AGGREGATION OF CONTENT AS A PERSONALIZED DOCUMENT

BACKGROUND OF THE INVENTION

1. Technical Field:

The present invention relates in general to the creation of an electronic document and in particular to the generation of an electronic document as a composite of reference documents. More particularly, the present invention relates to a method, system and program product for parsing reference materials or segments of reference materials and aggregating the segments in an ordered manner to create a personalized electronic document having navigational affordances. The present invention also relates to a method, system, and program product for indexing and formatting personalized electronic documents based on user suggested keywords.

2. Description of the Related Art:

Synthesis of personalized documents utilizing one or more other sources of information as reference material is a common function in today's academic and business environment. Students and professionals alike often desire to compile information (or data) from already published sources to create their personal work. For example, a history student writing a comprehensive paper on the American Revolution may utilize history books, newspaper or journal articles, and more recently written accounts found in electronic databases, such as the

Internet, as references. In a typical document synthesis process, only relevant component parts of the reference material are included in the created document. The created document generally is completed with a title, an index, reference section, and the writer's personal commentary linking the data together in a cohesive format, etc.

The fast growth of the Internet, due in part to the vast amounts of information currently available on it, has led to the Internet becoming one of the most utilized resource for data retrieval. Present use of the Internet allows for a user to enter a search query and receive, in response to the query, hypertext links to sites on the Internet where relevant information to the search query exists. In the present Internet environment, most of this information is found at web sites created with hypertext markup language (HTML). The information is found in web documents, which tend to exhibit an article or page level granularity, i.e., the entire document remains as a single block during editing and or combining with other documents. Conducting searches on sites created with HTML generally results in hits, some of which are based solely on an occurrence of a single word within the web site associated with the hypertext link. The documents containing the single word hits are generally immaterial to the search but are returned nonetheless. Also, if a user desires to access the section of a HTML document which contains a term in the search query, the user has to download (i.e., open within his computer system) the entire document and search through it. Also, if a user desires to create a composite of two documents containing a search query, the

user has to link the documents in their entirety or complete a manual cut and paste of sections of the documents in a word processing application.

During manual document synthesis, the document drafter reads through the entire reference material, selects or highlights sections of interest, copies those sections into his notebook or on his computer. The drafter then repeats the process for the next reference material. In some instances, all reference materials are first read, and then the drafter carefully selects individual portions from each reference as he creates his document.

A similar process is undertaken in the electronic medium, except that the reference material is available electronically. The drafter reads through several sources of online information and selects relevant portions to include in his personalized document. The relevant portions may then be cut and pasted or copied in some other manner into a word processing application where they are manually linked by the drafter.

Both of the above methods of personalized document synthesis are extremely time consuming and inefficient, particularly when the user merely wishes to have a grouping of relevant information, which he can easily reference by accessing a single document at a later date.

The general category of information delivery sites has grown in popularity on the Internet; however no resource correlation, aggregation and reuse is completed at these sites. Some prior works have discussed the

concept of aggregating multiple documents on the web into a single document. For example, U.S. Pat. No. 5,924,090 discloses a classification system by which items are placed into categories and subcategories based on content using meta-data attributes. The classification system utilizes an apparatus which searches a database and organizes search results into a set of most relevant categories to enable a user to obtain only those records which are relevant.

In "As We May Think", *The Atlantic Monthly*, July 1945, pp 101-108, which is reprinted using http protocol at sloan.stanford.edu/mousesite/Secondary/Bush.html, the author describes the concept of a memex device. The memex device facilitates information discovery and information synthesis for reuse. The article also discusses associative indexing by which two or more information items are connected by a user-determined key, and the creation of a trail representing a section of pertinent information items.

Neither of the above references disclose navigational affordances (i.e., physical indicator having perceived properties that indicates how something should be utilized/completed) which permit a personalized document synthesis from relevant reference material.

The present invention recognizes that it would be advantageous to have a method, system, and program product for parsing electronic reference material into component parts and efficiently synthesizing a personalized electronic document from the component parts of the electronic reference material. A method, system,

and program product that allows users to conduct a search for reference material and then automatically generate a composite document containing only relevant information from the selected reference material based on user guidelines entered in a document generation utility would be a welcomed improvement. It would also be advantageous if such a method, system and program product permitted automatic formatting and indexing of a document, such as the generated personalized document. These and other benefits are recognized in the present invention.

