

Mapping Output of Topical Searches in the *Science Citation Index*,[®] *Social Sciences Citation Index*,[®] *Arts and Humanities Citation Index*[®]

Eugene Garfield

Chairman Emeritus, ISI

3501 Market Street

Philadelphia, PA 19104

garfield@codex.cis.upenn.edu

A. I. Pudovkin

Institute of Marine Biology

Russian Academy of Sciences

Vladivostok 690041, Russia.

aipud@online.ru

V. S. Istomin

Center for Teaching, Learning, and Technology

Washington State University

Pullman, WA 99164-4550.

vi@mail.wsu.edu

Abstract

The authors have developed a system called *HistCite*[™] which generates chronological maps of subject (topical) collections resulting from searches of the *ISI Web of Science*[®] or *SCI/SSCI/AHCI* on CD-ROM. *WoS* export files are created in which all cited references for source documents are captured. These bibliographic collections are processed by *HistCite* which generates chronological tables as well as historiographs which highlight the most-cited works in and outside the collection. Several topics including small world networks, gene flow, and the structure of DNA will be demonstrated. Real time dynamic genealogical historiographs will be shown. *HistCite* also includes a module for detecting and editing errors or variations in cited references. Export Files of 5,000 or more records are processed in minutes on a PC. Ideally the system will be used to help the searcher quickly identify the most significant work on a topic and enable the searcher to trace its year-by-year historical development.

The overall theme of this symposium is needs assessment in libraries. The *HistCite*[™] system resulted from a long-term needs assessment of users of bibliographic databases. On the one hand, there is the librarians and users need to identify the key works on a particular subject. On the other hand, scholars and editors desire rapid historical reviews of new topics. *HistCite* is designed to satisfy both of these requirements, whether for reference service, or in writing review articles or historical introductions to new manuscripts.

HISTORICAL INTRODUCTION

Even before the advent of the *Science Citation Index* in print the use of citation data to help write the history of science was discussed in the 1964 report on "The use of Citation Data in Writing the History of Science"¹ which included an historiograph sketching history of DNA from Gregor Mendel in 1865 to Marshall Nirenberg in 1961 through various stages including Avery-McCleod-McCarty in 1944 through Watson-Crick in 1953. Flow charts of the papers were created manually based solely on the references cited in a set of core source papers identified in a book by Isaac Asimov on the genetic code. Then, in subsequent years I periodically speculated on the potential use of citation indexes for historiography.²

Similar maps were later created by Tony Cawkell.³ When I taught information retrieval at the University of Pennsylvania Moore School of Electrical Engineering, students were required to create similar topical historiographs. In all these mapping exercises it was explicitly assumed that scholars would use *ISI's* citation indexes to identify the needed citation links by manual searches. I believed that these historiographs would aid in studying contemporary history of science. Since history and bibliography were intimately linked the term "historio-bibliography" was coined.⁴

During the DNA mapping project, we often discussed the idea of writing computer programs that would create such maps directly from the electronic files of the *Science Citation Index*. We thought this would require random access to *ISI's* massive files so that cited and citing documents could be retrieved in real time. In the 1960s, however, low cost gigabyte memories were still a dream. The implementation of real time mapping had to wait for the time when computer memories were large and cheap enough to handle retrospective files covering many decades of literature. While on-line searches were possible in the 1970s, mapping in real time was still not feasible and because the PC had not yet come along. Only when the output of a completely linked large file of thousands of records could be handled by today's PCs did the creation of historiographs in real time become feasible.

There have been many different types of "mapping" exercises performed on a relatively small scale. In the past, co-citation clustering required main frame computers⁵ and in most cases still does. These ideas were later extended to creating small cluster maps on-line as in the *SciMap* system developed by Henry Small at *ISI*. In that system a starting paper is used to seed the creation of a co-citation cluster map⁶. In spite of the many mapping and visualization reports in the literature, none were applied to the creation of historiographs. Further, none of the many authors on co-citation mapping considered the potential significant relationship between historical displays and the need of reference librarians and users to evaluate the output of literature searches with *Science Citation Index (SCI)*, *Medline*, *Chemical Abstracts*, etc.

Until quite recently, we thought of creating historiographs primarily by seeding one or two primordial papers. The *Science Citation Index on CD-ROM* was used to trace forward in time papers that had cited the starting papers. This is the essence of the now traditional "cited reference search." Since the basic purpose of an historiograph is to display the chronological development of a topic or field year by year from the earliest papers forward in time we searched the annual *SCI* in the same way. Once we identified the initial group of citing papers, we then did further cited reference searches on those citing papers. This process is sometimes called citation chaining. We iterated this process for as many years of the literature as was necessary. This will be illustrated later when I discuss the work of Watson and Crick.

For each year of the literature searched there would be 0 to N papers retrieved. The full *SCI* Source record for each of the N papers would be captured including authors, titles, journal, volume page and years, and the list of cited references. If there are 20 cited references per source paper then there would be 20N cited references collected. Thus, if the collection involves 500 source papers, the uniquely cited references will total 5,000 to 10,000 items. If 1,000 different papers or books have been cited then the average frequency will be 10. However, the range of citation frequency would run from 1 to 1,000.

The *HistCite* system creates a series of tables and matrices which list all the 500 source documents in chronological order and shows their citation frequencies both within the collection and globally. The *HistCite* program also includes a variety of sort keys which create ranked directories of authors, journals, and citation frequencies.

***HISTCITE* PROGRAM**

In *HistCite*, we distinguish two types of frequency – LCS and GCS. The local citation score or frequency is the number of times an item is cited in the retrieved collection. The GCS is the global score, the number of cites in the entire *SCI/SSCI*.

The record for each source document contains both its LCS and GCS. Once *HistCite* sorts the papers by citation score, the user will select a group cited above an arbitrarily chosen threshold to be mapped. If there are 500 source papers then a 5% selection threshold would produce 25 core papers. These core papers should be of prime interest especially to a searcher who is not familiar with the subject matter. Ordinarily, one would examine this core list first in reviewing a new topic. The coordinates of these papers are used to create an historiograph of the topic which displays the papers and their citation links chronologically.

Identifying Core Literatures

As stated earlier, it was initially assumed we would begin with one primordial paper. However, it became apparent that one could feed in groups of papers by one or more authors – and by extension, larger clusters of papers by institution or by key word. Thus the output of any conventional search or a combination of citation and key word searches could be input to the system. Once the input bibliography is created, the core papers on the “topic” can be rapidly identified.

Visualization

The production of the various tables or lists from these procedures is separate from the problem of visualizing these data. Time does not permit me to go into the details of how these artifacts aid in the perception of the interrelationships between citing and cited papers. Using your PC screen, you can scroll a map of almost unlimited size. With *HistCite* you use mouse clicks and pop-up windows to navigate whatever map you create. However, in most situations 25 or [more] nodes are easily visible on the ordinary PC screen. Thus a map involving several hundred nodal papers can be created, but it is best to first view a condensed version containing 25 to 50 of the core papers in the local or global collection. In this way, you can quickly perceive their historical connections.

Outer References

HistCite also produces frequency ranked tables of “outer references,” that is, cited papers and books that fall outside the starting retrieved collection. These are works which do not turn up in the original *WoS* search but, significantly, are cited frequently in the papers that are retrieved. You can examine these candidate references and decide whether to add them to the initial collection. For example, a highly cited book or patent might be cited which is not part of the original *SCI* source database. For each of these a source record would have to be created manually. Many of these items will have been published prior to the starting reference and may inform the earlier history.

For each outer reference, *HistCite* provides a hot link to the *WoS* and executes a “cited reference search.” By clicking on the hot link to the full source record, it can be added to the marked list of the collection. Once a new .txt file is created, the *HistCite* software is invoked to create a new set of tables. The size of the outer reference list can be specified by the user. Generally, I add those outer papers which are cited at the frequency threshold used in the mapping procedure.

Missing References

It is well known that authors cite references with many variant spellings or make errors in one or more parts of the reference such as volume or page. These incomplete or “missing” references are identified in a separate table. As part of the procedures invoked, the program will seek out and display the closest matched paper in the collection and suggest changes. In many references, the initial page of the document is not cited. In chemical papers the exact page where a specific chemical compound is discussed will be cited. The cited page number can be edited so that it can be included in the citation frequency score for the fully paginated reference. Fortunately the number of such errors or variations is quite manageable. The reader can choose to edit them or not. Most are singletons that ordinarily will not affect the overall ranking or mapping results. However, to study historical impact, I have found it useful to unify citations to books.

At present, the program accepts export records from either the *SCI* CD-ROM, the *Web of Science*, or *SciSearch* on Dialog. However, Dialog records do not include the GCS. Since the LCS is computed by *HistCite*, it is still possible to create a very useful map of the local collection!

In the examples which follow, the collections were created by searching the *Web of Science* using either a title word, cited reference, or a combination search. The resulting “marked lists” were downloaded in the WOS EXPORT format which is in plain ASCII text. The plain ASCII .txt file is then processed by the program. *HistCite* output is presented in html. The *HistCite* program can reside on the user’s hard drive or on a central server.

