Wendy A Warr, Claus Suhr

Chemical Information Management

Wendy A. Warr, Claus Suhr

Chemical Information Management

Weinheim · New York · Basel · Cambridge

Dr. Wendy A. Warr
ICI Pharmaceuticals
Mereside
Alderly Park Macclesfield
Cheshire SK10 4TG
Great Britain

Dr. Claus Suhr
Berlinerstr. 3
6701 Dannstadt
Federal Republic n v

Published jointly by
VCH Verlagsgese lschaf mbH We nhe m (Federal Republ c of Germany)
VCH Publishers Inc New York NY USA)

Editorial Directo I Han F Ebel and Dr Wolfgang Ge hartz
Production Manag My am No hacke

A CIP catalogue record for this book is available from the British Library

LOC Card No. applied for.

Deutsche Bibliothek Cataloguing-in-Publication Data:

Warr, Wendy:
Chemical information management / Wendy Warr ; Claus Suhr.
Weinheim ; New York ; Basel ; Cambridge ; VCH, 1992
ISBN 3-527-28366-8 (Weinheim ...)
ISBN 1-56081-180-3 (New York)
NE: Suhr, Claus:

Printed on acid-free paper.

Composition, printing and bookbinding: Graphischer Betrieb Konrad Triltsch, D-8700 Würzburg
Printed in the Federal Republic of Germany

Preface

This book arose from an article written in 1989 and the first part of 1990 for Ullmann's Encyclopedia of Industrial Chemistry, Volume B1, pages 12-1 to 12-119. The chapter on patent information in that article has now been expanded and moved to form Part Two of this book. Changes to the other chapters, which now form Part One, have been minimal because we felt that currency and speed of publication were particularly important. We have managed to insert some very recent references by removing older ones, but some more general ones have to be mentioned here.

In the writing of Chapters 6 and 7 on Information Technology, notes from the Information Technology Primers run by the Association of Information Officers in the Pharmaceutical Industry proved very useful. Much of the material from these primers has recently been incorporated into a book [0.1]. A recent review on storage technology [0.2] should be read in conjunction with Chapter 6.

In Chapter 8 we have cited a well known directory of text retrieval software [8.80] but failed to mention a directory with a wider scope which is now into its fourth edition [0.3]. A lengthy review of online searching aids (front ends, gateways and other interfaces) has also appeared [0.4].

A recent tutorial article on spectral databases [0.5] is pertinent to Chapter 9.

There have been several new publications in the field of chemical structure handling [0.6]–[0.10], all of which could contribute to Chapter 10. Two of these [0.7], [0.8] deal with distributed computing environments and electronic publishing and therefore have relevance to other chapters. The book by Ash et al. will supersede an earlier one [9.1], [10.3] and the software directory [0.10] is more up-to-date, and has wider scope than the book cited at [10.149].

It is almost impossible for two authors alone to cover the wide range of topics in this book, so thanks are due to many people who helped in the writing and researching. Dr. Peter Willett provided a copy of the reading list used by students at the Department of Information Studies at Sheffield University and a bibliography and MSc Thesis on Chemical Structure Processing by Karen Libscomb. Lipscomb's short descriptions of Chemical Abstracts tautomerism and of HTSS are reproduced almost verbatim in Chapter 10, with apologies, but with compliments for their accuracy and succintness. Dr Willett and Dr John Barnard offered helpful advice on chemical structure systems. Dr David Bawden traced various useful references. He. Karen Blakeman, Dr Trevor Devon and Dr Willett are to be thanked for their contributions to the IT primers mentioned above. The Chapter on Business and Economic Information was essentially written by Gillian Swash and Elizabeth Walton of ICI Pharmaceuticals. Janet Ash wrote Chapter 9 using material gathered, or earlier published, by Dr Wendy Warr. A lecture on hypertext by Jane Martin of ICI

Paints was used in the compilation of Section 11.7. John Barber. Section 6.4 on Microcomputers was contributed by Dr. William Town. Derek Bowler, Graham Cousins, Madeline Gray and Pat Holohan of ICI Pharmaceuticals checked many of the references. We thank the Beilstein Institute, Chemical Abstracts Service, and the Institute of Scientific Information for checking relevant sections of this book.

