Introduction

The preliminary proposal by Harold Varmus, the recent head of the National Institutes of Health (NIH) in the USA, to develop a centralized electronic preprint server for the biomedical sciences, generated much controversy.\(^1\) The biomedical sciences have not been early adopters of preprint servers, even though, from 1961 to 1967, an early experiment involving the dissemination of paper-based preprints, via the Information Exchange Groups (IEGs), was supported by the NIH.\(^2,3\)

In 1965, before the experiment with the biomedical IEGs was terminated,\(^4\) Michael J. Moravcsik proposed the establishment of a central preprint registry in particle physics.\(^5\) The US Atomic Energy Commission became interested in the proposal, and also in a more far-reaching scheme of centralized preprint distribution.\(^6\) This experiment, the Physics Information Exchange (PIE),\(^7\) was to be limited, for a trial period, to papers in the field of high energy physics (HEP).\(^6,7\) It was (like the IEGs) a controversial proposal,\(^8,9\) and the far-reaching scheme was not undertaken in the 1960s.\(^6\) It was eventually started in 1991, when, through the efforts of Paul Ginsparg, a server for the centralized distribution of electronic preprints was successfully established.\(^10\) This server was based at Los Alamos National Laboratory (LANL), and had an initial focus on theoretical HEP. It has evolved into the arXiv server,\(^11\) which currently archives preprints in several areas of physics, mathematics, nonlinear sciences, and computer science.

In an article about Ginsparg’s initiation of an electronic ‘bulletin board’ (e-print archive) for HEP, two major enabling factors were noted, the ‘preprint culture of physics’ and the ‘technology that made the bulletin boards possible’\(^12\). Odlyzko has suggested that the rapid acceptance of Ginsparg’s preprint server was a case of simple sub-
author and reader communities essentially coincide

institution: ‘His research community in high energy theoretical physics had, during the 1980s, developed a culture of massive preprint distribution’.13 Ginsparg himself has suggested that a feature of his own research discipline that may account, at least in part, for its early adoption of such a culture was because ‘the author and reader communities (and consequently as well the referee community) essentially coincide’.14 An obvious question is: why didn’t an analogous ‘culture of massive preprint distribution’ also flourish in those areas of the biomedical sciences where overlapping ‘invisible colleges’ of authors, readers, and referees also may exist?

The present article has two purposes. The first is to identify some of the major factors that led to the termination of the IEG experiment on the centralized dissemination of biomedical preprints. The second is to outline a number of factors that appear to have played a significant role in the early adoption of a preprint culture by researchers in HEP. In considering these various factors, it should be borne in mind that those interested in sociological perspectives on the diffusion of technological innovations may ‘challenge the demand for causal sociological explanations of multifaceted historical developments’.15 The article will conclude with a historical vignette selected to illustrate a premise that the shift from ‘atom-based’ to ‘bit-based’ distribution of information16 will have as profound an effect on the dissemination of research results as did the shift from handwritten manuscripts to printed ones.

Some terminology

The term ‘server’ is applied in computer science to a computer or program that controls a central repository of data that can be downloaded and manipulated in some manner by a client.17

Traditional (paper) ‘preprints’ can be defined in several ways, depending on their stage in the conventional publication process.18 They can be manuscripts that are being circulated among peers for comments prior to submission for publication. They can have been submitted for publication, but no decision to publish has yet been made (usually referred to as ‘submitted to . . .’ ). Or they can be manuscripts that have been reviewed and accepted for publication (usually referred to as ‘in press . . .’ ).18

Of particular interest from the perspective of electronic preprints are those preprints that are intended for publication, and are self-archived by their authors either prior to, or after, acceptance for publication. This is one definition of the term ‘e-print’ (see, for example, Ginsparg19). However, ‘e-print’ (or ‘e-print’) can be defined in other ways. Here, the focus will be on ‘e-preprints’, with a particular emphasis on those that have been self-archived by their authors.