Over the past few years, we have created a large number of test collections on a variety of topics. Figure 1 contains a partial list of those topics. The entire list of the collections can be accessed at my website

Index of /Histcite:

<http://garfield.library.upenn.edu/histcomp/>

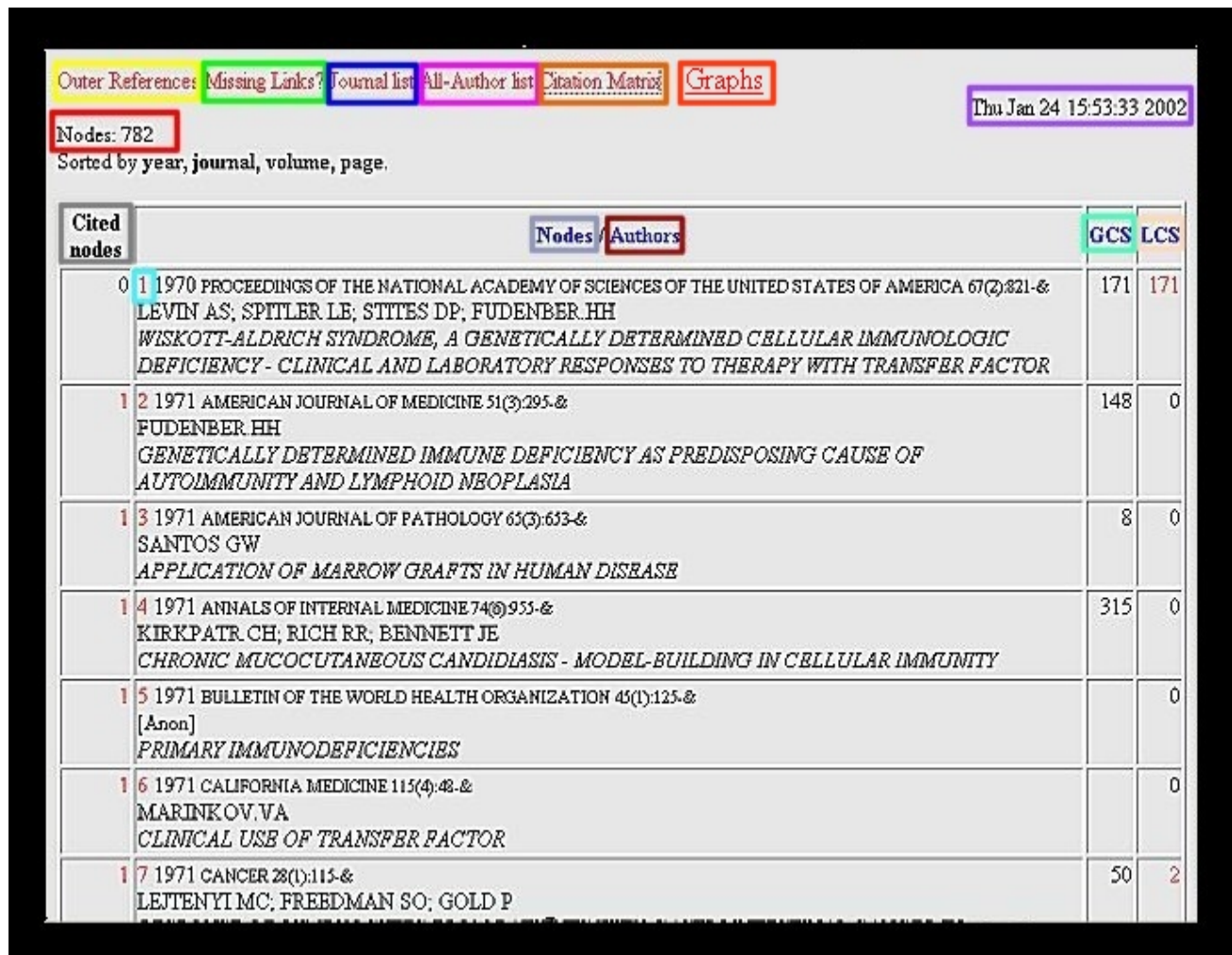
Guide to Database Created by Histcite

FOLDER	LAST MODIFIED	NODES	MAPS	DESCRIPTION
annualreviews/	26-Jun-2002	---	---	HistCite Listing of Annual Reviews - All Editions
bibliographic-coupling/	24-Aug-2001	223	G1, G2	Citations to Kessler's Bibliographic Coupling
CellularAutomataWolfram/	06-Jan-2003	1359	G1, G2	Titles containing "cellular automata"
cocitation-1973-2001/	08-Nov-2001	1059	G1, G2	Citations to Small & Griffith and Descendants
evidence based medicine/	20-Jan-2003	873	G1, G2	Evidence based medicine 1992-02
geneflow47-02/	28-Jun-2002	3552	G1	Geneflow 1947 - 2002
j info sci/	21-Jan-2003	1306	G1, G2	Articles from <i>Journal of Information Science</i> 1979-2002
lander venter/genome	16-Oct-2002	1746	---	Articles citing sequencing human genome, Nature 2001; Science 2001
scientometrics 19782002/	09-Dec-2002	3543	G1	Articles from or citing <i>Scientometrics</i> , 1978-2002
small world/	24-Jan-2003	442	G1, G2	S Milgram and Citing Papers and Small-World, 1967-2002
watson-crick45-58upd/	21-Jan-2003	210	G1, G2	Articles in 1953-1958 citing Watson & Crick's 1953 paper, "Molecular Structure of DNA"
watsoncrick chained/	28-Jan-2003	975	G1, G2	Articles in 1953-1958 citing Watson & Crick's 1953 paper, and citations to those articles

Figure 1: Selected List of Sample *HistCite* Collections

On the website, there is a detailed guide (help screen) which explains each element (see Figure 2).

HistCite Guide



The COLORED boxes are components of the Historiograph Compilation. Clicking on a box will describe in more detail about the specific component.

Figure 2: HistCite Guide

SMALL WORLD PROBLEM (STANLEY MILGRAM)

In 1967, in *Psychology Today*,⁷ Stanley Milgram published his primordial paper on the "Small World Problem." A simple Cited Reference search of the *SSCI/SCI* initially produced about 148 citing papers. Note that the cited page is page 61. (see Figure 3)



Cited Reference Search

10 references matched query: Cited Year=1967; Cited Author=milgram s; Databases=SCI-EXPANDED, SSCI, A&HCI; Timespan=All Years

STEP 2: CITED REFERENCE SELECTION

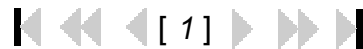
The table lists all of the cited references that match your search request and the number of times each variation has been cited. Select all desired references (including variants) and then press Search.

[Set language and document type limits and sort option.](#)

SELECT PAGE or select specific references from list.

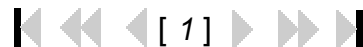
SEARCH to find articles that cite selected references.

References 1 -- 10



	Hits	Cited Author	Cited Work	Volume	Page	Year	
<input type="checkbox"/>	1	MILGRAM S	MEGAMOT	15	31	1967	
<input type="checkbox"/>	9	MILGRAM S	PATTERNS PREJUDICE	1	3	1967	
<input type="checkbox"/>	1	MILGRAM S	PATTERNS PREJUDICE	1	5	1967	
<input type="checkbox"/>	3	MILGRAM S	POLITISCHE GESELLSCH		170	1967	
<input type="checkbox"/>	10	MILGRAM S	PSYCHOL TODAY	2	60	1967	←
<input type="checkbox"/>	148	MILGRAM S	PSYCHOL TODAY	1	61	1967	←
<input type="checkbox"/>	2	MILGRAM S	PSYCHOL TODAY	1	62	1967	
<input type="checkbox"/>	6	MILGRAM S	PSYCHOL TODAY MAY		61	1967	
<input type="checkbox"/>	1	MILGRAM S	PSYCHOLOGY TODAY MAY			1967	
<input type="checkbox"/>	1	MILGRAM S	READINGS SOCIAL PSYC			1967	

References 1 -- 10



SELECT PAGE or select specific references from list.

SEARCH to find articles that cite selected references.

Figure 3: Cited Reference Search on Milgram's 1967 *Psychology Today* article.

However, another ten papers cited page 60. It was an easy error to make because the “first” page of the article is a photograph. By adding this variant and others, we obtain a somewhat larger collection of citing papers.

ISI Web of SCIENCE® Powered by ISI Web of Knowledge_{SM}

HOME

HELP

DATE & DB LIMITS

GENERAL SEARCH

CITED REF SEARCH

LOG OFF

General Search Results--Summary

Title=small world; DocType=All document types; Language=All languages; Databases=SCI-EXPANDED, SSCI, A&HCI; Timespan=All Years; (sorted by latest date)

SUBMIT MARKS

MARK PAGE

MARK ALL

Page 32 (Articles 311 -- 315):

◀◀◀

[31 | 32]

▶▶▶

Use the checkboxes to add individual articles to the Marked List. Be sure to click SUBMIT MARKS button before leaving page.

☐

VARMA BN
[SMALL WORLD OF KHANH-HAU - HENDRY,JB](#)
AM J SOCIOL 72 (4): 422-423 1967

☐

AUCHTER EL
[SMALL WORLD OF KHANH HAU - HENDRY,JB](#)
J DEV AREAS 1 (4): 547-549 1967

☐

MILGRAM S
[SMALL-WORLD PROBLEM](#)
PSYCHOL TODAY 1 (1): 61-67 1967

☐

BRANT CS
[THE SMALL WORLD OF HAU,KHAHN - HENDRY,JB](#)
AM ANTHROPOL 67 (1): 198-199 1965

☐

GITTINGER JP
[THE SMALL WORLD OF KHANH HAU - HENDRY,JB](#)
ANN AM ACAD POLIT SS 358 (MAR): 240-241 1965

☐

ROBINSON W
[ITS A SMALL WORLD](#)
AM J NURS 62 (11): 10-10 1962

SUBMIT MARKS

MARK PAGE

MARK ALL

Page 32 (Articles 311 -- 315):

◀◀◀

[31 | 32]

▶▶▶

315 of 32023526 documents matched the query.

Figure 4: WoS General Search for “Small World” in Title.

In addition to the cited reference search we also searched *WoS* for papers with the term “small world” in the title. This is illustrated in Figure 4. This added another 260 or so papers, bringing the total to 424 papers. A marked list was created and then saved as an ISI Export File. This .txt file was submitted to the *HistCite* program.

The initial output of the *HistCite* system is the master chronological table (see Figure 5). The first three papers in the table are by Varna, Evers, and Auchter. These are not relevant to the search but the fourth is the key paper by Milgram. Note that the Global Citation Score (GCS) is 148. The Local Citation Score (LCS) is 167. The 148 is the score noted in the *WoS* for citations to page 61. But we added an additional 19 for the variants noted earlier.

