Knowledge Space Revisited: Challenges for Twenty-First Century Library and Information Science Researchers

Pauline Atherton Cochrane

Abstract

This paper suggests writing a companion work to the Bourne and Hahn book, *History of Online Information Services, 1963–1976* (2003), which would feature milestone improvements in subject access mechanisms developed over time. To provide a background for such a work, a 1976 paper by Meincke and Atherton is revisited wherein the concept of Knowledge Space is defined as "online mechanisms used for handling a *user's knowledge level* while a search was being formulated and processed." Research that followed in the 1980s and 1990s is linked together for the first time. Seven projects are suggested for current researchers to undertake so they can assess the utility of earlier research ideas that did not get a proper chance for development. It is just possible that they may have value and be found useful in today's information environment.

When Charles Bourne had finished working on his 2003 coauthored (with Trudi Bellardo Hahn) book, *A History of Online Information Services, 1963–1976*, I suggested that he begin work on a companion volume that would feature the work done to improve subject access mechanisms in online information services. Such a book could document the various systems and projects and provide a list of milestones for those features that appeared for the first time, as he did in the book he had finished. As a framework, I told him he could use what members of the Classification Research Group in England (B. C. Vickery and E. J. Coates, among others) were writing about as they followed S. R. Ranganathan's lead to analyze what vocabulary control mechanisms existed and how they were being adapted for use in the new computer-based indexes and catalogs as well as new online systems. Bourne did not accept the challenge, and to this day historians have not done the task.

This paper presents the beginning of such a history from a personal point of view, but then I go on to challenge current LIS (Library and Information Science) researchers and historians to help with the task. Many of these early systems were filled with innovative ideas, but those ideas could not be fully developed or tested because the state of technology at the time was crude and cumbersome. Some of the milestones in Bourne & Hahn showed subject access and search features that remain standard to this day, but it also highlighted promising features that were never deployed in any operational system. To pique the interest of present-day designers in these "lost" features I will discuss some of them and suggest that these projects be re-examined, even resuscitated. In those old projects, I believe, are ideas for features that could improve user access to today's digital libraries, e-book collections, and institutional repositories. This paper is a memoir as well as a partial historical review because I will go beyond Bourne & Hahn and show how projects in the 1980s and 1990s continued the quest for improvement.

In Bourne & Hahn, the milestones are gathered on pages 414–418 as well as placed in the text where the system that incorporated this feature for the first time is documented. Here is an abridged list from that book, which includes some firsts in subject access mechanisms:

1963
- SRI demonstrated the first *online bibliographic search* system.
- SRI demonstrated the first *online full-text search* system.
- SRI was the first to demonstrate an online search system that *retrieved records on the basis of bibliographic citation elements* such as author, publication date, title words, or abstracts.
- SRI provided the first demonstration of *a stop list* in an online search system.

1964
- TIP was the first . . . to *retrieve citations on the basis of cited references or bibliographic coupling*.
- TIP was the first . . . to allow a searcher to *save search output* for later searching or use.
- TIP was the first . . . to implement *stem searching*.
- TIP was the first . . . to implement *left-truncation searching*.

1965
- BOLD was the first . . . to *provide posting counts* of the number of records associated with each search term.
- TEXTIR was the first . . . to *incorporate synonyms automatically* into the search formulation.
- TEXTIR was the first . . . to demonstrate *search term weighting*.

1966
- MULTILIST was the first . . . to permit a searcher to *retrieve automatically all the records that were either hierarchically subordinate to or hierarchically superior* to a given search term.
- TIRP was the first . . . to *display automatically alternate search terms and their posting counts* when a query found no records for a term.

1968
- AUDACIOUS was the first . . . to demonstrate an *online decimal classification authority file as an online search aid.*
- NASA/RECON was among the first . . . to operate in a *multifile access mode.*
- SUNY BCN was the first . . . to offer a database with *in-depth indexing to books and other monographs.*

1970
- SUPARS was the first . . . to allow users to *search for and examine search strategies created by other searchers*

1973
- CAN/OLE (or RETRO) may have been the first . . . to *offer a bilingual interface option.*

These milestones indicate how the search process was gradually improved as system designers learned more about the characteristics of the items in their databases and heard about user problems.