The Information Exchange Groups

The six-year experiment of the NIH with IEGs began in 1961. The IEGs were conceived and initiated by Errett C. Albritton of the Division of Research Grants of NIH,20 and were terminated on 1 March 1967.4

By 1966, a total of seven IEGs had been established. In a ‘Report to Members’ by Albritton (reviewed in Nature in 1966), it was suggested that these seven groups could be regarded as a sufficient pilot run out of the 250–300 possible scientific groups that the NIH might come to support. In the same report, Albritton estimated that 80% of the ‘memoranda’ circulated at the beginning of the IEG experiment eventually appeared in print in ordinary journals, and suggested that this percentage had risen to nearer 90% by 1966. At the time that the editorial in Nature was written, there were 2,500 members in the seven IEGs, and 1,200 ‘memos’ had been transmitted.3

In the same editorial in Nature in 1966,4 several concerns about the circulation of preprints were noted. These included issues about their differentiation from media of communication already serving science, cost-effectiveness, speed of communication versus the excellence of what is communicated, priority and quoting material in such ‘memos’, access to the preprints and to membership in the IEGs, and others. Examples of such criticisms are provided in a 1966 editorial by Philip H. Abelson in Science,21 and in letters from some reputable scientists published in 1966 in Science.22
Abelson concluded his editorial by attributing the growth of the IEGs in part to a mass protest against the inefficiency of the conventional publication process and in part to a desire by some researchers to avoid conventional quality review, ‘a discipline essential to the integrity of science’. Among the letters about the IEGs published in the 21 October 1966 issue of Science was one from Thomas H. Jukes of the Space Sciences Laboratory, University of California, Berkeley. He offered a point-by-point refutation of the disadvantages summarized by Dray. Of Issue No. 4 (the preprints may ultimately supersede existing journals), Jukes commented that ‘This process is commonly termed “evolution”’, and that ‘The current tendency is for journals to multiply rather than to disappear’. Of Issue No. 7 (preprints infringe on copyright), Jukes noted that: ‘IEG communications are not sold’.

The two primary reasons given by Confrey of NIH for the termination of the IEG experiment were: first, that the original purpose of the experiment had been achieved, and second, that the rapid growth of the IEGs had reached the threshold limit for the NIH facilities to accommodate. Confrey concluded that the IEG concept is workable, provided that the chosen research area is focused on an easily described and identifiable research phenomenon or problem around which the group can be built. He also concluded that the IEG experiment must either expand to a larger number of areas or be suspended.

In November 1966, Nature published an editorial about the decision (already taken) to terminate the IEGs. The editorial included a comment that ‘even the generosity of the National Institutes of Health had begun to seem inadequate’. In January 1967, another editorial about the termination of the IEGs was published in the New England Journal of Medicine. It was suggested that two major issues were delays in distribution because of numerous contributions (e.g. in the areas of immunology and nucleic acids) and attempts by authors to establish priority by submitting papers to an IEG immediately upon their completion. It was suggested in this editorial that scientific groups or societies should learn
from the IEG experiment, and then take up the idea again.\textsuperscript{26}

One of the conclusions reached by Green was that it was not the failure of the IEG experiment, but its success, that finally spelled its doom. He commented on the costs of the experiment, and argued that the cost of his IEG (No. 1) was miniscule, relative to the dividends received by the members and to the total costs of their research.\textsuperscript{20} He suggested that opposition from scientific journals had a crucial influence on the decision to terminate the IEGs. He noted that the editors of Nature\textsuperscript{27} ‘spilled the beans prematurely’ about a meeting in Vienna, on 10 and 11 September 1966, of the editors of several major biochemically oriented journals. Five of the editors voted to propose to their editorial boards not to accept articles or other communications previously circulated through IEGs.\textsuperscript{28,29} Also, papers could not be submitted simultaneously to a journal and an IEG, nor could papers already accepted for publication in a journal be released through an IEG.\textsuperscript{28} This policy, which, in effect, banned the inclusion of preprints into the scholarly literature, was soon adopted by several major biomedical journals.\textsuperscript{28,29} It was probably one of three major barriers to the further development of a ‘preprint culture’ in these sciences. The second was the termination, by NIH, of support for the IEGs, in part because of the costs that would be involved in any continuation or expansion of the IEGs. The third was the continuing opposition, by many respected and senior biomedical scientists, to the distribution of unrefereed papers.

