Astronomical Data Analysis Software and Systems VI
ASP Conference Series, Vol. 125, 1997
Gareth Hunt and H. E. Payne, eds.

The Sociology of Astronomical Publication Using ADS and ADAMS

Eric Schulman[1]

National Radio Astronomy Observatory,[2] 520 Edgemont Road, Charlottesville, VA 22903-2475, E-mail: eschulma@nrao.edu

James C. French, Allison L. Powell

Department of Computer Science, School of Engineering and Applied Science, University of Virginia, Charlottesville, VA 22903-2442, E-mail: french@cs.virginia.edu, alp4g@cs.virginia.edu

Stephen S. Murray, Guenther Eichhorn, Michael J. Kurtz

Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138, E-mail: ssm@cfa.harvard.edu, gei@cfa.harvard.edu, kurtz@cfa.harvard.edu

Abstract. We use the NASA Astrophysics Data System database of astronomical abstracts in seven major astronomy journals to study trends in astronomical publication over the last twenty years. Two of the most interesting trends are the decreasing fractions of papers with one author and the increasing number of authors per paper.

1. Introduction

The sociology of astronomical publication has traditionally been performed by looking for publication trends using every paper published in a few selected journals within a few selected years. For example, Abt (1981) examined the papers published in ApJ, ApJS, AJ, and PASP during the first year of each decade from 1910 to 1980.

By analyzing the NASA Astrophysics Data System[3] (ADS) database of astronomical abstracts we can study a large number of issues in the sociology of astronomical publication while including every paper published in a number of refereed journals during the past twenty years. Here we present preliminary results of a study of astronomical publication trends using papers published in A&A, A&AS, AJ, ApJ, ApJS, MNRAS, and PASP between 1975 and 1995.

---

[1] Jansky Fellow.

[2] The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

[3] http://adswww.harvard.edu/

© Copyright 1997 Astronomical Society of the Pacific. All rights reserved.

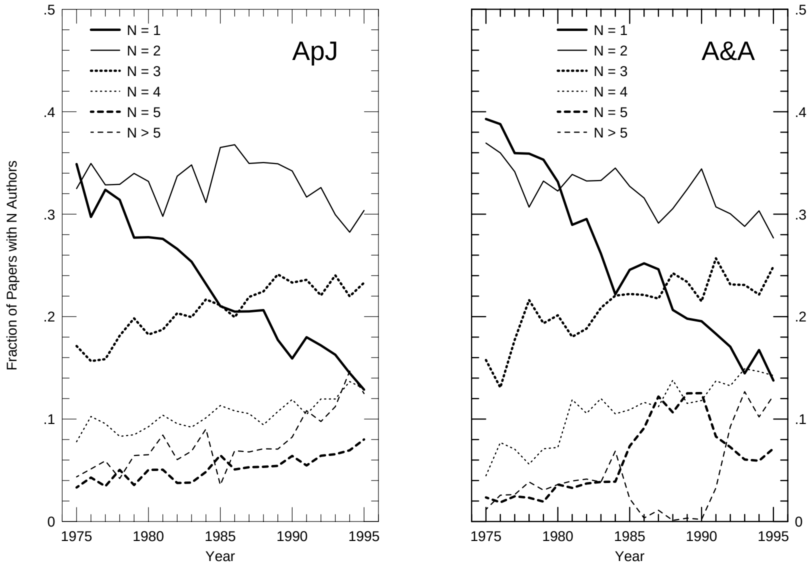

Figure 1. Fraction of ApJ and A&A Papers with N Authors.

2. The Astrophysics Data System Database

The ADS abstract service contains approximately 240,000 abstracts of astronomy and astrophysics papers from more than 1000 journals (Accomazzi et al. 1997). Most ADS abstracts of papers published between 1975 and 1995 were obtained through the NASA Scientific and Technical Information[4] (STI) Program, which compiled papers from the majority of astronomical journals. Although the database is at least 95% complete, there are some systematic errors in the data. For example, the STI author lists were truncated at the tenth author until about 1986 (e.g., Cohen et al. 1975 has fourteen authors, but ADS only lists the first ten). Between 1986 and 1990, ApJ and ApJS author lists were not truncated, but author lists in the other five journals were truncated at five authors, and from 1991 to 1994 the author lists in these five journals were truncated at ten authors. The number of author list truncations can be substantially reduced by comparing the ADS database with the Strasbourg Astronomical Data Center's SIMBAD[5] (Set of Identifications, Measurements, and Bibliography for Astronomical Data) database, a process that is currently underway. The SIMBAD database includes all papers since 1983 that mention at least one astronomical object (excluding Solar System bodies; papers published since 1950 that mention individual stars are also included). SIMBAD currently has information on 85,000 papers from 90 journals and conference proceedings.