It is hard to believe that there was a time when author, title, date of publication, and language identification could not be used in searching, or when the searcher had no access to a display of index vocabulary.[1] It is hard to believe that individual search histories were not stored until 1964 in the TIP project at MIT. This feature is now ubiquitous on Google's opening screen and other places on the Internet.

Not until 1970 were *all the searches* on a system deconstructed and made searchable in a file that could serve as a list of suggested search terms. This was done in 1970 at Syracuse University in the SUPARS Project where, by design, no controlled vocabulary or indexing records were made available. For test and evaluation purposes, the only searchable items available were from the free text of titles and abstracts from *Psychological Abstracts.*[2] As a user aid, the system designers decided to take the text of different users' searches and create a *search term database* that all users could consult. This "lost" feature (i.e., never became standard in operational systems) might have been what we today would call "an early social media mechanism."

To start my personal review, I decided to revisit a 1976 paper Peter P. M. Meincke and I (as Pauline Atherton) worked on over a three-year period after meeting at a NATO-sponsored conference in Wales. I consider it a good jumping-off place because it shows what we two learned from our re-

view of the early developments in subject search and retrieval mechanisms and how it impacted our thinking about the direction that had to be taken in retrieval system design. I will show that, by the 1990s, these 1970s ideas were beginning to have an impact on digital library development when the technology available was vastly improved. The work represented in the 1976 paper all began because Meincke, a physicist and a Provost, with responsibilities for Libraries and Information Technology at the University of Toronto, asked a LIS researcher how things could be improved.

Our discussions started with Vickery's list (published in Bourne, 1963, pp. 13–20) of "degrees of vocabulary control." He listed seven *mechanisms* indexers and catalogers had developed over time:

1. Words *chosen from title or text*, with *common words omitted*
2. Words *chosen from text . . .* with *consideration of variants*
3. Words *chosen from text . . .* with *generic relationships*
4. Words *chosen from text . . .* with consideration of *syntactical relationships between indexing terms*
5. Any of the above, with *addition of terms NOT used in text*
6. Assignment of index entries from a fixed *authority list or classification scheme*
7. Assignment . . . *representative of several viewpoints and aspects of subject*

I told Meincke that not all of these mechanisms were added to every online system at the time. Meincke could see how this list showed a progression of control and usefulness, but it was E. J. Coates's (1960) list of five "problems and mechanisms available in alphabetico-specific catalogs, classified catalogs, and alphabetico-classed catalogs" that gave our discussions a broader orientation. Coates's list may look like a mirror image of Vickery's list, but because Coates presented the *problems faced by a searcher,* Meincke saw the wider picture where new designs with new mechanisms needed to be developed.

Simply stated, Coates's viewpoint was that the catalog user had problems because he needed to know *how to*

- correctly *formulate a request,*
- *handle synonyms,*
- *handle* generic-to-specific or specific-to-generic *hierarchical relationships,* and
- *handle collateral relationships.*

Meincke could see that just incorporating Vickery's mechanisms up to and including the seventh level would not do the trick if the searcher's state of knowledge were not represented.

We both knew that books and journals were being digitized and library catalog records were becoming online catalogs, but we felt researchers were concentrating too much on statistical techniques for processing

and "indexing" full-text databases and not enough on searcher's aids. We knew there were efforts during the 1960s and 1970s that focused on the practicability of merging existing vocabulary control tools (NASA, AEC, DOD, CIA), but it was uncertain how those efforts could become an advanced searching tool if the searcher's state of knowledge was not known.[3] Meincke and I continued our discussion whenever I would visit Toronto where I was engaged in collaborative library education projects. We finally decided to write a theoretical paper that we hoped would stretch the boundaries of the standard system design, going beyond what was then called *man–machine interaction* to include mechanisms for handling a *user's knowledge level* while a search was being formulated and processed. We called this construct "Knowledge Space," and we asked, *How could a system match the user's level of understanding of the concepts in the field s/he is searching with the concepts presented in the file of documents and control vocabulary tools?*

