

## 化学物質データベース (WebKis-Plus) の紹介\*

白石 寛明\*\* · 西川 希\*\*

キーワード ①化学物質 ②データベース ③農薬 ④分析法 ⑤環境動態モデル ⑥生態影響

### 要 旨

化学物質データベース(以下「WebKis-Plus」と略す)では、複数のデータベースから収集した化学物質情報(一般、物性、毒性情報等)や環境分析手法や環境動態のモデルに関する情報を提供している。それらはできるだけCAS番号によって化学物質ごとに整理しており、データベース間の相互の関連づけを容易にすることで、その維持管理にかかわる労力の低減や効率化に考慮している。現在公開しているデータベースは①化学物質データベース(Kis-Plus)、②農薬データベース、③AQUIRE、④化学物質環境動態モデルデータベースおよび⑤環境測定法データベース(EnvMethod)である。さらにより多くの情報を掲載し、リスク評価のツールとして機能を充実させることを目的とし、新たなデータベースの作成や情報検索ページの追加、既存のデータベースのデータの更新作業を行っている。

### 1. はじめに

リスクコミュニケーション手法の1つとしてインターネットを介した化学物質情報の提供があげられる。WebKis-Plusでは、複数のデータベースから収集した化学物質情報(一般、物性、毒性情報等)や環境分析手法や環境動態のモデルに関する情報を提供している。それらはできるだけCAS番号によって化学物質ごとに整理しており、データベース間の相互の関連づけを容易にすることで、その維持管理にかかわる労力の低減や効率化に考慮している。現在公開しているデータベースはより多くの情報を掲載し、リスク評価のツールとして機能を充実させることを目的とし、新たなデータベースの作成や情報検索ページの追加、既

存のデータベースのデータの更新作業を行っている。

環境中に存在している化学物質によるリスクを算定するためには、生産・使用量、物理化学的性状、生態影響など多様なデータが必要となる。化学物質のデータベースは、多くの人がその必要性を感じながらも、一般の人や事業者が自由に利用できるデータベースはそれほど多くはない。一方で、インターネットの普及は急速に進んでおり、情報検索サイトを用いるとあたかも1つのデータベースのように化合物名、CAS番号などさまざまなキーワードでの検索が可能となっており、デスクトップコンピュータ上で利用できるデータベース環境の発達も目覚ましい。WebKis-Plusは、

\* Introduction of Chemical Substance Database (WebKis-Plus)

\*\*Hiroaki SHIRAIISHI Nozomi NISHIKAWA (国立環境研究所化学物質環境リスクセンター曝露評価研究室) National Institute for Environmental Studies Research Center for Environmental Risk Exposure Assessment Section

化学物質の安全性に興味のある人達の間でデータを共有しようという意図から、既存のデータベースやデータセットを骨格に、逐次データを追加整備するという方針で開発されている。

本システムは、科学技術振興調整費による「物質関連データ(化学物質, 食品成分, 表面分析)のデータベース化に関する調査研究(第1期 平成6～8年度)」の中の1課題として、さまざまな形態のデータ・情報を関連づけて統合的に利用する必要のある化学物質の生体影響データベースの調査研究と、その整備として神奈川県と兵庫県の研究機関の協力の下に始められたものである。本システムは7年度の後半より開発が始まり、9年から2年間の第Ⅱ期研究で拡張しつつ、その後は維持を主体に現在に至った。今後は、国立環境研究所・化学物質環境リスクセンターによりバックアップされることになっている。

## 2. データベースの詳細

一般に正規化されたデータベースは、独自のコードで化学物質を特定している場合が多いため、データベース間の化合物の関連づけをするのは困難である。唯一可能でもっとも手っ取り早い方法は、ケミカルアブストラクト番号(CAS番号)を共通のインデックス項目として互いに参照することである。CAS番号には、番号が設定されていない物質や異なる番号をいくつか持つなどの問題があり、完全な統合は困難であるが、とりあえずCAS番号で関連をとり、連携を深めていく中でCAS番号の統一を図ることがもっとも効率的と思われる。