SUMMARY OF THE INVENTION

A method, system, and program product is disclosed for electronically creating a personalized document from at least one electronic reference. The method first selects the electronic reference. Then, the electronic reference is parsed into sub-components. The method then aggregates similar items from among said sub-components, to create the personalized document, which is outputted with a set of navigational affordances based on a selection by a user prior to document generation.

In a preferred embodiment, the method, system, and program product also annotates the personalized document with pre-document generation commentary enter by the user and creates an index for the personalized document based on a user selected index option and index depth. The index is created utilizing key terms from within the sub-components.

The above as well as additional objects, features, and advantages of the present invention will become apparent in the following detailed written description.

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself however, as well as a preferred mode of use, further objects and advantages thereof, will best be understood by referencing the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

Figure 1A is a diagram of a data processing system utilized to implement a preferred embodiment of the present invention;

Figure 1B is a block diagram of a client-server-database network utilized to implement a preferred embodiment of the present invention;

Figure 2 is a diagram of a graphical user interface (GUI) of an infocenter application, within which a user may select options for generation of an electronic document and indexing in accordance with one embodiment of the present invention;

Figure 3A is a search GUI for retrieving reference material in accordance with one embodiment of the present invention;

Figure 3B is an annotation GUI for annotating the newly created personalized document in accordance with one embodiment of the present invention;

Figure 4 is a logic flow chart of the process of generating a personalized electronic document in accordance with one embodiment the present invention;

Figure 5 is a logic flow chart of the process of electronically indexing a document according to one embodiment of the present invention; and

Figure 6 is a block flow diagram of the process of generating a personalized electronic document in accordance with one embodiment the present invention.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENT

With reference now to the figures, and in particular with reference to **Figure 1A**, there is depicted the basic structure of a data processing system **20** utilized in the preferred embodiment of the invention. Data processing system **20** has at least one central processing unit (CPU) or processor housed in system unit **22**. System unit **22** is connected to several peripheral devices, including input/output devices such as a display monitor **96**, keyboard **82**, graphical pointing device **84**, and printer **94** for user interface. Also housed in system unit **22** are a permanent memory device (such as a hard disk) for storing the data processing system's operating system and user programs/applications, and a temporary memory device (such as random access memory or RAM) that is utilized by CPU to implement program instructions. System unit **22** communicates with the peripheral devices by various means, including a bus or a direct channel (more than one bus may be provided utilizing a bus bridge).

Data processing system **20** may have many additional components which are not shown such as serial, parallel, and USB ports for connection to, e.g., modem **92** or CD ROM **78**. In the preferred embodiment of the invention, communication to the data processing system **20** is made possible via a modem **92** connected to a land line or wireless cellular telephone system, which is in turn connected to a local network provider such as an Internet Service Provider (ISP). Additionally, data processing system **20** may be connected to a network via a network adapter **90**. Communicated data arrives at the modem or

network card and is processed to be received by the data processing system's CPU or other software application. In the preferred embodiment, Internet service providers offer reference data that can be downloaded into data processing system **20** via modem **92**. Modem **92** may also provide a connection to other sources of reference data, such as a server, an electronic bulletin board (BBS), or the Internet (including the World Wide Web).

Those skilled in the art will further appreciate that there are other components that might be utilized in conjunction with those shown in **Figure 1A**; for example, a display adapter connected to processor might be utilized to control a video display monitor **30**, and a memory controller may be utilized as an interface between temporary memory device and CPU. Data processing system **20** also includes firmware whose primary purpose is to seek out and load an operating system from one of the peripherals (usually a permanent memory device) when data processing system **20** is first turned on. In the preferred embodiment, data processing system contains a relatively fast CPU along with sufficiently large temporary memory device and space on permanent memory device, and other required hardware components.