[4] http://www.sti.nasa.gov/

[5] http://cdsweb.u-strasbg.fr/Simbad.html

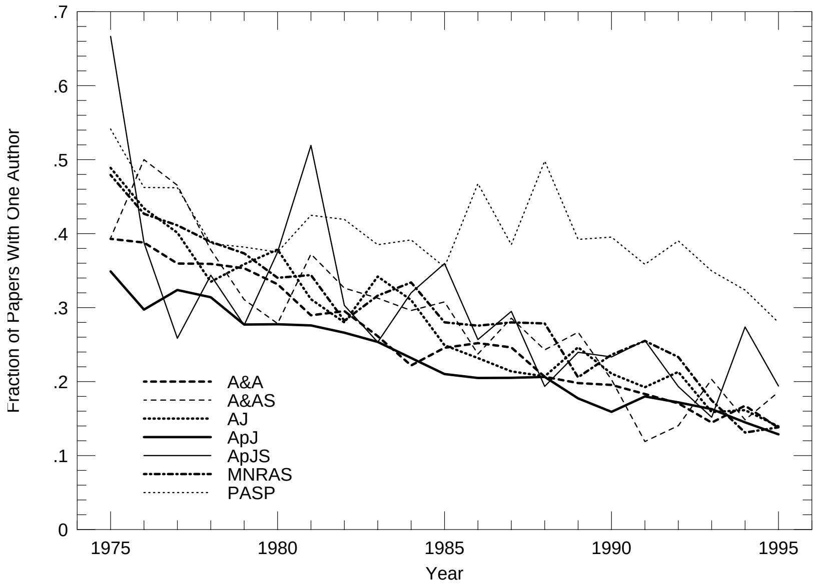

Figure 2. Fraction of Papers with One Author.

3. Number of Authors Per Paper

The fraction of ApJ and A&A papers with one to five authors is shown in Figure 1. The most striking change is the decrease in the fraction of single-author papers from more than $^1/_3$ to about $^1/_8$. The fraction of two-author papers remained fairly constant, while papers with larger numbers of authors became more frequent. The NASA STI truncation of A&A author lists is obvious between 1985 and 1990. The decrease in the fraction of single-author papers has occurred in all seven journals (Figure 2), although the fraction of single-author papers in PASP remains fairly large. Even though the NASA STI truncations make it difficult to analyze the mean number of authors per paper, it is still obvious that there are on average more authors per paper now than there were twenty years ago.

There are a number of possible reasons for decreasing fractions of single-author papers and increasing numbers of authors per paper. One is the growth of multiwavelength astrophysics (Abt 1993), which requires astronomers to be proficient in multiple wavebands or to collaborate with experts in other wavelengths. Another is an increase in the number of papers that present both observations and theoretical interpretations. A third is increasing competition for jobs and grants, which encourages astronomers to write as many papers as possible. Also, in the last five years there has been more research requiring large collaborations, such as using *HST* to determine Cepheid distances to nearby galaxies (e.g., Kelson et al. 1996, with 18 authors). There were nine papers with 50 or more authors in our sample, all of which appeared in the ApJ or ApJS after 1990. Five were a series of papers reporting on intensive *HST*, *IUE*, and ground-based optical and near-IR spectroscopic monitoring studies of Seyfert galaxies (e.g., Clavel et al. 1991, with 57 authors). The other four papers, including the 124-author Ahlen et al. (1993), were reports of high-energy cosmic ray experiments.

4. Future Work With ADAMS

In future work we will use the Advanced Data Management System[6] (ADAMS; Pfaltz & French 1993; Pfaltz 1993), an object-oriented database language that supports a single shared, distributed data space that can be accessed by applications programs coded in C, C++, FORTRAN, or Pascal. The class hierarchy supports multiple inheritance and user-defined data types. Unlike many object-oriented database languages, attributes in ADAMS are first class objects so schema evolution is particularly easy.

We will be analyzing the ADS database within ADAMS to develop improved methods of searching document collections. One of the goals of this work is to provide a sophisticated browse facility as an adjunct service to ordinary keyword searches in document retrieval systems. The idea is to use a keyword search to establish an initial focus and then let the searcher access additional documents by identifying a document of interest and asking for more documents "like this one." An underlying topical map will be maintained to support this kind of browse mechanism.

Acknowledgments. We thank Ellen Bouton for her invaluable assistance in finding and describing many different journals, and for very useful conversations about astronomical publication. This research is based on data obtained through NASA's Astrophysics Data System Abstract Service, and is supported by grants from NASA (NCCW-0024), the DOE (DE-FG05-95ER25254), and the NSF (CDA-9529253).

References

Abt, H. A. 1981, PASP, 93, 269

Abt, H. A. 1993, PASP, 105, 437

Accomazzi, A., Eichhorn, G., Kurtz, M. J., Grant, C. S., & Murray, S. S. 1997, this volume, 357

Ahlen, S., et al. 1993, ApJ, 412, 301

Clavel, J., et al. 1991, ApJ, 366, 64

Cohen, M. H., et al. 1975, ApJ, 201, 249

Kelson, D. D., et al. 1996, ApJ, 463, 26

Pfaltz, J. L. 1993, The ADAMS Language: A Tutorial and Reference Manual, Technical Report IPC-93-03, University of Virginia Institute for Parallel Computation

Pfaltz, J. L., & French, J. C. 1993, Data Engineering, 16, 14

[6]http://www.cs.virginia.edu/~adams/