The Physics Information Exchange

As early as the 1960s, the number of preprints in HEP had become large, and ways were being sought to deal with the resulting problems and inequities.\textsuperscript{6} As noted above, in March 1965, Moravcsik\textsuperscript{5} proposed the establishment of a central registry of preprints in particle physics, soon followed by the more ambitious plan for PIE.\textsuperscript{7} PIE would be analogous to an IEG, initially only in theoretical HEP.\textsuperscript{8} Moravcsik’s interest in preprints as an important form of communication was a continuing one, in part because of his desire to foster communication with scientists in resource-poor countries.\textsuperscript{10}

Like the IEG experiment, the PIE proposal generated controversy.\textsuperscript{8,9} For example, Simon Pasternack, editor of the major journal Physical Review, was concerned that journals would have only a secondary role in the scheme of physics communication. An editorial critical of PIE, published in Nature\textsuperscript{9} in 1966, expressed concerns about the ‘clubbiness’ of existing preprint networks, about the difficulties involved in recognizing and controlling plagiarism, and about an apparent insensitivity to ‘those qualities which distinguish good literature from bad’.\textsuperscript{9} The editorial concluded with the hope that ‘PIE will be stillborn’.\textsuperscript{9}

The US Atomic Energy Commission (AEC) subsequently financed a study by the American Institute of Physics to examine the feasibility and desirability of PIE. One of the conclusions of the report of this study, issued in August 1967, was that the first phase of such an experiment should involve a weekly preprint announcement service for high energy physicists, to be operated for 6 months.\textsuperscript{6} In April 1968, the newly formed Division of Particles and Fields of the American Physical Society (APS) agreed to sponsor a preprint accession list. This list would permit currently available and relevant theoretical and experimental preprints to be announced, but would avoid the expensive central dissemination of all preprints.\textsuperscript{6} The preprint list would be a more cost-effective approach to preprint distribution because, in the early 1970s, only 20–30\% of preprints were ever requested by anyone,\textsuperscript{6} and only about 3–5\% were regarded as ‘winners’.\textsuperscript{11} The preprint list would also allow control of the actual distribution of preprints to remain where it had traditionally been, with the authors who received requests for preprints and their host institutions.

Preprints in Particles and Fields

The result of these plans to develop a preprint announcement service was ‘Preprints in Particles and Fields’ (PPF), which began weekly publication at the Stanford Linear
Accelerator Center (SLAC) in January 1969. It received financial support from the AEC Division of Technical Information for an experimental 18-month period. The further contributions of the PPF to the establishment of a preprint culture in HEP and its subsequent transformation into an e-preprint culture, have been described in some detail by authors familiar with the history of the SLAC Library.

For example, in a brief history of the SLAC Library, Addis identified a number of important events. An initial one was the decision in 1962 by SLAC’s first director, W. H. K. Panofsky, to acquire and catalogue preprints in HEP. A second, after 1969, was the willingness of hundreds of physicists to pay an annual subscription fee to get PPF weekly by airmail. A third was the establishment of a PPF section called ‘Anti-preprints’, which recorded the preprints that subsequently were published in journals. This list is especially noteworthy because it was popular with journal editors, who could match references to preprints with the published article.

An example of another noteworthy event identified by Addis was the development, in 1979, of a new text formatting system called TeX. This system provided a way for authors to produce high quality mathematical text (including physics notation) that could easily be widely distributed via email. Then, in August 1991, Paul Ginsparg (a theoretical physicist) started the first e-preprint archive, at hep-th@xxx.lanl.gov, and invited fellow string theorists to deposit the TeX source for their new preprints by email. New preprints were announced and distributed via email. By 1994, features of the World Wide Web (WWW) had been exploited to provide more convenient access to e-preprints, and use of the WWW had expanded to the world beyond physics.

Why theoretical high energy physics?