データベースの開発でもっとも考慮しなければならないことは継続性であり、利用するシステムは世の中の大きな流れに逆らわないことが肝要である。Kis-Plusをはじめ現在公開しているデータベース(表1)のすべてが、MicrosoftのACCESSデータベースで作成されており、Windows NTに付属するインターネットインフォメーションサーバー(IIS)とODBC(Open Database Connectivity)ドライバーにより、インターネット上で公開されている。研究所のセキュリティ担当からは、Windowsはセキュリティ上問題が多いとの指摘を受けており、この選択には一言のある人もおら

表1 Webkis-plus 公開データベース

データベース	説明
化学物質データベース (Kis-plus)	KIS-NETを主体に、さまざまな文献情報、環境省測定結果等からKow, 水溶解度, 許容摂取量, 発がんリスク, 毒性値等のデータを統合したデータベース
農薬データベース	兵庫県立健康環境科学研究所で作成された農薬の構造式データベースに、農薬の県別出荷量, 毒性情報等を登録した出荷量データベースを統合したデータベース
AQUIRE データベース	米国EPA/ORDとNHEERL/MEDが作成し、無償で提供している、1915年から現在までの文献調査により得た水生生物毒性試験情報のデータベース
化学物質環境動態モデルデータベース	化学物質の環境動態予測モデルの情報を複数の機関から収集し、またNTISによる文献検索結果を利用して作成したモデル情報データベース
環境測定法データベース (EnvMethod)	公定法や環境省公表資料等に示されている分析法を集約し、化学物質名等からの検索を可能にしたデータベース

れるかもしれない。しかし、簡単にできるだけ手間をかけずデータベースを作成し、システムを維持していくという目的には、少なくともいままでは最適な選択であったと考えている。

最近では、LinuxでもオープンソースのデータベースとPHPのようなスクリプト言語で同様のシステムが構築できるようになっており、有料のデータベースサーバーを用いなくとも、ビジネスソフトの一部あるいは無料のデータベースソフトとインターネットのブラウザだけで簡単に動的なページを扱えるようになってきている。また、普段、誰でも利用しているエクセルのような表計算ソフトでも動的なページを作ることが可能であり、自宅のパソコンからでも発信が可能な時代である。ちなみに、WebKis-Plusはパソコン(現在、Slot Iのペンティアム III 700MHz)で動作している質素なシステムであり、アクセスの多いときは制限をかけてご迷惑をおかけしている。現在、システムや維持・管理体制の移行時期に差しかかっているように見え、将来を見据え対応したい。

データベースが専門の技術者はSQL(Structured Query Language)のコマンドベースのデータベースを好む傾向にあると思われるが、銀行のATM

のようなデータの更新が頻繁には起こらない、検索や閲覧が主体の化学物質のデータベースでは開発のスピード、保守管理の容易さ、設計の柔軟さからみると、はじめから大掛かりなシステムを構築するのは効率的でないように思える。貴重な時間は、身の丈にあったシステムで、中味の充実に割いたほうがより有効であろう。

## 2.1 Kis-Plus

Kis-plus は、KIS-NET を主体に、NORDBAS<sup>1)</sup> データベース、約5,000データの水・オクタノール分配係数(Kow)<sup>2,3)</sup>、環境庁環境安全課が測定している約200物質の水溶解度とKow、IRIS(Integrated Risk Information System, 米国EPA)<sup>4)</sup>から約600物質の許容摂取量(Reference Dose: RfD)や物質 mg/kg day 当たりの生涯にわたる発がんリスクの上限の傾き(Cancer Potency Slope Factor: CPS)の値、約1,300物質の簡便な水質の毒性試験法として知られる発光細菌を用いたマイクロトックス試験<sup>5)</sup>による毒性値などを統合したACCESS データベース(Kis-Nies.mdb:KIS-NET データを元にデータを追加して作成されたデータベースの意)である。

