document and a graphical representation of a second document from said first plurality of documents;
  - creating a collection of documents comprising a second plurality of documents having said first document and said second document;
  - displaying a graphical representation of said collection of documents on said display device;
  - viewing at least one of the documents in said collection by displaying an indicia of said at least one document.
2. A method as in claim 1 further comprising the step of displaying a base for said collection of documents and wherein said collection of documents is created by one of said user and said computer filing system.
3. A method as in claim 2 further comprising a step of selecting said collection of documents for further action by positioning said cursor over said base and by indicating to the computer system that said collection of documents has been selected.
4. A method as in claim 1 further comprising the step of:
  - creating a plurality of collections of documents from said collection of documents by comparing an internal representation maintained by said

77

computer system for each document in said collection of documents with an internal representation created by said system for each of said collections of documents.

5           5.     A method as in claim 1 wherein the step of creating said collection of documents comprises positioning said cursor over said graphical representation of said first document and selecting said first document and moving said first document such that said first document overlaps said second document and then signalling to said computer system that a collection of  
10 documents is to be created.

15           6.     A method as in claim 1 wherein the step of creating said collection of documents comprises selecting both said first and said second documents and signalling to said computer system that a collection of documents is to be created.

20           7.     A method as in claim 1 further comprising creating a representation of said collection of documents based on an internal representation of each of the documents in said collection of documents.

25           8.     A method as in claim 7 wherein said computer filing system determines said representation of said collection of documents and wherein said internal representation of each of the documents in said collection of documents comprises a representation of words within each of the documents.

          9.     A method as in claim 8 wherein said internal representation of each of the documents comprises, for a particular one of the documents, a



c  
vector containing a representation of words within said particular one of the documents and a representation of the uniqueness of the words in said particular one of the documents relative to the words <sup>in</sup> said first plurality of documents.

5

*add c4* 10. A method as in claim 8 wherein said representation of said collection of documents comprises a vector containing a representation of words within at least some of the documents in said collection of documents.

10

11. A method as in claim 1 further comprising viewing the contents of said collection of documents by displaying a graphical representation of each document in said collection within a region on said display device.

15

12. A method as in claim 3 further comprising viewing the contents of said collection of documents by selecting said collection and signalling to the computer system to display a graphical representation of each document in said collection within a region on said display device.

20

13. A method as in claim 1 further comprising viewing the contents of said collection of documents by displaying in an overlapping manner the graphical representations of each document.

25

14. A method as in claim 1 wherein said step of creating a collection of documents comprises creating a base for said collection and selecting said first and said second documents and signalling to said computer system that said first and said second documents are in said collection.

15. A method as in claim 1 further comprising adding a third document having a graphical representation to said collection by selecting said third document and moving said third document so as to position said graphical representation of said third document over said graphical representation of said collection and signalling to said computer system to add said third document to said collection.

16. A method as in claim 15 further comprising removing said first document from said collection by positioning said cursor over said graphical representation of said first document and selecting said first document and moving said graphical representation of said first document away from said graphical representation of said collection.

17. A method as in claim 1 wherein said step of viewing comprises pointing a cursor on said display device at a graphical representation of one of said documents in said collection for a predetermined period of time to reveal an indicia of said document in said collection.

18. A method as in claim 17 wherein said graphical representation of said collection is displayed on said display device while said indicia is displayed on said display device.

19. A method as in claim 18 wherein a means for indicating the location of said document within the graphical representation of said collection is displayed when said indicia is displayed.

20. A method as in claim 19 wherein said means for indicating the location comprises a graphical representation having a cone shape which has an apex and a base, said apex being adjacent to the location of said document.

5 21. A method as in claim 17 wherein said collection of documents comprises mail documents which originated from an electronic mail message and wherein said indicia for each of said mail documents comprises a list of words in each of said mail documents.

10 22. A method as in claim 17 wherein said indicia is a miniature of said document.

23. A method as in claim 17 wherein said indicia is a reproduction of said document.

15 24. A method as in claim 17 wherein said indicia has multiple pages for a multiple page document and wherein the user views the multiple pages of said indicia by signalling to the computer to move from page to page of said indicia.

