CURRENT COMMENT

Also see Reprint on p.288, Essays, Vol:3

Everything You Always Wanted To Know About ISI Data Bases But Were Afraid To Ask

Number 45, November 7, 19

The variety of information services provided by the Institute for Scientific Information® sometimes leads to confusion. Because we offer both broad and specific information services for the sciences and social sciences, some people assume that we have several separate data bases. Others believe that ISI® has a single, monolithic data base and that we merely flip a switch to generate services for any disciplinary field.

Neither assumption is entirely correct. Whether we should be considered to have one data base or several depends on which service you are talking about and at what point in its production cycle it happens to be.

For example, the data bases for the *Science Citation Index*® and the *Social Sciences Citation Index* ™ start out in our production cycle together, but eventually they split apart. *ASCA*® and *ASCATOPICS*® are always derived from the combined single data base. The data for the *Author Address Directory* and *Weekly Subject Index* sections of the six editions of *Current Contents*® are kept together through editing and keying procedures. An extraction program later replicates and separates the addresses and index terms for use in the different *CC*® editions. This process is complicated by the fact that many journals covered in *CC* are not yet covered by one of our three citation indexes. (We have gradually been working towards the elimination of this confusing distinction.)

The raw materials for the data bases—the journals themselves—are not physically segregated by subject area during their processing flow through ISI. Social science journals lie contentedly next to physics and biology journals while waiting to be processed. And the same is true in our library storage area. Our *Original Article Tear Sheet* (*OATS*®) service makes no disciplinary distinctions among journals.

What is this all leading to? Recently I was invited to speak at the Third Institute of Electrical and Electronics Engineers (IEEE) Conference on Scientific Journals. Although I was unable to deliver the prepared paper in person, it is reprinted on the following pages. In it, we try to answer many of the questions frequently asked about ISI's operations. While there have been other articles about ISI as an organization,[1,2] this is the first one to describe our procedures in detail. While no single article of reasonable length can describe every nu

sentation of what actually hap-

e article does not include a dis-
on of the data base from which
roduce *Current Abstracts of*
istry and Index Chemicus™,
natic New Structure Alert® ,
he *Index Chemicus Registry*
m® . All of these services
journal articles that announce
organic syntheses and com-
ds. For all intents and pur-
, all journals, editing, keying,
lata processing for this data
are handled separately from
ISI services. Descriptions
appeared in *CC* and
here.[3]

roughout the article the enor-
size of the ISI operation is
clear. Over 2,000 new source
are processed each day. This
res over two million key
es daily. I recently discussed in
detail how we use our Key-
system to cut costs and im-
quality for this massive file.[4]
her example of an innovative
f the computer, not discussed
e article, is our Project ZIP.
ted simply, Project ZIP is de-

consistent or incomplete author addresses entered into the data base. Every organizational address processed is verified manually against an authority list of "correct" addresses. If the incoming address is not on the authority list, it is edited, assigned a number, and added to the list for future use. If an incoming address is already on the authority list, the editor writes the appropriate code number beside it. When the article is processed, the data entry operator simply keys in the code number rather than the whole address. The correct and complete address is selected later from the computer's memory and automatically entered as part of the bibliographic record for that source item.

Like much published scientific research, the article reprinted here was not completely up to date even at the time it was presented. We are making improvements all the time. And next year, when we start our *Arts & Humanities Citation Index*[5] and our *Index to Scientific & Technical Proceedings*[6], further refinements will be made.

Garfield E. The who and why of ISI. *Current Contents* No. 1, 6 January 1975, p. 5-14*

Lazerow S. Institute for Scientific Information. *Encyclopedia of Library and Information Science.* Vol. 12. (A. Kent et al., eds.) New York: Marcel Dekker, 1974, p. 89-97.

Garfield E. We've added a *Weekly Subject Index* to *Current Abstracts of Chemistry and Index Chemicus. Current Contents* No. 5, 3 February 1975, p. 5-6*

———. Project Keysave—ISI's new on-line system for keying citations corrects errors. *Current Contents* No. 7, 14 February 1977, p. 5-7.

———. Will ISI's *Arts & Humanities Citation Index* revolutionize scholarship? *Current Contents* No. 32, 8 August 1977, p. 5-9.

———. ISI's new *Index to Scientific & Technical Proceedings* lets you know what went on at a conference even if you stayed at home. *Current Contents* No. 40, 3 October 1977, p. 5-10.

inted in: **Garfield E.** *Essays of an information scientist.* Philadelphia: ISI Press, 1977.