データベース間での化合物の関連づけには、CAS 番号を用いている。これは本システムの長所でもある。CAS 番号が付与されていない化学物質に欠落が生じるという欠点ともなっている。データベース内の全化合物のリストを作成した後、CAS 番号と各データベース固有の化合物番号とをクロス集計することでCAS-個別管理番号の対応表(CAS\_pointer)を作成してある。このKis-Nies.mdb に検索フォーム、クエリーなどを完備し、デスクトップ上で動作するデータベースとして完成し、いくつかの機関にデータの共有という目的のため配布したが、その後、積極的な活動も提供先からの応答もないまま現在に至っている。このような状況を教訓にすると、データベースはファイル自体を提供するのではなくデータベースの管理は一元化し、インターネットを通じて提供する方式が現在の通信環境には向いているように思われる。

いくつかのファイルをKis-Nies.mdb という1つのデータベースに統合したことで、各ファイルのリレーションづけなどのデータベース構造の変

更が容易となった。エクセルなどの表計算ソフトとの相性もよく、またODBCドライバーなどによりネット上の他のデータベースやデータファイルへの参照も可能であるため、保守管理がしやすい長所が生まれた。このKis-PlusをIISにのせ、Active Server PageによりWEB上に公開するようにaspコードを記載したシステムがWebKis-Plusである。平成9年11月より稼動しており、基本的な構成は現在も変わっていない。その後、パソコン通信であったKIS-NETもインターネット版として整備され、公開されている。

Kis-PlusのCAS-物質の対応表(CAS\_pointer)に対して、/kis-plus/cas\_link.asp?CAS=xxx-xx-x(xxx-xx-xはCAS番号)により指示を与えれば、データベースの内容を参照できる各種の番号を入手できる。大気や水質の環境基準など各種の法規制、地方自治体の化学物質の指針値、発がん性の評価、国際機関や外国の物質リスト、内分泌攪乱作用の疑われる物質、PRTR(環境汚染物質排出移動登録)の特定化学物質などの物質リストも物質をcas\_link.asp?CAS=xxx-xx-xの形式でハイパーリンクを持った表により提示することで、各種の情報を参照できるようになる。ハイパーリンク付の化合物リストはエクセルのマクロなどで容易に作成できるので、リンクを作成する手間はほとんどかからない。

このようにして、その時々で注目される物質を表形式で作成すれば、ほとんど手をかけずにインターネット上でさまざまな情報を閲覧できるようなサイトにすることができる。WebKis-Plusでは、画面の左側にこのような表を載せてある。したがって、これらの表は必ずしもデータベース化されているのではない。手抜きともいえるかも知れないが、当面の処置が必要な場合は利用者の便から考えると十分目的は達せられる。

インターネット上にはCAS番号などによりリンクできるように設計されたホームページがUSEPAのIRIS, the NIST Chemistry WebBook<sup>6)</sup>(米国通商部技術局の全米標準・技術研究所)やThe National Toxicology Program<sup>7)</sup>(the U.S. Department of Health and Human Services)などのようにいくつか存在する。これらのサイトは、他サイトからの直接リンクを考慮しているわけではないので、

サイトのアドレスが頻繁に変わることが多い。

WebKis-Plus では、/kis-plus/caskis-plus.asp で CAS\_pointer にある全 CAS 番号を公開している。この情報を用いれば、外部から直接リンクを張ることが可能となる。この方法は検索サイトからの「攻撃」を受けやすいという欠点もある。WebKis-Plus では意識的にリクエストボタンでなく CAS によるリンクを記載した表を多く載せており、アドレスも変更せず保っているが、今のところ致命的なトラブルはない。ただし、検索ロボットなどからのアクセスが多くなる時に応答が遅くなる場合があり、本来の利用者が利用できない局面がある。検索サイトがアクセスできる形ですべてを完全に公開し、民間企業に貢献することが果たしてよいのかどうか疑問でもあり、検索サイトや検索ソフトとの付き合い方は考え直す必要が生じるかもしれない。

