# Historiographic mapping of knowledge domains literature

**Eugene Garfield**

*Chairman Emeritus, ISI®*

**Abstract.**

**To better understand the topic of this colloquium, we have created a series of databases related to knowledge domains (dynamic systems [small world/Milgram], information visualization [Tufte], co-citation [Small], bibliographic coupling [Kessler], and scientometrics [*Scientometrics*]). I have used a software package called *HistCite*™ which generates chronological maps of subject (topical) collections resulting from searches of the *ISI Web of Science*® or ISI citation indexes *(SCI, SSCI, and/or AHCI)* on CD-ROM. When a marked list is created on *WoS,* an export file is created which contains all cited references for each source document captured. These bibliographic collections, saved as ASCII files, are processed by *HistCite* in order to generate chronological and other tables as well as historiographs which highlight the most-cited works in and outside the collection. *HistCite* also includes a module for detecting and editing errors or variations in cited references as well as a vocabulary analyzer which generates both ranked word lists and word pairs used in the collection. Ideally the system will be used to help the searcher quickly identify the most significant work on a topic and trace its year-by-year historical development. In addition to the collections mentioned above, historiographs based on collections of papers that cite the Watson-Crick 1953 classic paper identifying the helical structure of DNA were created. Both year-by-year as well as month-by-month displays of papers from 1953 to 1958 were necessary to highlight the publication activity of those years.**

*Correspondence to:* Chairman Emeritus, ISI®, 3501 Market Street, Philadelphia, PA 19104, USA. E-mail: garfield@codex.cis.upenn.edu

I was reluctant to accept Katy Borner's invitation to give this keynote talk since I had never heard the term 'Knowledge Domains' before. Furthermore, I am not an expert on the subject of visualization. Her misperception on that point was probably due to a paper I recently published in the special issue of the *Journal of the American Society for Information Science and Technology* on visualization [1]. The issue editor, Chaomei Chen of Drexel University, had roped me into that contribution since he had heard about my interest in mapping from colleagues Howard White and Kate McCain at Drexel [2].

Over a period of several months, my staff worked with Katy to identify various literature sub-sets she perceived as being relevant to the knowledge domain literature. To facilitate that process, we used a software package still in development called *HistCite*. This system has evolved over the past several years and traces its roots to a project in 1964 conducted by me and Irving Sher [3], who died several years ago, and sponsored by Harold Wooster of the U.S. Air Force. *The uses of citation data in writing the history of science*, is available at my web page at www.eugenegarfield.org. We interested Wooster in the idea when we completed our NIH-sponsored work on the *Genetics Citation Index* project. The *GCI* eventually led to publication of the 1961 volumes of the *Science Citation Index* (*SCI*) in 1964. Sher and I had speculated on the possibility that the cited references in scholarly papers could be used to create topological maps of science. To test this theory, we used Isaac Asimov's book *The Genetic Code* [4] as a model. Asimov, a professor of biochemistry, better known as the prolific science

fiction writer, identified the 40 key scientific events in the development of DNA science from the time of Gregor Mendel until the 1961 Nobel work of Marshall Nirenberg at NIH. We used about 60 published papers mentioned by Asimov to create a mini citation index from the 1000 odd references they cited.

From these data we were able to draw the first citation-based historiograph shown in Figure 1 (http://garfield.library.upenn.edu/papers/finaloverlay.pdf). Our interest in the graph theoretical aspects of citation networks was further reflected in a thesis by Ralph Garner at Drexel University in 1967 [5], an ISI employee at that time.

Each box in this historiograph is a key event. The colored connecting lines indicate various levels or strengths of citation linkage. Two decades later, the DNA project data were used as a model in a paper by two social networks researchers at the University of Pittsburgh, Norman P. Hummon and Patrick Doreian [6]. Except for their work, the original idea was basically ignored until a few years ago when my long-time colleague, geneticist Alexander I. Pudovin, and I discussed the possibility of reviving the original idea of writing a program that would create historio-graphs algorithmically. This led to the *HistCite*™ software described below. The process was first publicly discussed at a University of Pittsburgh conference [7] and then at the ASIST Annual Meeting in November 2002 [8]. The ASIS&T paper includes, among others, a *HistCite* analysis for 'gene flow', an area in population genetics of interest to Pudovkin. From those initial trials, the software has evolved to its present form.

To create a topical *HistCite* collection, one must first conduct a search of the *ISI Web of Science*® or similar database. In the following example a search was conducted in the *WoS* for literature on the 'Small World' problem, by using a combination of cited reference search, that is, papers which cite Stanley Milgram's 1997 paper in *Psychology Today,* and a general title word search on 'Small World'. This led to the creation of a marked list. (See Figure 2.) The arrows indicate several of the variants we included to retrieve about 160 citing papers.

Figure 3 shows one of the retrieved tagged records in the ISI Export Format – a paper by Egghe and Rousseau recently published in *JASIST* which cited the Milgram 1967 paper. Once a marked list is created from *WoS*, it can be exported as a text file in the ISI Export Format. The address for the saved file is used by the software to retrieve the relevant text file.

Within a few moments the system opens up a chronological table.

The intial *HistCite* display is chronological and alphabetical by journal (see Figure 4). The 1967 Milgram paper is the fourth one listed. On the right, the *Local Citation Score* of 167 indicates how often it is cited in the collection. The *Global Citation Score* of 148 is the count reported in *WoS* for the first variant as shown in Figure 2. By adding the other variants in the text file and rerunning the program, the LCS score becomes 167.

Using the mouse, one can then proceed to view a series of tables sorted by local or global citation score, by journal, author, or 'Outer References'. The journal display in Figure 5 shows how this topic is dominated by physics. This was not the case when Stanley Milgram first published. A current examination of the entries from 1967 to the 1990s shows this topic was primarily of interest to psychologists and social scientists. Then from about 1997, most of the literature is dominated by physicists.

Figure 6 ranks authors by number of publications. For example, early pioneers Manfred Kochen and Stanley Milgram only appear as 27th and 31st. It also includes the GCS and LCS totals. If the GCS sort key were used, then Barabasi and Albert would move to the top of the list.

The outer works in Figure 7 are not included in the local collection because they have neither cited Milgram nor have they used 'Small World' as a title word. These cited references are outside the retrieved collection. The sorted list of outer nodes is a virtual citation index of everything cited by the papers in the local or inner collection. The outer works will include any papers or books cited whether or not they are covered in *WoS* as sources. The full bibliographic data for each item can be looked up by clicking on *WoS*. This link takes one into the cited reference section of *WoS*. Note that numbers 3, 5, 7, 8, and 10 are cited books. To include them in the main collection and historiograph, their text records would have to be created manually. This would also be true for any cited papers not included in *WoS*.

In the latest version of the software, it is also possible to sort this file not only by citation frequency but also by journal, author or year. In Figure 8, we see that 110 papers from *Physical Review Letters* have been cited by 77 of the papers in the collection. The journal number is the alphabetical rank.

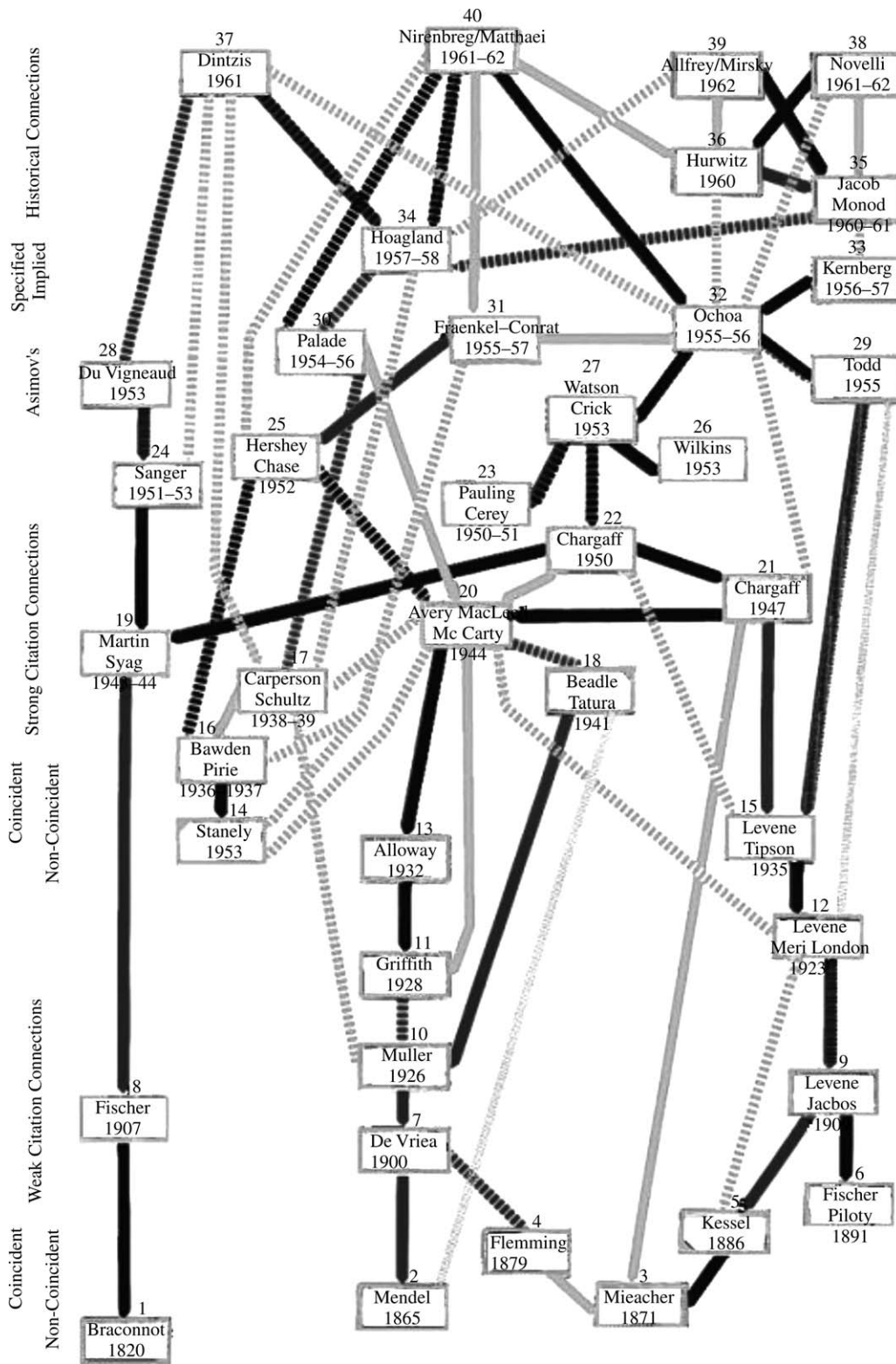Figure 9 shows the outer references to papers in *Nature* sorted chronologically.

Fig. 1. Historiograph: from Mendel to Nirenberg.

**Cited Reference Search**

11 references matched query:
Cited Author=milgram s AND Cited Year=1967
*Database(s)=SCI-EXPANDED, SSCI, A&HCI; Timespan=1945-2003*

**STEP 2: CITED REFERENCE SELECTION**
The table lists all of the cited references that match your search request and the number of times each variation has been cited. Select all desired references (including variants) by clicking the checkboxes or SELECT PAGE. Then press SEARCH. The search is added to the Search History.

Set language and document type limits.

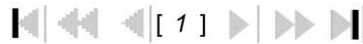SELECT PAGE   or select specific references from list.