20 25. A method as in claim 21 wherein indicia for documents in said collection other than said mail documents include miniatures of said documents other than said mail documents.

25 26. A method as in claim 21 further comprising creating a representation of said collection of documents based on an internal representation of each of the documents in said collection and wherein said list

of words for each mail document is obtained from said internal representation of each mail document.

5 27. A method as in claim 26 wherein said internal representation of each of the documents in said collection comprises, for a particular one of the documents, a vector containing a representation of words within said particular one of the documents and a representation of the uniqueness of words in said particular one of the documents relative to the words in said first plurality of documents.

10

15 28. A method as in claim 15 further comprising providing a first zone of a second zone on said graphical representation of said collection and wherein when said graphical representation of said third document is positioned over said first zone, the user of said computer system controls the orientation of the graphical representation of said third document relative to said collection so that the user controls the appearance of said graphical representation of said collection to the extent of the orientation of said graphical representation of said third document relative to said graphical representation of said collection, and wherein when said graphical representation of said third document is  
20 positioned over said second zone, the computer systems controls the orientation of the graphical representation of said third document relative to said collection so that the computer system controls the appearance of said graphical representation of said collection.

25 29. A method as in claim 1 wherein said graphical representation of said collection conveys visual information about the contents of said collection.

30. A method as in claim 29 wherein said graphical representation of said collection is dynamic such that adding or removing documents to said collection changes the graphical representation of said collection.

5  
Sub C 31. A method as in claim 1 wherein said step of creating a collection comprises indicating to said computer system that a collection of documents is to be created using said first document as a sample document and wherein said computer system searches said first plurality of documents to find said second document.

10

32. A method as in claim 31 wherein said computer filing system creates said collection by modifying an attribute of documents within said collection, wherein said attribute is used by said computer filing system for said first document and said second document.

15

33. A method as in claim 31 further comprising creating a representation of said collection of documents based on an internal representation of said first document.

20

34. A method as in claim 8 wherein said graphical representation is in multiple colors such that said computer filing system provides a color for a particular document in said collection based on a measure of the similarity between the representation of said collection and the internal representation of said particular document.

25

35. A method as in claim 34 wherein one of the hue and saturation of a color is varied according to said measure of the similarity.

*Sub C2*

36. A method as in claim 7 wherein the user of said computer filing system provides a third document to said filing system and instructs said computer filing system to file said third document and wherein said computer filing system determines whether to add said third document to said collection based on a measure of similarity between the representation of said collection and an internal representation of said third document.

*Sub A3 C8*

37. A method for organizing information in a computer system having a display device, said method comprising:  
providing a representation of a first document and a representation of a second document;  
creating a collection of documents comprising said first and said second documents;  
displaying a representation of said collection of documents;  
determining a representation of said collection when a third document is added to said collection based on one of: (a) an internal representation of each of the documents in said collection, and (b) a user defined specification.

*Sub C9*

38. A method as in claim 37 wherein said computer system has a filing system having a graphical user interface and said method is performed in part by said filing system and wherein said representations of said first document, said second document, and said collection of documents are graphical representations.

39. A method as in claim 38 wherein said user defined specification is a sample document designated by said user to act as an internal representation of said collection.

5 40. A method as in claim 38 wherein said user defined specification is a collection of words specified by said user.

10 41. A method as in claim 38 wherein internal representations of documents in said collection are used to determine relevancy comparisons between said collection of documents and a third document added by one of said user and said computer system.

15 *Sub C10* 42. A method as in claim 38 wherein said user defined specification comprises a programmable means for causing the execution of a series of instructions and wherein said programmable means is activated by selecting a graphical representation of said user defined specification on said display device.

20 43. A method as in claim 38 wherein said step of determining an internal representation of said collection occurs each time a document is added to or removed from said collection.

25 44. A method as in claim 43 wherein said step of determining a representation of said collection occurs each time a modified document is stored to replace a preexisting document in said collection.

45. A method as in claim 38 further comprising viewing at least one of the documents in said collection by positioning a cursor provided by said computer system on said display device on a graphical representation of said at least one document in said collection for a predetermined period of time to  
5 reveal an indicia of said document in said collection.

