lem, for "books" includes all the literature there ever was. This omission was not dealt with in any of the datacases discussed this far. Now is the time.

The Problem of References to Literature

Analytical works with a strong empirical foundation tend to be so absorbed in their own detail that they reach out less to the works of other scholars. George Goethals, Robert White's successor to the chair of Life Span Psychology at Harvard, produced with Dennis Klos, a volume of autobiographies, *Experiencing Youth* (Goethals & Klas, 1986). His focus in the interpretive section of that book is making sense of the case material (though limited to autobiographical studies) through the interpretive stances of three primary theorists in psychiatry. On reading *Experiencing Youth*, I was able to analyze the references in Goethals' interpretations for their relative centrality to his interpretations of the meaning of the individual cases. Figure 2.12 is a graphical summary of the relative centrality of references to the five theorists in *Experiencing Youth*.[37] It also includes suggestions for what materials should be made available online for a datacase that might be built around the collection of stories in *Experiencing Youth*.

If one has selected materials available online as references, would one want them to be included in the RING file assembly? That would lead to direct modification of those files, something not especially appropriate for shared resources. And yet, any analyst might find some elements of text in related literature certainly worthwhile keeping available, perhaps of directly copying into his own journal file. Journal links, to be described in more detail later along with the other features of the RING file assembly, provide this functionality and others. The analyst should be the person who decides where to draw the line between what is included in his journal and what remains less central for his own sense-making efforts.

This section has summarized the lessons learned from several attempts to construct datacases. We turn now to a description of my best current design ideas for such a facility, based on the experiences of the past projects.

A FILE ASSEMBLY FACILITY

File assemblies are sets of machine-readable files related primarily by the analyst's purpose in making sense of how elements of each file relate to elements in the others. A central structure of RING is the journal, typically another file, in which the analyst records his understanding about relationships between elements of the files. The centrality of the journal is represented in Figure 2.13, where the journal is shown as connected to other files by different kinds of links.

---

[37] The more complete analysis of these references is in Appendix D. In a personal discussion with Goethals, he essentially agreed to the connectness of this analysis.

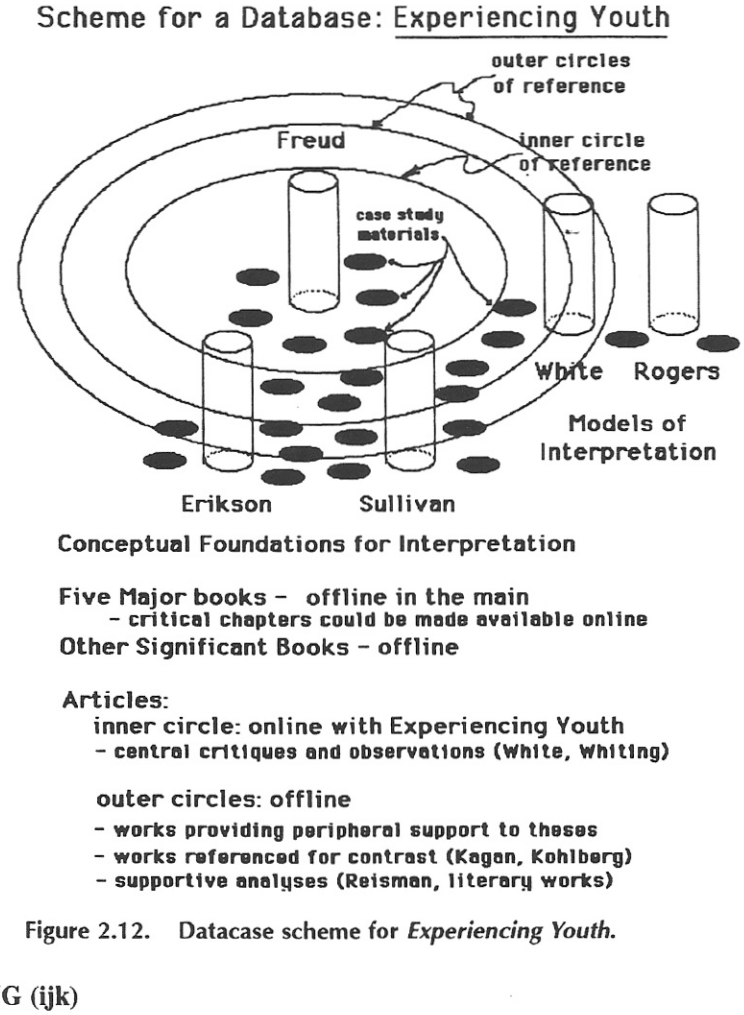

Figure 2.12. Datacase scheme for *Experiencing Youth.*

RING (ijk)

RING is a database shell designed for applications whose essence is the comparison and relating of information.[38] RING is designed to function in an environment of file assemblies. The RING shell[39] contains sets of utilities to load and save machine readable files; ~~it includes utilities to merge graphics files into existing files where appropriate~~. RING also contains a coherent set of programs

[38] The current prototype runs under Apple Macintosh Hypercard. The RING facilities are programmed primarily in hypertalk. The design and implementation of RING are an evolutionary development of the preceding case analysis support environments.