SEARCH   to find articles that cite selected references.

References 1 -- 11

| | Hits | Cited Author | Cited Work | Volume | Page | Year |
|---|---|---|---|---|---|---|
| ☐ | 1 | MILGRAM S | MEGAMOT | 15 | 31 | 1967 |
| ☐ | 9 | MILGRAM S | PATTERNS PREJUDICE | 1 | 3 | 1967 |
| ☐ | 1 | MILGRAM S | PATTERNS PREJUDICE | 1 | 5 | 1967 |
| ☐ | 3 | MILGRAM S | POLITISCHE GESELLSCH | | 170 | 1967 |
| ☐ | 1 | MILGRAM S | PSYCHOL TODAY | | | 1967 |
| ☐ | 16 | MILGRAM S | PSYCHOL TODAY | 2 | 60 | 1967 ← |
| ☐ | 148 | MILGRAM S | PSYCHOL TODAY | 1 | 61 | 1967 ← |
| ☐ | 2 | MILGRAM S | PSYCHOL TODAY | 1 | 62 | 1967 ← |
| ☐ | 6 | MILGRAM S | PSYCHOL TODAY    MAY | | 61 | 1967 ← |
| ☐ | 1 | MILGRAM S | PSYCHOLOGY TODAY MAY | | | 1967 ← |
| ☐ | 1 | MILGRAM S | READINGS SOCIAL PSYC | | | 1967 |

*Note: Hits are for all references -- not just for the current database and year selections.*
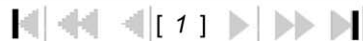
References 1 -- 11

Fig. 2. Cited reference search on Milgram's 1967 *Psychology Today* article.

```
PT Journal
AU Egghe, L
   Rousseau, R
TI A measure for the cohesion of weighted networks
SO JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND
   TECHNOLOGY
NR 34
CR BARABASI AL, 1999, SCIENCE, V286, P509
   BJORNEBORN L, 2001, P 12 ACM C HYP HYP, P133
   BOTAFOGO RA, 1992, ACM T INFORM SYST, V10, P142
   BRIN S, 1998, P 7 INT WORLD WID WE, P107
   CHAKRABARTI S, 1999, SCI AM, V280, P54
   DEBRA P, 2000, SCIENTOMETRICS, V47, P227
   DEVOCHT J, 1994, THESIS EINDHOVEN U T
   EGGHE L, 2002, IN PRESS MATH COMPUT
   FANG Y, 2001, SCIENTOMETRICS, V50, P273
   GARNER R, 1967, COMPUTER ORIENTED GR
   GIBBONS A, 1985, ALGORITHMIC GRAPH TH
   HARARY F, 1969, GRAPH THEORY
   HENZIGNER MR, 2001, IEEE INTERNET COMPUT, V5, P45
   JOHNSON S, 1995, COMMUN ACM, V38, P87
   KHAN KS, 1998, J AM SOC INFORM SCI, V49, P176
   KLEINBERG JM, 1999, J ACM, V46, P604
   KOCHEN M, 1989, SMALL WORLD
   LEAZER GH, 1999, P 62 ANN M AM SOC IN, P345
   MENDES E, 1998, J UNIVERSAL COMPUTER, V4, P4
   MILGRAM S, 1967, PSYCHOL TODAY, V1, P61
   NEWMAN MEJ, 2000, J STAT PHYS, V101, P819
   NEWMAN MEJ, 1999, PHYS REV E B, V60, P7332
   NG AY, 2001, P 17 INT JOINT C ART, P903
   PINSKI G, 1976, INFORMATION PROCESSI, V12, P297
   PRICE DJD, 1965, SCIENCE, V149, P510
   PRITCHARD A, 1984, THESIS POLYTECHNIC N
   RANDIC M, 1975, J AM CHEM SOC, V97, P6609
   RIVLIN E, 1994, COMMUN ACM, V37, P87
   ROUSSEAU R, 1987, SCIENTOMETRICS, V11, P217
   SMALL H, 1973, J AM SOC INFORM SCI, V24, P265
   TRINAJSTIC N, 1992, CHEM GRAPH THEORY
   WATTS DJ, SMALL WORLDS
   WIENER H, 1947, J AM CHEM SOC, V69, P17
   WILSON RJ, 1972, INTRO GRAPH THEORY
TC 0
BP 193
EP 202
PG 10
JI J. Am. Soc. Inf. Sci. Technol.
PY 2003
PD FEB 1
VL 54
IS 3
GA 642YJ
J9 J AM SOC INF SCI TECHNOL
UT ISI:000180834700001
ER
```

Fig. 3. .txt Export record from 'small world' collection.

**Papers citing "Small World Problem" by Milgram S., 1967, *Psychology Today*,
V1, P61 and papers with "Small World" in the title from 1967 to present**

Nodes: 424
Sorted by **year, journal, volume, page.**
Page **1**: 1

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 1 1967 AMERICAN JOURNAL OF SOCIOLOGY 72(4):422–423<br>**VARMA BN**<br>*Small World of Khanh-Hau – Hendry,JB* | 2 | 1 |
| 2 | 1 | 2 1967 AMERICAN JOURNAL OF SOCIOLOGY 73(1):115–115<br>**EVERS HD**<br>*Small World of Khanh-Hau – Comment* | 0 | 0 |
| 3 | 0 | 3 1967 JOURNAL OF DEVELOPING AREAS 1(4):547–549<br>**AUCHTER EL**<br>*Small World of Khanh Hau – Hendry,JB* | 0 | 0 |
| 4 | 0 | 4 1967 PSYCHOLOGY TODAY 1(1):61–67<br>**MILGRAM S**<br>*Small World Problem* | 148 | 167 |
| 5 | 0 | 5 1969 FORTUNE 80(4):121-&<br>**ZALAZNICK S**<br>*Small World Of Big Washington Lawyers* | 0 | 0 |
| 6 | 1 | 6 1969 SOCIOMETRY 32(4):425–443<br>**TRAVERS J; MILGRAM S**<br>*Experimental Study Of Small World Problem* | 64 | 43 |
| 7 | 1 | 7 1970 BRITISH JOURNAL OF SOCIAL PSYCHIATRY 4(2):83–87<br>**HART JW**<br>*Sociometry of Poverty* | 0 | 0 |
| 8 | 0 | 8 1970 ENGLISH JOURNAL 59(3):416–420<br>**SOFFER RS**<br>*Its A Small World* | 1 | 0 |
| 9 | 2 | 9 1970 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 15(2):101-&<br>**KORTE C; MILGRAM S**<br>*Acquaintance Networks between Racial Groups – Application of Small World Method* | 38 | 31 |
| 10 | 2 | 10 1970 SOCIAL FORCES 49(2):259-+<br>**WHITE HC**<br>*Search Parameters for Small World Problem* | 22 | 14 |

Fig. 4. Chronological *Histcite* file of 'Small World' collection.

Small World
Ranked Journal List for Local Collection

Ranked Journal list.
Total: 253
Sorted by **pubs**

| # | Title | Pubs |
|---|---|---|
| 1 | PHYSICAL REVIEW E | 45 |
| 2 | SOCIAL NETWORKS | 13 |
| 3 | PHYSICAL REVIEW LETTERS | 10 |
| 4 | AMERICAN JOURNAL OF SOCIOLOGY | 9 |
| 5 | NATURE | 9 |
| 6 | PHYSICA A | 8 |
| 7 | EUROPEAN PHYSICAL JOURNAL B | 7 |
| 8 | PNAS | 7 |
| 9 | EUROPHYSICS LETTERS | 6 |
| 10 | SOCIAL FORCES | 6 |
| 11 | NEW YORK TIMES BOOK REVIEW | 5 |
| 12 | SCIENCE | 4 |
| 13 | FORBES | 4 |
| 14 | AMERICAN SOCIOLOGICAL REVIEW | 3 |
| 15 | NATION | 3 |
| 16 | PHYSICAL REVIEW B | 3 |
| 17 | JOURNAL OF CHEMICAL PHYSICS | 3 |
| 18 | FORTUNE | 3 |
| 19 | INTERNATIONAL JOURNAL OF BIFURCATION AND CHAOS | 3 |
| 20 | TLS-THE TIMES LITERARY SUPPLEMENT | 3 |
| 21 | PHYSICS LETTERS A | 3 |
| 22 | NEW REPUBLIC | 3 |
| 23 | ENVIRONMENT AND PLANNING A | 3 |
| 24 | PMLA-PUBS OF MODERN LANGUAGE ASSOC OF AMERICA | 2 |
| 25 | AAA-ARBEITEN AUS ANGLISTIK UND AMERIKANISTIK | 2 |
| 26 | CONTEMPORARY SOCIOLOGY-A JOURNAL OF REVIEWS | 2 |
| 27 | LIBRARY TRENDS | 2 |
| 28 | NEW YORK REVIEW OF BOOKS | 2 |
| 29 | HISTOIRE | 2 |
| 30 | LANDSCAPE ARCHITECTURE | 2 |

Fig. 5. Ranked journal list for 'Small World' local collection.

Small World All-Author List for Local "Inner" Nodes

Ranked All-Author list.
Total: 494
Sorted by **pubs**

| # | Name | TGCS | TLCS | Pubs |
|---|------|------|------|------|
| 1 | [Anon] | 0 | 0 | 15 |
| 2 | Newman MEJ | 357 | 176 | 12 |
| 3 | Blumen A | 71 | 31 | 9 |
| 4 | Bernard HR | 85 | 24 | 8 |
| 5 | Killworth PD | 85 | 24 | 8 |
| 6 | Barabasi AL | 624 | 109 | 6 |
| 7 | Kim BJ | 18 | 0 | 6 |
| 8 | Watts DJ | 555 | 238 | 6 |
| 9 | Albert R | 621 | 109 | 5 |
| 10 | Amaral LAN | 243 | 98 | 5 |
| 11 | Kuperman M | 31 | 16 | 5 |
| 12 | Mccarty C | 59 | 11 | 5 |
| 13 | Barthelemy M | 206 | 90 | 4 |
| 14 | Bochner S | 23 | 10 | 4 |
| 15 | Choi MY | 9 | 0 | 4 |
| 16 | Elgazzar AS | 2 | 0 | 4 |
| 17 | Jasch F | 9 | 1 | 4 |
| 18 | Jeong H | 263 | 48 | 4 |
| 19 | Kertesz J | 16 | 0 | 4 |
| 20 | Zanette DH | 6 | 0 | 4 |
| 21 | Abramson G | 31 | 16 | 3 |
| 22 | Ahmed E | 3 | 1 | 3 |
| 23 | Holme P | 14 | 0 | 3 |
| 24 | Hong H | 8 | 0 | 3 |
| 25 | Hunter Je | 17 | 7 | 3 |
| 26 | Jespersen S | 50 | 28 | 3 |
| 27 | Kochen M | 4 | 1 | 3 |
| 28 | Lai YC | 0 | 0 | 3 |
| 29 | Latora V | 11 | 4 | 3 |
| 30 | Marchiori M | 11 | 4 | 3 |
| 31 | Milgram S | 250 | 241 | 3 |

Fig. 6. All-author list ranked by publications.

Figure 10 provides a ranked author listing for outer references. 61 papers by Milgram have been cited, including variants. Only 18 citing papers are involved.

In Figure 11, the outer references have been sorted by year and then alphabetically by author. The 1967 paper by R. P. Abelson has been cited in three papers but the first of them has omitted the page number.