46. A method as in claim 45 wherein said graphical representation of said collection is displayed on said display device while said indicia is displayed on said display device and wherein said indicia is a graphical  
10 representation of said document.

47. A method as in claim 46 wherein a means for indicating the location of said document within the graphical representation of said collection is displayed when said indicia is displayed.  
15

48. A method as in claim 45 wherein said computer filing system provides a call to the program which created said at least one document and said program responds to said computer filing system by providing said indicia.

49. A method as in claim 38 wherein said graphical representation of said collection includes a means for selecting said collection.  
20

50. A method as in claim 49 further comprising moving said collection by selecting said collection and indicating to said computer system to move said  
25 collection.



51. A method as in claim 50 wherein said means for selecting comprises at least one of a base and a top means and a wrapping means for identifying the documents within said collection as a collection and wherein said step of moving comprises pointing a cursor provided by said computer system on said display device at said means for selecting and selecting said collection and moving said cursor while said collection is selected.

52. A method as in claim 45 further comprising selecting a selective viewing mode such that only a user specified type of document is selected for viewing wherein only indicia for said user specified type of document is displayed when viewing documents in said collection during said selective viewing mode.

53. A method as in claim 52 wherein said user specified type of document is one of (i) documents containing pictures, (ii) electronic mail documents, (iii) documents containing a "To" field, and (iv) documents containing a chart.

54. A method as in claim 38 further comprising viewing at least one of the documents in said collection by selecting said at least one document to reveal an indicia of said document in said collection.

55. A method as in claim 54 further comprising moving said graphical representation of said document to a location near said collection, said step of moving being performed by said computer system after selecting said document.

56. A method as in claim 55 wherein said step of moving is an animation provided by said computer system and wherein said animation includes flipping said graphical representation of said document.

5 57. A method as in claim 55 wherein each time the user selects a document for viewing from said collection, a graphical representation of said document so selected moves to said location.

10 58. A method as in claim 55 wherein said indicia is a full size reproduction of said document, said indicia being displayed behind the graphical representation of said collection and the graphical representation of the document at said location.

*Sub C''* → 15 59. A method as in claim 38 wherein said representation of said collection is a user defined specification and said user defined specification is modified and wherein said computer system adds new documents to said collection on the basis of said modified user defined specification.

20 60. A method as in claim 59 wherein said computer system adds to said collection on the basis of a comparison between the internal representation of each of said new documents and said modified user defined specification.

25 61. A method as in claim 38 wherein said representation of said collection is said user defined specification and wherein the user indicates to said computer system to add a fourth item to said collection, which fourth item has an internal representation which does not match said user defined

specification, and wherein, after the user has indicated to add said fourth item to said collection, said computer system prompts the user to determine whether the fourth item is to be added to said collection without modifying said user defined specification or the fourth item is to be added and said user defined specification is to be modified.

62. A method as in claim 61 wherein said computer system prompts the user by displaying a window means for editing the user defined specification and wherein said fourth item is one of a document, a pile and a folder.

63. A method as in claim 38 further comprising creating a plurality of collections of documents from said collection of documents by comparing an internal representation maintained by said computer system for each document in said collection with an internal representation created by said computer system for each of said collections of document.

64. A method as in claim 38 further comprising displaying a window means having user commands for viewing said collection.

65. A method as in claim 63 wherein said collection of documents is a subdirectory of documents and wherein said graphical representation of said collection is one of a folder and a container representation means for indicating that said container representation means contains documents.

Sub A  
C12

5 66. A computer filing system for organizing information in a computer system having a processor, a bus, and memory for storing information including a plurality of documents, said computer filing system comprising:  
a display means for producing a display of graphical  
5 representations, said display means coupled to said processor;  
a cursor control means coupled to said processor, said cursor control means for controlling the position of a cursor on said display;  
a switch means for indicating a selection of an object displayed on  
said display, said switch means coupled to said processor and having a first  
10 and a second position;  
a graphical representation of a first document and a graphical representation of a second document, said cursor control means and said switch means being used with said graphical representations of said first and said second documents to create a collection of documents comprising said first  
15 and said second documents;  
a means for creating a graphical representation of said collection of documents comprising said first and said second document; and  
a means for determining a representation of said collection, said means for determining providing a representation based on an internal  
20 representation of said first and said second documents.