## 2.2 農薬データベース

WebKis-Plus には、兵庫県立健康環境科学研究所センターで作成された Microsoft ワードの図として、約3,000の農薬としての薬効がある物質の構造式を図示した農薬の構造式データベースがある。これに農薬の県別の出荷量、毒性情報などを登録した出荷量データベースを統合した「農薬データベース」(Noyaku.mdb)がある。農薬データベースでは、原体名(約3,300件)、農薬名(約2,600件)、商品名(約6,200件)、IUPAC 名、CAS 番号などから検索でき、WebKis-Plus 内で CAS 番号により Kis-Plus と統合されている。

農薬の都道府県別の出荷量は、1992年農薬年度以降の農薬要覧が最新版まで入力されている。環境への負荷量の推定には農薬製剤の出荷量でなく、有効成分である原体の量が必要であるため、農薬製剤の出荷量と対応する農薬製剤中の原体の含有量から、原体ごとの都道府県別の出荷量を計算し表示するようになっている。また、農薬の名称や商品名から使用している原体の含有量や原体からそれを含む農薬の名称と含有量、商品名、物性情報などを閲覧することができる。さらに、農薬の空間的・時間的な変動を考慮したばく露評価の基礎とするため、92年以前のデータの入力を昭和38年度版より行っており、15年分の出荷量の入力がほぼ完了している。今後、原体への換算に必要な

製剤中の含有量のデータベース化が必要である。

また、リスクコミュニケーションの観点からは、商品名からの検索の充実も必要と思われるが、研究的な要素が少なく手間ばかりかかるため、入力してくれるボランティアが現われないかぎり当面このままの状態であろう。

本来、情報の発信は情報作成者が行うのがもっとも適切であると思うが、農薬などの情報が公的機関から一般市民に発信されることは少なかった。このため、農薬データベースも作らざるを得なかったのだが、農薬の違法販売の問題を機に、農薬の情報の公開へ方向性が一段と強まっているように見える。商品名に関しては独立行政法人農薬検査所の HP<sup>8)</sup>が参考になる。

## 2.3 AQUIRE データベース

AQUIRE データベースは、1915年から現在までの文献調査により得た水生生物毒性試験情報のデータベースであり、米国 EPA の研究開発部門 (Office of Research and Development, 以下「ORD」と略す)と国立健康環境影響研究所 (National Health and Environmental Effects Research Laboratory, 以下「NHEERL」と略す)の大陸環境部門 (Mid-Continent Ecology Division, 以下「MED」と略す)が作成し、無償で提供されている。このデータベースでは、毒性試験ごと、試験化学物質、生物種、媒体、実施場所、曝露期間、エンドポイント、参考文献等の情報をまとめており、そのデータ数は毒性試験として22万3,853件、収載されている化学物質は7,304物質、生物種は4,191種に及んでいる(2003年5月現在)。

AQUIRE データベースは上記のように豊富な水生生物毒性試験情報を有しており、化学物質の水生生物への影響を予測する上で非常に強力なツールと成り得る。しかし、インターネットを介した情報検索ページでは、毒性データの統計処理などの一括処理が困難である。そこで、データベースとしてより簡単でユーザーが使いやすいものを提供することを目的として AQUIRE データベースの MS ACCESS 2000(データベース管理システム)への移植を行い、EPA の許可を得て、この一部を活用して情報検索ページの作成を行っている。

AQUIRE データベースのデータセットは ASC II ファイル形式でホームページ<sup>9)</sup>より入手可能であ

る。これを MS ACCESS 2000 に移植した。ローカルなデータベースとすることで、データの追加、変更、検索などが自由に行えるようになった。AQUIRE データベースは前述のように非常にデータ量が豊富であるため、複雑な検索は Web サーバーに多大な負荷をかけることが予想されたため、検索条件は化学物質、生物種の 2 つとし、まず化学物質から検索し、その後、生物種で絞り込むという形式をとった。また、既存の Webkis-plus と連動させるため、化学物質の CAS 番号で化学物質データベースと連結している。