Figure 12 presents the final product of the *HistCite* program, namely the historiograph for the most-cited papers in the local collection. The citation links between papers appear as a line with an arrow. Thus, paper 6 cited paper 4. The threshold number of papers included is set by the user; in this case, 17. The sizes of the circles are proportional to citation frequency. The Milgram paper is at the top (node #4) while the 1998 paper by Watts (node #194) is at the center. Had we mapped these papers based on global score, the Watts paper would appear even larger due to its higher global citation score.

Having illustrated the basic format of *HistCite*, let us now examine a collection on 'Knowledge Domains'. We conducted a *WoS* general search for papers on 'knowledge and domain*' and found 280 papers.

The first paper listed appeared in 1977 in the journal *Communication and Cognition*. The author is J. B. Grize (see Figure 13). It is interesting that this paper has neither been cited in the collection nor by the author himself. The gray areas indicate papers that are not cited in the collection, nor do they cite anything in the collection, which may be a criterion for eliminating papers of low relevance.

The paper most-cited within this small collection is the one by Alexander and Judy in *Review of Educational Research* (see Figure 14). In addition to the nine local cites, the paper has been cited globally in 114 papers.

When the file is sorted by GCS a paper by J. M. Berg turns up but it is really not related to the main knowledge domain theme and should be deleted (see Figure 15). Note that it is not cited at all in the local collection nor does it cite anything – hence the gray area. The next paper by Spilich et al is indeed relevant and is cited by seven papers in the file.

In the latest version of the software it is possible to obtain a vocabulary analysis of the title words in the collection (see Figure 16). This provides clues for expanding the *WoS* search.

Depending upon the size of the files involved, it may be more useful to view the permutations of title word pairs (see Figure 17). This is reminiscent of the *Permuterm Subject Index* in the print editions of the *Science Citation Index*.

Cited references outside of this network
Total: 6719 (top 100 shown)
Sorted by **LCS**.

| # | **LCS** | **Reference** |
|---|---|---|
| 1 | 45 | ALBERT R, 1999, NATURE, V401, P130 WoS |
| 2 | 45 | NEWMAN MEJ, 2001, PHYS REV E 2, V64 WoS |
| 3 | 41 | BOLLOBAS B, 1985, RANDOM GRAPHS WoS |
| 4 | 39 | GRANOVETTER MS, 1973, AM J SOCIOL, V78, P1360 WoS |
| 5 | 30 | WATTS DJ, 1999, SMALL WORLDS WoS |
| 6 | 26 | MOUKARZEL CF, 1999, PHYS REV E, V60, P6263 WoS |
| 7 | 25 | WASSERMAN S, 1994, SOCIAL NETWORK ANAL WoS |
| 8 | 23 | KOCHEN M, 1989, SMALL WORLD WoS |
| 9 | 22 | LAGOFERNANDEZ LF, 2000, PHYS REV LETT, V84, P2758 WoS |
| 10 | 22 | WATTS DJ, 1999, SMALL WORLDS DYNAMIC WoS |
| 11 | 21 | JEONG H, 2000, NATURE, V407, P651 WoS |
| 12 | 21 | ERDOS P, 1960, PUBL MATH I HUNG, V5, P17 WoS |
| 13 | 19 | REDNER S, 1998, EUR PHYS J B, V4, P131 WoS |
| 14 | 15 | GUARE J, 1990, 6 DEGREES SEPARATION WoS |
| 15 | 14 | FALOUTSOS M, 1999, COMP COMM R, V29, P251 WoS |
| 16 | 13 | POOL ID, 1978, SOC NETWORKS, V1, P5 WoS |
| 17 | 12 | DEMENEZES MA, 2000, EUROPHYS LETT, V50, P574 WoS |
| 18 | 12 | HUBERMAN BA, 1999, NATURE, V401, P131 WoS |
| 19 | 11 | BRODER A, 2000, COMPUT NETW, V33, P309 WoS |
| 20 | 10 | MITCHELL JC, 1969, SOCIAL NETWORKS URBA WoS |
| 21 | 10 | POOL ID, 1978, SOC NETWORKS, V1, P1 WoS |
| 22 | 10 | MOLLOY M, 1995, RANDOM STRUCT ALGOR, V6, P161 WoS |
| 23 | 9 | GRANOVETTER MS, 1974, GETTING JOB STUDY CO WoS |
| 24 | 9 | WILLIAMS RJ, 2000, NATURE, V404, P180 WoS |
| 25 | 9 | DOROGOVTSEV SN, 2001, PHYS REV E 1, V63 WoS |

Fig. 7.  Outer nodes – most-cited works outside local collection.

**Small World**
**Ranked Outer Journal List**

Ranked Outer Journal list.
Total:1087
Sorted by **PUBS**
50 shown.

| Num | Journal NUM | Title | Cited Pubs | TLCS | TLCS/t | CITING NODES |
|---|---|---|---|---|---|---|
| 1 | 838 | PHYS-REV-LETT | 110 | 220 | 46.17 | 77 |
| 2 | 67 | AM-SOCIOL-REV | 109 | 145 | 6.60 | 38 |
| 3 | 59 | AM-J-SOCIOL | 91 | 173 | 8.03 | 55 |
| 4 | 746 | NATURE | 91 | 212 | 42.66 | 85 |
| 5 | 942 | SCIENCE | 66 | 108 | 16.60 | 57 |
| 6 | 480 | J-AM-SOC-INFORM-SCI | 45 | 58 | 6.87 | 11 |
| 7 | 943 | SCIENTOMETRICS | 43 | 49 | 7.30 | 9 |
| 8 | 960 | SOC-FORCES | 35 | 45 | 1.89 | 22 |
| 9 | 963 | SOC-NETWORKS | 35 | 72 | 3.77 | 40 |
| 10 | 607 | J-PERS-SOC-PSYCHOL | 34 | 36 | 1.50 | 13 |
| 11 | 792 | P-NATL-ACAD-SCI-USA | 34 | 45 | 9.03 | 26 |
| 12 | 832 | PHYS-REV-E | 28 | 64 | 12.25 | 45 |
| 13 | 503 | J-CHEM-PHYS | 27 | 31 | 3.24 | 13 |
| 14 | 526 | J-DOC | 27 | 31 | 4.84 | 2 |
| 15 | 1005 | SOCIOMETRY | 27 | 43 | 1.08 | 23 |
| 16 | 831 | PHYS-REV-B | 24 | 34 | 3.34 | 15 |
| 17 | 64 | AM-PSYCHOL | 22 | 25 | 1.01 | 11 |
| 18 | 840 | PHYSICA-A | 22 | 28 | 5.99 | 17 |
| 19 | 17 | ADMIN-SCI-QUART | 20 | 22 | 2.51 | 9 |
| 20 | 836 | PHYS-REV-E-A | 19 | 29 | 5.93 | 19 |
| 21 | 766 | NUCLEIC-ACIDS-RES | 18 | 23 | 5.44 | 6 |
| 22 | 16 | ADM-SCI-Q | 17 | 18 | 0.60 | 8 |
| 23 | 582 | J-MATH-SOCIOL | 17 | 33 | 1.53 | 15 |
| 24 | 892 | PUBLIC-OPIN-QUART | 17 | 20 | 0.55 | 12 |
| 25 | 643 | J-SOC-ISSUES | 16 | 18 | 0.56 | 9 |

Fig. 8. 'Small World' outer references ranked by cited journal publications.

**Small World**
**Outer References to *Nature***

Cited references outside of this network.
Total: 6713 (top 100 shown)

| Num | OR Num | LCS | Reference/Journal/Year | LCS/t |
|---|---|---|---|---|
| 1 | 1716 | 1 | DOMB-C-1959-NATURE-V184-P509 | 0.02 |
| 2 | 2453 | 1 | GOFFMAN-W-1966-NATURE-V212-P449 | 0.03 |
| 3 | 6648 | 1 | ZAIKIN-AN-1970-NATURE-V225-P535 | 0.03 |
| 4 | 5020 | 1 | PIMM-SL-1977-NATURE-V268-P329 | 0.04 |
| 5 | 5214 | 1 | REJMANEK-M-1979-NATURE-V280-P311 | 0.04 |
| 6 | 179 | 1 | ANDERSON-RM-1982-NATURE-V318-P323 | 0.05 |
| 7 | 181 | 1 | ANDERSON-RM-1985-NATURE-V318-P323 | 0.05 |
| 8 | 2584 | 1 | GRAY-CM-1989-NATURE-V338-P334 | 0.07 |
| 9 | 6174 | 2 | VANRAAN-AFJ-1990-NATURE-V347-P626 | 0.14 |
| 10 | 520 | 1 | BENTLEY-GA-1990-NATURE-V348-P254 | 0.07 |
| 11 | 5024 | 4 | PIMM-SL-1991-NATURE-V350-P669 | 0.31 |
| 12 | 4861 | 2 | PAINE-RT-1992-NATURE-V355-P73 | 0.17 |
| 13 | 4770 | 4 | NOWAK-MA-1992-NATURE-V359-P826 | 0.33 |
| 14 | 2729 | 1 | HAMERSCASTERMAN-C-1993-NATURE-V363-P446 | 0.09 |
| 15 | 4769 | 1 | NOWAK-M-1993-NATURE-V364-P56 | 0.09 |
| 16 | 1783 | 2 | DOUGLASS-JK-1993-NATURE-V365-P337 | 0.18 |
| 17 | 2533 | 2 | GOULD-SJ-1993-NATURE-V366-P223 | 0.18 |
| 18 | 5818 | 1 | STEINBOCK-O-1993-NATURE-V366-P322 | 0.09 |
| 19 | 852 | 1 | BRAUN-HA-1994-NATURE-V367-P270 | 0.10 |
| 20 | 2819 | 1 | HASSELL-MP-1994-NATURE-V370-P290 | 0.10 |
| 21 | 6473 | 1 | WIESENFELD-K-1995-NATURE-V373-P33 | 0.11 |
| 22 | 1389 | 2 | COLLINS-JJ-1995-NATURE-V376-P236 | 0.22 |
| 23 | 584 | 1 | BEZRUKOV-SM-1995-NATURE-V378-P362 | 0.11 |
| 24 | 842 | 2 | BRAIMAN-Y-1995-NATURE-V378-P465 | 0.22 |
| 25 | 2206 | 1 | FRETTE-V-1996-NATURE-V379-P49 | 0.12 |

Fig. 9. 'Small World' outer references to *Nature* ranked by cited year.

