
In memoriam

**From vanity fair to scientific research:
The place of genealogy in contemporary science.
A scientometric approach**

SÁNDOR ZSINDELY[†]

*Information Science and Scientometrics Research Unit (ISSRU), Hungarian Academy of Sciences,
Budapest (Hungary)*

The place of genealogy in present scientific research has been investigated by scientometric methods. The term “genealogy” and related words were searched in the title, keywords, and abstracts of science journals for the period 1975–2006. It was concluded that 1991 onward the number of articles about “applied” genealogy has increased dramatically, whereas that of classical (or “pure”) genealogy only modestly. In contemporary science, the fields medicine and genetics are those who profit most from human genealogy. More than forty percent of the medical articles containing the search terms were from the neurology and oncology in the period investigated.

Introduction

One of the most frequent questions in philosophy is where do we come from and where are we going to? To solve at least partially the first half of the problem, the people had been made family trees since the earliest days to register their ancestors.

[†] Sándor Zsindely worked for the Editorial Office of *Scientometrics* from the very day of the launching of the journal until his day of death on May 8, 2008. He was 76 years old. This manuscript was set aside for completion a few months ago, but has been and will now be never finished. We tribute to the memory of our dear friend and colleague by publishing his last contribution.

Genealogy is the study or investigation of ancestry and family histories. The search for descendents includes also the lifestyle, biography and motivations of the individuals. In this sense, it contributes to a better understanding of history, of contemporary political and economical situations, as well as of property status and financial condition of the family in question. In its original form, genealogy was mainly concerned with the ancestry of rulers and aristocrats, often arguing or demonstrating the legitimacy of claims to wealth and power. Therefore, it was by no means a rarity that diplomas and certificates were falsified or non-existent but very highborn ancestors had been fabricated due to pure family interest or mere human vanity. Despite that, the archives of higher nobility which preserved the most detailed genealogical documents are now invaluable sources mainly for historians [1].

Until recently, genealogy, together with heraldry and sigilography, remained one of the auxiliary sciences of history. But with the advent of the Internet the situation has dramatically changed. At first, genealogy became an extreme popular hobby and the number of resources available to genealogists (for professionals as well as to hobbyists) has vastly increased, despite the fact that not all of them supply reliable information. At second, genealogy has found also its way into contemporary science after the 90's. With the discovery that a person's DNA contains information that has been passed down relatively unchanged from our earliest ancestors, analysis of DNA has begun to be used for genealogical research. Genetic methods are being applied to trace human migratory patterns and to determine bio-geographical and ethnic origins; in medicine, they can be employed for the investigation of some genetic predestined features and diseases; and so on, to mention only some of science fields in which genealogical methods have been found utilization in our days.

In this paper we will demonstrate the continuously increasing diffusion of genealogy into the sciences during the last three decades. For this reason the content of scientific articles was analysed searching for the term "genealogy" and related expressions in the text.

This method for looking for new trends in the sciences is based upon the assumption that "the present system of basic research in the sciences and scientific communication depends almost entirely on the primary journal literature. Modern science has developed a particular mechanism of communication which began with the appearance of the first scientific journals in the 17th century, and which remained practically the same ever since. This mechanism is based on the selective publication of fragments rather than complete treatises. [...] The fact that a paper has been accepted for publication in a well known refereed journal is probably the best immediate indication that it reports worthwhile research." [3]

Experimental

Databases

For the investigation of the role of genealogy in nowadays science, we used the Web of Science of Thomson Scientific, Philadelphia, USA. This includes the databases Science Citation Index (SCI) (extended), Social Sciences Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI).

Results and discussion

For getting the most relevant result as possible, a search using two terms coupled with Boolean AND was employed. In this sense the terms “genealog* AND famil*”, as well as “genealog* AND ancest*” were used to search for relevant information in the text of title, abstracts and keywords of the journal articles registered in the abovementioned databases for the period between 1975 and 2006. For the term “genealog* AND famil*” a total of 857 articles were found (Table 1), for “genealog* AND ancest*” this figure was 552 (Table 2).

