An assessment of systems and software engineering scholars and institutions (1994–1998)

Robert L. Glass *

1416 Sare Road, Bloomington, IN 47401, USA

Received 1 January 1999; accepted 22 March 1999

Who are the most published scholars in the field of Systems and Software Engineering (SSE)? Which are the most published institutions?

This paper is the sixth in an annual series whose goal is to answer those questions. The first such paper was (Glass, 1994); subsequently such studies have been published each year, in the October issue of this journal (beginning this year, the study is being published in the twelfth issue of the year, since JSS is now published 15 times per year rather than monthly). This is the second year in which the study has included five years’ worth of data (in the previous years, 1, 2, 3 and 4 years were covered). In future years, the study will continue to cover the most recent five-year period. This paper reports on the top scholars and institutions for the five-year period 1994–1998.

The methodology of the study and its limitations will be discussed later in this article. It is important to note two things at the outset, however:

1. The study findings are based on frequency of publication in the leading journals in the SSE field.
2. The study focuses on the field of SSE, and not, for example, on computer science or information systems.

Here are the findings.

1. Leading scholars

The leading scholars in the field are shown in Table 1. To be included in the table, scholars must have achieved a score of 4.67 or more during the years covered by this study. (The scoring scheme is discussed under the topic “Study Methodology” below, but a 4.67 roughly represents participation in somewhat more than four published papers during the study period.) The table lists 15 such scholars, with scores ranging from 4.67 to 11.4.

With five years’ data included in the study, the scholar totals are stabilizing with respect to those of previous years. That is, many of this year’s top scholars were also present in last year’s list (often in the same position). In the earlier years of the study, which contained fewer years of data, there was a great deal of scholar turnover in the top ranks due to the smaller amount of data available.

The top scholar in this year’s study is Richard Lai of LaTrobe University, Australia. This represents a significant change from previous years, in which Johnny S.K. Wong of Iowa State University was the perennial leader. Lai scored 11.4; Wong is second in this year’s study, with 10.3. Lai’s score derives from six single-authored, and eight co-authored papers during the five-year period of measurement. (Wong has been a co-author, during that period, of 17 published papers). Lai has moved up quickly in recent years, from second last year and tenth the year before.

Robert L. Glass of Computing Trends is third in this year’s study, with a score of 9.2 (he was fifth last year). Vic Basili of the University of Maryland is fourth with 7.3 (he was fourth last year as well). Elaine J. Weyuker of AT&T Research Labs is tied for fifth in this year’s study with Michael Jackson, an independent consultant (she was third last year, and he was unranked), with a score of 5.8.

Jen-Yen Jason Chen of National Chiao Tung University, Taiwan, is seventh this year’s study, with a score of 9.2 (he was fifth last year). Vic Basili of the University of Maryland is fourth with 7.3 (he was fourth last year as well). Elaine J. Weyuker of AT&T Research Labs is tied for fifth in this year’s study with Michael Jackson, an independent consultant (she was third last year, and he was unranked), with a score of 5.8.

Jen-Yen Jason Chen of National Chiao Tung University, Taiwan, is seventh this year, with a score of 5.5 (he was tied for sixth last year). Kassem Saleh of Kuwait University is eighth, with 5.3 (he was eighth last year as well). Shari Lawrence Pfleeger of Systems/Software is ninth with 5.2 (she was tied for sixth last year). Jeff Tian of SMU is tenth with 5.1 (he was fourteenth last year).

In Table 1, note that last year’s ranking for each ranked scholar is also included. The table also lists the
journals in which each scholar was published. Note that some scholars actually lost points this year, since papers published in 1993 were no longer counted in this year’s study.

2. Leading institutions

The leading institutions in the field are shown in Table 2. To be included in this table, institutions must have achieved a score of 10.92 or more during the years covered by the study. The institutional scoring scheme is similar but not identical to that for authors; the scheme and its differences are discussed under the study methodology section below. There are 15 leading institutions, with scorings ranging from 10.92 to 26.34.

Carnegie Mellon University (CMU) once again tops the list this year, with a score of 26.34. CMU passed the perennial leader, Bell Labs (Lucent), last year, and in fact this year Bell Labs has fallen to fifth in the study (after leading most of the first several years). The University of Maryland, with 24.59 remains in second place this year; the National University of Singapore moves up to third (from fifth last year) with 23.58; and National Chiao Tung University of Taiwan remains at fourth with 21.4. Bell Labs is fifth with 18.5.