Most operational systems at that time avoided handling or incorporating into search mechanisms the thesauri or classification schedules used for controlled vocabulary indexing. Without such tools available to searchers during the online search process, it was difficult if not impossible to match and mark vocabulary areas where the search terms and indexing terms in the database coincided or could be expanded. There was a reluctance to make these "mechanized vocabularies" available because so many researchers were critical of these tools, considering them either out of date or inadequate representations of fields of knowledge. We could see this was a multifaceted problem.[4]

We reviewed what many eminent critics were suggesting:

- Recognize the irremovable limitations imposed on library classification's single dimension when *a multidimensional continuum is needed* (Ranganathan, 1951)[5]
- Provide "spatial tags" and "semantic road maps" (phrases used by Doyle, 1972)
- Assign *representational vectors as a spatial organization of information resources* that is more compatible with the structure of the information itself (George Miller's [1968] notion of "obvious improvement")

Miller went on to say,

> There is a user who has a system of concepts, and there is an information store that also has a system of concepts. . . . In order to recognize the indicated location, the system being queried should have a similar conceptual organization. This interaction occurs as a dialog, and performance is determined largely by the adequacy of the concept indexing, so that the user's gaps, however characterized, can be recognized and filled quickly and precisely. (Cochrane, 1985, p. 38)

These critical but suggestive comments spurred us on to devise a theoretical model of a Knowledge Space. Meincke provided the illustrations

suggesting a *multidimensional space* with three sets of vectors for all the concepts coming from different sources: *the field, the searcher, and the representations of items in the system.* He wrote,

> Scientists have found it useful to visualize the actual state of physical objects . . . in terms of a multidimensional vector space. The actual description of such states is given in mathematical equations or other notations (similar to words in an indexing vocabulary), but the multidimensional space construct contributes enormously to the ability to visualize these states and the relationships between them. (Cochrane, 1985, p. 39)

After writing the paper, neither one of us was available to undertake writing a research proposal to develop such a system. Within a few years, our careers went off in different directions. Meincke became President of the University of Prince Edward Island, and I went off to Papua, New Guinea in 1984.

Not until 1994 did I see some of the new technological developments, one of which I thought could possibly be the beginning of the Knowledge Space Meincke had imagined in our 1976 paper. What this system almost allowed for was a fluid communication between the "searcher's knowledge" and the representations of the indexing vocabulary and related files. Imbedded in this system being developed at the University of Illinois was a hypertextual thesaural browser. Eric Johnson was the designer. He was a student at GSLIS/UIUC at the time, with a Computer Science degree from Northwestern and experience at developing a full-text retrieval system for the Sociology department. His research effort was part of the four-year Digital Library Initiative project at Illinois funded by the National Science Foundation. Johnson called the system IODyne. By the mid-1990s, he had demonstrated the browser with various vocabularies such as INSPEC, ERIC, LCSH, and DDC to show its flexibility. In IODyne the searcher would enter a term (word or phrase) and first see a display of that term with its hierarchical and collateral relationships (levels 1–6 of Vickery) and/or see the term in a KWIC display. By pointing and clicking, the terms chosen were placed in a search window with no additional keyboarding. The system allowed for the opening and closing of various levels of hierarchical relationships and alternate hierarchies (level 7 of Vickery), thereby selecting more terms. The searcher moved effortlessly between the thesaurual displays and the search window, between the retrieval result and the other windows as needed. Everything was hypertextual.[6] As a research consultant to the Library of Congress, I was able to arrange demonstrations of the IODyne hypertextual thesaural browser there. The demonstrations were part of a larger study to involve Congressional Research Service (CRS) researchers in the improvement of the Legislative Index Vocabulary (LIV) that was used in the CRS search and retrieval system (for these demonstrations LIV was loaded into IODyne).