[Outer References](#) [Missing Links?](#) [Journal list](#) [All-Author list](#) [Citation Matrix](#) [Graphs](#)

[HistCite Guide](#)

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 VARMA BN <i>Small World of Khanh-Hau - Hendry,JB</i>	2	1
2	1	2 1967 AMERICAN JOURNAL OF SOCIOLOGY 73(1):115-115 EVERS HD <i>Small World of Khanh-Hau - Comment</i>	0	0
3	0	3 1967 JOURNAL OF DEVELOPING AREAS 1(4):547-549 AUCHTER EL <i>Small World of Khanh Hau - Hendry,JB</i>	0	0
4	0	4 1967 PSYCHOLOGY TODAY 1(1):61-67 MILGRAM S <i>Small World Problem</i>	148	167
5	0	5 1969 FORTUNE 80(4):121-& ZALAZNICK S <i>Small World Of Big Washington Lawyers</i>	0	0
6	1	6 1969 SOCIOMETRY 32(4):425-443 TRAVERS J; MILGRAM S <i>Experimental Study Of Small World Problem</i>	64	43
7	1	7 1970 BRITISH JOURNAL OF SOCIAL PSYCHIATRY 4(2):83-87 HART JW <i>Sociometry of Poverty</i>	0	0
8	0	8 1970 ENGLISH JOURNAL 59(3):416-420 SOFFER RS <i>Its A Small World</i>	1	0

Figure 5: Chronological File of Papers citing S. Milgram's 1967 Psychology Today Paper, or using Small World in Title, 1962-2002

In addition to the basic chronological “home” table several sorts can be called out. These are activated by clicking on the hot links. The first is the Local Citation Score (LCS) (see Figure 6) which is based on the citation frequency within the basic collection. The paper by Milgram has now moved up to position #1 and its LCS is 167. The second most-cited paper in LCS is by Watts and Strogatz.

[Outer References](#) [Missing Links?](#) [Journal list](#) [All-Author list](#) [Citation Matrix](#) [Graphs](#)

[HistCite Guide](#)

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 **LCS**.

Page 1: 1

#	Cited nodes	Nodes / Authors	GCS	LCS
1	0	4 1967 PSYCHOLOGY TODAY 1(1):61-67 Milgram S <i>Small World Problem</i>	148	167
2	1	194 1998 NATURE 393(6684):440-442 Watts DJ; Strogatz SH <i>Collective Dynamics Of 'Small-World' Networks</i>	387	134
3	3	224 1999 SCIENCE 286(5439):509-512 Barabasi AL; Albert R <i>Emergence of scaling in random networks</i>	269	49
4	3	218 1999 PHYSICAL REVIEW E 60(6):7332-7342 Newman MEJ; Watts DJ <i>Scaling and percolation in the small-world network model</i>	74	47
5	4	219 1999 PHYSICAL REVIEW LETTERS 82(15):3180-3183 Barthelemy M; Amaral LAN <i>Small-world networks: Evidence for a crossover picture</i>	73	47
6	1	6 1969 SOCIOLOGY 32(4):425-443 Travers J; Milgram S <i>EXPERIMENTAL STUDY OF SMALL WORLD PROBLEM</i>	64	43
7	5	229 2000 EUROPEAN PHYSICAL JOURNAL B 13(3):547-560 Barrat A; Weigt M <i>On the properties of small-world network models</i>	58	39
8	2	9 1970 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 15(2):101-& Korte C; Milgram S <i>Acquaintance Networks Between Racial Groups - Application of Small World Method</i>	38	31

Figure 6: Most-cited papers from the starting bibliography ranked by Local Citation Score (LCS)

By clicking on GCS, the file is sorted by the global citation frequency, that is, how often each paper is cited in the entire WoS collection (see Figure 7). Now the Watts paper moves up to position #1 and is followed by HC White et al. Milgram is now in the fifth position. Now the Watts paper moves up to position #1 and is followed by HC White et al. Milgram is now in the fifth position. By juxtaposing GCS and LCS, we become aware of the extent of the literature both within and outside the starting collection. In the more recent literature, physics papers begin to dominate what was formerly a topic of sociological interest

[Outer References](#) [Missing Links?](#) [Journal list](#) [All-Author list](#) [Citation Matrix](#) [Graphs](#)

[HistCite Guide](#)

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 **GCS**.

Page 1: 1



#	Cited nodes	Nodes / Authors	GCS	LCS
1	1 194	1998 NATURE 393(6684):440-442 Watts DJ; Strogatz SH <i>Collective dynamics of 'small-world' networks</i>	387	134
2	5 25	1976 AMERICAN JOURNAL OF SOCIOLOGY 81(4):730-780 White HC; Boorman SA; Breiger RI <i>Social-Structure From Multiple Networks .1. Blockmodels Of Roles And Positions</i>	339	7
3	3 224	1999 SCIENCE 286(5439):509-512 Barabasi AL; Albert R <i>Emergence of scaling in random networks</i>	269	49
4	6 249	2000 NATURE 406(6794):378-382 Albert R; Jeong H; Barabasi AL <i>Error and attack tolerance of complex networks</i>	154	20
5	0 4	1967 PSYCHOLOGY TODAY 1(1):61-67 Milgram S <i>Small World Problem</i>	148	167
6	13 300	2001 NATURE 410(6825):268-276 Strogatz SH <i>Exploring complex networks</i>	128	23
7	3 26	1976 AMERICAN JOURNAL OF SOCIOLOGY 81(6):1384-1446 Boorman SA; White HC <i>Social Structure from Multiple Networks .2. Role Structures</i>	117	4
8	2 64	1981 AMERICAN JOURNAL OF SOCIOLOGY 86(5):1015-1035 Feld SL <i>The Focused Organization of Social Ties</i>	111	4
9	4 265	2000 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 97(21):11149-11152 Amaral LAN; Scala A; Barthelemy M; Stanley HE <i>Classes of small-world networks</i>	111	28
10	2 30	1976 SOCIAL FORCES 55(1):93-122 Burt RS <i>Positions in Networks</i>	108	2

Figure 7: Small World Collection ranked by Global Citation Score (GCS)

By clicking on the all-author hotlink, we find the most-published author on the Small World problem (see Figure 8). Hotlinks also permit display of the authors by global or local citation score. Thus the most-cited authors, are distinguished from the most-published ones. The individual citation frequencies for these papers are totaled.

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

Figure 8: Authors ranked by number of publications.

Figure 9 displays the journals which published on the Small World problem. The number of papers for each journal is shown on the right. This topic, originally mainly of interest to social network researchers, is now permeated by physics papers. Such drastic change in the history of this literature indicates that it might be better to split the historical exercise into the periods 1967-1997 and then 1998 to the present.

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-PUBLICATIONS OF THE MODERN LANGUAGE ASSOCIATION 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

Figure 9: Journals ranked by number of papers published

The “outer nodes” link lists the thousands of references that are cited in the main collection (see Figure 10). These “outer” references include books as well as journal articles not included as sources in the *Web of Science*. They are sorted by citation frequency in the local network. The user can retrieve these items from *WoS* so that their full bibliographic data and GCS can be added to the initial collection. When the *HistCite* program is run again they will be integrated into the collection. A semi-automatic look-up of each item can be activated to *Web of Science* by clicking on WOS. A quick scan of these references will reveal some of the historical precedents to this topic, e.g., the work of Manfred Kochen, a mathematician qua information scientist who edited a book on the small world problem in 1989.

ISI Web of Science Location

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

Figure 10: Outer nodes – most cited works outside original collection

The system identifies questionable or “missing” citations where there is reason to believe there is an error or variation that prevents unification (see Figure 11). This expert system permits the reader to correct errors of omission in year or volume or pagination. Once corrected, these items can be fed back into the file so as to complete the citation counts. In the example shown for Gould’s paper, he has cited a paper by Korte but has omitted the page number.

When we first assembled the data on the Milgram 1967 paper, all the occurrences of page 60 turned up provided we captured the citations in the original *WoS* search. We then edited all those records, by using a search and replace command.

Potentially missed citations...

40 nodes have citations that may potentially refer to other nodes.

1 | [22](#) 1975 INTERNATIONAL SOCIAL SCIENCE JOURNAL 27(2):303-327

GOULD P

Mathematics in geography - Conceptual revolution or new tool



KORTE C, 1970, J PERSONALITY SOCIAL, V15 may refer to [9](#) KORTE-C-1970-V15-P101

2 | [95](#) 1985 SOCIAL NETWORKS 7(4):323-339

KOCHEN M

The structure of acquaintance nets and rates of societal development

KILLWORTH PD, 1978, SOCIAL NETWORKS, V1, P1959 may refer to [46](#) KILLWORTH-PD-1978-V1-P159

3 | [107](#) 1987 SOCIAL NETWORKS 9(2):153-169

HO E; KOCHEN M

Perceived acquaintanceship and interpersonal trust - The cases of Hong Kong and China

KILLWORTH PD, 1978, SOCIAL NETWORKS, V1, P195 may refer to [46](#) KILLWORTH-PD-1978-V1-P159

4 | [136](#) 1992 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 25():277-329

BLASS T

The social-psychology Of Milgram, Stanley

MILGRAM S, 1967, PATTERNS PREJUDICE, V1, P3 may refer to [4](#) MILGRAM-S-1967-V1-P61

5 | [224](#) 1999 SCIENCE 286(5439):509-512

Barabasi AL; Albert R

Emergence of scaling in random networks

BARTHELEMY M, 1999, PHYS REV LETT, V82, P15 may refer to [219](#) BARTHELEMY-M-1999-V82-P3180

BARTHELEMY M, 1999, PHYS REV LETT, V82, P15 may refer to [220](#) BARTHELEMY-M-1999-V82-P5180