The preparation of Part Two has been greatly supported by some colleagues and organizations. We should like to express our sincere thanks to BASF Aktiengesellschaft, Ludwigshafen (Rhein), and in particular to the staff of its Documentation Department who were extremely helpful in contributing relevant literature and useful advice, further to Philip Higham and Raymond Andary of the World Intellectual Property Organization, Geneva, for providing valuable printed material, and to the management of Derwent Publications Limited, London, for permission to reprint Figure 16.

Abbreviations Used in this Book

ACS	American Chemical Society
ADMIS	Agfa Document Management Information System
ADONIS	Article Delivery Over Network Information Systems
AGLINET	Worldwide Network of Agricultural Libraries
AGRIS	Agricultural Sciences and Technology
AI	Artificial Intelligence
AIMB	Analogy and Intelligence in Model Building
ANSI	American National Standards Institute
API	American Petroleum Institute
AQUIRE	Aquatic Information Retrieval
ARGOS	Automatically Represents Graphics of Chemical Structures
ARIPO	African Regional Intellectual Property Organization
ARS	Agricultural Research Service
ASCA	Automatic Subject Citation Alert
ASCII	American Standard Code for Information Interchange
ASIDIC	Association of Information and Dissemination Centers
ASTM	American Society for Testing and Materials
BASIC	Basel Information Center
BIOSIS	BioSciences Information Service
BLDSC	British Library Document Supply Centre
CAD/CAM	Computer-aided design and manufacture
CAMEO	Computer Assisted Mechanistic Evaluation of Organic Reactions
CAN/SND	Canadian Scientific Numeric Database Service
CAR	Computer-assisted retrieval
CAS	Chemical Abstracts Service
CASDDS	Chemical Abstracts Service, Document Delivery Services
CASE	Computer Assisted Structure Elucidation
CASE	Computer-aided software engineering
CASP	Computer-assisted synthesis planning
CASR	Chemical Activity Status Report
CASSI	Chemical Abstracts Service Source Index
CCC	Copyright Clearance Center
CCD	Charge coupled device
CCDC	Cambridge Crystallographic Data Center
CCITT	Comité Consultatif International Télégraphique et Téléphonique
CCRIS	Chemical Carcinogenesis Research Information System
CDA/DDIF	Compound Document Architecture/Digital Document Interchange Format
CDC	Centers for Disease Control
CD-I	Compact disc interactive
CDIF	Crystal Data Identification File
CD-ROM	Compact disc — read only memory
CESARS	Chemical Evaluation Search and Retrieval System
CHEMICS	Combined Handling of Elucidation Methods for Interpretable Chemical Structures
CHIRON	Chiral Synthon
CHRIS	Chemical Hazard Response Information System
CIN	Chemical Industry Notes
CIS	Chemical Information System
CISTI	Canadian Institute for Scientific and Technical Information
CJO	Chemical Journals Online
CLF	Current Literature File
CNMR	Carbon-13 NMR Spectral Search System
CNRS	Centre National de la Recherche Scientifique
CODATA	Committee on Data for Science and Technology
COM	Computer-output microfilm
COUSIN	Compound Information System
CPI	Chemical Patents Index
CPI	Conference Papers Index