25 67. An apparatus as in claim 66 wherein said collection of documents has user manipulatable graphical user interface features allowing said user to operate on said collection as a group of documents.

*Sub C<sup>13</sup>*

68. An apparatus as in claim 66 further comprising a means for determining a representation of said collection each time a document is added to said collection.

5 69. An apparatus as in claim 68 wherein said means for determining includes a table showing the number of times a preselected group of words is used at least once in each of said documents stored in said computer system.

10 70. A method as in claim 1 wherein said graphical representation is in multiple colors such that said computer system provides a color for a particular document based on an attribute of said particular document.

71. A method as in claim 70 wherein one of the color parameters in a color space is varied according to said attribute.

15

*Sub C<sup>14</sup>*

72. A method as in claim 1 wherein said step of creating a collection of documents comprises creating a means for selecting a collection of documents and wherein said means for selecting includes one of a representation of said collection and a user defined specification.

*Add C<sup>15</sup>*

*Add D<sup>2</sup>*



08/289,108

ABSTRACT OF THE DISCLOSURE

A method and apparatus for organizing information in a computer filing system. The method and apparatus include the creation of a pile comprising a collection of documents and displaying a graphical representation of the collection of documents. The method and apparatus further include browsing the collection of documents by pointing a cursor at a particular item in the collection of documents to reveal an indicia for the particular item in the collection of documents. The filing system can automatically divide a pile (e.g. a collection of documents from an electronic mail network) into subpiles on the basis of the content of each document in the pile, and the filing system, at the user's request can automatically file away documents into existing piles in the computer system on the basis of a similarity match between the content (or other internal representation) of the document and the content (or other internal representation) of existing piles in the computer system. The filing system can also create a pile from a sample document by using the internal representation of the document as the internal representation of the new pile. The computer filing system provides various interfaces in connection with piles to the user of the system to provide feedback and other information to the user, including information concerning the documents and piles in the computer's filing system.

Attorney's Docket No.: 04860.P624

Patent

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below, next to my name.

I believe I am the original, first, and sole inventor (if only one name is listed below) or an original, first, and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

METHOD AND APPARATUS FOR ORGANIZING INFORMATION IN A COMPUTER SYSTEM  
the specification of which

X  is attached hereto.  
\_\_\_\_\_ was filed on \_\_\_\_\_ as  
Application Serial No. \_\_\_\_\_  
and was amended on \_\_\_\_\_  
(if applicable)

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above. I do not know and do not believe that the same was ever known or used in the United States of America before my invention thereof, or patented or described in any printed publication in any country before my invention thereof or more than one year prior to this application, that the same was not in public use or on sale in the United States of America more than one year prior to this application, and that the invention has not been patented or made the subject of an inventor's certificate issued before the date of this application in any country foreign to the United States of America on an application filed by me or my legal representatives or assigns more than twelve months (for a utility patent application) or six months (for a design patent application) prior to this application.

I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119, of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

<u>Prior Foreign Application(s)</u>			<u>Priority Claimed</u>	
<u>(Number)</u>	<u>(Country)</u>	<u>(Day/Month/Year Filed)</u>	<u>Yes</u>	<u>No</u>
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

_____ (Application Serial No.)	_____ Filing Date	_____ (Status -- patented, pending, abandoned)
_____ (Application Serial No.)	_____ Filing Date	_____ (Status -- patented, pending, abandoned)