**Small World**
**Outer Authors Ranked by Pubs Cited**

Ranked Outer author list.
Sorted by **Cited pubs**
4266 Total
Top 50 shown

| Num | Name | TLCS | TLCS/t | CITED PUBS | CITING NODES |
|---|---|---|---|---|---|
| 1 | COLLINE-P | 70 | 0.75 | 70 | 1 |
| **2** | **MILGRAM-S** | **72** | **2.39** | **61** | **18** |
| 3 | NEWMAN-MEJ | 88 | 20.57 | 32 | 45 |
| 4 | DOROGOVTSEV-SN | 59 | 12.83 | 31 | 23 |
| 5 | BURT-RS | 48 | 2.70 | 26 | 21 |
| 6 | GARFIELD-E | 34 | 1.70 | 25 | 9 |
| 7 | LODGE-D | 38 | 0.83 | 19 | 24 |
| 8 | COLEMAN-JS | 31 | 1.07 | 18 | 19 |
| 9 | WELLMAN-B | 23 | 1.52 | 18 | 14 |
| 10 | BERNARD-HR | 22 | 1.20 | 17 | 14 |
| 11 | GRANOVETTER-M | 33 | 1.45 | 17 | 26 |
| 12 | BARRAT-A | 26 | 2.52 | 16 | 18 |
| 13 | WHITE-HC | 21 | 0.56 | 16 | 10 |
| 14 | CRONIN-B | 15 | 2.17 | 15 | 2 |
| 15 | LAUMANN-EO | 29 | 1.06 | 15 | 18 |
| 16 | ALBERT-R | 66 | 12.15 | 14 | 50 |
| 17 | BLAU-PM | 21 | 0.83 | 14 | 10 |
| 18 | FREEMAN-LC | 28 | 1.35 | 14 | 20 |
| 19 | ZANETTE-DH | 20 | 5.54 | 14 | 14 |
| 20 | HOLLAND-PW | 24 | 0.81 | 13 | 12 |
| 21 | KLEINBERG-J | 14 | 3.26 | 13 | 9 |
| 22 | GOULD-P | 12 | 0.36 | 12 | 2 |
| 23 | MARSDEN-PV | 26 | 1.49 | 12 | 15 |
| 24 | SCHWEIZER-T | 12 | 1.15 | 12 | 1 |
| 25 | SMALL-H | 24 | 1.91 | 12 | 5 |

Fig. 10. 'Small World' outer authors ranked by publications cited.

**SMALL WORLD**
**Cited references outside of this network**
**Sorted by Year 1967**

Total: 6713 (top 25 shown)

| Num | OR Num | LCS | Reference/Journal/Year | LCS/t |
|---|---|---|---|---|
| 1 | 7 | 1 | ABELSON-RP-1967-ADV-EXPT-SOCIAL-PSYC-V3 | 0.03 |
| 2 | 8 | 2 | ABELSON-RP-1967-ADV-EXPT-SOCIAL-PSYC-V3-P1 | 0.05 |
| 3 | 196 | 1 | ANTONOVSKY-A-1967-J-NEUROL-NEUROSUR-PS-V30-P1 | 0.03 |
| 4 | 477 | 1 | BECKER-E-1967-BEYOND-ALIENATION | 0.03 |
| 5 | 535 | 2 | BERGER-PL-1967-SOCIAL-CONSTRUCTION | 0.05 |
| 6 | 617 | 1 | BIRKHOFF-G-1967-LATTICE-THEORY | 0.03 |
| 7 | 659 | 1 | BLAU-PM-1967-AM-OCCUPATIONAL-STRU | 0.03 |
| 8 | 673 | 1 | BLISHEN-BR-1967-CANADIAN-REV-SOCIOLO-V4-P41 | 0.03 |
| 9 | 904 | 1 | BROWN-D-1967-MOBILE-PROFESSORS | 0.03 |
| 10 | 947 | 1 | BULLOUGH-BL-1967-AM-J-SOCIOL-V70-P469 | 0.03 |
| 11 | 1095 | 1 | CARZO-R-1967-FORMAL-ORGANIZATIONS | 0.03 |
| 12 | 1287 | 1 | COLE-S-1967-AM-SOCIOL-REV-V32-P377 | 0.03 |
| 13 | 1419 | 1 | COOPERSMITH-S-1967-ANTECEDENTS-SELF-EST | 0.03 |
| 14 | 1503 | 1 | CROSS-HA-1967-PSYCHON-SCI-V7-P233 | 0.03 |
| 15 | 1571 | 8 | DAVIS-J-1967-HUM-RELAT-V20-P181 | 0.22 |
| 16 | 1718 | 1 | DOMHOFF -GW-1967-WHO-RULES-AM | 0.03 |
| 17 | 1719 | 1 | DOMHOFF-GW-1967-WHO-RULES-AM-P54 | 0.03 |
| 18 | 1794 | 1 | DREW-EB-1967-ATLANTIC-MONTHLY-V220-P75 | 0.03 |
| 19 | 1807 | 1 | DUFFIN-R-1967-GEOMETRIC-PROGRAMMIN | 0.03 |
| 20 | 1844 | 1 | EASTON-D-1967-AM-POLIT-SCI-REV-V61-P25 | 0.03 |
| 21 | 1887 | 1 | EISENSTADT-SN-1967-ISRAELI-SOC-BACKGROU | 0.03 |
| 22 | 2096 | 1 | FISHER-CS-1967-ARCH-EUR-SOCIOL-V8-P216 | 0.03 |
| 23 | 2098 | 1 | FISHER-ME-1967-REP-PROGR-PHYS-V30-P615 | 0.03 |
| 24 | 2150 | 1 | FOUCAULT-M-1967-MADNESS-CIVILIZATION | 0.03 |
| 25 | 2207 | 1 | FRIED-M-1967-J-AM-I-PLANNERS-V33 | 0.03 |

Fig. 11. 'Small World' outer references for 1967 sorted by author.
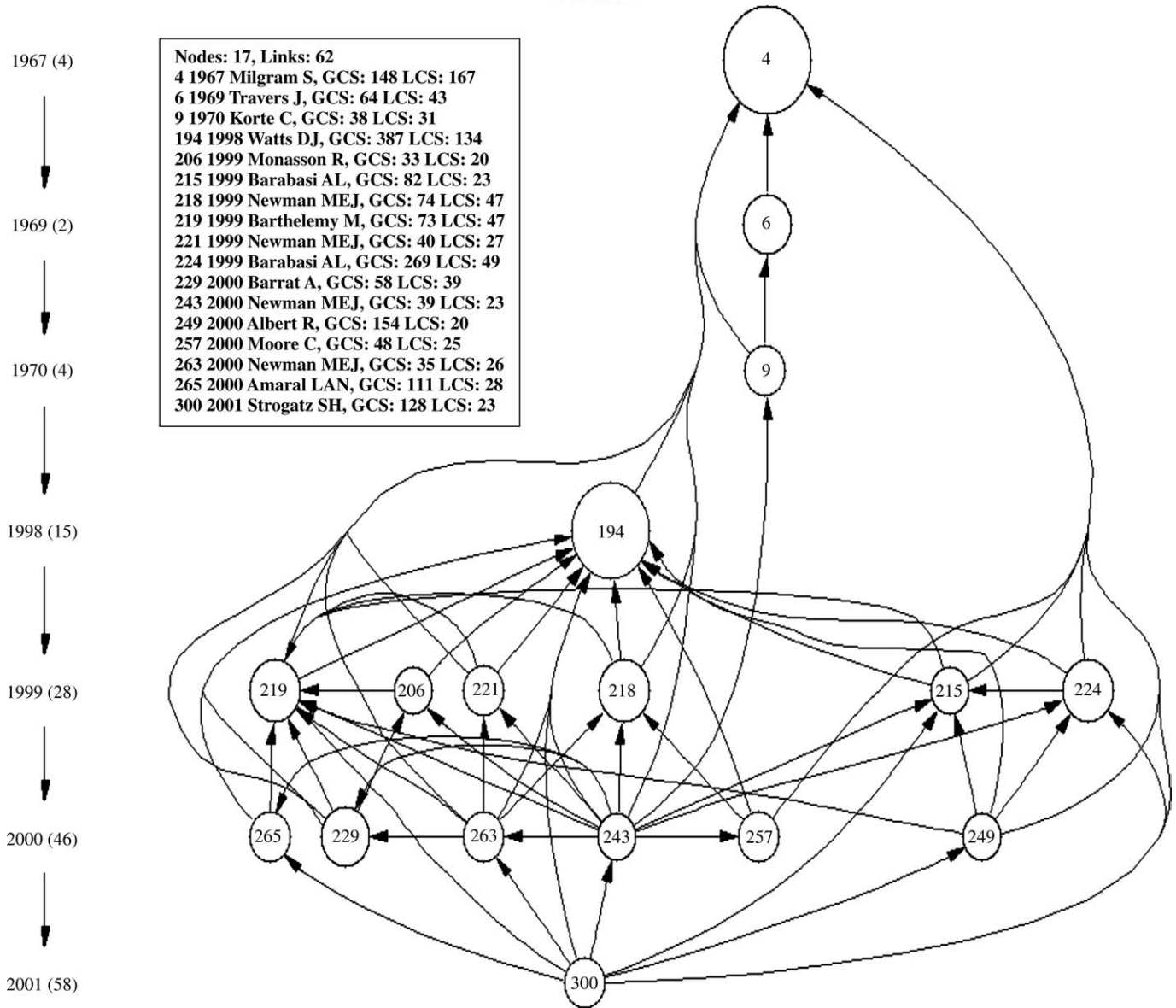
HISTORIOGRAPH OF SAMLL WORLD LCS PAPERS
1967–2002

1967 (4)

**Nodes: 17, Links: 62**
**4 1967 Milgram S, GCS: 148 LCS: 167**
**6 1969 Travers J, GCS: 64 LCS: 43**
**9 1970 Korte C, GCS: 38 LCS: 31**
**194 1998 Watts DJ, GCS: 387 LCS: 134**
**206 1999 Monasson R, GCS: 33 LCS: 20**
**215 1999 Barabasi AL, GCS: 82 LCS: 23**
**218 1999 Newman MEJ, GCS: 74 LCS: 47**
**219 1999 Barthelemy M, GCS: 73 LCS: 47**
**221 1999 Newman MEJ, GCS: 40 LCS: 27**
**224 1999 Barabasi AL, GCS: 269 LCS: 49**
**229 2000 Barrat A, GCS: 58 LCS: 39**
**243 2000 Newman MEJ, GCS: 39 LCS: 23**
**249 2000 Albert R, GCS: 154 LCS: 20**
**257 2000 Moore C, GCS: 48 LCS: 25**
**263 2000 Newman MEJ, GCS: 35 LCS: 26**
**265 2000 Amaral LAN, GCS: 111 LCS: 28**
**300 2001 Strogatz SH, GCS: 128 LCS: 23**

1969 (2)

1970 (4)

1998 (15)

1999 (28)

2000 (46)

2001 (58)

Fig. 12. Historiograph of 17 'Small World' key LCS papers most cited in the 1967–2002 collection.

Outer References Missing Link? Journal list All-Author list Citation Matrix Graphs          HistCite Guide

**Articles containing "knowledge" and "domain*" in the title**

Nodes: 280
Sorted by **year, journal, volume, page.**
Page **1**: 1

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 1 1977 COMMUNICATION AND COGNITION 10(2):93–95 **GRIZE JB** *Domains Of Knowledge* | 0 | 0 |
| 2 | 1 | 2 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):257–273 **CHIESI HL; SPILICH GJ; VOSS JF** *Acquisition of Domain-Related Information in Relation To High and Low Domain Knowledge* | 199 | 8 |
| 3 | 1 | 3 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):275–290 **SPILICH GJ; VESONDER GT; CHIESI HL; VOSS JF** *Text Processing of Domain-Related Information for Individuals with High and Low Domain Knowledge* | 222 | 7 |
| 4 | 0 | 4 1979 SCIENTOMETRICS 1(2):171–193 **CHUBIN DE; STUDER KE** *Knowledge and Structures of Scientific Growth–Measurement of a Cancer Problem Domain* | 11 | 0 |
| 5 | 0 | 5 1982 BULLETIN OF THE PSYCHONOMIC SOCIETY 20(3):148–148 **POST T; BRUDER G; GREENE T; VOSS JF** *Domain Knowledge and Priming Effects on Sentence Recognition Time* | 0 | 0 |
| 6 | 0 | 6 1983 BULLETIN OF THE PSYCHONOMIC SOCIETY 21(5):346–346 **VOSS JF; POST TA; FINCHER RH; GREENE TR** *Relation of Domain Knowledge and Working Memory in Text–Processing* | 0 | 0 |
| 7 | 0 | 7 1983 SCANDINAVIAN JOURNAL OF PSYCHOLOGY 24(1):89–91 **OHLSSON S** *On Natural and Technical Knowledge Domains* | 4 | 0 |

Fig. 13. *HistCite* chronological table 'Knowledge and Domain*' papers.

