The CIS (Chemical Information System): An Overview

CHEMICAL INFORMATION SYSTEMS, INC.*
7215 York Road
Baltimore, Maryland 21212
(301) 321-8450 or (800) CIS-USER
Fax: (301) 296-0712

* A wholly owned subsidiary of Fein-Marquart Associates
TABLE OF CONTENTS

The CIS: An Overview  1
Analysis and Modeling  3
Chemical and Physical Properties  4
Environment  5
Hazardous Materials Handling and Disposal  7
Manufacturing Data  8
Material Safety Data Sheets  9
Pharmaceuticals  9
Regulations  10
Spectroscopy  11
Structure and Nomenclature  13
Toxicology and Carcinogenicity  14
The CIS is a collection of computerized data storage and retrieval components for chemical information. Originally developed by Fein-Marquart Associates during the 1970s under contract to the US Environmental Protection Agency and National Institutes of Health, the system has been made available to the public since 1984 by Fein-Marquart's wholly owned subsidiary, Chemical Information Systems, Inc. In the years since the system first began operating, it has grown to one of the world's largest sources of online chemical information, made up (at present) of approximately forty different components. Accessible worldwide via terminals or personal computers operating on telephone lines, the system today has subscribers on five continents and its users are numbered in the thousands -- including municipal, state, and national governmental agencies in the United States and abroad; chemical and petrochemical companies; pharmaceutical manufacturers; universities; public interest groups; fire departments; hospitals; and many others.

Each component of The CIS is essentially a "standalone" system dealing with a particular aspect of chemistry, toxicology, environmental pollution, or whatever. However, all components are prepared according to a standard set of CIS guidelines so that they can share utility software and communicate with each other. This makes it relatively easy to conduct composite searches dealing with various subjects, and, similarly, to display information on substances not only from the component in which the substances were initially identified, but from other components as well.

Because the software for each component conforms to The CIS guidelines, each of the components can perform general "universal" functions such as:

Displaying NEWS messages informing users of changes or additions to the system, helpful suggestions, warnings, etc.;

displaying HELP messages to provide detailed information, on request, on the use or interpretation of specific system functions;

performing Boolean operations on query results lists, regardless of the component generating the list, i.e., merger, intersection, or negated intersection of lists;

sending comments to The CIS system managers for action and reference;

estimating the cost of the current session;

saving and playing back the sequence of operations performed during the session; and

displaying within one component the data contained within another.

The following information is available for most substances in The CIS:

CAS Registry Number;

CAS Index Name(s) and other names and synonyms by which the substance is commonly known in commerce and manufacturing;

chemical structural diagram; and

a list of sources of further information.

In addition, as described in the following sections, much additional information (e.g., mass spectrum, toxicity data, regulations) is also available.

The following sections discuss briefly the features of various components of The CIS. (The descriptions of the contents of the components are as of approximately 20 March 1991.)
ANALYSIS AND MODELING

Pattern Recognition (ARTHUR)

ARTHUR, from Infometrix, Inc., is a state-of-the-art pattern recognition and data analysis program. It contains algorithms for data manipulation, display, classification, dimensionality analysis, and a variety of regression techniques; its primary use is the formulation and evaluation of models for incompletely understood data sets, particularly when multi-variate data sets are involved. It is a complement to such statistical systems as SPSS and BMD. The version introduced on The CIS in September 1987 is 4.1.

The Chemical Modeling Laboratory (CHEMLAB)

CHEMLAB is a package of molecular structure calculation routines (options) developed or adapted by Case Western Reserve University, Department of Macromolecular Science. CHEMLAB is a completely interactive program that allows features to be selected and changed as often as desired with a minimum of effort. This allows the user to perform many different kinds of calculations on any number of molecules (one at a time) during a single CHEMLAB session. The user defines a molecule in the CHEMLAB Molecular Workspace. Data in this work area can be examined, modified, and extensively manipulated. Molecules can be entered via keyboard, or carried into CHEMLAB after extraction from SANSS (see entry for Structure and Nomenclature below). Either Linear Free Energy (LFE) calculations or full three-dimensional energy optimizations can be performed.

