

FOREWORD

The readers of these essays, reprinted from *Current Contents*, are in for a treat, if they like knowledge. Gene Garfield is a polymath. Besides culling the *Citation Classics* from various journals, Gene reports on Japanese science, the anomie-deviant behavior connection, science careers at undergraduate institutions, hypnosis and pain control, fraud and dishonesty in science, the secrets of wine, and Derek de Solla Price and Vartan Gregorian, among other topics.

I first met Gene Garfield in the Cohen Library at The Rockefeller University when Joshua Lederberg introduced us. Gene was a “general” in the war against junk information, the first information scientist of substance I had met. It was not immediately clear to me how he had attained such a high rank with his unconventionally long hair. But he proved to be an interesting character, a chemist turned information entrepreneur with a wide ranging, restless mind. And did he know about science! His skills in managing the ever growing flood of scientific data are a good part of the answer to both understanding and managing the information pollution problem in science. If one pauses to reflect that *Automatic Subject Citation Alert* identifies articles from about 5,000 journals, the *Science Citation Index* monitors 3,500 journals and indexes (600,000 items yearly with 9 million citations), the *Index to Scientific and Technical Proceedings* identifies 100,000 conference papers each year and *Automatic New Structure Alert* keeps chemists informed of 16,000 new compounds each month, one can see, even from this partial inventory, the magnitude of ISI’s achievement. Others, in different dominions of the empire of knowledge, ought to take note.

Citation analysis, of course, has its limitations—well known to information scientists like Gene. Such analysis is based on rules that are indifferent to the actual meaning of scientific articles. The presumption is that the authors of papers are sensitive to the meaning and importance of the papers that they cite. Citation analysis identifies the important papers better than one might have reason to expect—it checks out, if compared to the expert’s judgment. Like democracy, citation analysis is imperfect, but better than other systems. The various indices used by ISI also bear useful information about the relative importance of journals, reviews, and articles.

I’d like to add one more index to Gene’s list, what I will call the “Glashow Index,” because it was created by Sheldon Glashow, a theoretical physicist at Harvard. He provides a brutal criterion for evaluating one’s articles. Here is how it works: Upon writing a scientific article, an author (let’s say you) gets ten points. Every time the article is cited by someone else, add one point to your index; every time you cite your own previous article, you get minus five points. The goal of the game is to come out with a positive index. It’s not so easy. Citation Classics pass with flying colors.

As a high school student, I remember history texts that sometimes characterized historical periods in terms of their technical progress—the march of civilization in

terms of new goods, services, and a more equitable standard of justice. They assured the reader that he or she was living at the edge of modern technology, presumably in the best of times, in spite of what is written in the morning newspaper.

As an antidote to the religion of progress, I would suggest, as a kind of a thought experiment, examining historical periods not in terms of their "progress," but in terms of the trash they produce, whether it be organic disease, industrial pollution, or junk information. A history of garbage, if you please.

During the long period in which agriculture dominated all other forms of production, the primary forms of pollution were organic wastes. Medieval towns and villages reeked. Wastes produce bacteria and these, passed via food and vermin, caused devastating plagues. Demographic scholars claim, with justification, that such outbreaks limited the size of ancient cities to, at most, a few hundred thousand people.

In the nineteenth century, with the emergence of industrial civilization, sanitary engineering, sewers, and the germ theory of disease, cities grew and some achieved populations in the millions. The organic pollution of a previous age became understood and brought under some control. Humanity, it seemed, won a small battle in its eternal evolutionary war against the microbes. But industrial civilization, in turn, brought its own forms of contamination. A decade ago, a river near Cleveland caught fire, and today we have acid rain and ozone holes.

Yet, even as we move out of the "industrial era" into "the age of information" a new form of pollution arises—junk information. With dialectical certainty, it seems, nothing gets done without creating trash. The information age is no exception.

In dealing with any kind of trash, the first step is to understand it, the second is to manage it and, if possible to eliminate it. As a graduate student in theoretical physics, I saw an overwhelming number of publications in my field. I made efforts to read the abstracts of all interesting papers while studying a new subject. But in the research phase of work, particularly on a hot topic, the preprints arriving in the mail were the only items of interest, and the existence of these I learned about from phone conversations with colleagues.

Mathematical physicists pride themselves on being able to work starting from simple first principles. Biologists, chemists, and medical researchers aren't so lucky. Their domain of knowledge is far more vast, intrinsically complex, and is likely to remain so for the foreseeable future. The problem of junk information is, accordingly, aggravated. I have to admit I did not appreciate the problem my colleagues in the bio-medical disciplines had, until I took an interest in identifying topics for new conferences, mostly in the bio-med area, for the New York Academy of Sciences. The scientific fields were fast and rapidly changing, and trying to spot a prospective nascent conference topic was like trying to hit a moving target. What to do? I decided to consult my local guru.

When I told Joshua Lederberg about my problem of trying to identify rapidly growing fields in the bio-med sciences, he said, "Have a look at this" and placed in my hands an ISI cluster analysis of a new topic, based on citation frequencies. This indeed, was what I was looking for. A few years later, down at the ISI headquarters, Henry Small showed me the clusters on a computer, color coded for display clarity. One could watch scientific subjects grow like cell colonies. Furthermore, the cluster citation analysis, when applied to fields in physics that I knew first hand, accorded with my expectations, and thus I had confidence in them. They were effective.

Gene's great and unique talent is taking information and converting it into knowledge. In reading this collection of essays for both edification and pleasure, we all become the beneficiaries of that talent. Enjoy yourself; they are fun to read.

Heinz R. Pagels
Executive Director
The New York Academy of Sciences