Conventional data processing systems often employ a graphical user interface (GUI) to present information to the user. The GUI is created by software that is loaded on the data processing system, specifically, the data processing system's operating system acting in conjunction with application programs. The preferred embodiment of the invention is implemented with a GUI-based application having several user interfaces and

underlying functional components stored as program code on a medium connected to and readable by the processor.

The implementation of the present invention occurs on a data processing system as described above. It is understood however, that other types of data processing systems are possible, which may have some or more of the basic components described above. For example, a single purpose document synthesis system may be utilized in place of a general purpose data processing system.

The invention may be implemented within a network environment as illustrated in **Figure 1B**. Network environment comprises of a client on which the invention is implemented as an infocenter application **151** and a server **153**, which serves as a source of or conduit to reference data **155** for personalized document synthesis. Network environment may be a local area network (LAN) or wide area network (WAN), such as the Internet. The preferred embodiment is implemented on a WAN connected data processing system, which has Internet browser capabilities for searching the Internet for relevant reference material. The invention will be described herein with reference to a WAN and connected data processing system.

The World Wide Web (Web) is a graphic, interactive interface for the Internet and the term Internet is utilized interchangeably with Web throughout this specification. There are different computer program applications (i.e., web browser clients, referred hereinafter as web browser) on a data processing system connected to the web that are utilized to access servers

connected to the Web. Information is stored on a web server as web pages. A web page contains one or more graphic and/or textual displays, which may be linked together, and may be downloaded to a client data processing system utilizing a web browser. Each web page has a unique address, or Uniform Resource Locator (URL) within the Web that is accessible by utilizing Transfer Control Protocol/Internet Protocol (TCP/IP) transactions. The web page is often represented within a client browser as a corresponding hypertext link, which may also provide information on the page content.

Current web page design has shifted from the use of HTML format, which exhibits page level granularity to the Extensible Markup Language (XML) format, which exhibits a dynamically extendible mechanism for describing document content, more refined granularity, and other functional elements not available in HTML. XML was developed by the World Wide Web Consortium (W3C) in 1996. It is a file specification for placing structured data in a text file, which then permits access to individual components of the text file/data. A text file prepared using XML format can be later viewed without the program by which the file was produced. The text formats for the XML file are easy to generate and read by a computer in an unambiguous manner and are platform independent. XML utilizes tags (i.e., words bracketed by '<' and '>') and attributes to delimit pieces of data. XML includes syntaxes for pointing to parts (pieces of data) of an XML document. XML allows web authors to add tags to the web documents to specify a meaning to a search query to make searches more precise. XML also provides customization of the views of information by manipulating the data

accordingly. The present invention leverages the functionality of XML in implementing several of the steps illustrated in the flowcharts of **Figures 4** and **5**.

The present invention utilizes the functionality of XML language to permit a creation or synthesis of a personalized document from multiple XML documents found within a database. For the purposes of the invention, the term database is defined to mean any collection of one or more reference material selected by a user in generating the personalized document. The present invention provides a system of information discovery and reuse that yields a relevantly scoped, fully-realized, personalized document. The invention relies on XML document type definitions (DTDs) to enforce semantic organization on data and utilizes XSL as a data filtering technology which provides transcoding services for the sharing of composite results.

The invention is primarily implemented within an infocenter GUI as illustrated in **Figure 3A**. For the purposes of the invention, an infocenter GUI means a portal that is product or domain oriented. Infocenter GUI may alternatively be referred to as a resource center or a document generation center. The elements present in infocenter GUI are created utilizing widgets, which add to the user interface and offers the user more utility choices. In the preferred embodiment, infocenter GUI **300** has browser functionality enabling it to conduct searches over the Internet based on a user entered query. Infocenter GUI **300** accesses the web utilizing search GUI **201** illustrated in **Figure 2**. In **Figure 2**, a search query is enter in query field **207** and may be supplemented with

contextual search terms entered in context field **205** and category field **203**. The latter two fields are utilized to pin-point the search by further defining the general area being referenced. This feature makes use of an enhancement of the search utility due to the fine grain search capability with XML formatted documents (e.g., web pages), and enables more accurate hits to occur. The user selects the submit query button **209** to transmit the search request out to the Internet. When a hit occurs, a notification window **211** alerts the user that his search has been successful.