Figure 11: Missing links

In order to help the user better visualize the inter-relations between the thousands of cited papers in the network, the software creates a citation matrix which displays the nodal numbers for citing and cited works. This matrix can then become input for the creation of co-citation maps of other kinds of maps. (see Figure 12)

Small World Citation Matrix

Nodes: 424

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

Page 1: 1

cited nodes	Cited nodes	Nodes	GCS	LCS	citing nodes
	0	1 1967 VARMA BN	2	1 2	
1	1	2 1967 EVERS HD	0	0	
	0	3 1967 AUCHTER EL	0	0	
	0	4 1967 MILGRAM S	148	167	6 7 9 10 11 12 13 14 15 16 18 19 20 22 23 24 25 26 27 28 29 30 31 32 33 38 43 44 45 46 47 48 49 51 52 53 54 55 56 58 61 62 63 64 66 71 72 76 78 80 81 82 87 90 95 96 103 107 110 112 126 128 129 136 137 143 148 149 155 172 174 178 182 188 193 194 195 196 202 210 211 215 217 218 219 221 224 225 229 240 243 249 250 253 254 256 257 258 261 263 268 269 270 272 273 275 276 278 280 286 289 296 298 299 300 301 302 305 308 309 310 311 313 317 319 322 323 324 326 328 329 333 334 337 342 344 345 346 347 350 353 357 359 368 369 370 373 374 383 385 387 389 400 403 407 408 409 411 412 413 414 416 417 420 421 423 424
	0	5 1969 ZALAZNICK S	0	0	
4	1	6 1969 TRAVERS J	64	43	9 10 11 13 16 18 19 21 23 24 27 28 29 33 38 43 44 45 46 47 48 49 51 53 54 55 56 61 66 76 78 80 82 96 103 136 148 196 269 270 326 417 418
4	1	7 1970 HART JW	0	0	
	0	8 1970 SOFFER RS	1	0	
4 6	2	9 1970 KORTE C	38	31	13 18 19 23 24 27 28 29 32 33 38 44 45 46 47 48 49 51 54 55 56 61 76 78 82 136 182 202 243 385 417
4 6	2	10 1970 WHITE HC	22	14	18 23 24 25 26 30 44 45 51 55 72 202 276 323
4 6	2	11 1971 JACOBSON D	8	0	
4	1	12 1972 LEVINE JH	93	3	16 25 149

Figure 12: Citation Matrix

Figure 13 is a map created on the Small World problem based on the data compiled by the *HistCite* program. The table briefly identifies each of the 15 most-cited papers in the local collection. The size of the circles are proportional to the citation frequency. Node #4, the paper by Milgram, has been cited 167 times while the much more recent paper by Watts, #194, has been cited locally 134 times. The large number of circles in the more recent years demonstrates how this field has mushroomed. And the Watts paper would be even more prominent if we created a map based on GCS.

HISTORIOGRAPH OF SMALL WORLD PAPERS 1967-2002

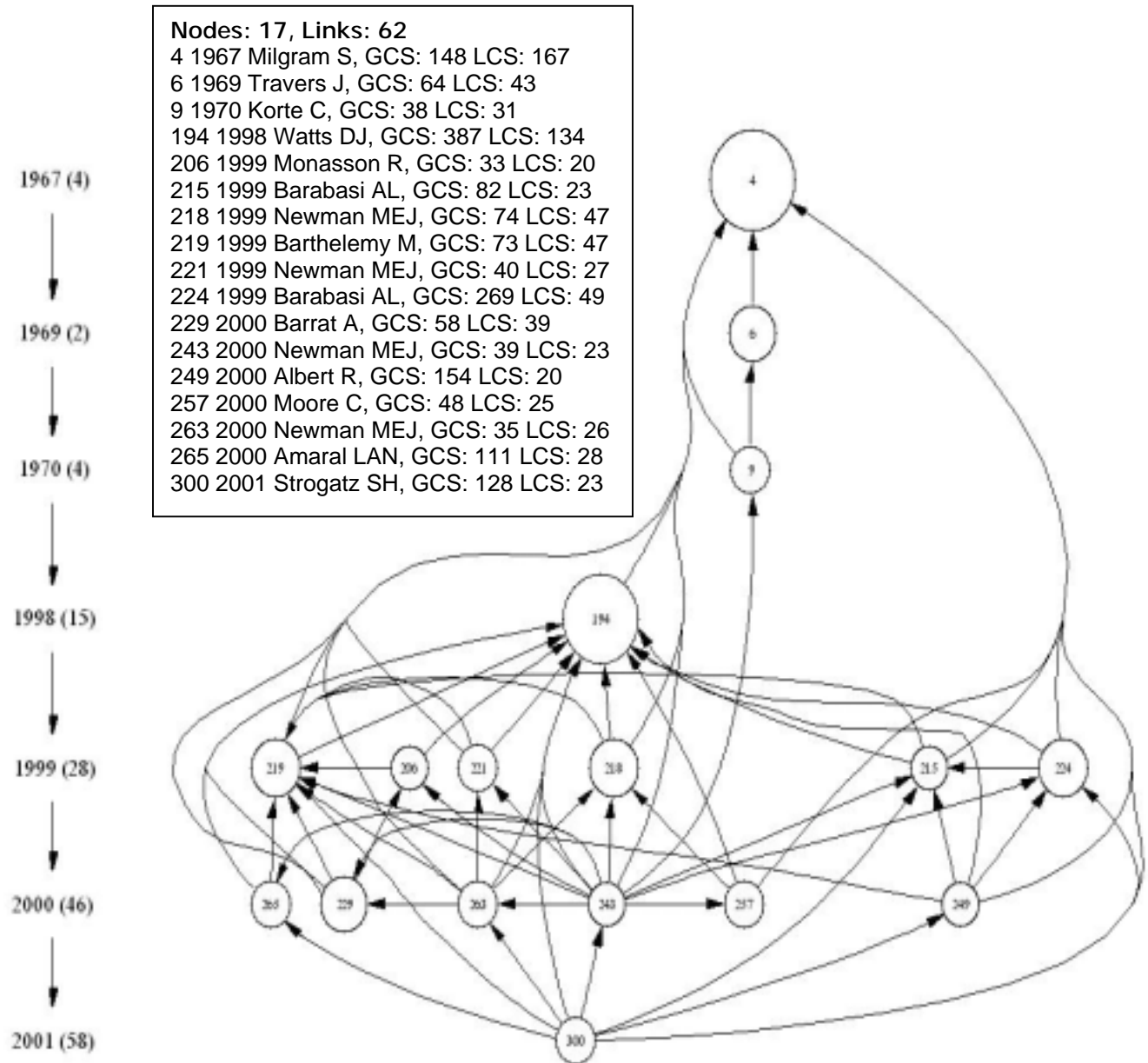


Figure 13: Historiograph of Small World Key Papers, 1967-2002

And when you are online you can click on each node to view the complete source record (see Figure 14).

4

Author(s)	MILGRAM S
Title	SMALL-WORLD PROBLEM
Journal	PSYCHOLOGY TODAY 1(1):61-67
Date	1967
Type	
Address	
Abstract	
GCS	148
LCS	167
cites	0
CR[3]	HARARY F, 1965, STRUCTURAL MODELS IN KEMENY JG, 1962, MATH MODELS SOCIAL S RAPOPORT A, 1963, HDB MATH PSYCHOLOGY, V2, PCH14

Figure 14: Source Record for Milgrams S, Psychology Today.

Figure 15 is an alternative method of displaying the information on this topic. If you were operating it on your PC, you could click on one node at a time and it would be highlighted in red. The green nodes are cited by node #30 and the blue are citing nodes. Cited papers are in green and citing papers are in blue. The next version of this software will permit use of circles which are proportional in size to citation frequency.

MAP OF 30 MOST-CITED PAPERS ON “SMALL WORLD” IN GLOBAL WOS

30 Nodes, 141 Links. Minimal LCS = 8. Sorted by LCS

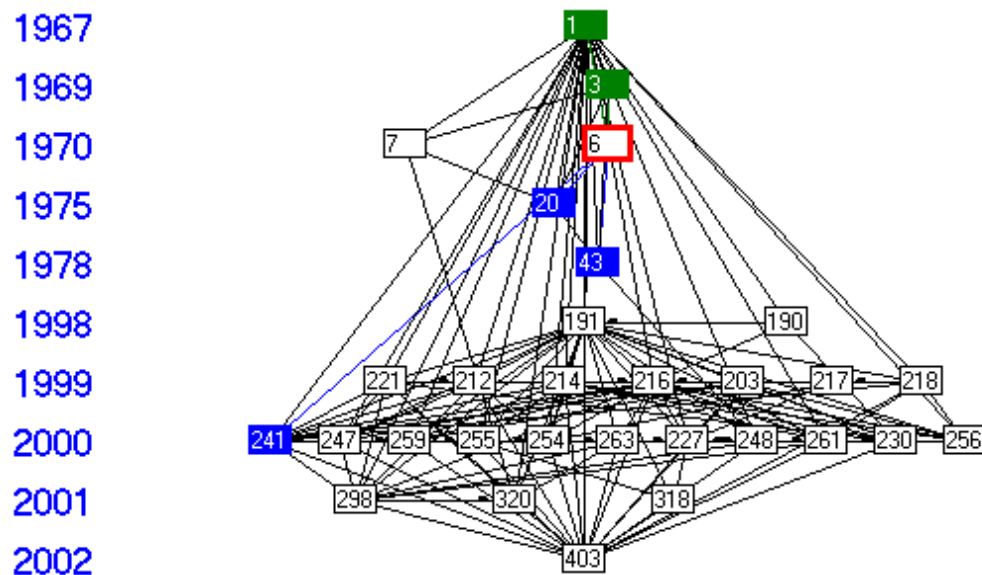


Figure 15: 2nd Small World Graph based on LCS

WATSON-CRICK 1953 DNA PAPER

Now I would briefly like to show you the use of *HistCite* in a truly historical mapping exercise. We are celebrating the 50th anniversary of the Watson-Crick discovery of the Double Helix structure of DNA. That 1953 paper was used to conduct a cited reference of the *SCI*. Since time is short, I am not going to dwell on the details of that search. In Figures 16, 17, 18, and 19, I simply want to show you the result of mapping the five years from 1953 to 1958.