At the Library of Congress we came upon a most startling finding. In private interviews the CRS staff members (all Ph.D.s in their respective fields) were asked to comment on the utility of the IODyne/LIV displays as a possible augmentation to the online retrieval system used by CRS. They made comments about currency and applicability of terms, relationships between terms (BT-NT and RT). They also commented on what they recognized immediately had utility for their personal files on topics they were researching for Congress (their files were on their individual microcomputers and were not yet part of what the Automation Office at the Library of Congress maintained). They wanted to know how they could add terms to the list, show some new relationships and erase some terms! What they wanted was their own version of the LIV thesaurus for use in both their private files and the CRS files!

We immediately saw that such a configuration of various versions of LIV would reorient the role of the LIV lexicographer. She would need a different maintenance system to handle such a dynamic vocabulary environment, a new kind of online search and retrieval system, and a new mode of communication between system and users. No longer would it suffice to have a delay of months before new terms were added to the official version of LIV. No one on the indexing side of CRS could see how the "other" versions of LIV would fit in the picture. From the point of view of trained vocabulary control workers (indexers, catalogers, and the lexicographer), this was an impossible reorientation, but from the point of view of CRS researchers (and Johnson and Cochrane), this was a promising and challenging new environment for search and retrieval operations and some new systems needed to be designed. Before the Research Director at CRS could break this "political" deadlock and possibly find funds for more dynamic tools, CRS and the entire Library of Congress had to decide what to do with microcomputers, the Internet, existing online search capacity, and related issues of database control. By the time they had answers to those questions, Johnson had moved on to other efforts and I had retired.

I always thought of this episode at CRS as a missed or lost opportunity to carry forward the ideas in the *Knowledge Space* article.[7] At CRS, researchers would have helped us plot the Basic and State concept vectors in a multidimensional Knowledge Space representing their field/s of interest. Representational vectors would start with LIV and CRS indexing records. A revised version of IODyne would have been developed to handle all these vectors together. For the search retrieval operations, "a user would be able to specify a *search volume* of manageable dimensionality," as Meincke (and Miller and Ranganathan) suggested. We would have the mechanism that could handle search volume. Coordinates on *reference axes* (for basic concept vectors) would be calibrated so the searcher could express his understanding of the field being searched.[8] This may sound

to some like *the cloud displays of terms* that had appeared in some online catalogs by then, but the IODyne displays and operations would have to be more sophisticated than those based on statistical computing of unstructured free text.

Besides my missed opportunities for advancing the state of the art, I wonder how many other promising developments had gone only so far and were now all but forgotten. I want to stimulate the readers of this paper to begin similar reviews, and I know where I would like them to start. Who knows what would happen if these projects were revisited, even resuscitated in the digital world of today? If system designers as well grounded as Eric Johnson could be found to work on new projects, who knows how far we could go? Historians could help locate the original sources that document these projects, much as Charles Bourne did.[9] Researchers could prepare proposals to create certain capabilities in new test-beds (for example full-text files specially processed) and use systems that would be versatile enough to demonstrate various capabilities. Evaluation tests similar to those done originally, when repeated, would determine the utility of these mechanisms for use in digital libraries and elsewhere.

The projects listed below contain ideas that need another airing because their research objectives are still relevant. Besides suggesting the review of specific projects, I have included some ideas I have had over the years. These ideas never resulted in research projects because we needed more complete digital collections covering a huge amount of literature. Those collections exist now in Google, Hathi Trust, institutional depositories, and Web sites.

Project Number (1)
Review the PRECIS/BNB project for books (index records were subject strings based on facet analysis). They maintained a thesaurus for the ten-year period of the British National Bibliography (BNB) when PRECIS replaced LCSH. To my knowledge these records were never incorporated into an online system like IODyne. Their work could now be coupled with a full-text file for the items indexed, thereby testing PRECIS's utility in today's digital library environment. (PRECIS/BNB was developed by Derek Austin and his staff, and the record of their work may be at the BNB.)