ユーザーは、まず化学物質の CAS 番号、名称、分子式のいずれかを入力・検索する。結果は化学物質一覧として表示され、そこから化学物質を選択すると、AQUIRE データベースの毒性試験情報の一覧が表示される。さらに一覧から毒性試験を選択すると、詳細が表示されるというしくみになっている。また、化学物質一覧では、さらに絞り込み検索をかけられるようにし、毒性試験情報一覧では対象生物種、媒体水質から絞り込み検索をかけられるようになっている。

こうして作成された AQUIRE データベースは、水生生物への化学物質の毒性を体系的に集めたものであり、インターネット上で簡単に情報検索を行うことができるツールとして、リスク評価の際の助力となると考えられるが、毒性データの信頼性の評価など不十分な点も見られ、日本独自のデータの追加を含め内容の充実を図っていく予定である。

日本でも化学物質の生態影響に関するリスクを考慮しなければならないという認識が持たれ、化審法においても生態影響に関するリスクを事前に審査する必要性が生じている。本データベースは生態毒性に関するもっとも大きなデータベースであり、今後、力を入れていかなければならない分野であると認識している。本データベースには構造式が付与されており、デスクトップ上では化学物質の類似性からの検索ができるようになっている。現在、この機能を用いて構造分類を行い、化学物質の生物に対する毒性の定量的予測(QSAR)ができるよう解析をしている。

#### 2.4 化学物質環境動態モデルデータベース情報検索ページ

このページは環境省受託業務「新規化学物質挙動追跡調査」業務の成果を活用して国立環境研究所化学物質環境リスク研究センターで作成されたものである。化学物質のリスク評価の 1 つとして、コンピューターモデリングシステム(モデル)を利用した環境動態の予測は有用であるが、世界中にはさまざまなモデルが点在しており、その中から適切なモデルを選択することは困難であった。そこで、このようなモデルを提供している複数の機関から情報を収集し、データベースの作成、インターネット上での情報公開、情報検索ページの作成を行った。

「化学物質環境動態モデルデータベース」(モデル情報.mdb)を作成するため、複数の機関からモデル情報を収集した。情報元として一般的に利用されるモデルを数多く提供している機関からの情報および NTIS による文献検索結果を利用した。また、モデル情報を系統立てて整理するため、インデックスとして名称、略名、オペレーティングシステム、記述コード、対象媒体、概要、関連リンク、参考文献等を設けた。収集した情報をインターネットでの情報公開を考慮してデータベース化した。Web ページのデザインを行い、このデータベースの情報をインターネット上で公開した。検索条件として提供機関、名称、対称媒体を選択し、情報公開画面からモデルのホームページへリンクできる形をとった。検索条件に合致した複数のモデル情報の一覧からモデル情報を選択して表示させることができる。化学物質環境動態モデルデータベース情報検索ページは、化学物質の環境中での動態を予測するために適切なモデルを選択する際の手助けとなり、リスクコミュニケーションの推進につながると考えられる。

#### 2.5 環境測定法データベース(EnvMethod)

環境測定法データベースはこれまで、農薬の分析法を中心にテキストとして保存されている分析法を対象に、インデックスサーバーによる全文検索と分析法のリストからの参照を行うものであったが、今年度に大幅に機能を拡充し環境測定法データベース(EnvMethod)として再構築された。

EnvMethod は、化学物質の分析法開発や環境分析に携わる方へ情報を提供することを目的として、公定法や環境省公表資料等に示されている分

析法を集約し、化学物質名等からの検索を可能にしたものであり、国立環境研究所の2号業務（情報提供）の機能を充実するためのモデルとして試行されたもので、環境省の協力を得て国立環境研究所環境情報センターと共同して作成しているものである。