Modeling Laboratory (MLAB)

MLAB is a tool for experimentation with, and evaluation of, mathematical models (functions). It is an interactive system, the heart of which is a curve-fitting program which will adjust the parameters of a model function to minimize the sum of the squared errors. A repertoire of mathematical operators and functions, a collection of routines for teletype and CRT plotting, and mechanisms for saving data between sessions provide a powerful and convenient environment for data manipulation, for arithmetic calculations, and for building and testing models. The user communicates with MLAB by typing commands. Most commands are executed at once; a few will prompt the user for additional information.
CHEMICAL AND PHYSICAL PROPERTIES

Information System for Hazardous Organics in Water (ISHOW)

ISHOW contains six types of data (melting point, boiling point, partition coefficient, acid dissociation constant, water solubility, vapor pressure) for more than 5,000 chemical substances. Not all properties are recorded for all substances.

Records can be retrieved not only by chemical name and CAS Registry Number, but also by ranges of values -- for example, one can search for substances whose melting points fall between 100 and 125 degrees.

Thermodynamic Property Values Database (THERMO)

THERMO provides thermodynamic property values for more than 14,000 chemical substances. The data are derived from two sources: The National Bureau of Standards (NBS) Tables of Chemical Thermodynamic Properties database and the thermodynamic data files of the Thermodynamics Research Center (TRC) at Texas A&M University. Each THERMO entry reports data for a specific substance in a particular state or degree of dilution, but the type of information contained in each of the data sources differs significantly. The NBS portion of the database is composed of entries from "Tables of Chemical Thermodynamic Properties," published in the Journal of Physical and Chemical Reference Data, volume 11, supplement 2, 1982. This information was originally presented in serial form, from 1965-81, as "NBS Technical Notes 270." It is the product of twenty years of research and data evaluation from 60,000 literature references.

Substances in the NBS portion of the database are primarily inorganic, with some one- and two-carbon organic materials included. Where available, values are given for up to six specific thermodynamic properties.

The TRC portion of the database has entries for several hundred organic compounds. It includes data from the TRC Hydrocarbon Project (formerly API Research Project 44) and the TRC Data Project. Coverage includes thirty-four distinct thermodynamic properties, although no single entry contains information for all these properties. Unlike NBS, the TRC data often apply over a range of temperatures and pressures. These variable properties can be calculated by interpolation for specified temperatures or pressures.
ENVIRONMENT

CERCLA Information System (CERCLIS)

The CERCLIS database contains one entry for each hazardous waste disposal site or spill that has either been listed by the Environmental Protection Agency (EPA) on the National Priority List (NPL) for cleanup under Superfund or nominated for consideration for the NPL. More than 34,000 sites all across the US were covered by the database at its last quarterly update.

Each entry in the database contains all the information provided to The CIS by EPA. This information generally includes:

One or more names for the site;

a variety of geographic locaters (street address, city, state, county, ZIP code, latitude/longitude, congressional district, SMSA, USGS hydrologic unit identifier);

a variety of codes classifying the site;

a series of tables indicating actions taken or proposed to be taken with regard to the site under the Superfund program; and

names and addresses of potentially responsible parties (PRPs) -- those who may be responsible for cleanup costs.

Individuals requiring more information about a site than is provided by CERCLIS can contact EPA regional offices to review the site dockets maintained in those offices.

All of the items above are searchable so that a user can retrieve the record(s) for a site or group of sites by name, by alternative names, by any of the geographic locaters (e.g., ZIP, county, etc.), by the available codes, or by the status entries under Superfund. A special search routine, GEOSEARCH, allows users to enter a latitude, longitude, and allowable distance (in miles) to retrieve all sites located within the allowable distance from the point defined by the user. A special file containing records known to include erroneous data has also been created so that users may identify sites on the file even before the data have been corrected by EPA.
Users can display individual whole records, sets of whole records, parts of individual records, or parts of sets of records, at the user's discretion. Records may also be displayed with or without PRP data.