Project Number (2)
Locate the ERIC Vocabulary Improvement Project's Play Thesaurus file and play with the idea of putting it into a hypertextual thesaural browser like IODyne that would allow for various hierarchical displays side by side based on different users' suggestions. This would be a simulation of the environment we envisaged for CRS at the Library of Congress. How tricky would such displays be?

Project Number (3)

In the Subject Access Project (SAP), an online catalog database of MARC records for two thousand books from the University of Toronto library was augmented with keywords and phrases from the tables of contents and back-of-the-book indexes (see pages 395–457 in my book of selected papers). The BOOKS file covered ten fields in the social sciences and humanities. The SAP retrieval tests could be compared with searching the e-book version for these two thousand titles and the LCSH syndetic structure for these ten areas. Surely full-text records for these two thousand books are now in digital collections somewhere. The files on the SAP project contain searches evaluated by actual users, so these could form the basis for comparative evaluations. Comparative evaluation with the PRECIS Project above might also be tried. (I believe the SAP files are in my archives at Syracuse University.)

Project Number (4)

With an eye to expanding on Vickery's 7th level of vocabulary control, representative of several viewpoints and aspects of a subject, I suggest a revisit to Lauren Doyle's *semantic roadmaps* and Loll Rolling's *Euratom thesaurus graphic displays*. The challenge would be to activate their ideas in today's online environment. (Euratom may have an archival copy of Rolling's graphic displays.)

Project Number (5)

Might dissertations in the History of Science be useful in projects designed to activate the idea of the Basic Concept vectors in Knowledge Space? As these efforts cover a long time period they could be matched with the literature from past and present centuries now in digital form. For other fields of knowledge maybe we could use subject bibliographies developed to cover fields of knowledge or important topics. A very large collection of these subject bibliographies is shelved near the Reference Desk at the Library of Congress. Has this collection (Class Z) been digitized by now? Would such a "filter" help us access the best in our digital collections? Would it give us a time dimension for concepts that reflect the eighteenth, nineteenth, twentieth, and twenty-first centuries? We will not know until we try.

Project Number (6)

Is the time right for a gigantic merge of all *extant controlled vocabularies* like thesauri, lexicons, and classification schemes? This was attempted in the 1960s and 1970s for the DDC and other federal agencies. Those early efforts at compatibility may have had the wrong objective, and they certainly had the wrong technology available, but there are reasons now that we may need such merged files. There were also many efforts to merge and compare LCSH with LCC, MeSH, and/or DDC, ERIC with PSYCHINFO,

etc. All these efforts should be reviewed. Unfortunately, most if not all of these efforts were printed, not digitized, but we could review them nonetheless. ClassWeb might be a starting point if different software were designed to analyze hierarchies and other features. Most vocabularies can now be found in digital form, thereby making new studies more practicable. What more has the European Community done in this area since the Euratom work by Rolling? Such a large project would probably take on the size of the Human Genome Project, and it might have a similar payoff. Look where that project has led. There is now ENCODE, the *Encyclopedia of DNA Elements*, and the new concept of DNA Junk. "A Road Map of DNA; A Key to Biology" was the headline in the *New York Times* for September 6, 2012. The article written by Gina Kolata reported on ENCODE as a huge federal project, a multi-institutional effort here and around the world.[10] Other "genome" projects are being developed. One was reported on the Arts page of the *New York Times* (October 9, 2012). Melena Rysik wrote that there is a new start-up called Art.sy that is a free "algorithmic guide for browsers and buyers of art." I wonder if the AAT vocabulary and facet analysis are playing their part in this development.

Project Number (7)
The TIP innovation that Mike Kessler at MIT in 1964 called *bibliographic coupling* deserves a review in a digital library environment. He wanted to quantify identical "entities" in different items in a file (for example, identical bibliographic references cited by various authors). He wanted to compute the strength of these "links" between papers as strong or weak depending on the numbers of identical references. TIP did not have a large enough database to properly test this, but proof of concept was there. Imagine matching references in the entire nineteenth-century digital file of books and journal articles. Would this ease the pain of free-text searching in such a file? Another possibility might be to start with one author's corpus of work (such as my book of selected papers), match all the bibliographic references used by that author over a twenty-five year period, and then go after the strongest linked references. Check what is retrieved and expand to compute references in that bundle of articles. This could be a technique for refining or expanding a search with little effort on the part of the searcher. It might result in a nice stack of relevant reading. Of course you would also work forward in time by using the accumulated references for more current literature by using the method we call citation indexing.