Iowa State University remains in sixth place with 17.28; Politecnico di Milano moves from unranked to seventh, with 15.22. The University of York (UK) is eighth, down from seventh, at 13.67. Ohio State is ninth at 13.6, as it was last year. City University of London is tenth with 13.32; it was fifteenth last year.

There are a couple of special circumstances associated with the top-ranked institutions. First, CMU’s score includes that for the Software Engineering Institute, which is located at CMU (that is not new in the study this year, but it does account for higher scores over the years than would have been achieved by CMU alone). Second, because of the spin-off of Lucent Technologies from AT&T, there is some controversy about how the scores for Bell Labs should be aggregated. The Bell Labs heritage now rests in both Lucent’s Bell Labs, and...
AT&T's own Research Labs (entering the list for the first time this year at fourteenth), and employees of the two have given us conflicting viewpoints on who should claim those historical scores. We have elected to go with the name, Bell Labs, and therefore have included what was formerly called AT&T Bell Labs with the Lucent, not AT&T, scores.

Because of the fact that 1993 scores were discarded this year in order to keep a running five-year total, the scores for many of the institutions are actually lower than they were last year. For example, CMU has fallen from a score of 27 to 26.34 (while remaining in first place), while IBM T.J. Watson Research has fallen from an eleventh place score of 11.06 to 7.62, no longer in the top 15 institutions (IBM researchers had been particularly active in 1993).

It should be noted that geographically separated branches of an institution are scored as unique institutions – for example, the campuses of the University of California are each scored separately, as are the various branches and laboratories of IBM (this rule is particularly harmful to the IBM score, since there are 18 separate IBM locations that have achieved scores). Note also that scores are not normalized for the number of researchers; for example, an institution with 100 researchers and one with two will be scored identically. Although this may be an advantage to institutions with a large number of researchers, it does not give special advantage to very small institutions, as normalization would do.

3. Other rankings

Recall that this study is specific to the field of Systems and Software Engineering (SSE). There are similar studies for the related fields of Computer Science (CS) and Information Systems (IS).

Regarding CS, the best and latest such study is Geist, 1996. That study examines CS research institutions only (not individual authors), and restricts itself to academic institutions (none from industry) in the US (none from outside the US). (The authors of this study tell us, however, that there is a Web version of the study that includes industrial as well as academic institutions, and that is international in scope. This raises one of the interesting dilemmas of the electronic age, specifically regarding what constitutes scholarly published work. We have chosen to utilize the printed, stable, referenceable version of the study, and we mention the Web version only in passing.) For its methodology (Geist, 1996) examined published papers in the journals of the ACM and IEEE Computer Society (excluding Communications of the ACM). As will be seen later in this paper, the CS study covers more journals than this one (and in that sense is broader), but it does not examine journals from outside the professional societies that we cover here.

Nevertheless, the findings of the CS study are interesting. It lists and ranks the top 100 US academic CS programs. The top 10 are the University of Maryland, MIT, the University of Illinois, the University of Michigan, the University of Texas, Carnegie Mellon, Stanford, the University of Wisconsin, the University of Southern California and Purdue. Some of those institutions are present in our SSE ranking, although the order is somewhat different. Given that SSE and CS are related but different fields, perhaps the biggest surprise here is the similarities among the lists. Three of the top ten in the CS list are in the SSE list. However, there are also significant differences. The top ranked institutions in our survey, Carnegie Mellon and the University of Maryland, are ranked sixth and first in the CS study. Bell Labs, fifth in our study, is not present in the CS list (because the published CS study omitted industry research organizations). The other top five institutions in our survey, Singapore and Chiao Tung, are also not included in the CS list (because it does not include international institutions). It is difficult to decide whether the studies differ because of the population of institutions they examine, or because of innate differences between the CS and SSE fields, but certainly those differences play some part.

There are four other rankings of CS programs of interest to our study. Three of them, those performed by US News and World Report (USNWR), the National Research Council (NRC) and Business Week (BW), are basically academic opinion polls. (USNWR and NRC evaluated academic computing programs, and BW ranked the nation's top computing research labs.) Those three subjective studies present similar findings – the top five CS schools are (USNWR/NRC/BW ranking) Stanford (1/1/1), Cal-Berkeley (2/3/6), MIT (3/2/3), Carnegie Mellon (4/4/2) and Cornell (5/5/20). It is interesting that these studies have findings quite unlike either ours or the (Geist, 1996) study; of the top five schools in these opinion polls, two are not present in either of our lists. This would suggest that even the informed opinions of fellow academics simply do not relate well to the publication frequency of the schools in question.