**Environmental Fate (ENVIROFATE)**

ENVIROFATE deals with the environmental fate or behavior (i.e., transport and degradation) of chemicals released in the environment. Its records are drawn from papers published around the world. They include data on environmental transformation rates (e.g., biodegradation, oxidation, hydrolysis) and on physical-chemical properties (e.g., water solubility, vapor pressure). The database currently includes more than 13,000 records on some 800 substances.

**PressNet Environmental Reports (PER)**

PER contains summaries of articles on environmental issues that have appeared in newspapers around North America, concentrating on local action in order to make the database a supplement to the national and international environmental coverage that is already readily available to everyone. More than 5,000 articles on state and local environmental issues and actions are referenced in the database as of this writing; they cover the period January-October, 1990, when weekly updates were suspended by PressNet, which supplies PER to The CIS. Resumption of weekly updates is expected later in 1991.

An individual interested in searching PER can input a word or phrase -- RECYCLING, for example, or DRINKING WATER -- and retrieve from the database all those records that contain the word/phrase. Retrievals can also be limited by date, by locality, or by a variety of other parameters. Once a set of records has been retrieved, it may be displayed in whole or in part, at the user's discretion. The user may also choose to display whole records or partial records. A given record typically includes the name of the newspaper where the article originally appeared; the date of the article's appearance; the headline accompanying the article; the byline of the author of the article; a summary of the article; and indications of the play given the article (its size, placement on the page, etc.).

PER is extracted from PressNet, the daily electronic news service for business and government operated by PressNet Systems, Inc., 400 E. Pratt Street, Baltimore, MD 21202, 800-666-3236.
Facilities Index Data System (FINDS) -- Added 4/18/91

The FINDS database contains entries for more than 450,000 sites or facilities regulated by the US EPA under a variety of statutes -- RCRA, the Clean Air Act, the Clean Water Act, TSCA, FIFRA, etc.

Each entry in the database contains all the information provided to The CIS by EPA. This information generally includes:

- The name of the site;
- A variety of geographic locaters (street address, city, county, ZIP code, latitude and longitude, congressional district);
- A variety of codes classifying the site; and
- A list of other EPA and state databases containing information about the site, including the code numbers used to retrieve pertinent records from those other databases.

Individuals using FINDS typically seek information about a given site by name or EPA Identifier, or information about all the sites in a given area (as defined, say, by ZIP code or city name). The retrieved records are then reviewed to see what other sources of information about a given site may be consulted. In the case of the CERCLIS database, which is already present on The CIS, users may simply type out CERCLIS records for records retrieved in FINDS. (The reverse is also true -- one can type out FINDS records for records retrieved in CERCLIS.) In the case of other EPA and state databases referenced in FINDS, appropriate Freedom of Information Act procedures must be used until such time as those other databases have also been added to The CIS.

Detailed search procedures in FINDS are virtually identical to those in CERCLIS; see the CERCLIS entry above for a description. As with CERCLIS, FINDS includes a special file with known erroneous or incomplete records; this gives users an opportunity to identify records of interest even if they could not be retrieved by ordinary search techniques.
HAZARDOUS MATERIALS HANDLING AND DISPOSAL

Chemical Hazard Response Information System (CHRIS)

CHRIS, produced by the US Coast Guard, provides information needed to respond to emergencies that occur during the transport of hazardous chemicals. CHRIS also provides information that can be used to design safety procedures aimed at preventing emergency situations. A variety of inputs may be used to retrieve records. CHRIS contains information on labeling, physical and chemical properties, health hazards, fire hazards, chemical reactivity, water pollution, and hazard classifications. At the present time, it covers 1,156 chemical substances; an update involving some sixty additional substances is expected in late spring of 1991. While geared toward chemicals transported over water, the information contained in CHRIS is useful for a wide range of chemical emergency situations.