[39] The collection of programs embodied in the RING Hypercard stack is a shell in the sense that every file in an assembly will be an instance of the stack; they will be different one from another because different ASCII files will have been loaded into those instances.

* it includes methods and facilities to manage picture files of different sizes.

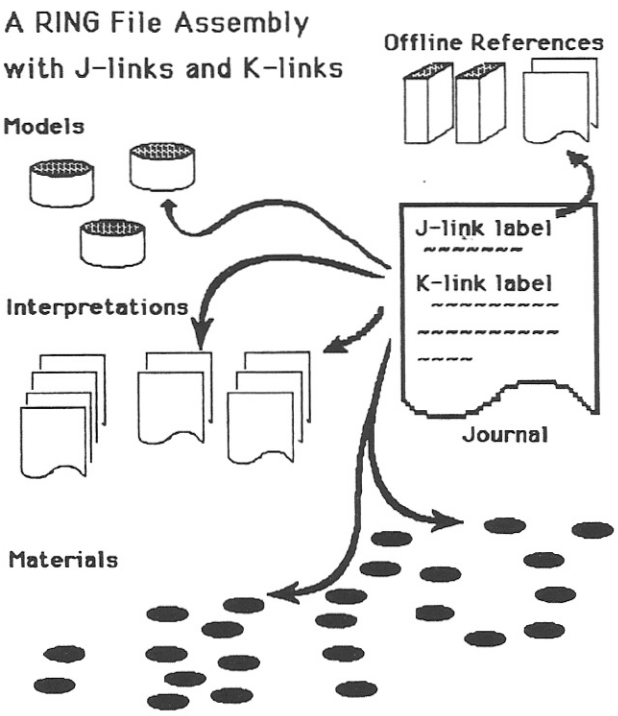

Figure 2.13. A RING file assembly.

for forming three kinds of links. The creation and manipulation of these links is RING's primary function.

What the RING file assembly facility Is. What is special about RING's structure is the interrelation of the instances of the RING shell during the course of execution. Each file in the file assembly (created by loading ASCII files with logical relations encoded symbolically) is an instance of the same shell. All the RING files in a file assembly are linked together through common routines in the shell. When the files of the assembly are selected for analysis in a specific session in the prototype, they are presented in the common Macintosh file loading selection interchange. When the file chosen to start the session (by this decision that file becomes de facto the journal file) is later chosen as the next file to be opened, the ring is closed. All the functions in each file instance are identical, except that the journal can recognize itself and execute routines specific to its special function within the RING structure.

What the RING file assembly facility does. RING presents three different but related kinds of links, described later, for the use of analysts in making sense of information in related machine-readable files. The primary goal of the RING

facility is to support the capturing of insights and the specification of interrelationships in knowledge links and in index entries. The RING linking routines create these links and indices. K-links are nontraditional information structures. To the extent that an analyst wishes to convert his insights for files in a more nearly standard form, reorganization is necessary. Other facilities of the RING shell support the reorganization of K-links as standard hypertext typed links—when the analyst has made that intellectual progress necessary for the more thorough integration of his ideas into a new order and new organization.

Why the RING file assembly facility is valuable. RING provides a facility for capturing insights and ideas about relationships before they are sufficiently well formed to fit neatly into a well-structured database. In this specific sense, RING is a link-authoring facility. It advances a solution for the difficulty people have found in creating hypertext databases. This is its primary value.

As distinct from other facilities advancing the creation of stable and well-formed structures, RING anticipates that flexibility and change are the central requirements for any use to which it might be applied. The tools of RING support the initial organization of the information selected as important by the analyst; it also supports the progressive reorganization of that information as the analyst more thoroughly understands the implications of the relationships.