I hereby appoint Paul A. Apffel, Reg. No. 35,427; Keith G. Askoff, Reg. No. 33,828; Aloysius T. C. AuYeung, Reg. No. 35,432; Bradley J. Berezna, Reg. No. 33,474; Roger W. Blakely, Jr., Reg. No. 25,831; Jeffrey Jay Blatt, Reg. No. 30,244; Stephen D. Gross, Reg. No. 31,020; David R. Halvorson, Reg. No. 33,395; George W. Hoover, Reg. No. 32,992; Tracy L. Hurt, Reg. No. 34,188; Eric S. Hyman, Reg. No. 30,139; Stephen L. King, Reg. No. 19,180; James D. McFarland, Reg. No. 32,544; George R. Meyer, Reg. No. 35,284; Ronald W. Reagin, Reg. No. 20,340; James H. Salter, Reg. No. 35,668; James C. Scheller, Reg. No. 31,195; Ira M. Siegel, Reg. No. 28,907; Maria McCormack Sobrino, Reg. No. 31,639; Stanley W. Sokoloff, Reg. No. 25,128; Edwin H. Taylor, Reg. No. 25,129; Lester J. Vincent, Reg. No. 31,460; Ben J. Yorks, Reg. No. 33,609; Philip K. Yu, Reg. No. 35,742; and Norman Zafman, Reg. No. 26,250; my attorneys; and Anthony C. Murabito, Reg. No. 35,295; and Edwin A. Sloane, Reg. No. 34,728; my patent agents; of BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN, with offices located at 12400 Wilshire Boulevard, 7th Floor, Los Angeles, California 90025, telephone (310) 207-3800, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected herewith.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

*1-00 et al*  
Full Name of Sole/First Inventor Richard Mander

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Palo Alto, California CA (City, State) Citizenship New Zealand (Country)

Post Office Address 899 Webster #2  
Palo Alto, California 94301

Full Name of Second/Joint Inventor Daniel E. Rose

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence San Jose, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 1084 South Blaney  
San Jose, California 95129

Full Name of Third/Joint Inventor Gitta Salomon

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Palo Alto, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 816 Cowper Street #3  
Palo Alto, California 94301

Full Name of Fourth/Joint Inventor Yin Yin Wong

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Menlo Park, California (City, State) Citizenship U.S.A. (Country)

Post Office Address 970 Alice Lane #4  
Menlo Park, California 94025



Full Name of Fifth/Joint Inventor Timothy Oren

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Sunnyvale, California Citizenship U.S.A.  
(City, State) (Country)

Post Office Address 961 Glenbar Avenue  
Sunnyvale, California 94087

Full Name of Sixth/Joint Inventor Susan Booker

Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence San Carlos, California Citizenship U.S.A.  
(City, State) (Country)

Post Office Address 147 Highland Avenue  
San Carlos, California 94070

Full Name of Seventh/Joint Inventor Stephanie Houde

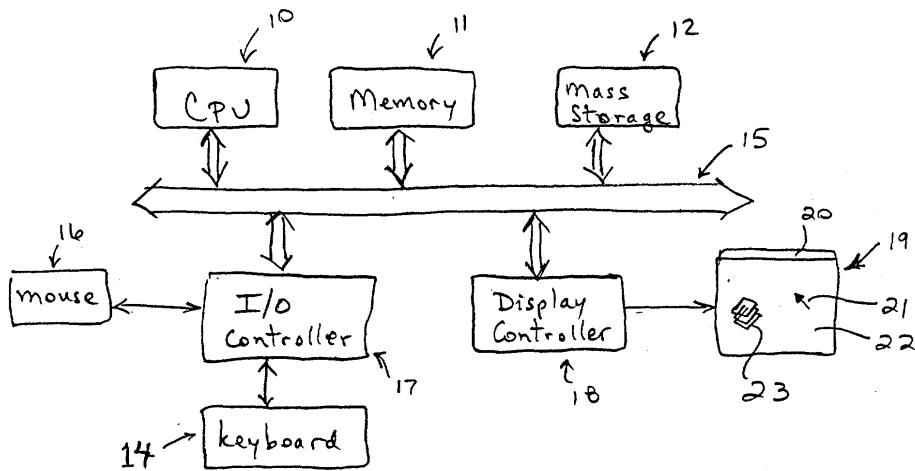
Inventor's Signature \_\_\_\_\_ Date \_\_\_\_\_

Residence Cambridge, MA Citizenship U.S.A.  
(City, State) (Country)

Post Office Address 279 Harvard Street #33  
Cambridge, MA 02139

08/28/08  
~~11/07/09~~

Figure 1



(1)