Returning now to **Figure 3A**, the web browser functionality of infocenter GUI **300** is enabled when the user enters a search term in search field **321**, which opens the search GUI **201** of **Figure 2**. Relevant hits are returned as hypertext links in the first frame **323** of infocenter GUI **300**. From here, a shopping cart selection of articles may then be completed by the user. The user selects those articles which he believes to contain good reference material for the generation of the personalized document and copies them (via a drag-and-drop operation or double-click selection, etc.) into the reference section **303**. Links to selected references are mirrored in reference section **303**. When the links are selected, the actual text of the documents (i.e., not just the hypertext links) are downloaded into the reference storage area of infocenter GUI **300** and temporarily stored while the parsing and synthesis steps occur. For illustration, three references are shown as having been selected for document synthesis. Once the desired references have been selected, the user is able to enter

the formatting, indexing, and annotating information desired to be reflected in the personalized document to be generated.

In the formatting, indexing and annotating areas of infocenter GUI **300**, the user is able to enter customization information, for example, navigational affordances and annotations, for the new personalized document. For the purposes of the invention, navigational affordances is defined to mean the title, index, reference section, headings and/or sub-headings, hypertext links to the reference sources, etc. that are included in the personalized document. The customization information thus includes the title of the new personalized document, entered into title field **305**, the level of indexing desired for the index of the personalized document, entered at index depth selection area **311**, and the presentation format of the personalized document selected from a list of selectable formats in format pull-down menu **307**. It is understood that other types of customization information are available and the above list is not meant to be limiting on the invention. The formatting styles include, for example, Lotus-style and IBM style, etc.

The user may enter personal commentary to be included within the personalized document in commentary field **309**. Of course, further commentary may be later added to the personalized document and the invention merely provides a short descriptive summary or introduction of the composite material. In another embodiment, a user is permitted to annotate the entire personalized document in sections based on the indexing criteria. Thus, a user may annotate the beginning or end

of each major index heading within the finished personalized document. For example, the user may enter an introductory paragraph at the beginning of the personalized document, a summary sentence or paragraph at the beginning of each major section, and add a conclusion at the end of the personalized document. The input and selection of more than one user commentary occurs in a commentary GUI as depicted in **Figure 3B** described below.

Finally, the user selects the output method in output selection area **313**. The user selects from one of the output methods, which includes Browser, PDF, download new XML source, submit for printing, save as document file, etc. In one embodiment, the personalized document is outputted as a new XML document, which is exportable over the Internet.

In another embodiment, the user may also select to create a bibliography of the references to be included at the end of the personalized document. The references included within the bibliography are presented with hypertext links, to permit a later user to pull up the entire text or relevant portions of the references. Once all of the fields have been completed, the user selects the submit button **315** to begin the document generation.

Figure 3B illustrates a graphical user interface **350** utilized for entering user comments, which serve as annotations for selected parts of a soon-to-be created personalized electronic document. GUI **350** includes two frames, section frame **351** and annotation frame **353**.

Section frame **351** includes a contents list for the desired layout of the personalized document. Each item in the contents list is a selectable item which when selected couples the commentary entered in the annotation frame **353** with the particular item. Thus, for example, a user selects the Introduction-Summary item and then proceeds to enter written commentary in the annotation frame **353**. When he has completed his commentary, he then selects another item and enter corresponding commentary. Selecting the other item opens a new page in annotation frame **353**, or if the item has been previously selected, opens the previous frame corresponding to that item. Beneath the frame are location (or placement) icons **357** for determining the position of the commentary within the personalized document section. Three possible locations are illustrated having corresponding buttons. These are before button **359A**, after button **359B** and select button **359C**. Various modifications of this placement function are possible. The buttons may be individually selectable after each item selection or applied to the entire personalized document. Once the commentary is completed the user selects the return button **361**, which closes GUI **350**.