In conclusion I have to ask, "Are Google, the Hathi Trust, federal agencies, and foundations ready to fund any of the seven projects mentioned above?" If what I read in the *New York Times* in January 2013 is any indication, the field may be moving in the direction of Knowledge Space.

Somini Sengupta (2013) tells us that Facebook appears ready to learn how people seek information in today's digital world. They have gathered a team together (two linguists, a Ph.D. in psychology, and statisticians, along with a cadre of programmers) to teach Facebook's computers how to communicate better with people. They intend to "adjust to the demands of users." Will others follow suit? Will digital library managers?

In yet another *New York Times* article, in February 2013, which covered the "mice and medicine" scandal, I learned that billions of dollars in medical research may have been wasted because there seems to be little carryover from treatment studies of mice with a certain disease to humans who have that disease. Jim Dwyer (2013) wrote, "Having decoded the human genome, researchers are finding many mechanisms of life to be stubbornly inscrutable, much as each generation of astrophysicists who expand the map of the cosmos must reach for new theories to explain the universe. Whether in outer space or inner space, using giant telescopes or precision molecular assays, we find that the mysteries of how things work have gotten deeper." Maybe it will take a similar investigative report to expose the limitations of our current efforts at developing digital library services. I hope not. We should be doing the assessment and learning from our own findings. I hope that LIS researchers will continue to develop our digital library services into a Knowledge Space as a *mission* on as vast a scale as those who worked on the drawings of constellations of stars and the planets in the dark sky during ancient times and those who are working now on ENCODE or will work on the Brain Activity project just announced. If we can design and operate earth satellites, global weather monitors, and robotic observers on the Moon and Mars, surely we can design a Knowledge Space worthy of the name. Otherwise we may be working on digital library systems, but are only creating "dark holes."

NOTES

1. At a library conference in 1978, I presented a table of unheeded research findings in online catalog design. The second of ten in that list was, "[we had unheeded findings from studies which] recommended that we include the following to improve the success rate of searchers: the contents notes in books, [and] in-depth subject analytics" (Cochrane, 1985, p. 149).

2. This was the contractor's requirement.

3. Back in the early 1960s, at the American Institute of Physics, we conducted a random, stratified Survey of Search Requirements from a sample of physicists registered in the National Science Foundation Register. From them we accumulated characteristics about physicists' search requests "which they would ask of another physicist." This file was never incorporated into any vocabulary studies or test of retrieval system design beyond our analysis done to suggest improvements in the print version of *Physics Abstracts,* which later developed into INSPEC, a British-based development in the later 1960s with little input from the United States.

4. Freeman and Atherton made an effort in 1968, in the AUDACIOUS Project, to use the Universal Decimal Classification as an indexing and searching tool, but there was a very cool reception from the field partly because the online system used was too cumbersome and the displays were not hypertextual (Cochrane, 1985, pp. 312–370).

5. And the UDC: see Donker Duyvis (1951).

6. I told Eric that I had waited thirty years for someone to devise such an online tool. What I had seen in PaperChase (Atherton, 1982) was a glimmer in the dark sky of future online catalogs, and what I saw in IODyne was a fireworks display!

7. I am reminded that I felt the same way when we failed to carry forward with the Play Thesaurus for ERIC Descriptors (1979). This tool grew out of the Vocabulary Improvement Project. This version of the thesaurus incorporated suggestions for new terms and new relationships from twenty different sources and displayed them all until the lexicographer reviewed, approved, or disapproved them. Although this was in machine-readable form, there was no capability in the maintenance system equal to IODyne. It was only a printed file for the lexicographer at ERIC Central, and the twenty decentralized clearinghouses were not yet connected online. See Booth (1979) for background on VIP, and Cochrane (1985, p. 235) for an illustration.