Hazardous Chemicals Information and Disposal (HAZINF)

HAZINF is based on the book *Hazardous Chemicals Information and Disposal Guide* by M. A. Armour, L. M. Browne, and G. L. Weir of the Chemistry Department of the University of Alberta. The database provides information necessary to assess and respond to hazards associated with chemical substances, particularly as they are likely to arise in the laboratory. There are some 220 substances or classes of substances covered in HAZINF, with up to fifteen data fields for each record in the database. As with CHRIS and OHM/TADS, records may be retrieved by a variety of means, including chemical name, CAS Registry Number, and numeric or textual values in the records.

Oil and Hazardous Materials/Technical Assistance Data System (OHM/TADS)

OHM/TADS -- the Oil and Hazardous Materials/Technical Assistance Data System -- is the principal resource developed by the US Environmental Protection Agency for storing and retrieving information needed to deal with emergency situations involving substances that have been designated as oils or hazardous materials by the EPA. At present there is information in the component for 1,402 substances. While the primary function of this database is to provide information to emergency response team personnel, it can also be regarded as a general source of diverse information on hazardous substances.
OHM/TADS permits retrieval of information by a wide variety of inputs -- for example, by chemical name or synonym, by CAS RN, by numeric values or ranges of values in numeric fields (i.e., vapor pressure, flash point, etc.), and miscellaneous textual entries contained in the texts of the records. OHM/TADS records are organized into 126 fields, each of which contains some class of data about the substance to which a record pertains. One can list all of the information that is available for a particular chemical or restrict the listing to only the fields of interest.

(An update to the OHM/TADS database is in process as of this writing in March 1991; however, no information on the scope or date of installation of the update is yet available.)

MANUFACTURING DATA

Toxic Substances Control Act Plant and Production Data (TSCAPP)

The US Toxic Substances Control Act (TSCA) required the establishment of an inventory of many chemicals in US commerce during the period 1975 through 1977. This inventory was generated by a two-phase reporting process. In the first phase, manufacturers and importers of chemical substances in commercial quantities were required to report on those substances. In the second phase, processors and users of chemical substances were required to report on those substances which they processed or used, but which were not reported in the first phase of the inventory (determined from the publication of the results of the first phase).

More than 85,000 reports, itemizing more than 140,000 chemical substances (more than 55,000 unique substances) were received by EPA in this two-phase process. The current version of the TSCAPP database contains some 127,000 production citations for some 53,000 unique substances, representing the nonconfidential portion of the aforementioned reports. For each chemical substance report, corresponding plant site information is also available in the database.

The production and plant site information in the database can be searched on the basis of CAS Registry Numbers, manufacturer's name and address, production volume, and certain special production and usage codes. These searches result in the identification of chemical substance/plant site pairs satisfying the search criteria. Information in the TSCAPP database pertinent to these substance/site pairs may then be displayed in tabular form, or the identified substances may be referenced in other CIS components where the CAS Registry Number is used as a basis of identification.
MATERIAL SAFETY DATA SHEETS

J. T. Baker Company Material Safety Data Sheets (BAKER)
Mallinckrodt, Inc., Material Safety Data Sheets (MALLIN)

BAKER and MALLIN are collections of material safety data sheets (MSDSs) prepared in accordance with guidelines issued by the US Occupational Safety and Health Administration (OSHA) by the J. T. Baker Chemical Company of Phillipsburg, New Jersey, and Mallinckrodt, Inc., of St. Louis. BAKER contains 1,675 and MALLIN 1,492 MSDSs. One chemical substance is covered in each sheet/record. A given record can be retrieved by input of a chemical name or synonym, a chemical name fragment, a CAS Registry Number, or a variety of other inputs.

While the contents of the records vary (depending on the information available and appropriate for a given substance), the records generally contain: Chemical name; formula; molecular weight; CAS Registry Number; common synonyms; precautionary labeling data; physical properties data (e.g., boiling point, solubility, etc.); fire and explosion hazard data; health hazard data; reactivity data; spill and disposal procedures; information about appropriate protective equipment; and information on storage and handling.