In another preferred embodiment, the GUI **350** is utilized as a framework for creating the personalized document. The commentary entered for each item in the content list is utilized to conduct the search for relevant reference information. Thus, when the return button **361** is selected, key terms within each commentary are selected as the search terms or phrases entered into the search query area of **Figure 3A**. These search terms

are transmitted over the Internet and return hits specific to the particular search terms of the particular section. The user then selects relevant hits for each section and submits these to the document generation utility. In one embodiment, the index terms or headings are created from the key terms and the extracted information from the respective relevant reference material.

Figure 4 illustrates the process of user interfacing for conducting a search of a database and subsequent submission of relevant reference material to the document generation utility. The process begins at block **401** and then proceeds to block **403** where the user enters a search query. Once the search results returns, the user selects appropriate matches for use as references in block **405**. The user then enters the title, indexing, annotating, formatting, and output information at block **407**. Once all necessary information has been entered, the user submits the request for a personalized document generation at block **409** and the process ends at block **411**. As discussed with reference to **Figure 3B**, other implementations of the invention may be completed in a different process order. The order presented herein is presented merely for illustration.

Referring now to **Figure 6**, there is illustrated a block flow diagram of the main utility functions of the invention. As illustrated, three reference documents, doc.A **601**, doc.B **603**, and doc.C **605** are utilized as inputs for creating the personalized document, doc.D **615**. Each reference document has a plurality of sub-components

based on their XML format. The selection of similar items in the reference is based on their semantic tagging in the XML format. The information has metadata, which may be separated based on subject, title and author. A contextual search is therefore possible. Parser utility **607** separates out the individual components of the references, respectively, and sends these separated components to compiler utility **609**. Compiler utility **609**, combines like sections of the various references (i.e., W of Doc.A **601** with W of Doc.B **603**, etc.). The combination may be done in a user-determined order or in default order. Utilizing a combination of the XML format, the index headings and other underlying components of the reference documents, the personalized document is then compiled from the reference documents, where like areas are grouped under a particular indexed heading from all three sources. The compiled sections are then sent to indexer utility, which creates the index of the compiled sections, as described with reference to **Figure 6** below. Formatter utility **613** then formats the personalized document including insertion of index, personal commentary, title, reference section, etc. The personalized document, doc.D **615**, is then outputted based on the output method selected by the user. Although the various utilities have been described in a particular order and illustrated with connecting links, the order in which they have been described may be immaterial to various embodiments of the invention and the particular utility, such as indexer utility **611**, may operate as a stand alone utility in some applications. Further, other functional components are possible for implementing particular functions of the invention described herein.

Figure 5 illustrates the process of indexing a document according to the present invention. The indexing process begins at block **501** and then proceeds to block **503**. At block **503**, the user is prompted to enter the required depth of the index. Once the depth information is received, a category list of headings and subheadings, which may be provided by the user or taken from an indexing database based on topics of relevance, is loaded at block **505**. The document is then searched utilizing the heading and depth information at block **507**. The search may be completed at a paragraph level or section level depending on the complexity of the document. Next, a map of the major points is generated based on the search at block **509**. The minor points related to each major point are then mapped to their respective major point at block **511**. A determination is then made whether the required depth has been achieved at block **513**. If the required depth has not been achieved, the step of finding the next minor point at the next depth/level at block **511** is continued, with each iteration yielding a deeper index of more refined points. Once the required depth has been achieved, the index is outputted to the user at block **515**. The index is then incorporated at the beginning of the document at block **517**. The process thereafter ends at block **519**.

It is understood that other process blocks may be necessary to complete the indexing function or that the process blocks may be arranged in a different order within the scope of the invention. For example, each category of heading and subheadings may be evaluated

sequentially within the document beginning at the start of the document and proceeding to the end. Also, indexing may be completed on an original document as well as on a newly created composite document (i.e., a personalized document). The index selection area of infocenter GUI **300** may also be expanded to permit the user to enter desired sequence of presentation of information within the personalized document, which directly affects the layout of the index as well.

The invention thus operates on information formatted as semantic XML elements, which lends itself to subject-oriented discovery. The invention implements a query interface that maps a user's intent to the semantics of the information database. The invention also implements an interactive interface that enables the user to add annotations, select navigational affordances (table of contents, general or subject-constrained indexes, link sets to or from other sources). The interface also enables the selection of presentation styles (typically, enterprise business rules influencing the look and feel of republished information), and the selection of desired result formats (as a new information unit to re-introduce back into the literature or exported to another database, as a readable on-line format for personal use, as a page--formatted result suitable for printing, etc.).

