

## 9. 研究者のアクセス手法 II

筑波大学附属図書館長 中山伸一

### はじめに

私は現在の自分の専門分野を「応用情報学」といっております。大学の専門は有機合成化学ですが、タンパク質を対象に電子スピン共鳴装置を使ってスピンドラベル法という手法を用い、タンパク質の高次構造の研究を行ってまいりました。その後、化学の研究所でフッ素化合物の合成研究を一年間行い、図書館情報メディア系の前身である図書館情報大学に勤めました。そして、その後ウェットな研究からドライな研究への方向転換を行い、現在は化学の研究も続けながら図書館情報学や認知科学的な研究を行っております。

指導を行っている(いた)大学院生や学部生の研究テーマは、「三次元構造にもとづくタンパク質構造の類似度」、「化学構造表現におけるビットストリングの方法の比較」、「感情による図書検索システム」、「性格と感情の個人差に基づく図書検索システムの可能性の検討」、「創造性測定と思考動機」、「物語映像に享楽する心の構造に関する研究」、「知識表現としての説明の仕方の分類」など多岐にわたっております。これらの領域の全てに関してお話をするには時間が足りませんので、ここでは『情報化学』をキーワードとした研究をとりあげ、その研究の方法とそこでの情報の位置づけを紹介するとともに、大学図書館との関わりを考えてみることにします。

### 情報化学とは

情報化学がどのような研究を行う領域であるのかが分かる、という図書館員の方は少ないと思います。日本化学会の部会の一つに情報化学部会というのがありますが、この部会は化学研究への情報・計算機の広い活用を図ることを目的に 1983 年に設立され、現在の会員数は 400 名程と小さな所帯です。アメリカ化学会にも同様の領域に対応する部会がありますが、こちらは Division of Chemical Information (CINF) と Division of Computer in Chemistry (COMP) の二つに分かれており、それぞれ 1,000 名程と 2,200 名程の会員数です。CINF は化学の情報を(計算機で)どう扱うか、COMP は化学を計算機でどう扱うかということが研究の中心です。情報化学部会はその両者を包含する研究領域を扱っております。以下に、情報化学部会の HP に掲載されている情報化学の研究領域を示します。

- ・ 化学情報学（情報検索、データベース、グラフ理論、反応設計など）
- ・ 理論化学、計算化学（量子化学、分子軌道法、分子力学、分子動力学、各種シミュレーションなど）
- ・ ケモメトリックス（統計、数理科学など）
- ・ 構造活性相関、構造物性相関
- ・ 知識情報学およびそのためのソフトウェア（エキスパートシステム、ニューラルネットワーク、ファジィ、カオスなど）
- ・ コンビナトリアルケミストリー
- ・ 通信・システム（コンピュータネットワーク、並列化、専用機、コンピュータグラフィックスなど）
- ・ ラボラトリーオートメーション
- ・ 上記分野に関する教育

<http://cicsj.chemistry.or.jp/intro.html> (参照 2012-06-08)

なじみの無い単語が連なっていてチンプンカンプンだと思いましたが、授業の中では少し詳しく述べたいと思います。

### 情報化学研究の流れと情報

私が情報化学の研究を行おうとする学生を研究指導する場合、だいたい以下のような大まかな流れで行っています。

研究テーマの設定（問題点の明確化）

関連研究の論文講読

仮説（モデル）の設定

研究テーマ・仮説の新規性のチェック

実験（仮説検証）計画の立案

    実験方法の選択、開発

    分析方法の選択、開発

実験（計算機実験）

実験結果の考察（仮説の妥当性の検証）

研究成果の新規性の再チェック

このように情報化学研究といっても特別な研究の流れがあるわけではなく、一般的な科学研究の流れである仮説検証のプロセスをたどります。ただ、仮説（モデル）を考えた後、それをウエットな実験により実証するのではなく、理

論もしくは計算機実験等により実証していくというのが特徴です。

情報学の研究一般に言える事ですが、情報化学の研究は何らかの領域の研究を支援するようなものが多いようです。例えば、化学情報学における大きなテーマの一つとして化合物の情報検索があります。しかし、化合物の情報検索が社会に直接役立つかというところではなく、薬の開発や材料の開発などという研究を支援するのに利用されるのが一般的です。従って、化合物の情報検索について研究しようという場合、何のために使われるのかという想定をする必要があります。それによって検索システムのモデルが変わってきます。その意味で、情報化学研究においては、情報化学そのものに関する問題点の把握だけではなく、その応用領域での問題点の把握というのも必要になる場合が多いのです。