PHARMACEUTICALS

Drug Information Fulltext (DIF)

DIF contains the complete contents of two reference volumes published by the American Society of Hospital Pharmacists: American Hospital Formulary Service Drug Information and Handbook of Injectable Drugs. The first of these is a collection of drug monographs on virtually every single-drug entity available in the US. The second contains information on the stability and compatibility of both commercially available and investigational drugs. Together, these publications provide over 1,200 monographs offering extensive information on various specific drugs and classes of drugs. Monographs can be retrieved by nonproprietary drug names (in the style of USAN and the USP Dictionary of Drug Names), trade names, Chemical Abstracts Service Registry Numbers (CAS RNs), AHFS Classification Numbers, therapeutic classifications, and words, phrases, or values contained in the texts of the monographs. Although the contents of the monographs vary, they generally contain thorough discussions of pharmacokinetics, metabolism, and absorption of drugs; toxicity; drug interactions; uses; cautions; dosages; stability; chemistry; and other information important to therapeutic use.
The Merck Index Online was added to The CIS in early December 1984. This massive compendium of chemical and pharmaceutical information, an industry bible for decades, serves a variety of needs for a variety of CIS users -- information on the preparation, patent status, properties, trademarks and drug codes of compounds for chemists, biochemists, and pharmacists; information on the use, principal pharmacological action, and toxicity of substances for biologists, pharmacologists, and health professionals. The Merck Index Online includes all of the monographs contained in the eleventh printed edition of the text (the centennial edition). When regular semi-annual updates to the component begin later in 1991, new monographs and revisions to existing monographs will become available to users of the online system several years before they are available to those who rely on the printed volume. The full text of the database -- including CAS Registry Numbers -- is searchable online.

PDR On-Line®

PDR On-Line® contains the full text of all entries for prescription, over-the-counter, and ophthalmological drugs in the familiar reference works in the Physicians' Desk Reference® series -- more than 2,600 entries as of this writing. Published annually by the Medical Economics Company, Inc., with the cooperation of the manufacturers whose products are covered in the volume, PDR® contains essential information on major pharmaceutical and diagnostic products -- descriptions, pharmacology, uses, drug interactions, adverse reactions, contraindications, precautions, dosages and administration, and so forth. The information is provided by the manufacturers, edited and approved by their medical departments and/or medical consultants. Records can be retrieved by generic and/or proprietary drug names, by manufacturers' names, by therapeutic use, and by a variety of other types of inputs. The database is updated every two to three months to reflect supplements to the printed volume.

REGULATIONS

Federal Register Search System (FRSS)

The FRSS was established to accumulate regulations, rules, standards, and guidelines involving chemical substances, and to provide a means of access to information regarding the status of such regulations and standards. Although it is no longer updated due to the termination of funding by the EPA, the component is retained on The CIS because its citations have
shown useful to users for the period covered. Each citation includes: Substance name(s) and CAS Registry Number, if any; the FR volume and page number; the type of FR article (Notice, Proposed Rule, Final Rule); the agency and/or office; the CFR sections, if any, affected; references to prior FR articles; a descriptive string of "keywords" and ordinary English words that, when expanded for presentation to the user, form an abstract of the intent of the article; and annotated cross-references to citations for other, related substances.

It is important to note that the abstract contained in this system tells the user what the article is about -- not what it actually says. This system is a method of locating and identifying Federal Register articles; it is deliberately not an alternative to the Federal Register itself.

The FRSS contains data for 162,492 citations, beginning with the issue of 1 January 1978 through 29 November 1983, at which time the EPA ceased funding updates. There are no current plans to resume updating the component in the future.

**Regulations Database**

An as-yet-unnamed database dealing with US federal regulations relating to chemicals will become available on The CIS late in the spring of 1991. While details of the component are still being worked out as of this writing, it is known that the component will cover all pertinent regulations promulgated by the Environmental Protection Agency and most, if not all, regulations promulgated by the Food and Drug Administration and Department of Transportation. Regulations issued by other federal agencies, by various US states, and by various non-US entities may be added later. Users will be able to retrieve all pertinent entries by chemical name, by CAS Registry Number, by agency names, by section number from the Code of Federal Regulations, by program name or mnemonic (e.g., Clean Water Act or CWA), and by a variety of other inputs.