Outer References Missing Link? Journal list All-Author list Citation Matrix Graphs

HistCite Guide

**Articles containing "knowledge" and "domain*" in the title**

Nodes: 280
Sorted by **LCS.**
Page **1**: 1

| # | Cited nodes | Nodes/Author | GCS | LCS |
|---|---|---|---|---|
| 1 | 1 | 31 1988 REVIEW OF EDUCATIONAL RESEARCH 58(4):375–404 **ALEXANDER PA; JUDY JE** *The Interaction of Domain-Specific and Strategic Knowledge in Academic Performance* | 114 | 9 |
| 2 | 1 | 2 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):257–273 **CHIESI HL; SPILICH GJ; VOSS JF** *Acquisition of Domain-Related Information in Relation to High and Low Domain Knowledge* | 199 | 8 |
| 3 | 1 | 3 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):275–290 **SPILICH GJ; VESONDER GT; CHIESI HL; VOSS JF** *Text Processing of Domain-Related Information for Individuals with High and Low Domain Knowledge* | 222 | 7 |
| 4 | 0 | 14 1986 HUMAN LEARNING 5(2):75–90 **HASSELHORN M; KORKEL J** *Metacognitive Versus Traditional Reading Instructions – The Mediating Role of Domain-Specific Knowledge on Children's Text- Processing* | 24 | 5 |
| 5 | 0 | 18 1987 INTERNATIONAL JOURNAL OF MAN-MACHINE STUDIES 26(1):105–121 **MUSEN MA; FAGAN LM; COMBS DM; SHORTLIFFE EH** *Use of a Domain Model to Drive an Interactive Knowledge-Editing Tool* | 59 | 4 |
| 6 | 2 | 36 1989 JOURNAL OF EDUCATIONAL PSYCHOLOGY 81(3):306–312 **SCHNEIDER W; KORKEL J; WEINERT FE** *Domain-Specific Knowledge and Memory Performance – A Comparison of High-Aptitude and Low-Aptitude Children* | 48 | 4 |
| 7 | 5 | 67 1992 EDUCATIONAL PSYCHOLOGIST 27(1):33–51 **ALEXANDER PA** *Domain Knowledge – Evolving Themes And Emerging Concerns* | 43 | 4 |
| 8 | 2 | 16 1986 JOURNAL OF MEMORY AND LANGUAGE 25(4):431–444 **MCCUTCHEN D** *Domain Knowledge and Linguistic Knowledge in the Development of Writing Ability* | 55 | 3 |

Fig. 14. 'Knowledge Domain' papers sorted by LCS.

Outer References Missing Links? Journal list All-Author list Citation Matrix Graphs      HistCite Guide

**Articles containing "knowledge" and "domain*" in the title**

Nodes: 280
Sorted by **GCS.**
Page **1:** 1

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 39 1990 ANNUAL REVIEW OF BIOPHYSICS AND BIOPHYSICAL CHEMISTRY 19( ):405–421 **BERG JM** *Zinc Finger Domains - Hypotheses and Current Knowledge* | 246 | 0 |
| 2 | 1 | 3 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):275–290 **SPILICH GJ; VESONDER GT; CHIESI HL; VOSS JF** *Text Processing of Domain-Related Information for Individuals with High and Low Domain Knowledge* | 222 | 7 |
| 3 | 1 | 2 1979 JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18(3):257–273 **CHIESI HL; SPILICH GJ; VOSS JF** *Acquisition of Domain-Related Information in Relation to High and Low Domain Knowledge* | 199 | 8 |
| 4 | 0 | 183 1998 JOURNAL OF COGNITIVE NEUROSCIENCE 10(1):1–34 **CARAMAZZA A; SHELTON JR** *Domain-specific knowledge systems in the brain: The animate-inanimate distinction* | 158 | 2 |
| 5 | 1 | 31 1988 REVIEW OF EDUCATIONAL RESEARCH 58(4):375–404 **ALEXANDER PA; JUDY JE** *The Interaction of Domain-Specific and Strategic Knowledge in Academic Performance* | 114 | 9 |
| 6 | 0 | 38 1989 MERRILL-PALMER QUARTERLY-JOURNAL OF DEVELOPMENTAL PSYCHOLOGY 35(1):27–62 **CHI MTH; HUTCHINSON JE; ROBIN AF** *How Inferences about Novel Domain-Related Concepts Can Be Constrained by Structured Knowledge* | 71 | 1 |
| 7 | 0 | 18 1987 INTERNATIONAL JOURNAL OF MAN-MACHINE STUDIES 26(1):105–121 **MUSEN MA; FAGAN LM; COMBS DM; SHORTLIFFE EH** *Use of a Domain Model to Drive an Interactive Knowledge-Editing Tool* | 59 | 4 |
| 8 | 2 | 16 1986 JOURNAL OF MEMORY AND LANGUAGE 25(4):431–444 **MCCUTCHEN D** *Domain Knowledge and Linguistic Knowledge in the Development of Writing Ability* | 55 | 3 |

Fig. 15. 'Knowledge Domain' papers sorted by GCS.

**RANKED KNOWLEDGE DOMAIN
TITLE WORDS LISTING**

Words dictionary.
Total No.: 843
Sorted by: **frequency.**
Top 25 shown.

| Num | Word | raw/freq. % |
|---|---|---|
| 1 | KNOWLEDGE | 265/11.726 |
| 2 | DOMAIN | 172/7.611 |
| 3 | DOMAINS | 56/2.478 |
| 4 | DOMAIN-SPECIFIC | 30/1.327 |
| 5 | USING | 28/1.239 |
| 6 | ACQUISITION | 27/1.195 |
| 7 | SYSTEMS | 22/0.973 |
| 8 | EXPERT | 21/0.929 |
| 9 | SYSTEM | 20/0.885 |
| 10 | LEARNING | 19/0.841 |
| 11 | KNOWLEDGE-BASED | 19/0.841 |
| 12 | DEVELOPMENT | 14/0.619 |
| 13 | ANALYSIS | 14/0.619 |
| 14 | USE | 13/0.575 |
| 15 | TEXT | 13/0.575 |
| 16 | MODEL | 13/0.575 |
| 17 | COMPLEX | 13/0.575 |
| 18 | PROBLEM | 11/0.487 |
| 19 | REPRESENTATION | 10/0.442 |
| 20 | MEDICAL | 10/0.442 |
| 21 | DISCOVERY | 10/0.442 |
| 22 | DESIGN | 10/0.442 |
| 23 | APPLICATION | 10/0.442 |
| 24 | MODELS | 9/0.398 |
| 25 | MODELING | 9/0.398 |

Fig. 16. 'Knowledge Domain' title words listing ranked by frequency.

The historiograph (see Figure 18) for this collection illustrates the problem of conducting a search based solely on the terms knowledge and domain. In a collection of 280 papers, only eight were cited locally three or more times. Based on this simplistic approach, the terms used are inadequate to the task. Clearly the scope of the field had to be interpreted differently. That task was left to Katy Borner. Recognizing this need for subjectivity, in order to capture literature relevant to the 'mapping of knowledge domains,' Borner and her student Raghaveer Mukkamalla decided to create a multi-domain collection. They used five different *HistCite* data sets covering the topics of information visualization (Ed Tufte), dynamic systems (Stanley

Milgram), co-citation (Henry Small), bibliographic coupling (Kessler), and *Scientometrics* (the journal). Once these data sets were merged, about 6000 papers were collected. However, by filtering out papers based on citation and other criteria, the file was reduced to about 3600 papers.

Figure 19 shows the *HistCite* file, sorted by LCS, for the 5643 papers included in the multi-domain file. Note that manually created entries for the Tufte books include psuedo journal, volume, and issue numbers in cited nodes 1, 2, and 5.

The limitations of space preclude a detailed explanation here of how Borner compiled the collection of knowledge domain literature used to create the map in the next figure. The description of that process will be found at http://garfield.library.upenn.edu/histcomp/ multi-domain_master-file/k_borner_supplement.doc (four MB).

Figure 20 is but one of several maps at different citation thresholds which can be found at http:// garfield.library.upenn.edu/histcomp/multi-domai- n_extra-graphs/graph/list.html

In Figure 20, the paper at the top, #2, is Kessler 1963. The map is dominated by papers in bibliometrics. Directly below is #97, Henry Small's 1973 paper. To the right of the map is paper #39 by Milgram, followed by Travers, #60, Granovetter, #93, and Watts 1998, #3976. On the far right near the bottom is the work of Tufte, #480, and Cleveland, #540. In order to see additional papers, we must change the threshold. See the following URLs for larger maps at lower LCS citation thresholds:

LCS > 45 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/2.html

LCS > 39 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/3.html

LCS > 35 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/3.html

LCS > 31 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/4.html

LCS > 30 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/5.html

LCS > 24 - http://garfield.library.upenn.edu/histcomp/ multi-domain_extra-graphs/graph/6.html

Since *HistCite* was originally developed as a tool for historiographic analysis, the closing example is related to the original DNA project. However, in Figure 21, the focus is on the much discussed 1953 paper by Watson and Crick [9] which identified the helical structure of DNA. This paper has been explicitly cited 2316 times in the past 50 years [10]. We have found, however, that for extremely active fields with high immediacy, it is necessary to do the historical analysis in five-year

**KNOWLEDGE DOMAIN
TITLE WORD PAIRS**

Word pairs dicitonary.
Total No.: 1480
Sorted by: **frequency.**
Top 25 shown.

| Num | pair | raw/freq. % |
|---|---|---|
| 1 | DOMAIN & KNOWLEDGE | 98/4.949 |
| 2 | DOMAIN-SPECIFIC & KNOWLEDGE | 23/1.162 |
| 3 | DOMAINS & KNOWLEDGE | 22/1.111 |
| 4 | ACQUISITION & KNOWLEDGE | 20/1.010 |
| 5 | DOMAIN & USING | 12/0.606 |
| 6 | EXPERT & SYSTEMS | 10/0.505 |
| 7 | KNOWLEDGE & REPRESENTATION | 6/0.303 |
| 8 | KNOWLEDGE & PRIOR | 6/0.303 |
| 9 | DOMAIN & USE | 6/0.303 |
| 10 | DOMAIN & SPECIFIC | 6/0.303 |
| 11 | DOMAIN & PUBLIC | 6/0.303 |
| 12 | DOMAIN & MODEL | 6/0.303 |
| 13 | DISCOVERY & KNOWLEDGE | 6/0.303 |
| 14 | COMPLEX & SKILLS | 6/0.303 |
| 15 | COMPLEX & KNOWLEDGE | 6/0.303 |
| 16 | BASES & KNOWLEDGE | 6/0.303 |
| 17 | PLANNING & PUBLIC | 5/0.253 |
| 18 | INSTRUCTION & LEARNING | 5/0.253 |
| 19 | ELICITATION & KNOWLEDGE | 5/0.253 |
| 20 | CONTENT & DOMAIN | 5/0.253 |
| 21 | ACTION & KNOWLEDGE | 5/0.253 |
| 22 | PROBLEM & SOLVING | 4/0.202 |
| 23 | KNOWLEDGE & STRUCTURES | 4/0.202 |
| 24 | KNOWLEDGE & SHARING | 4/0.202 |
| 25 | KNOWLEDGE & MEDICAL | 4/0.202 |

Fig. 17. 'Knowledge Domain' title word pairs ranked by frequency.

segments. In that way, the significant works of the period under study can be seen in their proper chronological context.

