DATA/INFORMATION AVAILABILITY

Edited by
RALPH I. COLE

The American University
Technology of Management Series
PAUL W. HOWERTON, *General Editor*

THOMPSON BOOK COMPANY
WASHINGTON, D.C.

Library of Congress Catalog Card No. 66-25651

Printed in the United States of America

FOREWORD

In a recent Institute held under the auspices of the Center for Technology and Administration of the American University, the subject of the "Availability of Data/Information" was treated in appreciable depth. Those papers which tended to highlight the issues of information availability have been edited and related to one another in a forceful way by Ralph I. Cole, Director of Institutes and Special Programs in the Center.

The meeting has stimulated "thinking in depth" in the subject area as is demonstrated by the returns we had from the participants. By participants we mean both the distinguished group of speakers and the highly articulate attendees. The speakers brought a wide variety of well-informed backgrounds and interest to the consideration of the subject. We owe them a debt of gratitude for their meaningful contributions to the information sciences.

We do not represent this collection of opinions as an agreed position on how to manage the availability of Data/Information, because we feel that there is no one demonstrably superior method. We try to present the issues both existing and potential with some indication of tested solutions to identified problems.

Paul. W. Howerton
General Editor

PREFACE

The papers selected for presentation herein represent but a few of the more important aspects of data generation, format, storage retrieval and utilization which concern the management of technology.

The first two papers in Part I treat the history and growth of data handling, particularly those aspects which concern *structuring* explicit meanings to chemical and biological data Elements when used in conjunction with electronic methods for Filing and Retrieval. The third paper of Part I attempts to look into the future of "computerized" information handling and commands considerable attention.

Part II consists of two papers dealing with the science of management and the relationships that exist to both technological advances and decision-making, all referenced against the oncoming rush of knowledge.

Part III deals with Data/Information sources and their Utility and represents the experience of five distinguished speakers each selecting but one aspect to explore.

Part IV includes three papers each treating a different and difficult aspect of improving the structure, content, and usefulness of data.

Without question the field of Data/Information suffers from a lack of formal structure and guidance that will be benefited from a study of these discussions. Judging from the laments of managers, greatly increased effort in this particular area is sorely needed.

Ralph I. Cole
Editor

LIST OF CONTRIBUTORS

ALBERT N. ABAJIAN

Consultant specializing in diversified Management and Economic Research. Lecturer and Author. Among his credits are "A Proposed Coordinated System for International Commercial-Economic Communication in the Government," prepared for a subcommittee of the U.S. Senate, and "EPIC–Economic Progress Through International Communication." He is a fellowship graduate of the International Marketing Institute of the Harvard Graduate School of Business Administration, and Research Associate, the Center for Technology and Administration, The American University.

ROBERT L. BIRCH

Reference Librarian affiliated with the U. S. Patent Office. Specialist in data recovery and related techniques and fields. Frequent contributor to information science literature as a lecturer, teacher, and author.

RALPH I. COLE

Director, Institutes and Special Programs, The Center for Technology and Administration, and Professorial Lecturer, the American University, Washington, D.C. Well-known technology and management consultant in research and development administration in both industry and government. Frequent lecturer, author and contributor in the management-information effectiveness area. Awarded Legion of Merit for his contribution to Air Force and Signal Corps research administration. Engineering Manpower Commissioner. Fellow of the IEEE and a member of numerous professional and technical societies.

CHARLES DE VORE

Deputy Executive Assistant for Scientific Information, Office of Naval Research, Washington, D.C. Well-known author and executive in broad areas of scientific and technical information, both in government and in industry. Frequent lecturer, writer, and participant in national managerial-scientific meetings. Member, IEEE and AAAS.

C. STARK DRAPER

Director of Instrumentation Laboratory, and Professor, Massachusetts Institute of Technology. Distinguished inventor in fields of aeronautics, in control, and in space, in which he has made outstanding contributions.

He has often been honored for these achievements both here and abroad. Well-known lecturer and author in broad areas of scientific research and in research management. Substantial contributor to the scientific literature of our day. Honored member of many technical and professional societies. Listed in *Who's Who in America.*

PAUL W. HOWERTON

Director of the Center for Technology and Administration and Visiting Professor, The American University, Washington, D.C. Teacher, lecturer, writer and consultant in the technology of management and more particularly in cybernetics, electronic data processing, information sciences, chemistry, mathematics, and linguistics. Former Deputy Assistant Director for Central Reference of the CIA where he was awarded the Intelligence Medal for Merit for his work in developing information control systems and in computational linguistics. Member of many professional societies both here and abroad. Fellow of the AAAS. Member of Sigma Xi.

AUGUSTUS C. JOHNSON

Noted mathematician-consultant affiliated with Booz-Allen Applied Research, Inc., Frequent contributor in such areas as operations research-linear programming involving industrial and government distribution problems as well as the entire field of logistics. Also in the design and operation of data-collection systems for space-environment tests, as well as a wide variety of systems analysis programs. Lecturer and author. Member of the Society for Industrial and Applied Mathematics.

F. ELLIS KELSEY

Assistant to the Surgeon General, U.S. Public Health Services. Well-known authority, writer, teacher, and consultant in information sciences, particularly difficult aspects of information abstracting, indexing, storage, and retrieval of advanced chemical-biological literature interfaces.

HAROLD F. LANIER

Manager, Advanced Technology Division, Goodyear Aerospace Corporation. Noted technology and management-sciences executive with outstanding credits in guided-missile programs and in research-development administration, both in the government and in industry. Well-known lecturer and contributor with a deep involvement in the managerial implications of advancing technology, as well as in the information sciences.