The invention implements the searching within a document or document database and filtering for relevant items that correspond to a search query, and returns a ready made, highly articulated composite that cleanly annotates and correlates salient points. The composite may then be shared either as an addition to the existing

body of research information or as a reproduction for export to other databases.

As a final matter, it is important that while an illustrative embodiment of the present invention has been, described in the context of a fully functional data processing system, those skilled in the art will appreciate that the software aspects of an illustrative embodiment of the present invention are capable of being distributed as a program product in a variety of forms, and that an illustrative embodiment of the present invention applies equally regardless of the particular type of signal bearing media used to actually carry out the distribution. Examples of signal bearing media include recordable type media such as floppy disks, hard disk drives, CD ROMs, and transmission type media such as digital and analogue communication links.

While the invention has been particularly shown and described with reference to a preferred embodiment, it will be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the invention.

CLAIMS:

What is claimed is:

1. A method within a data processing system for electronically creating a personalized document from at least one electronic reference, said method comprising the steps of:

in response to a user input, selecting said at least one electronic reference;

automatically parsing said at least one electronic reference into sub-components;

in response to a completion of said parsing step, automatically aggregating similar items from among said sub-components, to create said personalized document; and

outputting said personalized document with a set of automatically generated navigational affordances based on a selection by a user prior to document creation.

2. The method of Claim 1, further comprising the steps of:

annotating said personalized document with pre-document generation commentary entered by said user; and

indexing said personalized document based on a user selected index option and index depth, wherein said

indexing step utilizes key terms extracted from said sub-components.

3. The method of Claim 1, wherein said selecting step includes the step of copying the content of said at least one electronic reference into a temporary storage location connected to said data processing system.

4. The method of Claim 1, further comprising the steps of conducting a contextual search for said at least one electronic reference, wherein said search is based on user entered query and context information.

5. The method of Claim 4, wherein said at least one electronic document in said conducting step is created with extended markup language having meta-tags for delimiting relevant sections of said at least one electronic document, and wherein said conducting step searches a database of materials for said at least one electronic document having said meta-tags.

6. The method of Claim 5, where said search is conducted over the Internet and said database of materials includes at least one web page, represented by an associated hypertext link, wherein said at least one electronic document is located at said at least one web page and said selecting step includes selecting the hypertext link of said at least one electronic document.

7. The method of Claim 5, wherein said parsing step utilizes said meta-tags and a collection of pre-document generation, user-entered criteria to delineate said sub-components of said at least one electronic document.

8. The method of Claim 7, wherein said aggregating step includes the steps of:

matching said sub-components with each other;

grouping said sub-components having similar elements to create sub-set groups; and

linking said sub-set groups to generate said personalized document.

9. The method of Claim 8, wherein said outputting step includes the step of:

applying formatting preferences to said personalized document that were selected by a user prior to document generation;

pre-pending a title and summary within said personalized document; and

placing said index in said personalized document at a pre-determined location.

10. A computer program product for electronically creating a personalized document from at least one electronic reference, said program product comprising:

a computer readable medium; and

program instructions on said computer readable medium for:

in response to a user input, selecting said at least one electronic reference;

automatically parsing said at least one electronic reference into sub-components;

in response to a completion of said parsing step, automatically aggregating similar items from among said sub-components, to create said personalized document; and

outputting said personalized document with a set of automatically generated navigational affordances based on a pre-programmed selection by a user.

11. The computer program product of Claim 10, further comprising program instructions for:

annotating said personalized document with pre-document generation commentary entered by said user; and

indexing said personalized document based on a user selected index option and index depth, wherein said

indexing step utilizes key terms extracted from said sub-components..

12. The computer program product of Claim 10, wherein said selecting program instructions includes program instructions for copying the content of said at least one electronic reference into a temporary storage location connected to said data processing system.