How the facility relates to others. RING is a set of ideas about how to create hypertext from masses of unstructured information. It is dependent on computing because of the scale of complexity to which it aspires. The RING prototype is currently embodied in hypertalk for use with Hypercard on Apple Macintosh computers. The ideas could be reconstituted as programs in other database environments as well. It is the intention of the RING design that the files created in one instantiation of RING should be vendor independent. It is a further intention that such files should be portable between personal computers and networks supporting client server workstation organizations.

File Construction and Reconstruction

File construction and maintenance is RING's second major function. The product created by the final reorganization at the end of an analyst's work with RING is intended to be a traditional hypertext database.[40] Insights captured as relationships embodied in knowledge links can ultimately be recast in a form that can be integrated with stable information structures in a traditional hypertext database. The central image for use of RING is that it is comprised of volatile, intermediate structures (knowledge links) from which elements are removed when they can be fit into a form consonant with integrated data in

[40] The Xerox Company's Notecards or Civilized Software's Storyspace may be taken as representative.

other files; the migration of information from ill-structured to well-structured databases is central. In such a volatile environment, backup is expected to be a frequent activity. It is appropriate to undertake reorganization after a backup has been made.

What the file reconstruction facility is. The RING file programs are a set of utilities for file maintenance. These include routines to import (load) files and to export (save) files. The source files to be used with RING are ASCII files. Within those files, delimiting characters mark data fields; the order of data fields and the specific delimiting character define which chunks of text are destined for which fields on the "card" (database record).

What the file reconstruction facility does. The file importing programs route the text to the appropriate field on the appropriate card record. Index information is loaded with the text fields and stored in the appropriate corresponding fields. Knowledge links are loaded in a similar fashion. When a decision is made to back up the set of files, utilities dump each file of the file assembly in a form compatible with the importing programs.

Why the file reconstruction facility is valuable. The datacase physical files at any time are an image of relationships in the mind of the analyst. As the analyst experiences new insights and decides to try out new hypotheses, the form and structure of the files will change. In periods of active use, the RING files will change significantly and quickly. But having changed one's mind, it would be nice to be able to change it back. The physical files can be embodied in different ways with different organizations so long as there exist programs to perform translation from one embodiment to another. Frequent backup would be reasonable, especially if it is easy and not costly. The storage required for ASCII files is not great.

Among files that are frequently reconstructed, cross-system portability can be a valuable side effect. Because file save programs are writing ASCII files with logically encoded relationships, the information should be portable to any system for which appropriate loading programs have been written. With files encoded with a standard structure representation language, such as the Hypertext Markup Language (HTML) extension to the Standard Generalized Markup Language (SGML), transition from personal computers to client-server systems will be possible whenever workstations have platform-specific interfaces written for them.

How the file reconstruction facility relates to others. The file reconstruction facility is a way of building in support for knowledge link to typed link reorganization.

In a system where file reorganization is commonplace, the maintenance programs can also be useful in file creation. To add new information to the database, it could be typed into one of the RING files. That information would be saved when the files are closed, and it would enter into the reconstruction cycle

when downloaded for backup. Alternately, text files from whatever source generated could have field-delimiting characters added in a word processing program, then be loaded into the database with the RING file import programs. A third alternative would include the importation of large unstructured text files into the RING file shell, with that text's distribution to appropriate records through use of database cut-and-paste word processing functions; the appropriate field-delimiting characters then are generated for the ASCII files at the next file downloading.

The Link Authoring Facility

The three types of links are index links, knowledge links, and journal links. Using the first two link types specifies relationships between elements of the various online files. The journal links have two functions. The first is to include citations from occasionally referenced files and pointers to them. The second function is to aid in the reorganization of the RING files and the journal when the analyst chooses to do that.

The indexing facility. The purpose of the indexing facility is to permit the analyst to mark the text of any record or any chunk of that text with a label of his own choosing. Indices are useful only if available when and where needed. The RING prototype included programs for collecting and compiling index entries and then distributing them throughout the database so they may be locally available to every record.

A central aspect of RING linking is that it is symbolic; that is, based on shared labels as markers of scoped chunks of text. The techniques are based on file search and text groups and not on directly attached pointers. Location of a text group is speeded by employing indexed searching techniques.

What the indexing facility is. The indexing facility has four components: *index terms* (attached to the text fields with which they are loaded), *index compilation programs* (to be used after the files and indices are loaded), *index distribution programs* (to distribute the indexed entries for each item from the common index onto the individual records where they occur), and the *index-use programs*, which (in the current prototype) step through the list of references for an index entry and display each record in turn.