In Figure 16, we have the usual HistCite table for the papers that cite Watson-Crick, 1953, but have added a few of the key outer references for Avery and Hershey.

[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 <i>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</i>	0	23
2	0	2 1952 JOURNAL OF GENERAL PHYSIOLOGY 36(1):39-56 HERSHEY AD; CHASE M <i>Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage</i>	747	23
3	2	3 1953 ACTA CRYSTALLOGRAPHICA 6(8-9):673-677 FRANKLIN RE; GOSLING RG <i>The Structure of Sodium Thymonucleate Fibres .1. The Influence of Water Content</i>	14	11
4	3	4 1953 ACTA CRYSTALLOGRAPHICA 6(8-9):678-685 FRANKLIN RE; GOSLING RG <i>The Structure of Sodium Thymonucleate Fibres .2. The Cylindrically Symmetrical Patterson Function</i>	10	8
5	1	5 1953 ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS 46(1):12-17 SMITH CL <i>The Breakdown of Desoxyribonucleic Acid Under Deuteron and Electron Bombardment</i>	5	1
6	2	6 1953 BIOCHEMICAL JOURNAL 55(5):774-782 WYATT GR; COHEN SS <i>The Bases of the Nucleic Acids of Some Bacterial and Animal Viruses - The Occurrence of 5-Hydroxymethylcytosine</i>	57	8
7	3	7 1953 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 18():123-131 WATSON JD; CRICK FHC <i>The Structure of Dna</i>	61	21
8	1	8 1953 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 8():133-134 WYATT GR <i>The Quantitative Composition of Deoxypentose Nucleic Acids As Related To the Newly Proposed Structure</i>	9	4
9	2	9 1953 COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY 18():171-183 LARK KG; ADAMS MH <i>The Stability of Phages As a Function of the Ionic Environment</i>	13	2

Figure 16: Citations to Watson-Crick

In Figure 17, we have the table for the 975 papers that were retrieved by virtue of chaining citations to the 210 papers that cited Watson-Crick. In other words, these are second generation citations to the citing papers in Figure 16.

[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 articles 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- SEVAG MG <i>[unknown]</i>	216	37
2	1	2 1944 JOURNAL OF EXPERIMENTAL MEDICINE 79():137-157 AVERY OT; MACLEON CM; MCCARTY M <i>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</i>	331	43
3	0	3 1945 JOURNAL OF BIOLOGICAL CHEMISTRY 161(1):83-89 SCHMIDT G; THANNHAUSER SJ <i>A Method for the Determination of Desoxyribonucleic Acid, Ribonucleic Acid, and Phosphoproteins in Animal Tissues</i>	696	34
4	1	4 1945 JOURNAL OF BIOLOGICAL CHEMISTRY 161(1):293-303 SCHNEIDER WC <i>Phosphorus Compounds in Animal Tissues .1. Extraction and Estimation of Desoxypentose Nucleic Acid and of Pentose Nucleic Acid</i>	952	30
5	2	5 1946 JOURNAL OF GENERAL PHYSIOLOGY 30(2):117-& MIRSKY AE; POLLISTER AW <i>Chromosin, a Desoxyribose Nucleoprotein Complex of the Cell Nucleus</i>	323	35
6	0	6 1947 JOURNAL OF THE CHEMICAL SOCIETY (SEP):1131-1141 GULLAND JM; JORDAN DO; TAYLOR HFW <i>Deoxypentose Nucleic Acids .2. Electrometric Titration of the Acidic and the Basic Groups of the Deoxypentose Nucleic Acid of Calf Thymus</i>	70	31
7	3	7 1951 BIOCHEMICAL JOURNAL 48(5):584-590 WYATT GR <i>The Purine and Pyrimidine Composition of Deoxypentose Nucleic Acids</i>	276	63
8	0	8 1951 JOURNAL OF BIOLOGICAL CHEMISTRY 189(2):597-605 MARSHAK A; VOGEL HJ <i>Microdetermination of Purines and Pyrimidines in Biological Materials</i>	136	30
9	0	9 1951 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 37(4):205-211 PAULING L; COREY RB; BRANSON HR <i>The Structure of Proteins - 2 Hydrogen-Bonded Helical Configurations of the Polypeptide Chain</i>	185	26
10	1	10 1952 BIOCHEMICAL JOURNAL 52(5):558-565 MARKHAM R; SMITH JD <i>The Structure of Ribonucleic Acids .2. The Smaller Products of Ribonuclease Digestion</i>	104	28
11	0	11 1952 JOURNAL OF GENERAL PHYSIOLOGY 36(1):39-56 HERSHEY AD; CHASE M <i>Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage</i>	206	

Figure 17: Chained Citations to Watson-Crick

In Figure 18, we have the year-by-year map of the 22 most-cited papers in the chained indexed file. Notice that in 1953 there were nine highly-cited papers. And in 1954, there are five. Using the typical reference citation, that is, only author, volume, page, and year, it is not possible to differentiate the month-by-month progression, after the April, 1953 Watson and Crick paper. However, the *HistCite* system can take into account the cover dates of the journals involved if they are included.

WATSON AND CRICK YEAR-BY-YEAR HISTORIOGRAPH

1938-1955

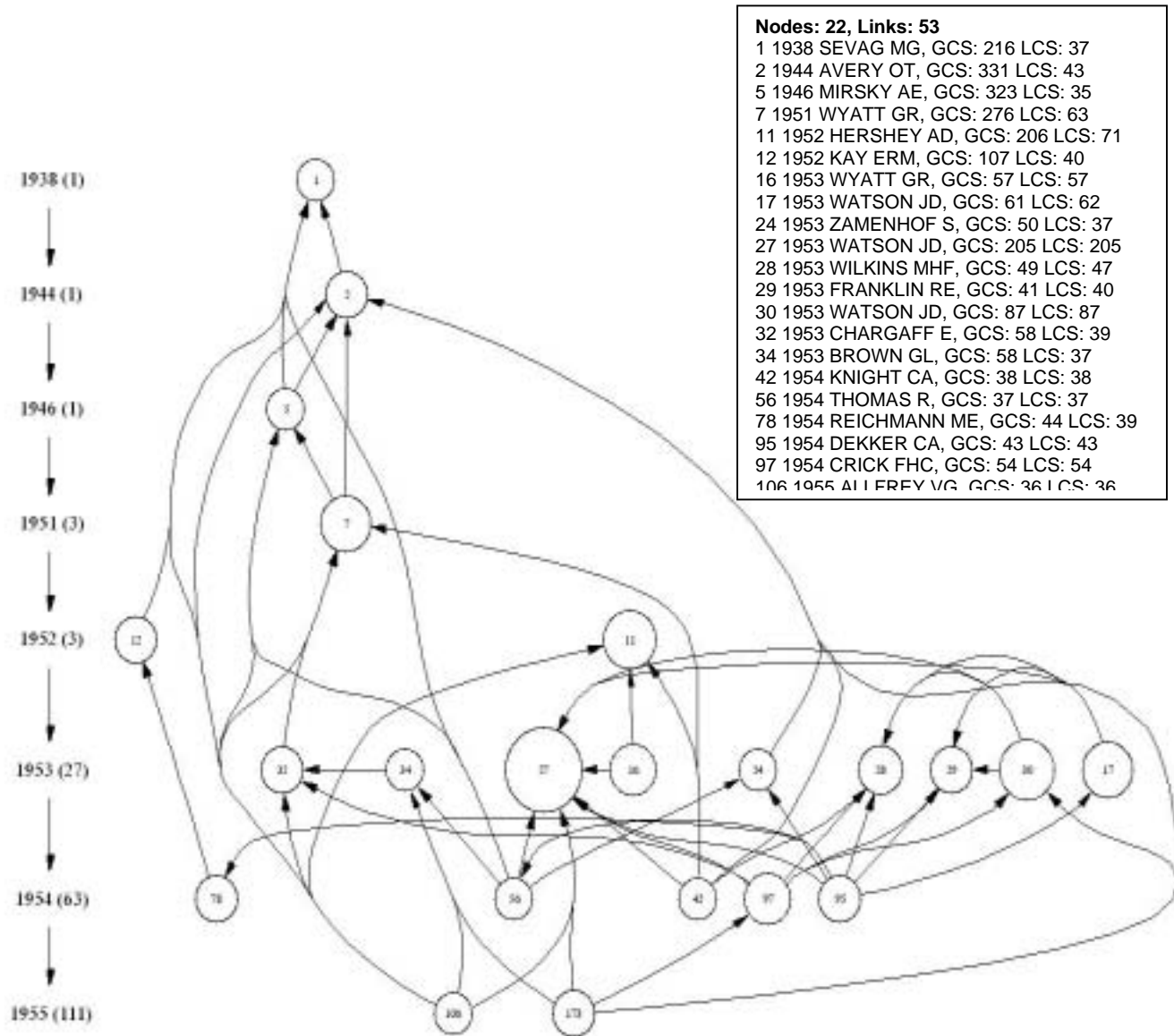


Figure 18: Watson-Crick – year-by-year historiography

In Figure 19, we see how the historiography changes not just year-to-year but month-by-month. Unfortunately, *WoS* does not contain cover dates until 1985 or there about. So we had to manually insert in the export files the cover dates for the few dozen papers involved in this example.