CPSS	Chemist&'s Personal Software Series
CPU	Central processing unit
CRDS	Chemical Reactions Document Service
CROSSBOW	Computerized Retrieval of Organic Structures Based on Wiswesser
CRT	Cathode ray tube
CSD	Cambridge Structural Database
CSO	Central Statistics Office
CTCP	Clinical Toxicology of Commercial Products
DAP	Distributed Array Processor
DARC	Description, Acquisition, Retrieval and Correlation
DCA/DIF	Document Content Architecture
DDE	Dynamic Data Interchange
DECHEMA	Deutsche Gesellschaft für chemisches Apparatewesen
DIF	Drug Information Fulitext
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DIP	Document Image Processing
DIPPR	Design Institute for Physical Property Data
DOC	Dictionary of Organic Compounds
DOE	Department of Energy
DRI	Data Resources Inc
DVI	Digital video interactive
EAGLE	European Association for Gray Literature Exploitation
EBCDIC	Extended binary coded decimal information code
ECDIN	Environmental Chemicals Data and Information Network
ECTR	Extended Connection Table Representation
EI	Engineering Index
EIIA	European Information Industry Association
EINECS	European Inventory of Existing Chemical Substances
EIU	Economist Intelligence Unit
ENVIROFATE	Environmental Fate
EPA	Environmental Protection Agency
EPI	Electrical Patents Index
EPIC	Thermosalt Estimate of Properties for Industrial Chemistry
EPO	European Patent Office
EPOQUE	EPO Query system
ERIC	Educational Resources Information Center
EROS	Elaboration of Reactions for Organic Synthesis
ESA	European System of Integrated Economic Accounts
ESA/IRS	European Space Agency Information Retrieval Service
ESPRIT	European Strategic Program for Research and Development in Information Technologies
EUSIDIC	European Association of Information Services
F*A*C*T	Facility for Analysis of Chemical Thermodynamics
FAO	Food and Agriculture Organization
FDA	Food and Drug Administration
FIZ	Fachinformationszentrum
FREL	Fragment Reduced to an Environment which is Limited
FRSS	Federal Register Search System
FT-IR	Fourier transform infrared
GAIIA	Global Alliance of Information Industry Associations
GCL	GENIE Query Language
GIABS	Gastrointestinal Absorption
GREMAS	Generic Retrieval by Magnetic Tape Search
GRUR	Gewerblicher Rechtsschutz und Urheberrecht
GUI	Graphical User Interface
HCI	Human -- computer interaction
HDS	Hampden Data Services
Heilbron&'s	Dictionary of Organic Compounds
HMSO	Her Majesty&'s Stationery Office
HODOC	Handbook of Data on Organic Compounds
HORD	Hierarchically Ordered Ring Description
HOSE	Hierarchically Ordered Spherical Description of Environment
HSDB	Hazardous Substances Data Bank
HTSS	Hierarchical Tree Substructure Search
IAEA	International Atomic Energy Agency
IARC	International Agency for Research on Cancer
ICC	Inter-Company Comparisons
ICP	Index to Conference Proceedings Received
ICR	Intelligent optical character recognition
ICSD	Inorganic Crystal Structure Database

ICSTI	International Council of Scientific and Technical Information
ICSUAB	International Council of Scientific Unions Abstracting Board
IDC	Internationale Dokumentationsgesellschaft für Chemie
IDIOTS	Infrared Spectra Documentation and Interpretation Operating with Transcripts and Structures
IFI	Information for Industry
IFIS	Industry File Index System
IFLA	International Federation of Library Associations
IGOR	Interactive Generation of Organic Reactions
IIA	Information Industry Association
IIB	Institut International des Brevets
IKBS	Intelligent knowledge-based systems
IMF	International Monetary Fund
INID	Internationally Agreed Numbers for the Identification of Data
INIS	International Nuclear Information System
INPADOC	International Patent Documentation Center
INPI	Institut National de la Propriété Industrielle
INSPEC	Information Service for the Physics and Engineering Communities
IPC	International Patent Classification
IPSS	International Packet Switching Service
IRDC	IR Data Committee of Japan
IRSS	Infrared Search System
ISDN	Integrated Services Digital Network
ISHOW	Information System for Hazardous Organics in Water
ISI	Institute for Scientific Information
ISO	International Standards Organization
ISTP	Index to Scientific and Technical Proceedings
ITC	International Translation Center
IUPAC	International Union of Pure and Applied Chemistry
JAICI	Japan Association for International Chemical Information
JANAF	Joint Army, Navy, and Air Force
JCAMP	Joint Committee on Atomic and Molecular Physical Data
JCPDS	Joint Committee on Powder Diffraction Standards
JICST	Japan Information Center of Science and Technology
JIPID	Japanese International Protein Information Databank
JPO	Japanese Patent Office
LAN	Local area network
LHASA	Logic and Heuristics Applied to Synthetic Analysis
MACCS	Molecular ACCess System
MCS	Maximal common substructures
MDL	Molecular Design Ltd.
MEAL	Media Expenditure Analysis Ltd.
MIMD	Multiple instruction stream, multiple data stream
MIPS	Martinsried Institute for Protein Sequences
MISD	Multiple instruction stream, single data stream
MMI	Man - machine interaction
Modem	Modulator — Demodulator
MSDC	Mass Spectrometry Data Center
MSDS	Material safety data sheet
MSI	Marketing Surveys Index
MSSS	Mass Spectral Search System
NASA	National Aeronautics and Space Administration
NBRF	National Biomedical Research Foundation
NBS	National Bureau of Standards
NFAIS	National Federation of Abstracting and Indexing Services
NFPA	National Fire Protection Hazard Rating
NIOSH	National Institute for Occupational Safety and Health
NIST	National Institute of Standards and Technology
NLM	National Library of Medicine
NLP	Natural Language Processing
NMRLIT	Literature Search System
NTIS	National Technical Information Service
OAPI	Organisation Africaine de la Propriété Intellectuelle
OCR	Optical character recognition
ODA	Office Document Architecture
ODIF	Office Document Interchange Format
OECD	Organization for Economic Cooperation and Development
OHM/TADS	Oil and Hazardous Materials/Technical Assistance Data System
OLPI	International Association of Producers and Users of Online Patent Information