8. See Meincke's drawings of these axes in our original 1976 paper or in Cochrane (1985, pp. 41–45).

9. Bourne still has all his original sources in storage.

10. The concept of DNA Junk reminded me of another time when, for a moment, I had another insight into a special feature for Knowledge Space. I was interviewing three high school students at University High on the UIUC campus. I asked them what they would like to see in an online version of *World Book Encyclopedia*. They agreed when one of them said he would like to be able to quickly "*skip the stuff I already know* and go immediately to something new about astronomy, the planets, and space travel." In a way the DNA junk is what researchers found new after arranging everything they knew. Could we say the same about everything we know about retrieval system design? Are multidimensional space, search volume, and reference axes part of the new "junk"?

References

Atherton, P. (Ed.). (1965) *Classification research; Proceedings of the Second International Study Conference of FID/CR.* Copenhagen: Munksgaard.

Atherton, P. (1982). Friendly catalog forgives user errors. *American Libraries, 13*(5), 303–306.

Booth, B. (1979). A 'new' ERIC thesaurus, fine-tuned for searching. *Online* (July), 20–29.

Bourne, C. P. (1963). *Methods of information handling.* New York: Wiley.

Bourne, C. P. & Hahn, T. B. (2003). *A history of online information services, 1963–1976.* Cambridge, MA: MIT Press.

Coates, E. J. (1960). *Subject catalogues: Headings and structure.* London: Library Association.

Cochrane, P. A. (1985). *Redesign of catalogs and indexes for improved online subject access; Selected papers.* Phoenix, AZ: Oryx Press.

Donker Duyvis, F. (1951). The UDC: What it is and what it is not. *Revue de la Documentation 18*(2), 99–105.

Doyle, L. B. (1972). Indexing and abstracting by association. *American Documentation, 23* (October), 378–90.

Dwyer, J. (2013, February 15). Athlete's Near-Death From Sepsis, and His Mysterious Recovery. *New York Times* .

Johnson, E. & Atherton, P. (1995). A hypertextual interface for a searcher's thesaurus. In *Digital Libraries '95 Proceedings. Austin, TX June 11–13, 1995* (pp. 77–86).

Kolata, G. (2012, September 6). Study Discovers Road Map of DNA; A Key to Biology. *New York Times.*

Meincke, P. P. M. and Atherton, P. (1976). Knowledge space. *Journal of the American Society for Information Science, 27*(1), 18–24.

Miller, G. (1968). Psychology and information. *American Documentation* (July), 286–9.

Ranganathan, S. R. (1951). *Philosophy of library classification.* Copenhagen: Munksgaard.

Rolling, L. (1965). The role of graphical display of concept relationships in indexing and retrieval vocabularies. In Atherton, Pauline (Ed.), *Classification Research: Proceedings of the Second Annual International Study Conference* (pp. 295–325). Copenhagen: Munksgaard.

Rysik, M. (2012, October 9). Online, a Genome Project for the world of art. *New York Times.*

Sengupta, S. (2013, January 29). For Search, Facebook Had to Go Beyond 'Robospeak.' *New York Times.*

Pauline Atherton Cochrane is professor emerita at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Professor Cochrane began her library science career with an M.A. from Rosary College (now Dominican University) in 1954. Her 50+ year career spans work at the Chicago Public Library and Chicago Teachers College, World Book Encyclopedia, American Institute of Physics, and Syracuse University. She consulted and lectured throughout the world (especially at the United Nations and agencies in Washington DC); conducted research projects to improve catalogs and indexes; created an online UDC-based retrieval system, the first campus-based retrieval system for Psychological Abstracts; strove to improve the ERIC Thesaurus; and as part of the Digital Library Initiative, helped to design a hypertextual thesaurus browser. As an educator she helped to introduce computer-based cataloging and online reference courses in the 1960s and early 1970s. The first twenty-five years of her career are documented in an Oryx Press 1985 book, *Redesign of Catalogs and Indexes for Improved Online Subject Access; Selected Papers.*