**SPECTROSCOPY**

**Carbon 13 Nuclear Magnetic Resonance Search System (CNMR)**

In this database, 13C NMR spectra pertaining to 11,693 compounds are currently available (representing a complete update of the component by addition of the latest available file in November 1985). The user may search this database for compounds having specific spectral characteristics, i.e., values of shift, multiplicity, or intensity. As the user adds more details to
his specification, the list of potential compounds narrows. At any time the
user may display the identity and measured spectrum of each of the
candidate compounds.

Conversely, by specifying a compound, the user may obtain its spectrum and
associated data, such as literature reference and means of measurement.

**Infrared Search System (IRSS)**

The Infrared Search System (IRSS) allows the retrieval of infrared spectra
for known substances as well as the comparison of the spectrum of an
unknown against a library of spectra. IRSS consists of over 4,500 spectra
from two primary sources, the Boris Kidric Institute in Yugoslavia and the
EPA. Several substances may have IR spectra from both organizations.
Data can be displayed in tabular form, or the actual spectra can be displayed
on an appropriate graphics terminal.

**Nuclear Magnetic Resonance Literature Search System (NMRLIT)**

NMRLIT is a component of The CIS which permits searching the index to
"Nuclear Magnetic Resonance Literature Abstracts and Index" published by
Preston Publications, Inc. Currently there are over 43,009 references
covering abstracts published from 1964 through December 1984; the
component is no longer being updated.

Searches in NMRLIT may be based on subject, nucleus, author, or general
reference. As with all other components of The CIS, Boolean logic may be
invoked during searches so that one might ask, for example, for all abstracts
by a given author, in a given journal, during a given year.

**Wiley Mass Spectral Search System (WMSSS)**

The Wiley Mass Spectral Search System (WMSSS) operates against the fifth
edition of the Registry of Mass Spectral Data from John Wiley & Sons, Inc.;
it currently contains some 140,000 spectra for around 120,000 compounds.
Note that the WMSSS component includes all the spectra previously found
on The CIS in the MSSS component, which was based on the mass spectral
file from NBS/NIST.

In addition to searches based on the molecular weight, molecular formula,
and partial formula of the compounds, the database may be searched on the
basis of individual peak and intensity values and on the basis of the
complete spectrum. Two complete spectrum searches, a Biemann search and
a probability-based matching technique, may be employed. The spectra retrieved by the searches may be displayed as tables of peak/intensity values or in graphic form if a suitable graphics device is available.

**STRUCTURE AND NOMENCLATURE**

Structure and Nomenclature Search System (SANSS)

The SANSS component is designed to contain an entry for each compound included in the databases associated with the other individual CIS components. Each such entry includes, as available: CAS Registry Number; systematic name (8th CI and/or 9th CI); synonyms and trade names; molecular formula; structural diagram and/or connection table; and references to other sources of information. Currently there are approximately 350,000 substances in the database.

Users can query SANSS for compounds having a wide variety of characteristics. For example, one can search for all compounds containing a specific structural fragment. The system can also be searched on the basis of name (complete or partial); ring system; functional groups; molecular formula (complete, partial, or ranged); molecular weight; and atom count.

(As an adjunct to SANSS on The CIS, Fein-Marquart markets SuperStructure, a microcomputer software package for chemical structure graphics. This simple but powerful package permits users to draw chemical structures on a CRT using a mouse, then use that structure as a query input to SANSS to search for similar or identical structures. SuperStructure can also be used standalone for private registry systems or in conjunction with word processors for technical report generation. SuperStructure supports:

- Complex compounds with up to 255 atoms; polymers; multiple attachment points; dot disconnected fragments (e.g., hydrates, copolymers); charged and isotopically labeled atoms;
- Screens of predrawn common ring systems and acyclic fragments, plus user-defined and saved moieties;
- All bonds, including stereo up, stereo down, and stereo unknown;
- Query mode multivalued atoms and bonds;
- Translation (movement), rotation, mirroring, deletion, and copying of whole molecules, moieties, and fragments;
saving/restoring moieties on diskette; and printing of structures on dot-matrix printer.