OPAC	Online public access catalog
ORAC	Organic Reactions Accessed by Computer
ORTEP	Oak Ridge thermal ellipsoid plot
OSAC	Organic Structures Accessed by Computer
OSI	Open System Interconnection
PAIRS	Program for the Analysis of Infrared Spectra
PATDPA	Patente des Deutschen Patentamts
PATOLIS	Patent Online Information System
PATOS	Patent-Online-System
PBM	Probability-based matching
PC	Personal computer
PCT	Patent Cooperation Treaty
PDG	Patent Documentation Group
PIR	Protein Identification Resource
POSSUM	Protein Online Substructure Searching - Ullman Method
PPDS	Physical Property Data Service
PROLOG	PROgramming in LOGic
PSS	Packet Switching Service
PSTN	Public Switched Telephone Network
QI	Quality Index
QSAR	Quantitative structure – activity relationship
RAM	Random access memory
RAPRA	Rubber and Plastics Research Association
REACCS	REaction ACCess System
ReSy	Research System
RISC	Reduced instruction set computer
ROM	Read only memory
ROSDAL	Representation of Structure Diagram Arranged Linearly
RTECS	Registry of Toxic Effects of Chemical Substances
S4	Softron Substructure Search System
SAHO	Spectral Appearance in Hierarchical Order
SANDRA	Structure and Reference Analyzer
SANSS	Structure and Nomenclature Search System
SCORE	Scan Conversion for Outline Representation of Images
SDI	Selective dissemination of information
SECS	Simulation and Evaluation of Chemical Syntheses
SEMA	Stereochemically Extended Morgan Algorithm
SGML	Standard Generalized Mark-up Language
SIGLE	System for Information on Gray Literature
SIMD	Single instruction stream, multiple data stream
SISCOM	Search for Identical and Similar Components
SISD	Single instruction stream, single data stream
SMD	Standard Molecular Data
SMILES	Simplified Molecular Input Line Entry System
SOLUB	Aqueous Solubility Data Base
SQL	Standard Query Language
SRD	Standard Reference Data
SSI	Small scale integration
STIRS	Self-Training, Interpretive, and Retrieval System
STN	Scientific and Technical Information Network
TDRS	Text Data Retrieval System
THOR	Thesaurus-Oriented Retrieval
TIB	Technische Informationsbibliothek
TLV	Threshold Limit Value
TORC	To Ring-Code
TOSCA	Toxic Substances Control Act
TRC	Thermodynamics Research Center
TSCA	Toxic Substances Control Act
TSCATS	Toxic Substances Control Act Test Submissions
ULSI	Ultralarge scale integration
UMI	University Microfilms International
USPTO	United States Patent and Trademark Office
VCI	Verband der Chemischen Industrie
VDU	Visual display unit
VINITI	All-Union Institute for Scientific and Technical Information
VLSI	Very large-scale integration
WDC	World Data Centers
WHO	World Health Organization
WIMPs	Windows, icons, mice, and pointing devices

WIPO	World Intellectual Property Organization
WLN	Wiswesser Line Notation
WMSSS	Wiley Mass Spectral Search System
WORM	Write-once - Read many (optical disc)
WPA	World Patent Abstracts
WPI	World Patents Index
WPIM	World Patents Index Markush
WYSIWYG	What You See is What You Get

Table of Contents