Using 50-year citation counts gives greater emphasis to more recent highly-cited papers. After a few decades, a landmark paper will become the common wisdom of the field and citations to it will rarely occur.

To overcome this distortion we created a *HistCite* file of the 200 odd papers published from 1953 to 1958 which cited the Watson-Crick paper. We added a few additional key papers from the outer references such as Avery-McCarty-McLeod 1944 [11] and Hershey 1952 [12].

Figure 22 provides the year-by-year mapping of the most-cited papers. This map makes it visually apparent that many well-cited papers published in 1953 were related to Watson-Crick. They appear on the map in a single row for 1953. This map includes a dozen of the earlier key papers, including Avery et al. 1944 which is so important in the history of DNA. In the first iteration, it appeared as an outer node as did Hershey 1952 and other key authors, all of whom were frequently cited by the authors who cited Watson and Crick. While this is interesting, the year-by-year display does not show the rapid month-by-month sequence of publishing events. So we modified the
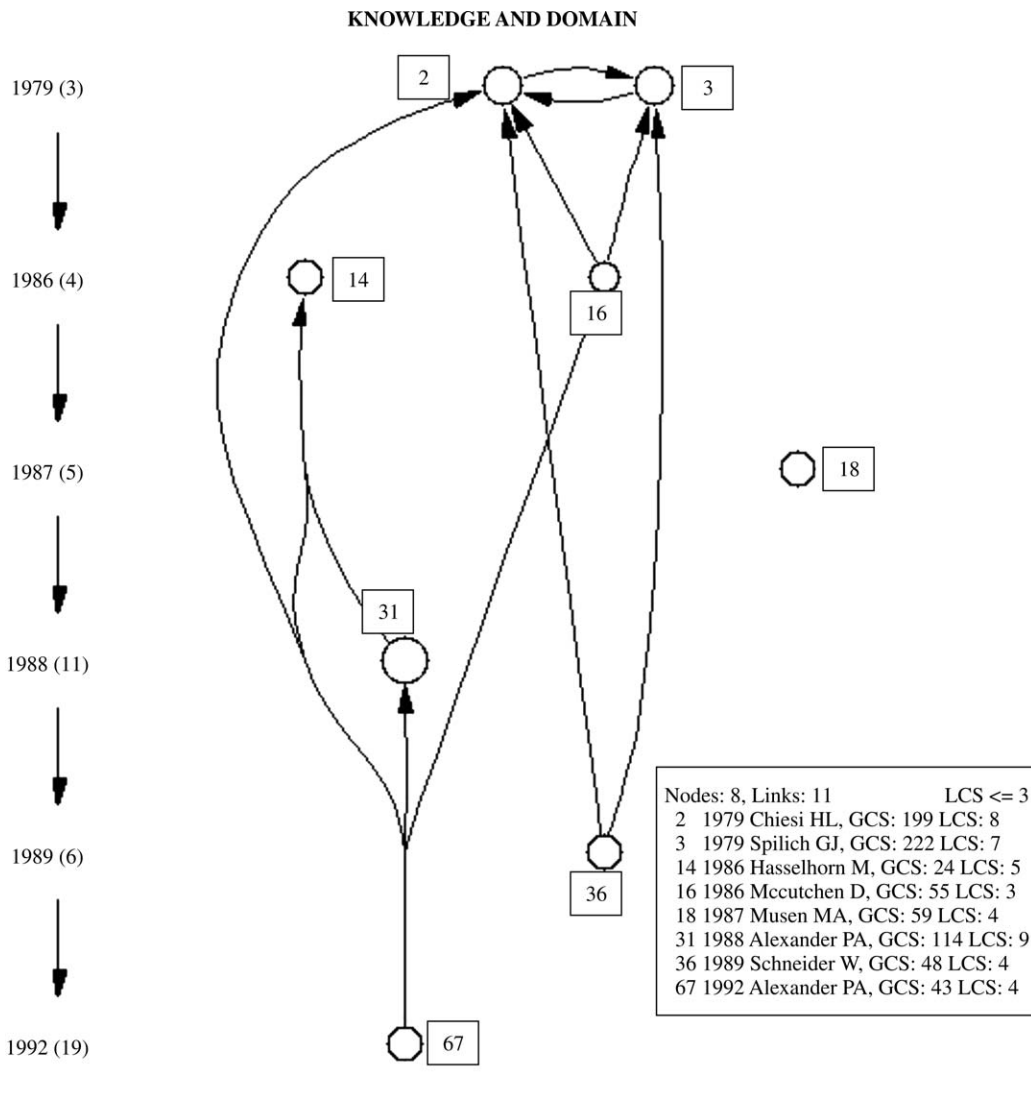
**KNOWLEDGE AND DOMAIN**



Fig. 18. Map of 'Knowledge Domain' papers cited three or more times.

*HistCite* graphical program to accommodate more precise publication dates.

In order to expand the scope of the collection, we created a new collection consisting of 975 papers that cited the Watson-Crick paper itself but also an additional 745 papers that cited the 210 papers in the previous example. We call this chained citation indexing. All of the papers listed chronologically in Figure 23 were published before 1953 and were added from the list of 'Outer References'.

Figure 24 provides a month-by-month display for 1953 papers.

In a recent interview [13], James Watson apologized for not having cited Avery in the primordial *Nature*

paper on the DNA helical structure. But as we have seen, this omission on his part made little difference in the algorithmic mapping of the development of DNA science. But in view of Watson's comment, I have added a dotted line from Watson-Crick to Avery indicating that key link. However, to obtain a more accurate picture of Avery between 1944 and 1945, we would have to create a citation index for the 1944 literature, since the *SCI* source data begins with 1945. There is some controversy as to how well the significance of the Avery work was appreciated by the scientific community at that time [14–19]. I believe our data indicate that it had a substantial impact before and after the Watson-Crick paper appeared.

Outer Reference Missing Link? Journal list All-Author list Citation Matrix Graphs    HistCite Guide

**Multi-Domain**

Nodes: 5643
Sorted by **LCS.**
Page **1**: 1 2 3 4 5 6 7 8 9 10 11 12

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 480 1983 VISUAL DISPLAY OF QUANTITATIVE INFORMATION 1(1):1–1 **TUFTE ER** *Visual Display of Quantitative Information* | 879 | 879 |
| 2 | 0 | 1389 1990 ENVISIONING INFORMATION 1(1):1–2 **TUFTE ER** *Envisioning Information* | 446 | 446 |
| 3 | 3 | 98 1973 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE 24(4):265–269 **SMALL HG** *Co-citation in Scientific Literature – New Measure of Relationship Between 2 Documents* | 235 | 228 |
| 4 | 3 | 122 1974 SCIENCE STUDIES 4(1):17–40 **SMALL HG; GRIFFITH BC** *Structure of Scientific Literatures .F. Identifying and Graphing Specialties* | 212 | 213 |
| 5 | 0 | 3706 1997 VISUAL EXPLANATIONS: IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVES 1(1):1–2 **TUFTE ER** *Visual Explanations: Images and Quantities, Evidence and Narratives* | 171 | 171 |
| 6 | 0 | 39 1967 PSYCHOLOGY TODAY 1(1):61–67 **MILGRAM S** *Small-World Problem* | 47 | 150 |
| 7 | 0 | 2 1963 AMERICAN DOCUMENTATION 14(1):10–& **KESSLER MM** *Bibliographic Coupling between Scientific Papers* | 128 | 144 |
| 8 | 1 | 3976 1998 NATURE 393(6684):440–442 **Watts DJ; Strogatz SH** *Collective dynamics of 'small-world' networks* | 260 | 106 |
| 9 | 4 | 141 1975 SOCIAL STUDIES OF SCIENCE 5(1):86–92 **MORAVCSIK MJ; MURUGESAN P** *Some Results on Function and Quality Of Citations* | 161 | 105 |
| 10 | 5 | 1322 1989 SCIENTOMETRICS 16(1–6):3-& **SCHUBERT A; GLANZEL W; BRAUN T** *Scientometric Datafiles – A Comprehensive Set of Indicators on 2649 Journals and 96 Countries in All Major Science Fields and Subfields 1981–1985* | 104 | 105 |

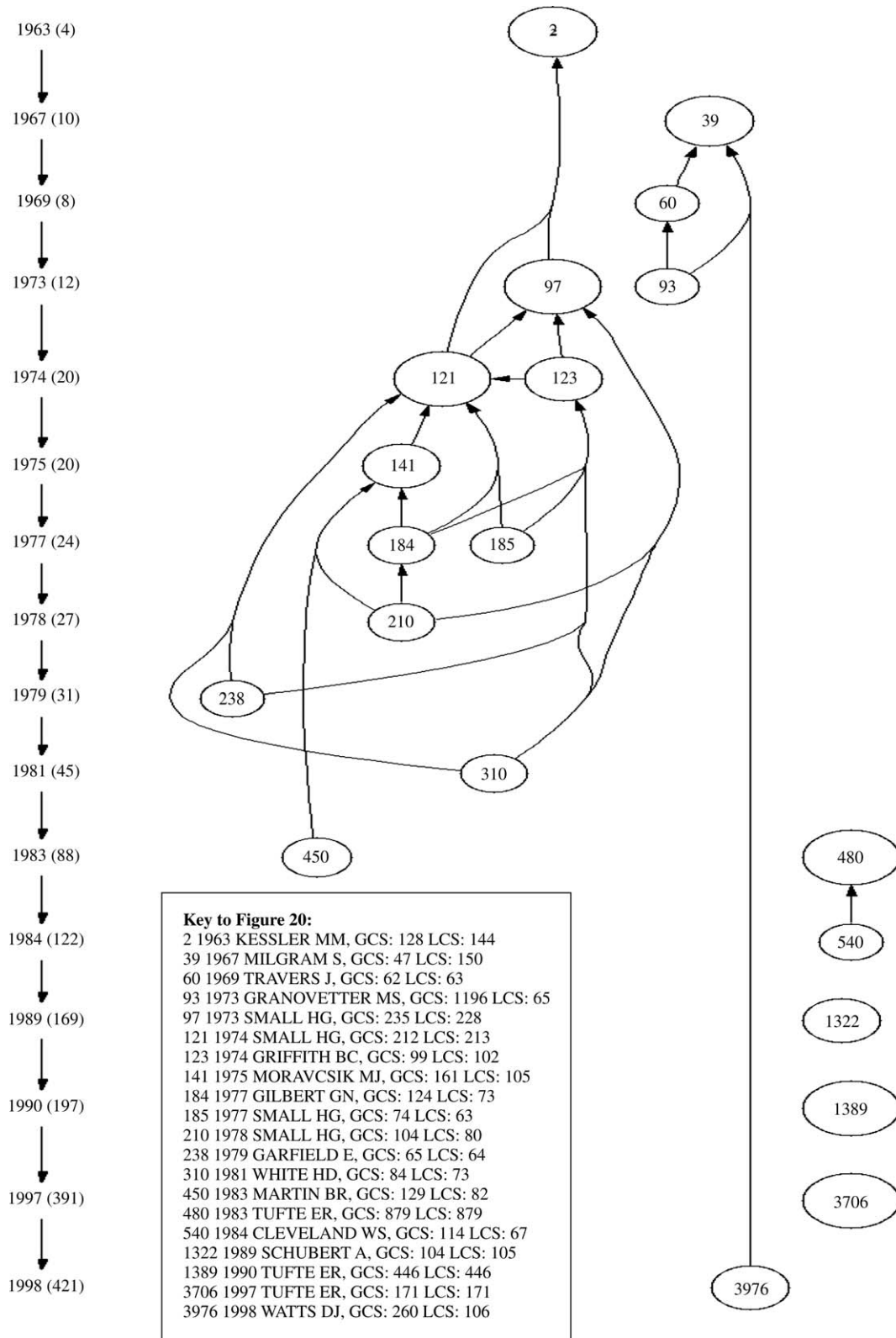Fig. 19. 'Multi-Domain' *HistCite* collection ranked by LCS.