WATSON AND CRICK MONTH-BY-MONTH HISTORIOGRAPH FOR 1953

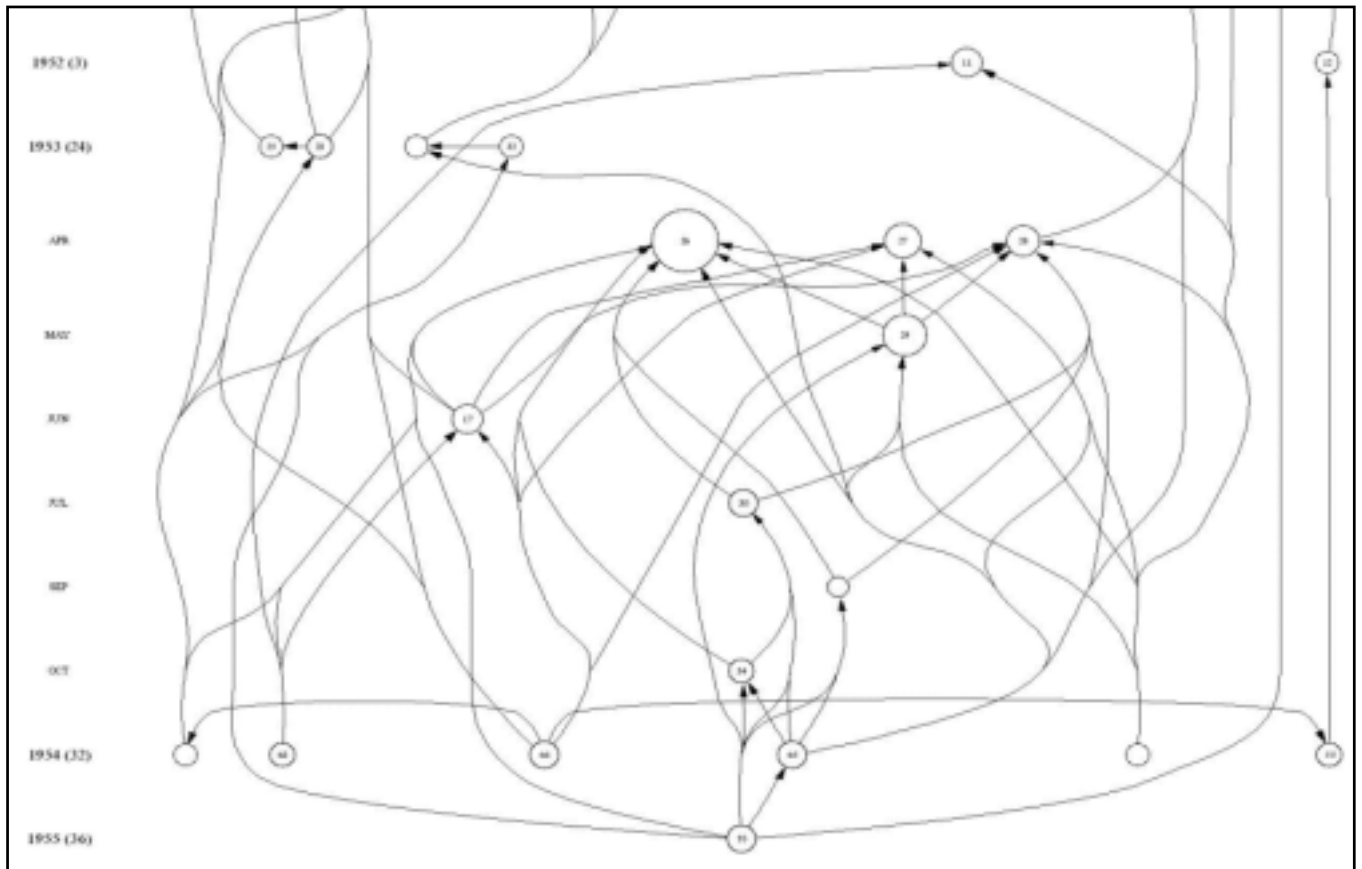


Figure 19: Watson-Crick Month-by-Month historiograph

We have described a tool which permits the user to manage the voluminous references produced in a comprehensive search of the literature. For those who are new to the subject, the mere juxtaposition of the most-cited papers for each five- or ten-year period of the literature will help identify the key literature to be used first. For those who are knowledgeable in the field, the system will help jog the memory to recall the key works which were associated with the development of the field. While the relevance of citing works may be apparent, the collective bibliographic coupling and co-citation of papers in and outside the basic bibliography should provide a comprehensive structure for completing a synoptic history of the topic.

To further illustrate the scope of the *HistCite* program, we include here several additional examples.

In this example in Figure 20, all papers published in the journal *Evolution* are listed. Note that the chronological sort key includes not only volume, page and year, but also issue number.

[Outer nodes?](#) [Missing links?](#) [Journal list](#) [Author list](#)

Citations to the journal *Evolution*

Nodes: 1975

Sorted by year.


Cited nodes	 Nodes / Authors	GCS	LCS
0	0 1947 EVOLUTION 1(4):249-262 HEISER CB <i>Hybridization Between the Sunflower Species Helianthus-Annuus and H-Petiolearis</i>	61	6
0	1 1947 EVOLUTION 1(4):221-236 WEIDENREICH F <i>The Trend of Human Evolution</i>	17	0
1	2 1947 EVOLUTION 1(3):143-153 KING JC <i>Interspecific Relationships Within the Guarani Group of Drosophila</i>	11	2
0	3 1947 EVOLUTION 1(1-2):89-102 HEUTS MJ <i>Experimental Studies on Adaptive Evolution in Gasterosteus- Aculeatus L</i>	92	7
0	4 1947 EVOLUTION 1(1-2):103-110 SPENCER WP <i>Genetic Drift In A Population Of Drosophila-Immigrans</i>	5	2
0	5 1947 EVOLUTION 1(3):217-218 MOODY PA <i>A Simple Model of Drift in Small Populations</i>	0	0
0	6 1947 EVOLUTION 1(3):191-216 DOBZHANSKY T; SPASSKY B <i>Evolutionary Changes in Laboratory Cultures of Drosophila-Pseudoobscura</i>	33	3
0	7 1947 EVOLUTION 1(1-2):32-41 STIRTON RA <i>Observations on Evolutionary Rates in Hypsodonty</i>	20	1
0	8 1947 EVOLUTION 1(1-2):48-62 KING JC <i>A Comparative Analysis of the Chromosomes of the Guarani Group of Drosophila</i>	14	3

Figure 20: Chronological file of papers published in *Evolution*

The WOS output file is normally arranged in reverse chronological order. The ISI production procedure is, of necessity, often asynchronous with the cover dates. Thus a December 2000 article may appear in WOS in January 2001. Also, timing variations occur when ISI adds back issues to its cumulative files. So it is not possible to rely on the WOS, at present, to produce listings in perfect chronological arrangement, even for a single journal.

While the *HistCite* program will produce a nearly perfect sort for a single journal's records, it is not always possible to do this for a collection covering articles from many journals since Volume numbers vary and are not correlated with exact publication dates. However, if the monthly or weekly dates are included in the original WOS records as was done for Figure 19, then the file will be in almost perfect order.

As another illustration of a keyword search, consider the topic of "gene flow." This is of considerable interest to one of us (Pudovkin). Instead of a cited reference search, we conducted a typical title word search in WoS. Between 1974 and 2001, 620 papers were published on this topic.

Figure 21 shows a portion of the full chronological collection. The gray areas indicate papers which are neither cited in the collection nor cite any other paper in the collection.

[Outer References](#) [Missing Links?](#) [Journal list](#) [All-Author list](#) [Citation Matrix](#)

Geneflow Papers - 1974 to August 2001

Nodes: 620

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

Cited nodes	Nodes / Authors	GCS	LCS
0	1 1974 GENETICS 78(3):961-965 SPIETH PT <i>Gene Flow and Genetic Differentiation</i>	43	9
0	2 1975 AMERICAN NATURALIST 109(969):597-601 SLATKIN M; MARUYAMA T <i>Influence of Gene Flow on Genetic Distance</i>	21	6
0	3 1975 AMERICAN NATURALIST 109(970):659-676 MAY RM; ENDLER JA; MCMURTRIE RE <i>Gene Frequency Clines in Presence of Selection Opposed by Gene Flow</i>	88	15
0	4 1975 AUK 92(3):493-510 COOKE F; MACINNES CD; PREVETT JP <i>Gene Flow Between Breeding Populations f Lesser Snow Geese</i>	71	3
0	5 1975 GENETICS 80(2):349-361 MCKENZIE JA <i>Gene Flow and Selection in a Natural Population of Drosophila- Melanogaster</i>	17	0
0	6 1975 GENETICS 81(4):787-802 SLATKIN M <i>Gene Flow and Selection in a 2-Locus System</i>	52	3
0	7 1975 HEREDITY 34(JUN):407-415 BRUSSARD PF; VAWTER AT <i>Population Structure, Gene Flow and Natural Selection in Populations of Euphydryas-Phaeton</i>	9	0

Figure 21: Chronological table of papers on gene flow from 1974 to 2001

When you click on LCS, the papers by Slatkin move to the top. This is not a surprising result as he is the leader in this field. 29 of these were used to create an historiograph. Only a portion of the 29 papers cited 10 or more times is shown.

[Outer References](#) [Missing Links?](#) [Journal list](#) [All-Author list](#) [Citation Matrix](#)

Gene Flow Papers – 1974 to August 2001

See the Historiograph of the 29 most cited papers in LCS by clicking [here](#)

Nodes: 620

Sorted by **LCS**.