What the indexing facility does. The text field for each record has associated with it an index field, containing terms that are indices to contents of the text field. After a file of logically defined text and index fields are loaded onto a sequence of card records of the Hypercard stack, the user initiates index compilation. The process collects the index entries from each card record and associates with each term the record identification; the index entries are then sorted and combined into single-line index entries for each unique term. This data structure is the common index for the specific RING file. When the user initi-

ates distribution, copies of the entries of the common index are installed on each card record on which the index entry occurs. These distributed indices are thus comprised of locally indexed terms and their pointers into card records related by the set of common index entry terms.

Index entries of two types each can be created in two ways. General entries are interpreted as applying to the entire piece of text on the card record. Alphanumeric strings keyed into the index field of the card record, when compiled and distributed, comprise a general index reference to the entire text of the card record. A program to perform this function can be activated by the mouse entering the screen area of a card record button. Specific index entries can similarly be created by using the mouse to select a string of text. When the mouse then enters in the index link button, the selected text string is grouped, highlighted in boldface, and a dialog box requests a label be keyed for an index entry. When the analyst keys that entry, the label is inserted into the card record index field with a copy of the selected text appended and a further specification of the index reference. Such entries can, of course, also be keyed directly into the card record index fields if the analyst so desires.

Why the indexing facility is valuable. The central virtue of the RING indexing facility is its *user constructibility.* An analyst who uses a RING file may choose to incorporate the index created by some other person and modify it as well. The choice may equally well be to make a new index entirely. A further choice might be to load and use different indices at different times. One reason for doing so could be that different people might share the same RING text files.

Another reason might be that the analyst is considering alternative perspectives on the meanings of various texts and wants to avoid the confusion of having overlapping and possibly incompatible index references. This amounts to purposefully building transient relationship markers through implicit linking by shared symbolic labels. It is even possible to imagine that the analyst will discover new relationships through the implicit linking of shared terms. If some item is later to be dropped from the web of related indexed texts, all that is required is deletion of the index entry from the specific card record on which the text appears (followed eventually by index recompilation).

Another possible use for this specific indexing facility is the construction of trial link types for use in a standard hypertext database. An analyst could look at his own use of index terms for candidates to create categories for link types in a traditional hypertext database. He could even rationalize his index terms and recompile them before converting the RING file assembly into a more traditional hypertext database form.[41]

---

[41] Ultimately, once one is happy with an index term seen as a link type, one would want to mechanize the process through which that indexed relation is converted into a standard hypertext typed link. That is, one would want to automate the construction of typed links from these index entries. This has not been done yet in the prototype.

The knowledge linking facility. A central contribution of RING is the definition of the knowledge link as a structure helpful in advancing integrated understanding of information in file assemblies (see Figure 2.13).

What the knowledge linking facility is. The knowledge link or K-link is, first, an idea about what an analyst needs to do when dealing with large bodies of computer-resident information. Second, the K-link is a series of techniques for helping the analyst pursue that activity. Third, the K-link is a series of programs that permit an analyst symbolically to link scoped chunks of text in an assembly of files which he sees containing information relating them to one another.

What the knowledge linking facility does. The programs implementing the K-link permit an analyst to link under a common symbolic label in turn any chosen and scoped chunk of text in each file of the file assembly. In each file, that scoped chunk of text is grouped and redisplayed in a boldface font when the mouse movement indicates that creation of a K-link is intended. Further, in the journal file, the analyst is expected to compose a statement describing the significance of the relationship perceived among the file elements linked together. When the analyst then selects some phrase or phrases to highlight, that phrase is also grouped and highlighted. The programs then request a label for the K-link. After the K-link label is supplied by the analyst, that text string is inserted at the head of the text body in which the grouped string is located. It is also inserted in the card record index field to enhance the speed of search.

Subsequently, when the analyst seeks to locate information related through a K-link, the programs cycle through all the files of the RING file assembly and for each file bring to the screen the appropriate card record and register the grouped text in the middle of the displayed text field. If there is enough video display space available, the analyst will see each record displayed in a window and the journal entry specifying what he previously considered most important about the relationship between the linked elements. That evaluation of what is significant in the relation can be changed whenever the analyst decides doing so is appropriate.

Why the knowledge linking facility is valuable. The essential values of the knowledge link are two. First, it provides a means of relating information directly between disparate computer-based files. There is a need for such facilities now and it is likely to increase in the future.

Further, the expectation that the analyst will specify an appreciation of the relationship between the elements provides a way of capturing half-formed insights that can subsequently be tested and modified as appropriate. This is a consequence of the label for the K-link being a free-form text string. The analyst can specify anything at all as a label. This argues that the importance of the label to the human as mnemonic is greater than its importance to the computer as a blind record-locator code (blind in the sense that the computer is not expected to appreciate what the label string might mean). The flexibility of

being able subsequently to locate the K-link labels by search permits success on an incompletely keyed label, or on one imperfectly remembered.