Fig. 20. 'Multi-Domain' historiograph of 20 LCS papers cited 63 or more times.

Key to Figure 20:
2 1963 KESSLER MM, GCS: 128 LCS: 144
39 1967 MILGRAM S, GCS: 47 LCS: 150
60 1969 TRAVERS J, GCS: 62 LCS: 63
93 1973 GRANOVETTER MS, GCS: 1196 LCS: 65
97 1973 SMALL HG, GCS: 235 LCS: 228
121 1974 SMALL HG, GCS: 212 LCS: 213
123 1974 GRIFFITH BC, GCS: 99 LCS: 102
141 1975 MORAVCSIK MJ, GCS: 161 LCS: 105
184 1977 GILBERT GN, GCS: 124 LCS: 73
185 1977 SMALL HG, GCS: 74 LCS: 63
210 1978 SMALL HG, GCS: 104 LCS: 80
238 1979 GARFIELD E, GCS: 65 LCS: 64
310 1981 WHITE HD, GCS: 84 LCS: 73
450 1983 MARTIN BR, GCS: 129 LCS: 82
480 1983 TUFTE ER, GCS: 879 LCS: 879
540 1984 CLEVELAND WS, GCS: 114 LCS: 67
1322 1989 SCHUBERT A, GCS: 104 LCS: 105
1389 1990 TUFTE ER, GCS: 446 LCS: 446
3706 1997 TUFTE ER, GCS: 171 LCS: 171
3976 1998 WATTS DJ, GCS: 260 LCS: 106

Outer References Missing Links? Journal list All-Author list Citation Matrix Graphs          HistCite Guide

**Articles from 1953–1958 citing Watson and Crick's 1953 paper, "Molecular Structure of DNA" and selected outer references**

Nodes: 210
Sorted by **year, journal, volume, page**
Page **1:** 1

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 1 1944 JOURNAL OF EXPERIMENTAL MEDICINE 79():137–157 **AVERY OT; MACLEON CM; MCCARTY M** *Studies on the Chemical Nature of the Substance Inducing Transformation of Pseumococcal Types. Induction of Transformation by A Deoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III* | 331 | 148 |
| 2 | 0 | 2 1952 JOURNAL OF GENERAL PHYSIOLOGY 36(1):39–56 **HERSHEY AD; CHASE M** *Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage* | 747 | 23 |
| 3 | 2 | 3 1953 ACTA CRYSTALLOGRAPHICA 6(8-9):673–677 **FRANKLIN RE; GOSLING RG** *The Structure of Sodium Thymonucleate Fibres .1. The Influence of Water Content* | 14 | 11 |
| 4 | 3 | 4 1953 ACTA CRYSTALLOGRAPHICA 6(8-9):678–685 **FRANKLIN RE; GOSLING RG** *The Structure of Sodium Thymonucleate Fibres .2. The Cylindrically Symmetrical Patterson Function* | 10 | 8 |
| 5 | 1 | 5 1953 ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS 46(1):12–17 **SMITH CL** *The Breakdown of Desoxyribonucleic Acid Under Deuteron and Electron Bombardment* | 5 | 1 |
| 6 | 2 | 6 1953 BIOCHEMICAL JOURNAL 55(5):774–782 **WYATT GR; COHEN SS** *The Bases of the Nucleic Acids of Some Bacterial and Animal Viruses – The Occurrence of 5-Hydroxymethylcytosine* | 57 | 8 |
| 7 | 3 | 7 1953 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 18():123–131 **WATSON JD; CRICK FHC** *The Structure Of DNA* | 61 | 21 |
| 8 | 1 | 8 1953 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 18():133–134 **WYATT GR** *The Quantitative Composition Of Deoxypentose Nucleic Acids As Related To The Newly Proposed Structure* | 9 | 4 |

Fig. 21. First page of *HistCite* collection of 210 papers, of which 208 cited Watson-Crick 1953, from 1953 to 1958. The first two papers are outer references added to increase the chronological depth of the historiograph in Fig. 22.

Fig. 22. Year-by-year map of Watson-Crick 1944–1957.

The figure contains the following labels and key:
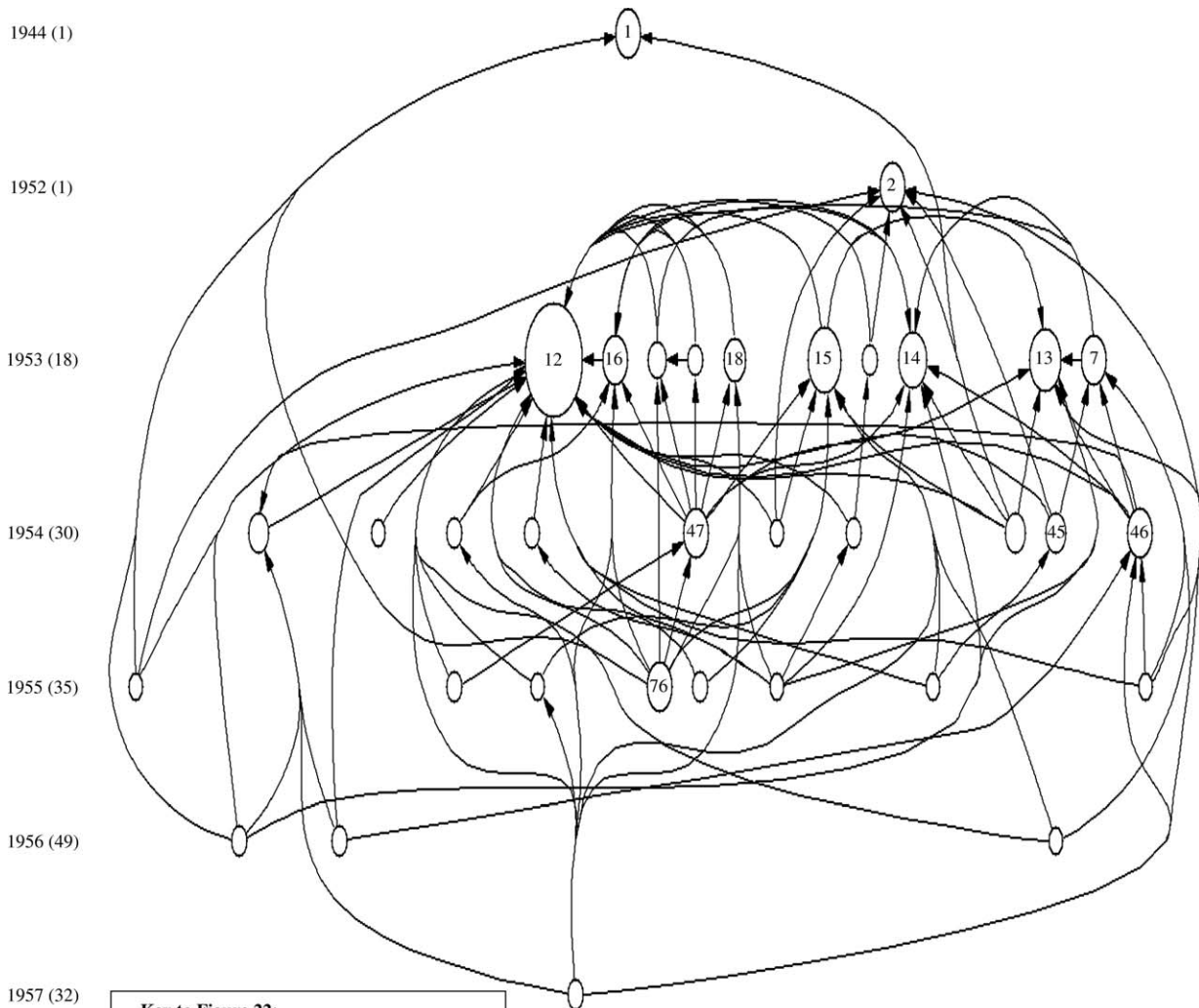
1944 (1)
1952 (1)
1953 (18)
1954 (30)
1955 (35)
1956 (49)
1957 (32)

**Key to Figure 22:**
Nodes: 34, Links: 98      LCS >= 5

1 1944 AVERY OT, GCS: 0 LCS: 23
2 1952 HERSHEY AD, GCS: 747 LCS: 23
3 1953 FRANKLIN RE, GCS: 14 LCS: 11
4 1953 FRANKLIN RE, GCS: 10 LCS: 8
6 1953 WYATT GR, GCS: 57 LCS: 8
7 1953 WATSON JD, GCS: 61 LCS: 21
12 1953 WATSON JD, GCS: 205 LCS: 205
13 1953 WILKINS MHF, GCS: 170 LCS: 41
14 1953 FRANKLIN RE, GCS: 216 LCS: 35
15 1953 WATSON JD, GCS: 87 LCS: 53
16 1953 FRANKLIN RE, GCS: 24 LCS: 21
18 1953 WILKINS MHF, GCS: 24 LCS: 19
29 1954 THOMAS R, GCS: 37 LCS: 12
30 1954 LALAND SG, GCS: 11 LCS: 7
31 1954 PEACOCKE AR, GCS: 15 LCS: 5
33 1954 LEVINTHAL C, GCS: 19 LCS: 5
36 1954 CRAMPTON CF, GCS: 28 LCS: 6
41 1954 ELSON D, GCS: 25 LCS: 8
44 1954 RICH A, GCS: 25 LCS: 13
45 1954 DELBRUCK M, GCS: 24 LCS: 15
46 1954 DEKKER CA, GCS: 43 LCS: 21
47 1954 CRICK FHC, GCS: 54 LCS: 21
52 1955 ALLFREY VG, GCS: 36 LCS: 5

59 1955 BLOUT ER, GCS: 12 LCS: 6
60 1955 ELSON D, GCS: 31 LCS: 5
72 1955 STENT GS, GCS: 27 LCS: 5
75 1955 DUNN DB, GCS: 28 LCS: 6
76 1955 FEUGHELMAN M, GCS: 39 LCS: 21
79 1955 COX RA, GCS: 10 LCS: 5
80 1955 PLATT JR, GCS: 7 LCS: 5
95 1956 PAULING L, GCS: 12 LCS: 6
119 1956 THOMAS CA, GCS: 12 LCS: 6
122 1956 CAVALIERI LF, GCS: 14 LCS: 6
154 1957 COX RA, GCS: 10 LCS: 7

Outer References Missing Links? Journal list All-Author list Citation Matrix Graphs     HistCite Guide