2003年7月1日現在では、EnvMethodに収載されている分析法は1,694件、対象化学物質は1,712物質であり、公定分析法、環境省公表資料および米国EPA公表資料等が含まれている（表2）。

### 3. 今後の拡張

公開している化学物質データベースは神奈川県KIS-NETをベースにしており、化学物質の種類も限定されていた。化学物質の審査などに当たっては、これをより体系化し、すべての既存化学物質や審査済みの新規化学物質のデータベース化が必要である。このため約3万7,500物質の名称、官報告示番号、CAS番号などを入力し基盤整備を行っている。

最近、構造式から視覚的に物質を特定すること

ができるサイトが増加している。2次元の構造式を利用した大きな構造式データベースを持つサイトとして、CambridgeSoft Corporationが提供しているChemFinderがある<sup>10)</sup>。より詳細な情報が得られる有料のChemIndexもあるが、ChemFinderでも一般的な物性情報や個別情報のリンク先が提供されており、多くの情報が得られる。

1次元で化学構造を表記する形式にSMILESがある。1次元表記であるためデータベース化しやすい利点があり、各種のフォーマットからbabel（フリーソフト）、ChemDrawなどの変換ソフトや分子エディターを利用すれば2次元の構造式に変換することができる。NCIから、約13万物質のCAS番号や約4万物質の名称とSMILES形式で記載された構造式のファイルを入手できる<sup>11)</sup>。WebKis-Plusも部分構造式から検索できるサイトへ脱皮を図る時期にきていると思われ、構造式の入力など少しずつ準備をしている。ただし、インターネット上での構造式検索は、パフォーマンスがあまり芳しくない傾向があり、どのような形で運用するかは未定である。

表2 EnvMethod 収載分析法

公定分析法
環境基準関係（過去の分析法も一部収載）
指定物質抑制基準
排水基準検定方法
土壌汚染対策法関連測定方法
農用地土壌汚染対策地域の指定要件に係る検定方法
農薬登録保留基準
食品、添加物等の規格基準（過去の分析法も収載）
環境省公表資料
有害大気汚染物質測定方法マニュアル
排出ガス中の指定物質の測定方法マニュアル
外因性内分泌攪乱化学物質調査暫定マニュアル
要調査項目等調査マニュアル（1999～2001）
底質調査方法
ダイオキシン類に係る大気環境調査マニュアル
ダイオキシン類に係る底質調査測定マニュアル
ダイオキシン類に係る土壌調査測定マニュアル
埋設農薬調査・掘削等暫定マニュアル
化学物質分析法開発調査報告書（1974～2001）

### —参考文献—

- 1) Nordic Council of Ministers, 'Environmental Hazard Classification-classification of selected substances as dangerous for environment(1)', TemaNord 1994: 643, Copenhagen 1994
- 2) Hancsh C. and Leo A., 'Expoloring QSSAR', ACS, Washington, DC, 1995
- 3) Leo A., Hancxh C., Elkins D., Partition coefficients and their uses, Chem. Rev., **71**, 525-615, 1971
- 4) IRIS(Integrated Risk Information System) <http://www.epa.gov/iris/index.html>
- 5) Kaiser, K.L.E., Palabrica, V.S., Photobacterium phosphoreum Toxicity Data Index, Water Poll. Res. J. Canada, **26**, 361-431, 1991
- 6) the NIST Chemistry WebBook: <http://webbook.nist.gov/chemistry/>
- 7) The National Toxicology Program (NTP), the U.S. Department of Health and Human Services: <http://ntp-server.niehs.nih.gov/default.html>
- 8) 独立行政法人農薬検査所: <http://www.acis.go.jp/>
- 9) ECOTOX Database, U.S. Environmental Protection Agency (U.S.EPA): <http://www.epa.gov/ecotox/>
- 10) ChemFinder, <http://chemfinder.cambridgesoft.com/>
- 11) National Cancer Institute (NCI): <http://www.nci.nih.gov/>