The package is available for either the IBM PC or compatibles. For details about SuperStructure, contact Fein-Marquart Associates, 7215 York Road, Baltimore, MD 21212, 301-821-5980.)

TOXICOLOGY AND CARCINOGENICITY

Aquatic Information Retrieval (AQUIRE)

AQUIRE contains information on acute, chronic, bioaccumulative, and sublethal effects data from tests performed on freshwater and saltwater organisms (bacteria, birds, and aquatic mammals are excepted). Data items included are chemical substance information, test organism details, study protocol, and test results. Each record contains experimental results from a single assay. As with most CIS databases, users can retrieve records by virtually all of the data elements in the records -- e.g., chemical name, CAS RN, species name (Latin or common), test results parameters, type of water body, etc.

The most recent update to AQUIRE was performed in March 1990: The database now includes information on more than 5,200 chemicals and more than 2,400 species in 104,330 records. Copies of articles from which data were drawn may be ordered directly through the component.

Chemical Evaluation Search and Retrieval System (CESARS)

The Chemical Evaluation Search and Retrieval System (CESARS) provides detailed information and evaluations on a group of chemicals of particular importance in the Great Lakes Basin. CESARS was developed by the Office of Materials Control of the state of Michigan's Department of Natural Resources.

CESARS presently contains extensive information on 194 chemicals; an expansion of the database (to approximately 250 chemicals) is anticipated some time in 1991. Each chemical represents one record, or "chemical evaluation," consisting of up to 185 data fields. These fields include toxicity, physical and chemical properties, carcinogenicity, environmental fate (where available), and many other categories of information. A wide variety of inputs may be used to retrieve records from the database. All CESARS records contain references to the studies from which the data in the records were drawn.
Chemical Carcinogenicity Research Information System (CCRIS)

CCRIS, the Chemical Carcinogenesis Research Information System, became an operational component of CIS in June, 1983. As of the most recent update (27 November 1989), CCRIS contained individual assay results and test conditions for 1,451 chemicals in the areas of carcinogenicity, mutagenicity, tumor promotion, and cocarcinogenicity. The data in the component have been extracted from studies that follow widely accepted scientific protocols, although the originator of CCRIS, the National Cancer Institute of the National Institutes of Health, has not endorsed, nor independently reviewed, the results. Generally speaking, studies were obtained from selected literature reviews, surveys of chemicals present in the environment, and NIH-sponsored activities. Bibliographic information is provided for each study for further reference.

Clinical Toxicology of Commercial Products (CTCP)

CTCP is an online interactive database based on the fifth edition of the book of the same title by Drs. Gosselin, Smith, and Hodge. It contains information about approximately 23,000 commercial products -- manufacturer, uses, and composition. For the chemicals comprising a product, CAS Registry Numbers, concentrations, and indication of toxicity (if applicable) are given. The information generally relates to current formulations of products, but information on prior formulations is retained whenever possible.

The database may be searched against essentially all of its fields. For example, one might inquire about all products containing phosphoric acid which are used as fertilizers. The capability which The CIS offers to combine search terms (searching for all of them simultaneously) makes searching the online CTCP much simpler and more flexible than searching the printed text -- and the online version of CTCP is available nowhere but on The CIS.

Dermal Absorption (DERMAL)

DERMAL contains records on toxic effects, absorption, distribution, metabolism, and excretion data related to the dermal absorption of chemicals. Data on toxic effects of chemical exposure through other routes, such as oral or inhalation, are included if they appear in articles along with dermal data. DERMAL contains about 3,000 test records for 655 substances. Records are retrievable by a wide variety of inputs -- for example, chemical name, CAS RN, study purpose, etc.
Genetic Toxicity (GENETOX)

GENETOX contains mutagenicity information on more than 4,300 chemicals as tested against a variety of biological systems. For each substance, the assay results are tabularly formatted and provide specific information on the type of assay performed, the biological host for the assay, the type of endpoint measured, and the final qualitative results. As with almost all CIS databases, records can be retrieved by a wide variety of inputs -- chemical name, CAS RN, test system, or test results, for example.

Gastrointestinal Absorption Database (GIABS)

GIABS contains references to articles in the scientific literature dealing with the absorption, distribution, metabolism, or excretion of specific chemicals in laboratory animals or humans. Each record in the database deals with a specific experiment involving a specific chemical. Chemical name and CAS Registry Number of the primary test substance are included in the record. Also included are information on species and strain of the test subject, route of application, duration of test, and purity of test substance. A full bibliographic citation is provided in each record; copies of articles referenced may be ordered directly through the component. At its most recent update on The CIS (October 1988), GIABS included 12,052 citations to a total of 4,941 unique literature references from 1967 to 1987; information on more than 3,100 unique chemicals is included.

Integrated Risk Information System (IRIS)

IRIS, a product of the US Environmental Protection Agency, is a risk-assessment database, providing information on levels of exposure to hazardous chemicals below which no adverse health effects are expected to occur in various segments of the human population (e.g., children, adults). Both carcinogenic and non-carcinogenic health effects are taken into account, for exposure by ingestion or inhalation. Substantial discussions of the rationales underlying the setting of acceptable exposure limits are included. Regulatory restrictions on chemical substances under federal statutes are also summarized, and some chemical/physical properties data are included. The database covered 448 substances as of the most recent update on 4 February 1991; additional substances are expected to be added with each regular quarterly update for the foreseeable future.

Plant Toxicity Data (PHYTOTOX)

PHYTOTOX contains some 70,000 records relating to the biological effects of the application of organic chemicals to terrestrial plants. Both natural
and synthetic organic compounds administered to native, crop, or weed plant species have been considered. The records in the database include information about the effects observed in the experiment and bibliographic references to the source documents from which the data were drawn. Data were compiled in the University of Oklahoma Department of Botany and Microbiology.

**Registry of Toxic Effects of Chemical Substances (RTECS)**

The Registry of Toxic Effects of Chemical Substances component (RTECS) is based on the publication of the same name from the US National Institute of Occupational Safety and Health (NIOSH). Currently there is information in the system for more than 106,000 substances; updates of the component are performed quarterly, and additional substances are usually added with each update.

RTECS can be used to display information about particular chemical substances. For this purpose, substances can be identified by means of CAS Registry Number, NIOSH number, chemical name or synonym, and a variety of other inputs. For each substance, a quantified summary of published toxicity, mutagenicity, teratogenicity, carcinogenicity, and irritation measurements can be displayed.

A special RTECS search capability also permits the identification of all substances for which specific end effects (e.g., eye irritation) have been measured on specific classes of animals (e.g., rabbits) for specific means of application, at dosages within a given range.

Bibliographic references are available for each study cited in the database. In addition, RTECS records include special fields dealing with review articles on toxicology, with evaluation of a substance for carcinogenicity by IARC or NTP, and with transport labeling requirements as promulgated by the US Department of Transportation and the UN International Maritime Organization.

**Toxic Substances Control Act Test Submissions (TSCATS)**

This database enables users to identify (and secure copies of) unpublished health and safety studies submitted to the US Environmental Protection Agency under various sections of the Toxic Substances Control Act. TSCATS, which is updated quarterly, presently includes some 36,000 citations on more than 4,200 chemical substances. There are a variety of ways of identifying studies of interest -- for example, one can search the database by CAS RN, by chemical name, by name of submitting
organization, by study purpose, by observed effects, or by a variety of other inputs. Once studies of interest have been identified, TSCATS provides a mechanism for ordering copies, which are distributed on microfiche or on paper printouts prepared from microfiche. In addition to offering copies of particular studies identified through TSCATS, CIS also offers for sale complete microfiche sets of the studies referenced by TSCATS.