How the knowledge linking facility relates to others. The knowledge link and index link are two complementary and supplementary tools for relating information within file assemblies. Index links are useful initially only within a single RING file. However, as the analyst proceeds with integration of his understanding of the material, the use of common labels for index entries across files will permit building up a structure of reference that will be more complex than what is found in the single file and will be more nearly suitable for components of traditional hypertext files.

The knowledge link initially is the primary facility for relating information across various files of a file assembly. But, as the analyst progresses in his understanding of relationships within the files being examined, for a variety of reasons—ranging from search efficiency to mental hygiene—there will come a point when the maintenance of K-links will become a problem. Whenever that occurs, file reorganization is called for. These two link types are then related by the third link type, the journal link.

The journal linking facility. The general purpose of the journal linking facility is to aid in the reorganization of the RING file assembly relationships into structures of a more traditional hypertext database. It helps turn possibly vague and transient relationships into well-articulated categories of relatedness that can provide long-term stable structures in traditional hypertext databases.

What the journal linking facility is. The journal link is a directed relationship between a specific chunk of text in a nonjournal RING file and the journal file. When in use, the journal link will connect with a symbolic pointer, such as is used in creating knowledge links, elements of one, and an entry in the journal file. As with the knowledge link, it is expected that the analyst will specify what of the linked element is of interest and why it is so. This specification will comprise the entry of the journal file.

When the RING file assembly is undergoing reorganization, the text of the nonjournal file will be copied directly into the journal file.

What the journal linking facility does. In everyday use, the journal link permits bidirectional card record and text element location between two files of the RING file assembly.

During reorganization, the journal link ~~collects within~~ will collect the journal information originally accessible only in nonjournal files of the RING file assembly. The journal link is also a tool for deconstructing the knowledge links of the RING file assembly so that the information can be reorganized into the categories of typed links in a traditional hypertext database. This collection of information into the journal file permits breaking links between files of the RING file assembly. This is a way of purging the complexities that are no longer needed from the web of relationships built up through the process of analysis.*

Why the journal linking facility is valuable. In the use of hypertext, it is important to distinguish between permitting complexity and managing it. Man-

* These features are not implemented in the current prototype.

aging complexity requires the ability to both increase and decrease it when appropriate. The K-link is the mechanism for increasing complexity in a RING file assembly. The journal link or J-link is the mechanism for decreasing complexity.

How the journal linking facility relates to others. One important example that highlights the relation between the J-link and others is the removal of a file from a RING file assembly. One would want to capture the information in that specific file relevant to the information in the others but may decide that the information can be adequately represented by citation from the file and commentary on it. The J-link permits, during reorganization, the information from one file to be copied into the journal file and the link to the source file to be broken. This functionality is also required in the case of K-links. The J-link facility is likewise able to go from the journal to a specific other file, extract the relevant file element, preserve a pointer to it, and delete the symbolic link from the specific other file. When subsequently, the RING file assembly is opened without the other specific file included in the RING, that file automatically disappears from the K-link (although a trace of it may remain in the K-link commentary of the journal) and the text extracted by the J-link appears as a citation to a remote (offline) file, appended to the K-link commentary in the journal.

USEFULNESS OF THE RING FACILITY

Datacase Studies Now Underway

The psychology of the particular revisited. The reasons for deciding to retreat from the Three Years and Talking datacase to working with my earlier study are still valid. I have decided to complete the Psychology of the Particular datacase in the new RING environment. A further incentive occurred with the two books involving that study going out of print. Bringing materials of the case study corpus into machine-readable form is now going forward both by scanning of typed texts and by keying of central documents that are not scannable.

The system environment. The prototype implementation of POP will be in Macintosh Hypercard 2.1. The RING prototype runs on Apple SE-class machines (even earlier Apple Plus machines with enough memory to run System 7), but for real usefulness, systems with large or multiple screens are essential.

The contents of the datacase. The contents would, if presented in a figure like Figure 2.5, appear almost identical. The contents are the same. Every chapter that appears in one of the two books is, however, an individual file. The logic of the structure and its organization within the RING files assemblies is radically different from the earlier version of Figure 2.5 wherein the texts appeared with an organization similar to that of the book form. The collection of vignettes, which comprise a major resources within the corpus, have been retyped and can be loaded into the RING facility. Other corpus documents, such