Cited nodes	Nodes / Authors	GCS	LCS
11	71 1985 ANNUAL REVIEW OF ECOLOGY AND SYSTEMATICS 16():393-430 SLATKIN M <i>Gene Flow in Natural Populations</i>	554	111
4	121 1987 SCIENCE 236(4803):787-792 SLATKIN M <i>Gene Flow and the Geographic Structure of Natural Populations</i>	646	104
2	76 1985 EVOLUTION 39(1):53-65 SLATKIN M <i>Rare Alleles as Indicators of Gene Flow</i>	536	100
4	153 1989 EVOLUTION 43(7):1349-1368 SLATKIN M; BARTON NH <i>A Comparison of 3 Indirect Methods for Estimating Average Levels of Gene Flow</i>	401	82
0	37 1981 GENETICS 99(2):323-335 SLATKIN M <i>Estimating Levels of Gene Flow in Natural Populations</i>	220	53
0	29 1980 NATURE 284(5755):450-451 SCHAAL BA <i>Measurement of Gene Flow in Lupinus-Texensis</i>	165	39
4	31 1981 ANNALS OF THE MISSOURI BOTANICAL GARDEN 68(2):233-253 LEVIN DA <i>Dispersal Versus Gene Flow in Plants</i>	190	37
5	112 1987 EVOLUTION 41(2):385-400 WAPLES RS <i>A Multispecies Approach to the Analysis of Gene Flow in Marine Shore Fishes</i>	198	30
4	64 1984 GENETICS 106(2):293-308 LARSON A; WAKE DB; YANEV KP <i>Measuring Gene Flow Among Populations Having High-Levels of Genetic Fragmentation</i>	119	28
0	16 1977 THEORETICAL POPULATION BIOLOGY 12(3):253-262 SLATKIN M <i>gene flow and genetic drift in a species subject to frequent local extinctions</i>	162	25
3	156 1989 GENETICS 123(3):603-613 SLATKIN M; MADDISON WP <i>A Cladistic Measure of Gene Flow Inferred from the Phylogenies of Alleles</i>	171	22

Figure 22: Gene flow collection sorted by Local Citation Score (LCS)

The genealogical graphical presentation is seen in Figure 23. Note that each rectangular node is hot linked to a full source entry (Figure 24).

GENE FLOW – 29 Most Cited LCS Papers

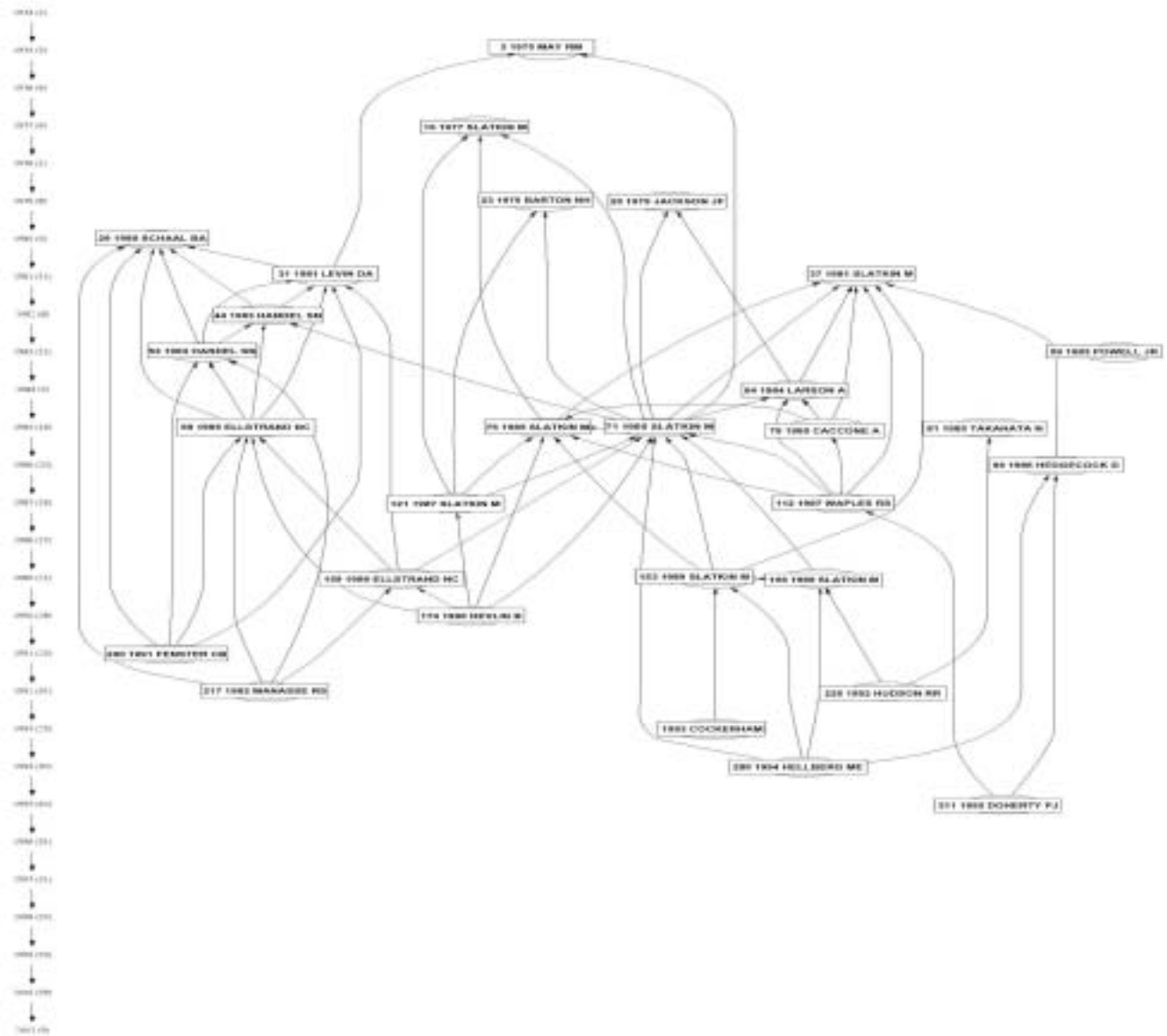


Figure 23: Computer generated historiograph of “gene flow” most-cited papers.

Node #3 at the top of Figure 23 is Richard May's 1975 paper in the *American Naturalist* (see Figure 24).

3

Author(s)	MAY RM; ENDLER JA; MCMURTRIE RE
Title	GENE FREQUENCY CLINES IN PRESENCE OF SELECTION OPPOSED BY GENE FLOW
Journal	AMERICAN NATURALIST 109(970):659-676
Year	1975
Type	Article
Address	PRINCETON UNIV,BIOL DEPT,PRINCETON,NJ 08540
Abstract	
WoS CS	88
LCS	15
cites	0
CR[17]	BARBER HN, 1965, HEREDITY, V20, P551 BARBER HN, 1957, NATURE, V179, P1267 BISHOP J, 1972, J ANIM ECOL, V4, P209 CROW JF, 1970, INTRO POPULATION GEN ENDLER JA, 1973, SCIENCE, V179, P243 ENDLER JA, 1976, SUBSPECIES SPECIES C FISHER RA, 1937, ANN EUGEN, V7, P355 FISHER RA, 1950, BIOMETRICS, V6, P353 HALDANE JBS, 1948, J GENET, V48, P277 HANSON WD, 1966, BIOMETRICS, V22, P453 JAIN SK, 1966, HEREDITY, V21, P407 KETTLEWELL HBD, 1969, HEREDITY, V24, P1 KETTLEWELL HBD, 1969, HEREDITY, V24, P15 KETTLEWELL HBD, 1961, HEREDITY, V16, P403 KIMURA M, 1958, 9 NAT I GEN ANN REP, P84 ROUGHGARDEN J, 1974, AM NAT, V108, P649 SLATKIN M, 1973, GENETICS, V75, P733

Figure 24: Full source entry for node #3, paper by R. M. May, *American Naturalist*, 1975.

In Figure 25, the same set of 29 LCS papers is shown in another format. Each document is represented by a circle whose area is proportional to its citation frequency.

This last figure provides yet another format (see Figure 26). It is necessary to see this dynamically since it is possible to navigate the network by moving the cursor from node to node. In this case, node #71 is in white and the lines show links in two directions – cited and citing.

Time and space do not permit us to discuss “Why Do We Need Algorithmic Historiography?” A paper on this theme has been published in a special issue of the *Journal of the American Society for Information Science and Technology* on “Visualization of Scientific Paradigms.”⁸