13. The computer program product of Claim 10, further comprising program instructions for conducting a contextual search for said at least one electronic reference, wherein said search is based on user entered query and context information.

14. The computer program product of Claim 13, wherein said at least one electronic document in said conducting program instructions is created with extended markup language having meta-tags for delimiting relevant sections of said at least one electronic document, and wherein said conducting program instructions searches a database of materials for said at least one electronic document having said meta-tags.

15. The computer program product of Claim 14, where said search is conducted over the Internet and said database of materials includes at least one web page, represented by an associated hypertext link, wherein said at least one electronic document is located at said at least one

web page and said selecting program instructions includes program instructions for selecting the hypertext link of said at least one electronic document.

16. The computer program product of Claim 14, wherein said parsing program instructions utilizes said meta-tags and a collection of pre-document generation, user-entered criteria to delineate said sub-components of said at least one electronic document.

17. The computer program product of Claim 16, wherein said aggregating program instructions includes program instructions for:

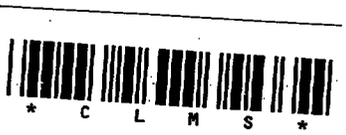
matching said sub-components with each other;

grouping said sub-components having similar elements to create sub-set groups; and

linking said sub-set groups to generate said personalized document.

18. The computer program product of Claim 17, wherein said outputting program instructions includes program instructions for:

applying formatting preferences to said personalized document that were selected by a user prior to document generation;



pre-pending a title and summary within said personalized document; and

placing said index in said personalized document at a pre-determined location.

* C L M S *

19. A data processing system for enabling personalized electronic document generation from online reference material, said processor system comprising:

a processor and a data storage area;

a connection medium for connecting said processor to a database of relevant reference materials;

a first graphical user interface (GUI), stored in said data storage area and executable by said processor, for entering a search term and conducting a search on said database;

a second GUI, stored in said data storage area and executable by said processor, for selecting relevant reference material returned from a search conducted on said first GUI and for entering navigational affordances related to a user desired formatting of a personalized document;

a document generation utility, linked to said second GUI, for synthesizing a personalized document from component parts of said relevant reference material, wherein said personalized document is presented with said navigational affordances, said document generation utility being stored in said data storage area and executable by said processor.

20. A method for electronically creating a personalized document comprising the steps of:

creating an annotated summary of a personalized document including key terms for inclusion in said personalized document;

submitting said annotated summary to a search engine to enable a search for and retrieval of reference material having component parts about said key terms;

receiving a set of said reference material; and

automatically synthesizing said personalized document utilizing said annotated summary and said component parts of said set of reference material.

21. The method of Claim 20, wherein said annotated summary in said creating step includes a title, an introduction, and commentary on one or more of said key terms.

22. The method of Claim 20, wherein said automatically synthesizing step includes the step of:

selecting a formatting style for said personalized document;

choosing an output method for said personalized document; and

creating an index of said personalized document after said personalized document has been synthesized.

23. The method of Claim 22, wherein said creating step includes the step of selecting a depth of said index.

* C L M S *

24. A system for electronically creating a personalized document comprising:

an annotating utility for entering user commentary about said personalized document;

a searching utility for searching within a database of electronic documents for reference material related to key terms from within said user commentary;

a user interface for allowing user selection of relevant reference materials and navigational affordances to include in said personalized document;

a document generation utility for automatically generating said document utilizing said relevant reference materials, said user commentary, and said navigational affordances; and

an output utility for outputting said personalized document.

25. The system of Claim 24, further comprising an indexing utility for automatically creating an index of said personalized document.

26. A computer program product for electronically creating a personalized document comprising:

a computer readable medium; and

program instructions stored on said computer readable medium comprising:

an annotating utility for entering user commentary about said personalized document;

a searching utility for searching within a database of electronic documents for reference material related to key terms from within said user commentary;

a user interface for allowing user selection of relevant reference materials and navigational affordances to include in said personalized document;

a document generation utility for automatically generating said document utilizing said relevant reference materials, said user commentary, and said navigational affordances; and

an output utility for outputting said personalized document.

27. The computer program product of Claim 26, further comprising an indexing utility for automatically creating an index of said personalized document.