G. S. SIMPSON

Battelle Memorial Institute Executive, concerned with major responsi-

bility for administration of specialized information centers. Noted lecturer, authority in the information sciences field, more particularly in the management of information centers. Frequent participant in the coordination of interdisciplinary research in such areas as aeronautical, aerospace, electronics, defense, and the biosciences—to name but a few.

JOHN I. THOMPSON

President, John I. Thompson and Company, Washington, D.C. Noted technology and managerial executive specializing in research and development management, including a deep involvement in the information sciences area. Well-known lecturer, and author, recognized as an outstanding authority in many facets of research and development administration.

MORRIS M. THOMPSON

Deputy Assistant Chief Topographic Engineer for Research and Technical Standards, U. S. Geological Survey. Noted aerial topographer, frequent lecturer, writer, and contributor in various aspects of photogrammetry. Fellow, AAAS and ASCE. Editor-in-chief, *Manual of Photogrammetry,* published by the American Society of Photogrammetry (1965).

ISAAC D. WELT

Deputy Director of the Center for Technology and Administration, and Research Professor, The American University, Washington, D.C. Noted teacher, lecturer, writer, and consultant in chemistry and information sciences, more particularly in abstracting, indexing and retrieval aspects. Active in the cardiovascular literature project of the National Academy of Sciences since its inception, and its Director, 1955-1961. Associate Director of the Institute for Advancement of Medical Communication prior to its Cardiovascular Literature Program becoming affiliated with the Center for Technology and Administration in 1964. Author of many scientific papers and frequent contributor and editor of information science volumes. Member of numerous scientific and professional societies.

THOMAS WINTERFIELD

Senior Analyst, National Oceanographic Data Center. Experienced geologist and oceanographer specializing in such matters as physical properties of sea water, etc. Prominent in information sciences, more particularly in the data format, data processing, and retrieval aspects.

CONTENTS

Part III. SPECIALIZED DATA/INFORMATION SOURCES AND THEIR UTILITY

PART I

CURRENT DEVELOPMENTS IN DATA SOURCE, TECHNIQUE, AND FUTURE TRENDS

BIOLOGICAL DATA/INFORMATION CENTERS AND THE FUTURE

ISAAC D. WELT, PH.D.
Deputy Director for Scientific and Technical Information
Center for Technology and Administration
The American University

A rigorous definition of the terms "data" and "information" is needed before we venture to talk about them. Very few words have been as badly misused as this popular word *information.* Information science, information theory, information storage and retrieval are used synonymously with a total lack of regard for precision of language. The most flagrant abuse of the term is probably in describing a document storage and retrieval system. Here, there is a confusion between ends and means.

Ideally, the aim of an information center is to provide information. In practice, however, the vast majority of these organizations provide nothing more than unevaluated documents which, presumably, contain information. It is up to the user to extract pertinent information from this pile of printed material.

Document storage and retrieval systems have been discussed quite amply in the published literature of documentation as well as at numerous meetings, symposia, conferences, and workshops. In recent years mechanized methods for the handling of documents have come to the fore, involving either magnetic tape or photographic film as the storage medium for indexes to documents. Photographic film possesses the added economic advantage of being available, in microform, as a storage medium for documents as well.

The relationship between true data and information storage and retrieval, which is the subject of the present Institute, and what is generally described in the literature under that name, is close enough to have provided opportunity for confusion. There is a spectrum of activities, of a

continuous nature, which embraces a good part, if not all, of the processes involved in scientific communication. It begins with data and ends with data.

First, we have the experimental data developed by the scientist and engineer in his laboratory, much of them in tabular or graphic form. Convention dictates that the results of research be written up for presentation at meetings or for publication in scientific journals. As a result, a superstructure of language is built around the data. Thus we have the historical introduction to the work, a description of apparatus and techniques, the actual data, a discussion of what the author thinks they mean, some polemics, a summary and a list of citations or references. This constitutes the conventional scientific paper.

Some time after its publication most scientists and engineers decry the time lag. The paper is abstracted by, we hope, a competent subject-matter authority. If the abstract is at all informative, it must contain something about methodology and conclusions. It does not usually contain the actual data. Still later on, the abstract, or perhaps the original paper itself, is indexed for accurate and efficient retrieval. Retrieval of what? Of the paper. In other words, this is indexing for document retrieval. An efficient document-retrieval system will greatly facilitate the writing of review articles, both critical and uncritical. To do so, the reviewer should read all of the pertinent documents, thereby becoming familiar with the data contained in them. He is actually synthesizing facts which are scattered among many journals in the literature. At this point in the spectrum, data collections become feasible and true information storage and retrieval appears. Laboratory data bearing on a particular problem can now be collected, published in the form of a handbook, or stored on magnetic tape or photographic film. Critical evaluation of the data by competent scientists and engineers is possible. We have arrived at the far end of the communication spectrum and once again are dealing with data rather than with documents. We can thus provide a service which is badly needed by almost all scientists and engineers—namely, specific answers to specific questions, such as the melting point of a chemical substance or the coefficient of expansion of a new alloy.

DATA/INFORMATION

The definition for data—or rather for datum—as given in Webster's *Second International Dictionary,* is perhaps adequate for our needs. Datum is defined as "a fact or principle granted or presented." In the field of science, it is most frequently of a quantitative nature, although not