Checking the relevancy of the selected journal articles it turned out that for the term “genealog* AND famil*” approx. 83 percent of the gained articles are really dealing with genealogy of human families, whereas for “genealog* AND ancest*” only approx. 62 percent of the articles were relevant. The topics of the irrelevant papers came from zoology, botany, mathematics and other fields.

For both terms, the distribution of the number of articles per year can be seen in Figures 1 and 2, respectively. It is noticeable that after a section of a modest increase; both curves begin to rise steeply, from approximately 1991 onward. During the first half of the period, i.e., between 1975 and 1990, only 81 articles appeared in our topic, but between 1991 and 2006, this figure increased to 775 in the case of “genealog* AND famil*”.

In the followings we used the articles selected by the term “genealog* AND famil*” without any corrections, because we suppose that this database is also in this form characteristic for the presence of genealogical methods in contemporary science. This is especially true for the period of sharp increase in the number of journal articles, i.e., between 1991 and 2006. Therefore, we have concentrated our efforts to the investigation of those articles, which were published between 1991 and 2006 and contain the term “genealog* AND famil*” in their text.

The classification of articles into science fields can be done e.g., according to the affiliation of the journal publishing the article in question. In this case, however, we found this method not to be appropriate, because many journals e.g., in genetics publish articles of medical interest, and vice versa. Moreover, in a veterinary journal, the family

tree of the founders of the veterinary school of Vienna can be read, and the *Russian Journal of Genetics* dealt with the genealogy of the members of the Russian Academy of Sciences, to mention some of the curiosities.

Table 1. The term “genealog* AND famil**” in journal articles of the databases SCI, SSCI and A&HCI for 1975–2006 by year

Year	Total		Pure genealogy	
	No. of articles	Cumulative	No of articles	Cumulative
1975	1	1	1	1
1976	4	5	3	4
1977	5	10	5	9
1978	6	16	3	12
1979	7	23	4	16
1980	2	25	1	17
1981	8	33	7	24
1982	6	39	3	27
1983	3	42	0	27
1984	3	45	3	30
1985	7	52	6	36
1986	5	57	4	40
1987	2	59	1	41
1988	7	66	6	47
1989	7	73	4	51
1990	8	81	4	55
1991	31	112	4	59
1992	34	146	1	60
1993	36	182	1	61
1994	30	212	1	62
1995	31	243	3	65
1996	49	292	4	69
1997	49	341	8	77
1998	49	390	4	81
1999	54	444	11	92
2000	44	488	8	100
2001	55	543	13	113
2002	65	608	8	121
2003	62	670	14	135
2004	63	733	8	143
2005	48	781	2	145
2006	76	857	4	149

The selection of the articles for the science fields and subfields were, therefore, done manually by the title of the articles. As a result we found three groups of items, i.e., publications (1) about genealogy in the classical sense (we named it “pure genealogy”, (2) about other researches using genealogical data (“applied genealogy”),¹ and (3) in which the term “genealogy” has not its original meaning and is used in a quite other

¹ Ad analogiam “pure and applied chemistry”

context. This later group represents only approx. 17 percent of the total and is of no interest in the present study.

The cumulative number of publications of the first group containing papers about pure genealogy can be seen in Table 1 and Figure 1, for the 1975–2006 period. Comparing the relevant curve with that of the total, i.e., the cumulative number of all the articles found by the term “genealog* AND famil*”, a remarkable difference can be observed. While the later raises sharply after 1991, the former shows a more flatten increase during the investigated years (Figure 1).