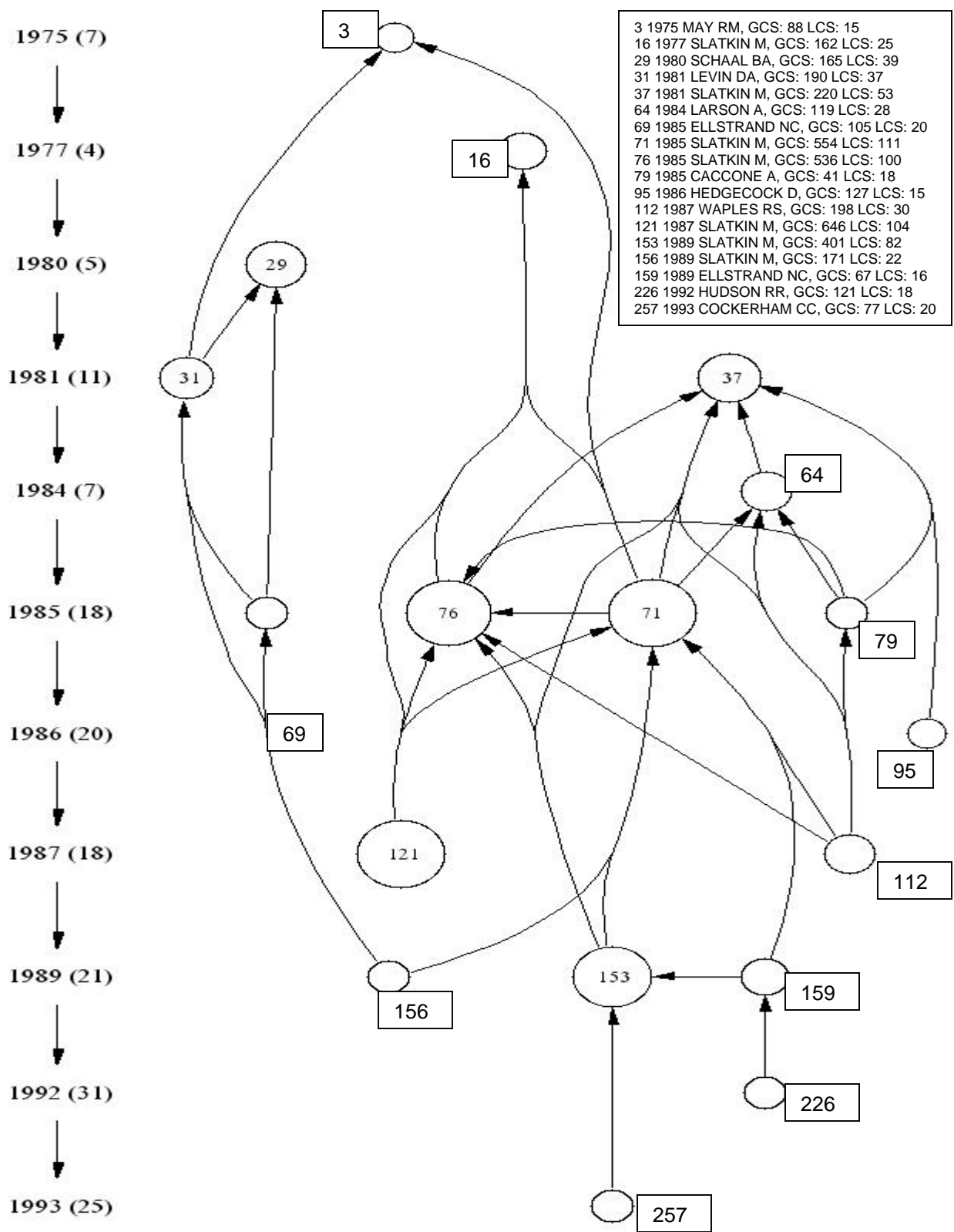


Figure 25: Gene flow, 1974-2001, collection with each paper represented by a circle proportional to citation frequency

GENE FLOW: 1974 TO AUGUST 2001

50 Nodes, 138 Links. Minimal LCS = 7. Sorted by LCS

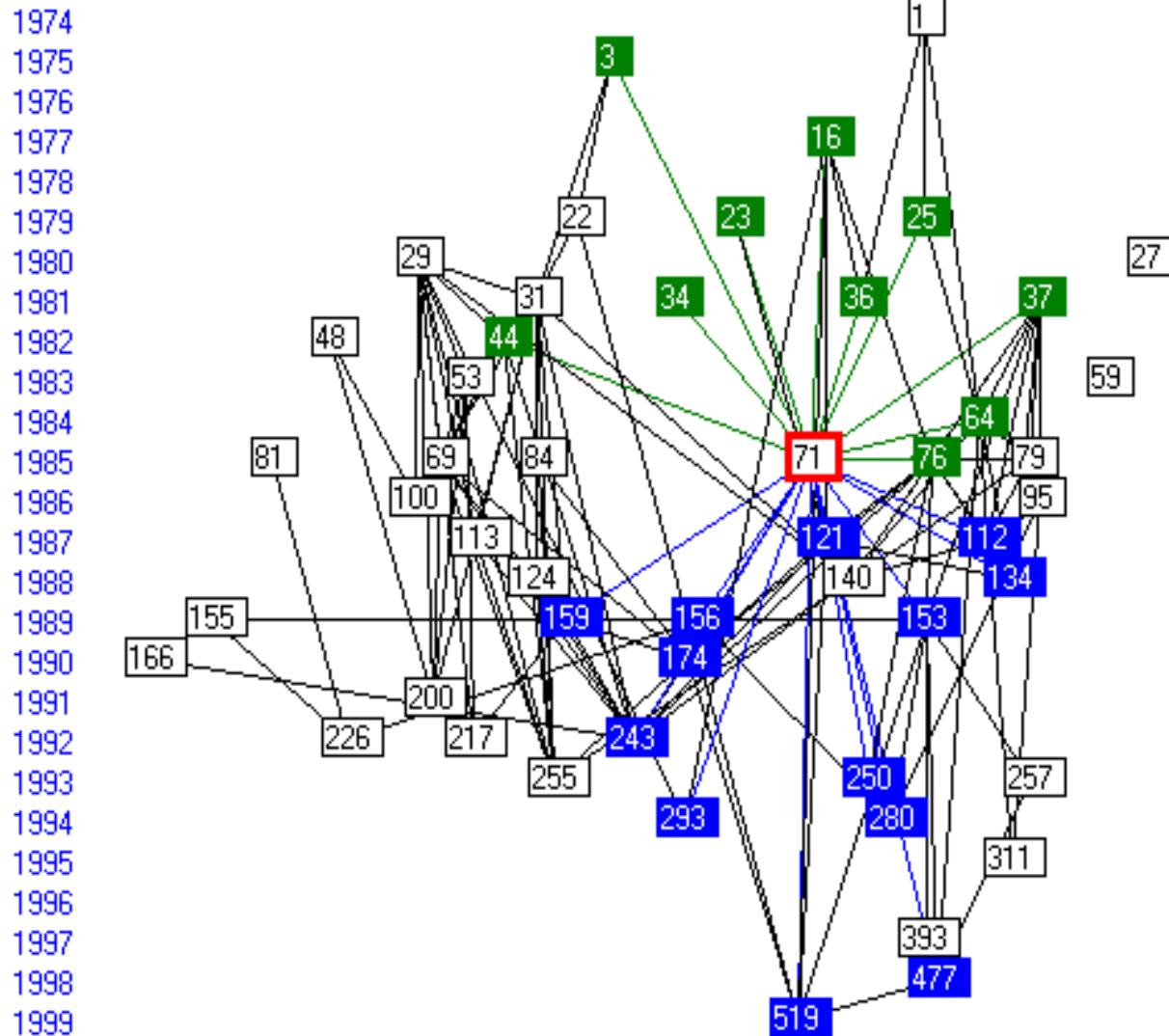


Figure 26: Gene Flow 1974 to August, 2001 in alternate format for PC use

In conclusion, let me emphasize that *HistCite* is a work in progress. For example, in the next version of the software a table of word frequencies will be provided so that the most-common vocabulary is identified.

REFERENCES:

1. Garfield, E., Sher I.H., & Torpie R.J. (December 1964). The Use of Citation Data in Writing the History of Science. Report of research for Air Force Office of Scientific Research under contract AF49(638)-1256. Philadelphia: The Institute for Scientific Information, December 1964. Available: <http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf>
2. Garfield, E. (April 14, 1971) Citation indexing, historio-bibliography and the sociology of science biography. Current Contents, No. 15, pages M25+. Reprinted from: K.E. Davis & W.D. Sweeney (Eds) Proceedings of the Third International Congress of Medical Librarianship 5-9 May 1969. (pp. 187-204). Amsterdam: Excerpta Medica. Reprinted in Garfield E. (1977). Essays of an Information Scientist (Vol 1, pp.158-174). Philadelphia: ISI Press (1977). Available: <http://www.garfield.library.upenn.edu/essays/V1p158y1962-73.pdf>
3. Cawkell, A.E. (October 30, 1989). Acoustic journals and acoustic research articles. Current Contents No. 44, 4-15. Reprinted in Garfield, E. (1977) Essays of an Information Scientist, (Vol. 12, pp. 4-15). Philadelphia: ISI Press. Available: <http://www.garfield.library.upenn.edu/essays/v12p301y1989.pdf>
- Cawkell, A.E. (2000). Visualizing Citation Connections. In B. Cronin & H.Barsky Atkins (Eds), The Web of Knowledge: A Festschrift in Honor of Eugene Garfield (ASIS Monographic Series) (pp. 177-194). Medford, NJ: Information Today, Inc.
4. Garfield, E. (April 14, 1971) Citation indexing, historio-bibliography and the sociology of science biography. Current Contents, No. 15, pages M25+. Reprinted from: K.E. Davis & W.D. Sweeney (Eds) Proceedings of the Third International Congress of Medical Librarianship 5-9 May 1969. (pp. 187-204). Amsterdam: Excerpta Medica. Reprinted in Garfield E. (1977). Essays of an Information Scientist (Vol 1, pp.158-174). Philadelphia: ISI Press (1977). Available: <http://www.garfield.library.upenn.edu/essays/V1p158y1962-73.pdf>
5. Garfield, E. (June 8, 1992). Contract Research Services at ISI—Citation Analysis for governmental, Industrial, and Academic Clients, Current Contents No. 23, 5-13. Reprinted in Garfield E. (1993). Essays of an Information Scientist (Vol. 15, pp. 75-83). Philadelphia: ISI Press (1993). Available: <http://garfield.library.upenn.edu/essays/v15p075y1992-93.pdf>
- Small, H. and Garfield, E. (1985). The geography of science: disciplinary and national mappings, Journal of Information Science, 11, 147-159 (1985). Reprinted in Garfield E. (October 27, 1986) Current Contents No. 43, 3-14. Reprinted in Garfield E. (1988). Essays of an Information Scientist. (Vol. 9, pp. 324-335). Philadelphia: ISI Press. Available: <http://www.garfield.library.upenn.edu/essays/v9p324y1986.pdf>
6. Small, H. (1994). A Sci-Map case study: Building a map of AIDS Research. Scientometrics, 30, :229-241 (1994)
Small, H. , Sweeney, E., and Greenlee, E. (1985). Clustering the Science Citation Index using co-citations. 2. mapping science. Scientometrics, 8, 321-34
7. Milgram S. "Small World Problem," Psychology Today, 1(1):61-67 (1967)
8. Garfield E, Pudovkin AI, Istomin VS. "Why do we need Algorithmic Historiography?" *Journal of the American Society for Information Science and Technology* (JASIST) 54(5):400-412 (March 2003) [http://garfield.library.upenn.edu/papers/jasist54\(5\)400y2003.pdf](http://garfield.library.upenn.edu/papers/jasist54(5)400y2003.pdf)