Why were the earliest adopters of an e-preprint culture those in theoretical HEP? It has been noted that theoreticians in HEP publish relatively frequently, and that their papers sometimes function as iterative discussions, with each other and with experimentalists. They want to stay up to date on one another’s research. However, perhaps the most concise summary of the multiple factors involved has been Ginsparg’s. The HEP physicists have been members of a well-defined and highly interactive community of voracious readers, with a pre-existing hard-copy preprint habit, a standardized text formatting system (TeX), and a generally high degree of computer literacy. They also have had little concern about patentable content, and have preferred to assign intellectual priority at the point of dissemination, rather than after peer review. The most crucial factor may have been the extent to which the author, reader, and reviewer communities have coincided, and thus have had an implicit agreement (or ‘scholarly consensus’) about standards of quality for research considered to be acceptable for publication.

What about opposition from journal editors?

It was claimed above that the decision of some very influential journal editors to ban the inclusion of preprints into the scholarly literature was a major barrier to the further development of a preprint culture in the biomedical sciences. Why were journal editors not an equally formidable barrier to the development of a preprint culture in HEP? A detailed discussion of this issue is beyond the scope of this article, but six points will be noted.

First, there was indeed opposition to the PIE proposal from the editor of a very influential journal in physics, the Physical Review. Secondly, the decision not to initiate PIE, but instead to develop a preprint accession list, was probably a crucial one, and the concomitant development of an ‘Anti-preprints’ list (of preprints that subsequently were published) helped to mollify dubious or hostile journal editors. Thirdly, a major publisher in HEP has been the APS, a society that must take into account its revenue stream, but also must evolve to meet the needs of its research community. Fourthly, editorial policies regarding prior electronic publication are still at an early stage of development.

editorial policies regarding prior electronic publication are still at an early stage of development
still at an early stage of development.\textsuperscript{38} Fifthly, publication and research habits probably will continue to vary, even across subfields of physics, and the current framework of publication in physics may already be flexible enough to accommodate a variety of behaviours on the part of both authors and evaluators.\textsuperscript{39} The sixth, and probably most important, point is that the standard of work archived at the LANL server has been very high, in that 70\% of papers are eventually published in journals and another 20\% in conference proceedings.\textsuperscript{39}

Some ancient history

Concerns about control of the quality of information disseminated by various means have a long and fascinating history. For example, the requirement that books be licensed for printing (by the privy council or other royal nominees) was introduced in England in 1538. There was concern that the introduction of printing ‘immensely increased facilities for the spread of sedition and heresy, grave dangers always threatening the life of governments and the church’.\textsuperscript{40} Although the infamous Star Chamber was abolished in 1641, this did not lead immediately to any liberation of the press. By this time, journalism, which had its origin in the manuscript newsletters, had come into existence. After the introduction of regular postal services, such newsletters could be dispatched to subscribers from a central source. However, for a long time, the licensing system and the Star Chamber decrees were an impediment to the publication of any printed journal.\textsuperscript{40} In these early attempts at quality control of information, the emphasis was sometimes on ‘quality’, but perhaps more often on ‘control’.

Concluding comment

Will a new revolution in scientific publishing, in which journals come to be regarded as an overlay on preprint databases, now overtake the biomedical sciences, following the lead of HEP?\textsuperscript{14,37} The most prudent prediction probably is: much more quickly in some areas of research than in others. The issues involved continue to be actively debated, on (for example) an online forum sponsored by Sigma Xi.\textsuperscript{41}

Acknowledgements

The author is very grateful to those who responded promptly to his emailed requests for pertinent information, and especially to those associated with the SLAC Library. Partial support for this work was provided by an NCIC/El Lilly Award from the National Cancer Institute of Canada (NCIC).

References

11. The arXiv archives. \url{http://arxiv.org/} (see also \url{http://xxx.lanl.gov/}).


31. Kirk, B. Preprints, publications and all that. SLAC Beam Line, July 1977, Stanford Linear Accelerator Center (Jean M. Deken, Archivist).


James E. Till
Joint Centre for Bioethics, and Department of Medical Biophysics, University of Toronto, c/o Ontario Cancer Institute, University Health Network, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada
Email: till@oci.uhnres.utoronto.ca.