Table 2. The terms “genealog* AND ancest*” in journal articles of the databases SCI, SSCI and A&HCI between 1975 and 2006 by year

Year	No. of articles	Cumulative
1975	1	1
1976	0	1
1977	1	2
1978	0	2
1979	0	2
1980	0	2
1981	0	2
1982	1	3
1983	1	4
1984	0	4
1985	0	4
1986	0	4
1987	0	4
1988	2	6
1989	0	6
1990	1	7
1991	10	17
1992	17	34
1993	17	34
1994	16	67
1995	29	96
1996	25	121
1997	30	151
1998	41	192
1999	33	225
2000	28	253
2001	41	294
2002	54	348
2003	35	383
2004	42	425
2005	55	480
2006	72	552

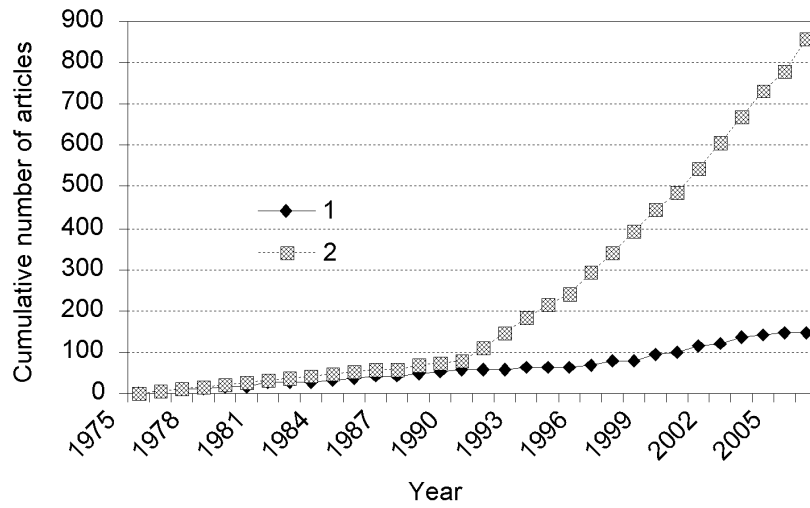


Figure 1. Cumulative number of articles found for the term “genealog* AND famil*” for the years 1975–2006. 1 – Pure genealogy, 2 – Total

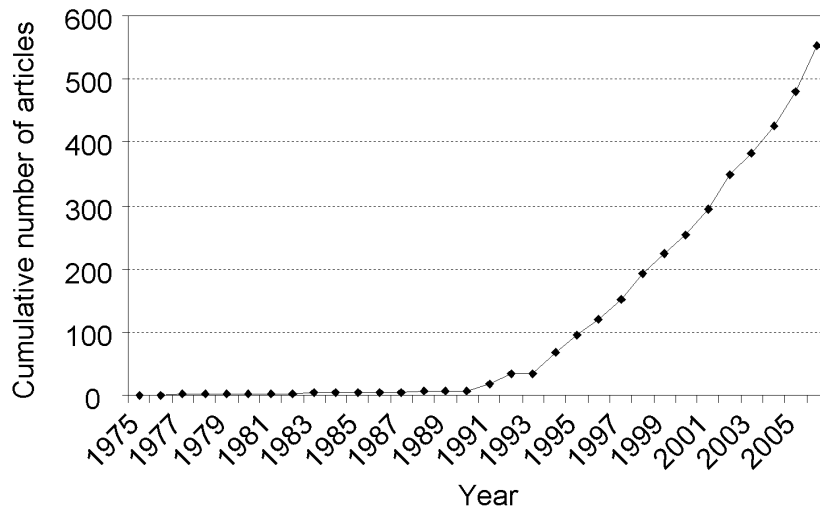


Figure 2. Cumulative number of articles found for the term “genealog* AND ancest*” for the years 1975–2006

But there is also a difference between the so-called pure genealogy and the (human) genealogy applied in the sciences. In the classical sense, the family trees register only the paternal line of the descent in the most cases, and the maternal lines were usually neglected. The cause of that originates in the right of inheritance of the older times. However, in the sciences both paternal and maternal lines were taking into account, moreover, the later are usually preferred.

The curve for the cumulative number of articles found by the term “genealog* and famil*” has two distinct parts, one for 1975–1990, and, an other one, for 1991–2006. For the first period, the distribution by science fields of the articles are summarized in Table 3, that for the second one in Table 4. As seen from the tables, in the first period, 55 of the total of 81 articles (i.e., 67.90 percent) deal with pure genealogy, and 21 (25.92 percent) with medicine. In the second, longer period these figures are 94 (12.13 percent) and 409 (52.78 percent), respectively. That is, the ratio between pure genealogy and medicine has decreased from 2.62 in the first period to 0.23 in the second one.

Table 3. Number of journal articles containing the term “genealog* AND famil*” between 1975 and 1980 by science fields

Field	No of articles	Percent
Genealogy	55	67.90
Medicine	21	25.92
Genetics	3	3.70
Mathematics	1	1.23
Sociology	1	1.23
Total	81	99.98

It is interesting to note that of the 146 genealogy items for the period 1975–2006, 60 are book reviews, and only 3 book reviews belong to other disciplines. This can be taken as evidence that until today, the publication channel of pure genealogy remained mainly the book form.

In the followings, we will investigate the second part of the curve for articles selected by the term “genealog* and famil*”, i.e., for the period 1991–2006, ranking the science fields by the number of articles published (Table 4). As seen, the first three places are occupied by medicine, pure genealogy and genetics, making approximately three-quarters of the total. In the medicine field we have ordered all the articles dealing with whatever subfields of medicine ever, irrespective of the publishing journal. In genetics, only relevant publications of the science fields unlike to medicine have remained.

Table 4. Number of journal articles containing the term “genealog* AND famil*” between 1991 and 2006 by science fields

Field	No of articles	Percent
Medicine	409	52.78
Basic genealogy	94	12.13
Genetics	76	9.81
Sociology	49	6.32
History	26	3.35
Biology	19	2.45
Botany	15	1.93
Mathematics	15	1.93
Chemistry	12	1.55
Zoology	12	1.55
Anthropology	7	0.90
Ecology	7	0.90
Entomology	5	0.65
Linguistics	5	0.65
Literature	4	0.52
Physics	4	0.52
Informatics	3	0.39
Veterinary medicine	3	0.39
Arts	3	0.39
Geography	2	0.26
Library science	1	0.13
Behaviour science	1	0.13
Microbiology	1	0.13
Pharmacology	1	0.13
Philosophy	1	0.13
Total	775	100.02

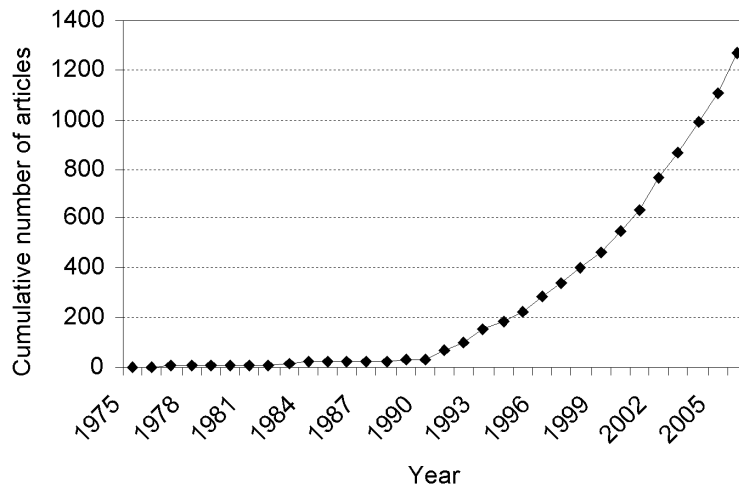


Figure 3. Cumulative number of articles found for the term “genealog* AND genetic*” for the years 1975–2006

Table 5. Number of journal articles about medicine containing the term “genealog* AND famil*” between 1991 and 2006 by subfields

Subfield	No. of articles	Percent
Neurology	112	27.38
Oncology	59	14.42
Ophthalmology	34	8.31
Cardiology, pulmonology	24	5.86
Urology and nephrology	22	5.38
Anatomic disorders	21	5.13
Hematology	20	4.89
Geriatrics	17	4.16
Dermatology	12	2.93
Rheumatism	11	2.69
Gastroenterology, diabetology	9	2.20
Psychiatry	7	1.71
Epidemiology	7	1.71
Endocrinology	6	1.47
Immunology	5	1.22
Psychology	5	1.22
Cytology	5	1.22
Gynecology and obstetrics	5	1.22
Anatomy	4	0.98
Bacteriology	3	0.73
Medical care	3	0.73
Otolaryngology	3	0.73
Pediatrics	3	0.73
Drug addictology	2	0.49
Forensic medicine	2	0.49
Pediatrics	2	0.49
General medicine	2	0.49
Preventive medicine	1	0.24
Anesthetics	1	0.24
Pathology	1	0.24
Surgery	1	0.24
Total	409	99.94

It goes without saying that in the sciences genealogy stays in a close contact with genetics. This can be illustrated very well by the search with the term “genealog* AND genetic*” for the 1975–2006 period (Figure 3). The raise of that curve begins in the same year, i.e., 1991, as that of the articles found for the term “genealog* AND famil*” and presented in Figure 1. During this period the number of articles published per year and found by the term “genealog* AND genetic*” had increased from 2 in 1975, to 158 in 2006, giving in sum 1267 items. However, this group of articles contains mainly less relevant articles than does that for the term “genealog* AND famil*”, and, therefore, we did not use it for searching.

The 409 medicine papers are grouped in subfields and ranked by the number of articles in Table 5. It came as no surprise that on the first two places of Table 5 neurology and oncology emerged with 112 (27.38 percent) and 59 (14.42 percent) papers, respectively. The inheritability of some diseases belonging to the abovementioned two medical subfields are quite well known. The third in the rank is ophthalmology, with 34 publications (8.31 percent). Cardiology, urology, anatomic disorders, as well as haematology have a share higher than or near to 5 percent in that ranking. These seven subfields represent more than 70 percent of the total. The other 24 share the rest.

Conclusions

It is not our task to draw far-reaching conclusions from that short scientrometric analysis. Anyway, it is a fact that at least from 1991, human genealogy is not merely an auxiliary science of history, still less a territory for family-researcher hobbyists, but a discipline whose results are used by researchers of various science fields, especially those of medicine and genetics. In medicine, the subfields neurology, oncology, ophthalmology, cardiology, and urology are the main beneficiaries of genealogical tables. The “applied” form of genealogy became more “democratic” and “emancipated”, than the classical (“pure”) one of the past. In medical genetics the whole community irrespective of the social stand of the individual is the subject of research and both lines of the family tree, the paternal as well as the maternal, are equally taken into account.

The number of articles per year containing the term “genealogy” is steadily increasing, a flattening of the slope is not in sight.

*

The author is indebted to Professor T. Braun and Dr. A. Schubert for their encouragement and advice during the preparation of this work.

References

1. A. W. KOVÁCS, Studies of Hungarian Genealogical Research, Paper presented on the 9th Congress of Genealogy and Heraldry organized by the Commission for Heraldry, Genealogy and Sigilography of the Romanian Academy of Sciences on May 7–10, 1998, in Iași. *Colloquia*, 10–11 (2003–2004) : 240–249.
2. <http://encyclopedia.thefreedictionary.com/genealogy>
3. T. BRAUN, A. SCHUBERT, S. ZSINDELY, Keeping the gates of science journals: A quantitative view, In: *Gondolatok a Könyvtárban – Thought in the Library*, Magyar Tudományos Akadémia Könyvtára, Budapest 1992, pp. 117–123.