The fourth ranking of CS (and other) programs may be found in the Gourman studies (e.g., Gourman, 1993). This is a ranking performed in terms of quality (a much more comprehensive concept than the scholarship we and the other CS studies cover, since it includes things of interest to, for example, potential students). The top five programs in that report are MIT, Stanford, Carnegie Mellon, Cal-Berkeley and Cornell. Because of the similarity between these rankings and those of USNWR, NRC and BW, it is clear that they are (once again) ranking something different from what we are attempting to rank here, scholarship (as measured by publication frequency).
Regarding the field of IS, there have been several studies of scholarship. The leading institutions, for example, were studied in Shin Im (1998) and Segars (1992). The top five in the first study were Arizona, Minnesota, MIT, Carnegie Mellon and NYU (in the second they were very similar – Minnesota, Arizona, MIT, Texas and NYU). Only one of these – CMU – is in the top five (or even fifteen!) of the SSE list. Clearly, the field of IS is distinct from that of SSE.

The Gourman reports do not explicitly rank IS programs. They do rank a program called “Information Science,” but that is a Library Science related field, different from IS.

4. Study methodology

In this sixth year of our study findings, only papers published during the calendar years 1994–1998 were tallied. We have achieved the goal set at the beginning of this study, to aggregate five years’ worth of data in order to provide credible and stable ranking of the top scholars and institutions. From now on, only the most recent five years of publications will be included in the study. Next year, for example, we will add data from 1999, but drop that from 1994. There may be minor discontinuities each year from now on, as the early year findings are discarded, but we anticipate that in general the future findings of this study will exhibit a satisfying stability.

The study findings are heavily dependent, of course, on the journals we include in our survey. The following journals were selected for their relevance to the field of SSE:

- Information and Software Technology (IST), Butterworth-Heinemann, UK;
- Journal of Systems and Software (JSS), Elsevier Science;
- Software Practice and Experience (SPE), John Wiley and Sons, UK;
- Software (SW), IEEE;
- Transactions on Software Engineering and Methodologies (TOSEM), ACM;
- Transactions on Software Engineering (TSE), IEEE.

These journals were chosen on the basis of a survey of the editorial board of the Journal of Systems and Software conducted in 1991, and there have been no changes in the list of journals since that time. We believe that the board represents a knowledgeable, active and unbiased source of judgment about the SSE field. (See the inside front cover of JSS for the current members of this board; board turnover has been minimal and the current list is reasonably close to the list at the time the survey was performed.)

Other journals suggested by board members but not included in this study are the following:

- Communications of the ACM. This journal frequently publishes special issues on a single topic. These theme issues would tend to skew the findings of a study of this kind. A special issue that focused on the design of a chip of one particular vendor, for example, would have caused that vendor to be near the top of our rankings that year even though there were no other papers from that vendor in our data. It is interesting to note that CACM was excluded from the CS study discussed above, as well.
- MIS quarterly. This is a leading journal of the IS field, but it was not thought sufficiently relevant to SSE.
- IEEE computer. This is a leading journal of the computing field in general, but again, it was not thought specifically relevant to SSE.
- American programmer (the name was recently changed to “Cutter IT Journal”). This journal presents important findings from the field of SSE practice, but was not thought sufficiently related to SSE scholarship to be included.

Several other journals have been suggested from time to time by others interested in this study. To keep the findings relatively stable, however, we have resisted those suggestions. Now that the study has achieved its goal of presenting five years of data, we will reconsider, in the future, the list of journals examined. However, we will keep any such changes evolutionary in order to minimize turmoil in the ratings.

Regarding counting schemes, we used the following methods:

Scholars:
- Single authors receive a score of one for each paper published.
- Authors of multiple-authored papers initially receive a score equal to their fractional representation on the paper. For example, co-authors each receive 0.5; if there were five authors, each would receive 0.2.
- For author totals, the initial scores for multiple authors are then modified as follows:
  - 0.5 becomes 0.7,
  - 0.33 becomes 0.5,
  - ≤ 0.25 becomes 0.3.

The reason for this transformation is (a) the belief that the original fractional representation unfairly penalized authors of multiple-author papers, and (b) the transformation was used in a previous similar study of authors in the IS field (Shim, 1991). Obviously, this kind of transformation can be misused by authors conniving to increase their scores, as a reviewer of this paper has suggested, but we see no reason to believe that is happening.

Institutions:

Author raw scores (excluding the above transformation) are attributed to the institution named on the paper.
A previous study in the IS field (Segars, 1992) included length in published pages as part of the score for a paper. We choose not to consider paper length in our scoring scheme. Some similar studies have ranked the participating journals, resulting in papers published in some journals being worth more than others. In our study, all journals included are ranked equally. Many of the journals examined contain regular columns that do not represent research papers and are not refereed. For example, JSS contains an Editor’s Corner, and SW contains a number of regular columns. None of this material is included in the totals of our survey.