**Articles citing Watson and Crick's 1953 paper, "Molecular Structure of DNA" the article citing them (1953–1958), and selected outer references**

Nodes: 975

Sorted by **year, journal, volume, page**

Page **1:** 1 2

| # | Cited nodes | Nodes/Authors | GCS | LCS |
|---|---|---|---|---|
| 1 | 0 | 1 1938 JOURNAL OF BIOLOGICAL CHEMISTRY 124():425– <br> **SEVAG MG** <br> *[unknown]* | 216 | 37 |
| 2 | 1 | 2 1944 JOURNAL OF EXPERIMENTAL MEDICINE 79():137–157 <br> **AVERY OT; MACLEON CM; MCCARTY M** <br> *Studies on the Chemical Nature of the Substance Inducing Transformation of Pseumococcal Types. Induction of Transformation by a Deoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III* | 331 | 43 |
| 3 | 0 | 3 1945 JOURNAL OF BIOLOGICAL CHEMISTRY 161(1):83–89 <br> **SCHMIDT G; THANNHAUSER SJ** <br> *A Method for the Determination of Desoxyribonucleic Acid, Ribonucleic Acid, and Phosphoproteins in Animal Tissues* | 696 | 34 |
| 4 | 1 | 4 1945 JOURNAL OF BIOLOGICAL CHEMISTRY 161(1):293–303 <br> **SCHNEIDER WC** <br> *Phosphorus Compounds in Animal Tissues .1. Extraction and Estimation of Desoxypentose Nucleic Acid and of Pentose Nucleic Acid* | 952 | 30 |
| 5 | 2 | 5 1946 JOURNAL OF GENERAL PHYSIOLOGY 30(2):117-& <br> **MIRSKY AE; POLLISTER AW** <br> *Chromosin, a Desoxyribose Nucleoprotein Complex of the Cell Nucleus* | 323 | 35 |
| 6 | 0 | 6 1947 JOURNAL OF THE CHEMICAL SOCIETY (SEP):1131–1141 <br> **GULLAND JM; JORDAN DO; TAYLOR HFW** <br> *Deoxypentose Nucleic Acids .2. Electrometric Titration of the Acidic and the Basic Groups of the Deoxypentose Nucleic Acid of Calf Thymus* | 70 | 31 |
| 7 | 3 | 7 1951 BIOCHEMICAL JOURNAL 48(5):584–590 <br> **WYATT GR** <br> *The Purine and Pyrimidine Composition of Deoxypentose Nucleic Acids* | 276 | 63 |
| 8 | 0 | 8 1951 JOURNAL OF BIOLOGICAL CHEMISTRY 189(2):597–605 <br> **MARSHAK A; VOGEL HJ** <br> *Microdetermination of Purines And Pyrimidines in Biological Materials* | 136 | 30 |
| 9 | 0 | 9 1951 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 37(4):205–211 <br> **PAULING L; COREY RB; BRANSON HR** <br> *The Structure of Proteins – 2 Hydrogen-Bonded Helical Configurations of the Polypeptide Chain* | 185 | 26 |
| 10 | 1 | 10 1952 BIOCHEMICAL JOURNAL 52(5):558–565 <br> **MARKHAM R; SMITH JD** <br> *The Structure of Ribonucleic Acids .2. The Smaller Products of Ribonuclease Digestion* | 104 | 28 |

Fig. 23. Opening page of *HistCite* collection of 'chained' citations to Watson-Crick.
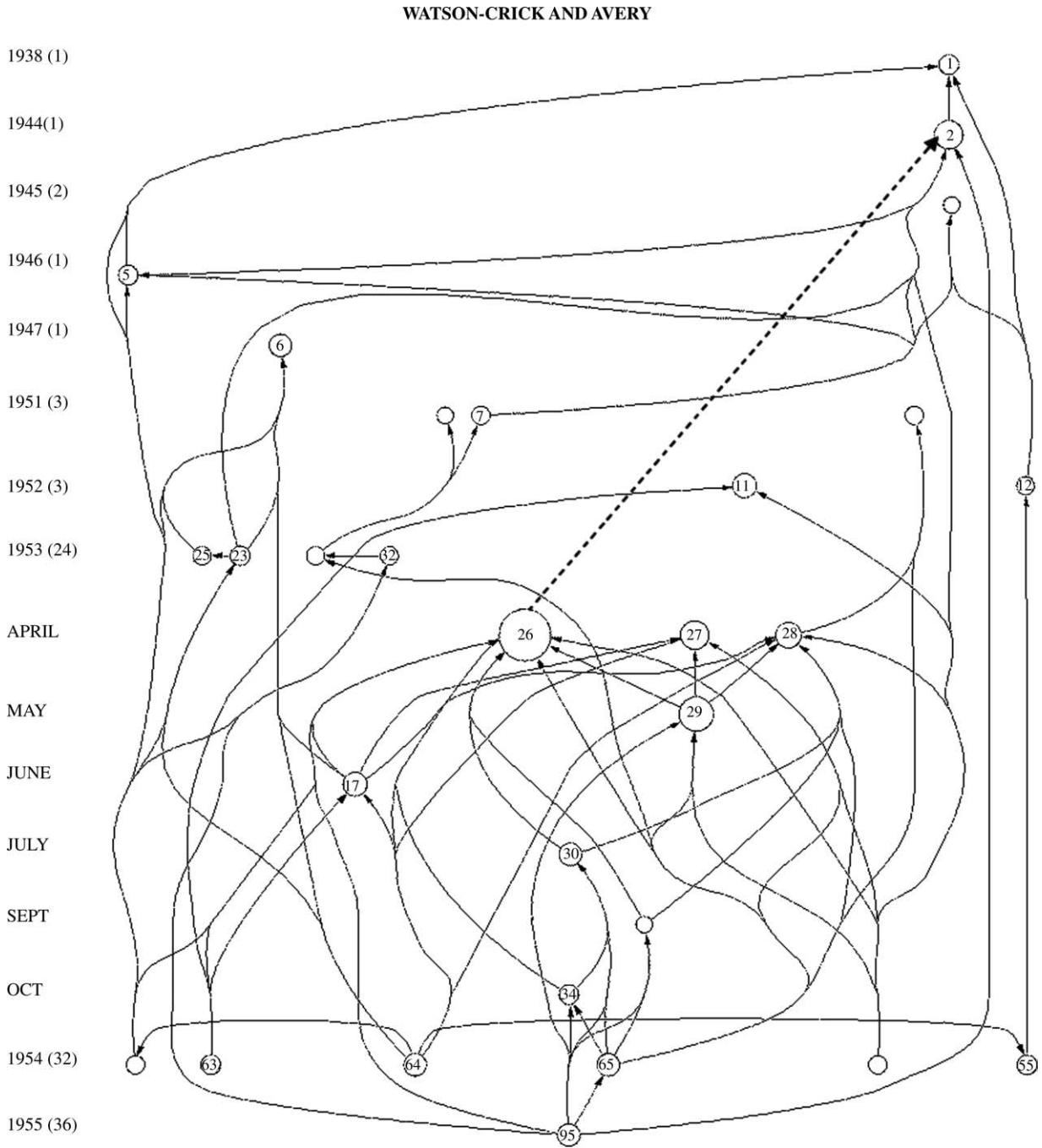
**WATSON-CRICK AND AVERY**



Fig. 24. Month-by-month historiograph linking Watson-Crick to Avery.

## References

[1] E. Garfield, A.I. Pudovkin and V.S. Istomin, Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology (JASIST)* 54(5)(2003) 400–412. Available at: http://garfield.library.upenn.edu/papers/jasist54(5)400y2003.pdf (accessed 27 January 2004).

[2] Over a 25-year period, I published a dozen or more essays on mapping science. See my website at http://garfield.library.upenn.edu/mapping/mapping.html (accessed 27 January 2004).

[3] E. Garfield, I.H. Sher and R.J. Torpie, The use of citation data in writing the history of science. (Unpublished report of research for Air Force Office of Scientific Research under contract AF49(638)-1256 undertaken by The Institute for Scientific Information, Philadelphia, December 1964.) Available at: www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf (accessed 27 January 2004).

[4] A. Asimov, *The Genetic Code* (New American Library, New York, 1963).

[5] R. Garner, *Computer-Oriented Graph Theoretic Analysis of Citation Index Structures* (Drexel University Press, Philadelphia, 1967). Available at: www.garfield.library.upenn.edu/rgarner.pdf (accessed 27 January 2004).

[6] N.P. Hummon and P. Doreian, Connectivity in a citation network: the development of DNA, *Social Networks* 11(1989) 39–63. Available at: http://garfield.library.upenn.edu/papers/hummondoreian1989.pdf (accessed 27 January 2004).

[7] E. Garfield, *From computational linguistics to algorithmic historiography* (Unpublished Lazerow Lecture held in conjunction with panel on ''Knowledge and Language: Building large-scale knowledge bases for intelligent applications'' presented at the University of Pittsburgh on September 19, 2001) Available at: http://garfield.library.upenn.edu/papers/pittsburgh92001.pdf (accessed 27 January 2004).

[8] E. Garfield, A.I. Pudovkin and V.S. Istomin, Algorithmic citation-). linked historiography – mapping the literature of science. In: *Proceedings of the 65th Annual Meeting of the American Society for Information Science & Technology (ASIS&T). 39* (November 2002) 14–24 Available at: http://garfield.library.upenn.edu/papers/asist2002proc.pdf (accessed 27 January 2004). *Abridged version* published in *Proceedings of the 65th Annual Meeting of the American Society for Information Science & Technology (ASIS&T)* 39:14–24 (November 2002). Available at: *http://garfield.library.upenn.edu/papers/asist2002proc.pdf*

[9] J.D. Watson and F.H.C. Crick, Molecular structure of nucleic acids – a structure for deoxyribose nucleic acid, *Nature* 171(4356)(1953) 737–738.

[10] B.J. Strasser Who cares about the double helix? *Nature* 42(6934) (April 24, 2003) 803–804. (The author includes a table showing citations by time periods.)

[11] O.T. Avery, C.M. Macleon and M. Mccarty Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III, *Journal of Experimental Medicine* 79 (1944) 137–157.

[12] A.D. Hershey, J. Dixon and M. Chase, Nucleic acid economy in bacteria infected with bacteriophage-T2 .1. Purine and Pyrimidine Composition, *Journal of General Physiology* 36(6) (1953) 777–789.

[13] Anonymous. Genes, Girls, and Honest Jim, *Bio-IT World* 2(4) (2003) 28.

[14] J. Lederberg, Reply to H. V. Wyatt, *Nature* 239(5369) (September 22, 1972) 234.

[15] J. Lederberg, Greetings (on the occasion of Symposium entitled DNA, The Double Helix, Perspective and Prospective at Forty Years) *Annals of the New York Academy of Sciences* 758 (1995) 176–179.

[16] H. Zuckerman and J. Lederberg, Postmature scientific discovery, *Nature* 324(6098) (December 18, 1986) 629–631.

[17] G.S. Stent, Prematurity in scientific discovery. In: E. B. Hook (ed.), *Prematurity in Scientific Discovery*, (University of California Press, Berkeley and Los Angeles, 2002) 22–33.

[18] G.S. Stent, The aperiodic crystal of heredity, DNA: the double helix, *Annals of the New York Academy of Sciences* 758 (1995) 25–31.

[19] G.S. Stent, Prematurity and uniqueness in scientific discovery, *Scientific American* 227(6) (1972)84 +

## Editor's Note: