ドイツの科学・技術情報活動

ドイツにおける化学情報活動

Chemical Information Activities in Germany

~~~~~ デプランク, ルネ*

[抄録] 本稿は、(社)化学情報協会主催の第10回化学情報講演会「ドイツの科学・技術情報活動」の講演を編集し、収録したものである。演者はドイツ連邦共和国の化学の専門情報センターFIZ Chemieの代表者であり、FIZ Chemie, Gmelin Inst., Beilstein Inst. 3機関の活動を紹介する形でドイツでの化学情報活動の現状および新しい動きを概括した。まずドイツにおける化学情報の歴史を概観した後、FIZ Chemieの活動をその主要分野である有機反応、ケミカルエンジニアリングでのデータベースの作成を中心に述べ、Gmelinでは所有しているデータの総合的な処理、提供システム、Beilsteinでは構造検索の社内データベースへの提供など、それぞれの新しい動きを中心にその活動を述べた。

[キーワード] ドイツ, 情報機関, FIZ Chemie, Gmelin Institute, Beilstein Institute, 化学, 情報活動, 情報提供, 情報システム, ファクトデータベース, STN

[Abstract] This is a lecture paper presented at the Tenth Chemical Information Seminar titled Scientific and technical information activities in Germany. The author is representatives of three organizations, a specialized chemical information center called FIZ Chemie, Gmelin Institute and Beilstein Institute. He outlined the current situation and new movements of chemical information activities in Germany by introducing activities of the above organizations. After briefing the history of chemical information in Germany he talked about activities of FIZ Chemie centered around database production of organic reaction, chemical engineering, comprehensive data processing and provision system conducted at Gmelin, and activities centered around new movements such as providing structure searching for inhouse databases carried out in Beilstein.

[Keywords] Germany, information organization, FIZ Chemie, Gmelin Institute, Beilstein Institute, chemistry, information activities, information provision, information system, fact database, STN

Deplanque, René : Director, FIZ Chemie

この講演は、1994年5月23日(月)、(社)化学情報協会の主催により、東京・千代田区の日本化学会7Fホールで開催された、「第10回化学情報講演会 ドイツの科学・技術情報活動」の講演を、主催者のご好意により再録したものです。

ドイツの化学情報の歴史

本日は皆様方に、ドイツにあります非常に大きな化学の情報センターについてお話しできることを大変うれしく思います。きょうは私はFIZ Chemieを代表すると同時に、Gmelin InstituteとBeilstein Institute, すべてを共通して代表する形になっています。

それではまず、ドイツの化学情報の歴史について簡単にお話し申し上げたいと思います。

1817年に初めて有機と無機と両方ですが、『Gmelinハンドブック』が刊行されました。これは3巻にわたるものでして、全分野についての情報をここに含むことができました。1830年に初めての抄録誌が発行されましたが、これは『薬学中央誌 (Pharmazeutisches Zentralblatt)』と言っています。この抄録誌が母体となって、1835年に『化学中央誌 (CHEMISCHES ZENTRALBLATT)』が発足いたしました。1881年になると、当初、無機と有機を合わせていた『Gmelinハンドブック』から、有機化学の分野が独立して、『Beilstein有機化学ハンドブック』が成立します。

1969年には、ドイツ全体をカバーする形での『化学中央誌 (CHEMISCHES ZENTRALBLATT)』は廃刊されて、これ全体がドイツ化学会の一つのセクションになり、化学情報ドキュメンテーション・ベルリン, CIDBというセクションになりました。1981年になって、FIZ Chemieという化学情報専門センターが正式に発足をします。

1983年にはFIZ ChemieがドイツにおけるCASの代理機関となります。同様にSTNの全体システムの構築にも関与します。1984年になると、先ほど申し上げた『化学中央誌 (CHEMISCHES ZENTRALBLATT)』の後継の情報であった『ChemInform』を電子化するプロジェクトが開始されます。1985年には、これがSTNの中のオンラインデータベースとなります。1988年になると、FIZ ChemieがCASのデータベースのドイツにおけるインプットを行うようになります。1991年、92年には、MDLとの協力関係が強化されて、化学

反応に関する情報の収載の強化が行われます。

FIZ Chemieの活動とは

FIZ Chemieは企業形態をとって、連邦政府並びに各州の政府からの助成を得ています。同様に経営に参加をしているのが、ドイツ化学会とDECHEMA、さらには合成樹脂研究協会です。

FIZ Chemieは国内外にわたる各種の機関と様々な契約関係を持っています。

FIZ Chemieが作成しているデータベースは、化学反応データベース、文献、ファクトおよび製品データベースなどがあり、定期的な抄録サービスや、専門文献などのスタンダードなプロフィールサービスを行っています。

FIZ Chemieはユーザのためにさまざまな検索調査を行っています。各種の専門雑誌における発表論文、あるいは各種の研究レポート、報告書のたぐい、特許文書、また化合物の[クラス]、部分構造などに関して調査を行っています。また物質データ、さらには経済的な側面からの情報、出典、そして製品に関するインフォメーション、こういったものも調査します。

経済的な情報というのは、経済にかかわるデータについての情報も検索をするということであり、それとの関連で最終的には化学の機械などの製品についての情報ならびにどこからそれを入手できるかなどの情報も提供します。

FIZ Chemieは、産業界ないしは研究者のために、世界中の主要な化学に関連するデータベースの検索をします。遡及的な検索も行いますし、またアボナマーの形での検索、これはプロフィールサービスといった形ですが、こうしたことも行っています。

このほかにFIZ Chemieは、FIZ Chemieの持っているデータベースへのオンラインの直接の接続、そして社内用のデータベースの構築、またすべての情報あるいはドキュメンテーションにかかわる問題についての情報を提供したり、案内を行ったり、文献を入手する際の援助を行ったり、

コンピュータ化を行うに際してのプログラムないしはプログラミングに際しての援助を行ったり、コンピュータ化において生じる様々な問題に対して助言を行ったりしています。もう一つ、ユーザのための研修会も、重点を置いて行っています。

FIZ Chemie が持っているデータベースの主要な分野ですが、まず最初に有機反応についての分野が挙げられ、ChemInform というデータベースを持っています。これは『化学中央誌』を引き継ぐものと見ていただいて結構だと思います。「ChemInform RX」というデータベースもありますが、これも反応のデータベースでして、社内データベースとしても利用していただけますし、同時に STN でオンラインの形での利用も可能です。

FIZ Chemie がもう一つ重点として見ている分野で、データベースの数も多いのはケミカルエンジニアの分野です。FIZ Chemie は、過去50年間にわたる東部ドイツ、旧東ドイツの化学産業のさまざまな測定データを把握しています。このような形ができたために、データベースとしては世界中のデータを把握した形で市場に提供することができるようになりました。

ポリマー化学の分野においては、三つのデータベースがあります。ポリマーの分野においてはドイツ合成樹脂研究所との協力によってデータベースを作成しています。KKF というデータベースがありますが、最初の K が合成樹脂、次の K がカウツェック、生ゴム、そして F はファーザー、繊維ということで、それを三つつなげたものです。その次にポリマーに関するデータベース POLYMAT, POLYMAT PC というデータを CD-ROM の形で提供しているものもあります。

このほかにも各種のデータベースがあります。例えば、バイオテクノロジー、分析の分野、安全技術の分野、労働医学の分野、潤滑学の分野、そして化学経済、化学産業に関係するデータベースなどがあります。

このような活動を FIZ Chemie ではおよそ100名の従業員で行っています。それによって達成される年間の総売り上げが2000万ドイツマルクであ

り、この売上高から見ると、ドイツ国内でも化学の情報分野で最大のセンターといえます。

ケミカルエンジニアリングデータをさらに充実

次に FIZ Chemie のエンジニアリング関係の部分と、ChemInform について詳しいご説明をしたいと思います。

FIZ Chemie がつくっている化学エンジニアリングに関するデータの内容は、まず分析に関する内容、そして経済性について、エンジニアリング自体に関する内容、また性質や安全性についての内容ですが、機械、器具の観点からの、方法の点からの、物質の点からの、データを取載しています。これらのデータがオンラインまたはインハウスの形で提供されています。

近代的なケミカルエンジニアリングのプロセスシミュレーションは、非常に熱物理学的あるいは熱力学的な性質のデータを必要としており、こうした目的に用いられるデータは、極めて質の高いものでなければいけません。このように極めて質の高いデータを保証するために、単にデータ一つ一つの品質チェックを行ってその高度な質を確保するにとどまらず、取載されたすべてのレコードに関して、実際に産業の中において機械デザインなどの現場で実証を行い、有益であると実証されたデータに基づいた情報を取載しています。それによって極めて高い質が維持されているのです。

我々の取載しているすべての物質は、産業界によって重要な物質、およびそれらの化合物、あるいはそれらを分解して得られるもの、すべてを含んでいます。産業界にとって関心のある物質は、現在のところ幅が非常に広くて、温度の指標でいえばゼロから1000ケルビン、あるいは圧力でいえばゼロから35メガパスカルの分野に広がっています。

主要な分野で把握された中の2000近くもの個々の性質を、データに把握しています。

多種多様なポイントのデータが含まれていますが、例えば可溶性だとか、運搬時の移動のときの

性格あるいは活性、その他相と相との間の平衡がどの程度か、こういった平衡を保つかという点についてのデータが含まれています。

今後ともこのデータベースの内容をさらに拡充する方向で、他のパートナーとの協力を行っていますが、特に DECHEMA との協力、また [グメーリン] 教授のいらっしゃる [オレンブルク] 大学との協力を続けてまいります。現在不足しているデータについては、[オレンブルク] 大学において、具体的な実験によって数値データを得、これを情報の中に補足していったり、あるいは他の文献から現在不足しているデータを発見し、それを収載してまいります。

反応データベース「ChemInform」

FIZ Chemie の活動の主要な分野の一つに有機反応分野があり、『ChemInform』を別の形で継続的に発行していく活動をしています。

『ChemInform』では一つ一つの部分的な反応だけを記載するのではなくて、文献の中に出ている最終的な段階になるまでの総合的な反応を掲載することに注意が向けられています。

アブストラクトならびに反応経路の掲載については、極めて厳密な選抜が行われています。掲載されるのは、本当の意味での新しい効果をもたらす新しい反応だけであり、厳しい選択の結果、全体の情報量は、ある程度に抑えられています。特に産業界との関係を密接に保つためにバイエル社との協力を行っており、『ChemInform』の中にはバイエル社のデータも把握されています。

『ChemInform』では全体の反応経路が、オンラインの形で、また社内データベースの形、どちらの形でも検索ができるようになっています。そのような形態のために、ChemInform の RX は、それまで市場にあった反応データベースとは性質を異にするものです。一つの物質から出発して次の物質に至る反応をワンウエーで書いてあり、そしてでき上がった産物からまた次の産物に至る反応を次々と示しています。したがって、すべての部

分反応に至るまでの総合的な反応の経路全体を『ChemInform』は含んでいることになるのです。

Gmelin Institute の活動

次に、Gmelin Institute との関連についてお話しいたします。

『Gmelin ハンドブック』は化学の無機化学と有機金属化合物についての情報を含んだ最大のハンドブックです。すでに Gmelin で160点以上もの有機金属化合物についての本が出版されていることをご存じの方も多いかと思いますが。その意味では、有機金属化合物の分野で Gmelin は、世界最大の情報機関だといえます。

先ほど申しましたように、『Gmelin ハンドブック』は、はじめ、1817年に3巻の形で発行されましたが、1924年以来、第8版の作成が進んでいます。現在のところでは既に第700巻目を超えて、700以上もの本が出ており、トータルのページは30万ページにもなっています。

こういう大変規模の大きなハンドブックですので、これを検索するためのファシリティーも充実しています。アルファベット順での検索もでき、完全なカタログもありますし「Gmelin Formula Index」を使って、STN を経由してオンラインで検索をすることができます。

ただ、Gmelin の仕事はこれだけではありません。最も新しい Gmelin の商品、提供する情報として Gmelin のファクトデータベースがあります。

ドイツの連邦研究技術省の助成を得て、1991年からこの Gmelin ファクトデータベースは、STN によってオンラインで検索をすることができるようになりました。現在、Gmelin ファクトデータベースの中には、無機および有機金属化合物の総数にして80万件以上ものデータが蓄積されており、それぞれの物質の400もの性質に関して記載されています。また32万もの有機金属化合物がその3次元的な構造のデータも含んで収載されています。そして20以上もの原子配位までを記録することができ、そのことによって部分構造をも検索で

きるようになっている唯一のデータベースです。

Gmelin の新しい方向性は、データベースとハンドブックとを結びつけていくことであり、その中で、フルテキスト情報も得られるようになっていきます。このような形によって、Gmelin の中に含まれていたさまざまな情報が、最終的に統合されたシステムになることが可能だったのです。

さらに Gmelin の製品をつくりあげていくために、GECI と呼ばれているグメーリン・エレクトロニック・カード・インデックスというものが提供されています。この GECI を使うと、いくつものデータベースを好きな数だけ同時に検索していくことができます。例えば Chemical Abstracts, Beilstein, あるいは Gmelin を全部一緒に検索することができるのです。ここで検索をした結果を直ちにドキュメンテーションシステムに持つていくことができます。これはコンピュータの専門家の方に申し上げればわかっていただけたと思いますが、SGML ベースでのプロセスです。

Gmelin のデータベースならびにハンドブックのもとになっているのは、すべて原論文です。Gmelin のために論文を書く著者には、直接 EHE システムに入力をしてもらいます。そしてその EHE システムに入力をしてもらったものが、さらに次の過程に進んでいきます。

同じ論文から、データベース用のインプットが行われます。このデータベースへのインプットは、たくさんの非常勤の人をお願いをしてやっています。この Gmelin のデータの全体のシステムの中には、データベースとハンドブックがインテグレートされた形で入っています。

オンラインのためのデータベースは、今では社内用のデータベースとしても使用可能な形で提供されています。このようなシステムを使って、フルテキストの CD-ROM 製品をつくっていくことが可能になります。それぞれの課題に応じた必要な各物質に関する情報を、フルテキストでユーザに提供することになります。

Beilstein での新しい動き

次に紹介します機関は皆様方もよくご存じの Beilstein でございます。1881年に初めての『Beilstein ハンドブック』が刊行されましたが、そのときに含まれていた化合物の数が1500, そしてページ数が2200ページでした。1885年の第2版では4080ページ、1906年の第3版では11000ページと版を重ねるごとに急速にページ数がふえていったのですが、この時期には有機化学の分野での発展が非常に目まじしかったわけです。

1984年には、ページ総数が38万ページ、460巻にも達しており、このときに初めて英語版も出版されました。1988年には、Beilstein が STN のオンライン上に収載されましたが、このプロジェクトに関しても連邦の研究技術省から大きな助成を受けています。1989年以降は、DIALOG からのオンラインでも検索することができるようになっています。現在 Beilstein の情報についてオンラインで検索できる対象は600万ほどの有機化合物で、それぞれの性質などを記載したデータを検索できるようになっています。

Beilstein での新しい動きは、これは今年のことですが、Cross Fire という Beilstein のファイルによって、600万近くの各種の構造を社内用のデータベースとして提供していくことです。Beilstein の Cross Fire というのは、PC をフロントエンドに使っていただけたという形での提供です。IBM の AXP を利用してこうした検索をしていただけます。このような形により、600万もの各種の物質構造の検索を、すばやくデスク上から行うことができるようになります。

収載されている600万件のデータは、それぞれの物質の構造、クロスレファレンスの指標、CAS の登録番号、物質を同定するためのデータ、およびこのデータファイルにそれ以外のより多くの情報が含まれているかについての情報などです。このデータを使えば、この検索後にさらにファクトデータベースの検索を行う場合、自動的に STN の検索につながっていく可能性が出てくるので

す。Beilstein の Cross Fire は、STN の ISIS のインタフェースを持っています。

もう一つ、Beilstein の提供している情報には、CD-ROM の形でのカレントファクトがあって、この中に含まれる情報は、100近くの専門雑誌の中に出ている有機化学または薬学の分野における最新の情報です。この情報の中には過去12か月の完全なデータが収載されていますので、その後、書誌データベースで必要な検索を続けることができますようになります。

このシステムのメリットは、非常に価格が安いことで、PC 1台当たりの価格は880マルクです。

さらにもう一つ Beilstein の提供する情報に、AUTONOM があります。AUTONOM の持っている特徴は、図形からコンピュータの上で IUPAC に適合された名前を、自動的につくり上げることができることです。これは特に環形の化合物で命名が非常に難しい場合など、大変役に立ちます。図形の形で示された化合物について、85%まで正確な命名をいたします。これによって、合成化学の分野で仕事をされていて、新しい物質、化合物をつくった場合など、その命名の際に非常に大きな助けとなります。このシステムは、PC あるいは [デックス] ワークステーション上で使用可能な形で提供されます。

以上ドイツにおける化学情報について、STN 以外の分野でこういったことがなされているかを少しでもおわかりいただけたかと思います。ご清聴ありがとうございました。

Q & A

Q FIZ Chemie が作成している化学エンジニアリングに関するデータの内容で、経済性ということがでてきましたが、それは物質の経済性なども評価するという意味でしょうか。もし、それが評価されるのであれば、どのような方法で評価されているのか、あるいはそういう内容の文献などが入っているのでしょうか。

A 物質の経済性に関するデータということで

すが、経済性の評価については、工業化学の分野で活躍をしている方をお願いして、例えばどの程度産業界に具体的な関心がありそうな物質であるかということと同時に、その物質があまりにエネルギーをたくさん使い過ぎるので経済的に成り立たないとか、価格、コスト、用途がどんな状態であるかといった情報を、盛り込んでいます。

Q コストということは、コスト・ベネフィットという意味も含めてのことでしょうか。

A コスト・ベネフィットということであれば、特定の産業のエンジニアリング装置の中のプロセスで使う特定の目的に対して、どのような反応で最も適切な物質をつくることができるか、そういった意味で言えるかと思います。

Q Beilstein の Cross Fire のことですが、どういう情報の提供の仕方をされるのか、CD-ROM の形で行われるのでしょうか。

A パワーステーションからの検索ということで、CD-ROM の形ではなくテープです。

Q ただ、かなりしんどいと思いますが、構造検索でやってスピードは速いのでしょうか。

A 数秒での検索が可能です。非常に検索のアルゴリズムが速いんです。この磁気テープが数ギガバイトにも及ぶハードディスクの上で作業が進む——実際には毎回、磁気テープ上で、探していくのではなくて、磁気テープと数ギガバイトに及ぶハードディスクとの間のインタフェースをつくって、選択のときにはハードディスク上で行う形になるので、非常に検索は速いと思います。

A (補足) Cross Fire というのは、要するにフロントエンドのエミュレータです。STN エクスプレスの類似品とお考えいただければいいと思いますが、ただ、パワーステーション上で構造検索ができるようになっている。つまり提供する媒体が磁気テープであるというだけのことです。それで構造検索しかできません——「しか」というのはちょっと言い過ぎですが、構造検索ができて、それで実際のデータを欲しいときには、STN につないで取り出す形になります。

(文責 本誌編集委員会)