5. Study limitations

There are practical limitations to the number of journals that can be studied for such a survey. This study is limited to six journals (the CS study included 17 journals, but only examined institutions and not scholars).

The publishers of JSS, Elsevier, believed at the outset of this study that it was important to limit the number of journals included. That turns out to have been a wise consideration. The conduct of this survey is extremely time-consuming, and in fact it went through three false starts, each of which foundered on the amount of clerical work to be done. The problem we have encountered is in counting institutions; to date, we have found no easy way to count institutions associated with particular papers (tabulating scholars, by contrast, can be automated through standard library techniques). We rejected the notion of assigning authors to institutions based on a current directory, on the grounds that the mobility of scholars would make such findings inaccurate. (In this study, we count a paper for the institution(s) for which the scholar worked at the time of publication.)

Another limitation of our study lies in the difficulty of measuring quality. What we are trying to measure here, of course, is the top quality scholars and institutions in the SSE field; that is, those who (which) have made the greatest contribution to the field. But what we are really measuring is the top quantity scholars and institutions, those who or which have published the most papers in the field. In defence of what we have done, we offer these thoughts:

1. Other similar studies have used the same approach. Both the (objective) CS and IS studies named above and referenced below have counted the number of papers published, and have not been able to measure the quality of that work.

2. Quality does enter into our measurement, in the following ways:
   (a) The papers counted have all been through a peer referee process. Our expectation is that this review process guarantees that the papers are of sufficient quality to have met acceptance criteria.
   (b) The choice of journals was made by the JSS editorial board, whom we consider to be a panel of experts. We believe that we have included only high-quality journals in our study. (One reviewer of this paper, however, noted that there are “huge differences” between the journals chosen; certainly there is a wide spectrum, ranging from deep theory to deep practice). We also believe that quality journals tend to choose quality referees to make the selection of high-quality papers. (Note that, in subject areas where quality measurement is difficult, the choice of quality is by a panel of expert judges (consider the fields of ice skating and diving, for example). We believe the same approach is valid here.)

3. Perhaps the best measure of quality would be the number of times papers under consideration were actually referenced in the literature of the field (that is, by citation counts). However, there are problems with that approach:
   (a) Such data is extremely laborious to obtain. We know of no standard library measurement of citations at the level of individual papers (there are, however, such studies for journals).
   (b) Even citation analysis has its opponents. For example, in Barrett (1997), some viewed academic citation analysis as “an uncomfortably silly fascination”, “a little bit of fun to see, but... not a really good thing to do with your time”, or (just because something has been cited presents) “no evidence... that anyone has read it”.

It is our belief that citation analysis would be a better scheme for measuring what we are trying to measure, in spite of the opposition noted in b, above; but citation analysis is simply too hard to do.

One final limitation of this study has been noted by some who find its results surprising, particularly in the list of top scholars. It is important to say here that this study does not attempt to identify the most influential or visible people in the SSE field. Such a list would no doubt contain easily recognizable names, which this list sometimes does not. However, any such list would need to be based on objective measures. There is a great danger, which we believe we have avoided here, of publishing a list resulting solely from a popularity contest. And it is often true – in fact, it is almost a truism of academe – that the most scholarly people are not necessarily the best known.

6. Conclusions

This study is one in an ongoing series whose goal is to identify the top scholars and institutions in the field of
SSE. Similar studies in related fields (CS and IS) convince us that such a study is meaningful and worthwhile.

By now, at the end of five years of conducting the study, we believe we can identify with some confidence those top scholars and institutions:

**Top Scholars:**
1. Richard Lai of La Trobe University,
2. Johnny S.K. Wong of Iowa State University,
3. Robert L. Glass of Computing Trends,
4. Victor R. Basili of the University of Maryland,
5. Elaine J. Weyuker of AT&T Research Labs.

**Top Institutions:**
1. Carnegie Mellon University and the Software Engineering Institute,
2. University of Maryland,
3. National University of Singapore,
4. National Chiao-Tung University of Taiwan,
5. Bell Labs, Lucent.

These ranking contain both predictable and surprising findings.

Regarding the relationship of the field with its collegial fields of CS and IS, we find:
- Similarities with CS in the list of top institutions, but still enough differences to be able to say that SSE is a different field from CS.
- Enough differences with the field of IS to say that they are clearly quite different fields. (A study of the curriculum differences between the fields may be found in Glass, 1992).

**References**