情報学と化学の複合領域である情報化学研究には、化学からこの領域に入った研究者と情報学から入った者がおり、どちらから入ったかにより研究のための情報の扱い方が異なります。研究の始まりである問題発見の段階で考えると、化学からこの領域に入った者は、応用領域である化学の問題発見は自分の専門なので知識があります。しかし、情報学からこの領域に入った者は、専門外の情報を扱わなくてはなりません。後者の問題発見は、化学の専門家からの依頼や、一般誌などを読んだの思いつき（この場合はどうすれば問題解決できるのかを合わせて思いつく場合が多い）がきっかけとなります。そこから明確な研究テーマを設定するまでは、情報化学の専門誌を読む事になります。

研究テーマが設定されると、情報学の知識を使ってどのようにその問題を解くことができるのかを考える仮説（モデル）設定の段階に入ります。もちろん、テーマ設定と仮説設定は綺麗に分離される訳ではなく、並列的に行われる場合も多くあります。仮説設定に必要な情報は、対象とする問題に関する先行研究であり、これはどのような研究領域においても必須です。ただ、情報化学研究の場合、その方法論として情報学や情報工学の方法を援用する機会が多いので、化学からこの領域に入った者は、それらの方法についての情報や知識を得る必要があります。私はタンパク質の高次構造の研究を行っておりますが、大学の同僚（情報工学の専門家）に誘われて情報工学の国際会議に出席し、ニューラルネットワークという新しいパターン認識手法を知って、高次構造予測にそれを適用する研究を行ったことがあります。残念ながら、この研究は他の研究者に先を越されましたが、同時期に私を含めて多くの研究者が同じような研究成果を発表しており、新たな手法の開発に注意することの重要性を実感しました。

実験計画の立案は、どのようにして仮説を検証するかを考える段階です。ここでは仮説に従ったシステム開発やデータ分析の方法を考えるため、プログラミングや分析手法等に関するような情報や知識を必要とします。化学からこの領域に入った者は、そのような情報や知識を新たに得なければなりません。

実験以降の検証プロセスは一般的な科学研究の流れと同じで、情報化学の専門知識が必要となります。

### 大学図書館および大学図書館職員に期待すること

情報検索の端末が研究室に置かれるようになり、図書館職員の情報検索への関与は現在ではほとんど認められなくなっております。それに加えて電子ジャーナルの導入により、図書館の器としての利用は減ってきております。一方、アクセスが多様化したことにより、全体的な図書館機能の利用度は向上していると考えられます。特にインターネットを介して電子的資料を読むという図書館機能を利用する研究者は急激に増大しております。資料の検索は研究者の能力の問題ですが、検索機能そのものの提供や、そこで得られた資料をできるだけ電子的に得られるように、適切で分かりやすい環境を作っていくことは図書館の問題です。端末の上に、いかに使いやすい図書館を構築できるかが図書館に求められている課題なのです。

情報化学のような複合領域の研究を行おうとすると、自分の専門としない領域の資料が大変重要になります。近年多くの複合領域の研究が盛んに行われるようになってきておりますが、その複合の度合いはますます高まり、研究者が研究対象に関わるすべての専門資料を扱うのが困難な状況が現れてきております。図書館職員が資料論の知識を豊富に持って対応してくれる事は、研究者にとって大変ありがたいことです。もちろん、専門分野の研究者の方が、知識を多く持っていることは間違いありません。図書館には、そのような人を紹介するというゲートキーパとしての役割も期待されると考えます。その意味で、図書館職員が研究者との接点を多様に持つ事が必要でしょう。

### おわりに

研究に関わる情報・資料は、日々変化しております。大学図書館と大学図書館職員は、それに常に追随して学ぶことにより、わがままな研究者の研究を支援していただければと思います。

## 研究者のアクセス手法II

筑波大学附属図書館長 中山伸一

## はじめに(研究領域)

- 三次元構造にもとづくタンパク質構造の類似度
- 化学構造表現におけるビットストリングの方法の比較
- 感情による図書検索システム
- 性格と感情の個人差に基づく図書検索システムの可能性の検討
- 創造性測定と思考動機
- 物語映像に享受する心の構造に関する研究
- 知識表現としての説明の仕方の分類

## 情報化学

- 化学情報学
- 理論化学、計算化学
- ケモメトリックス
- 構造活性相関、構造物性相関
- 知識情報学およびそのためのソフトウェア
- コンビナトリアルケミストリー
- 通信・システム
- ラボラトリーオートメーション
- 上記分野に関する教育

## 情報化学関係の学会

- 日本
  - 日本化学会情報化学部会
  - 日本コンピュータ化学会
  - 情報計算法学生物学会
- 米国
  - The American Chemical Society
    - Division of Computers in Chemistry
    - Division of Chemical Information

## 情報化学

- 化学情報学
- 理論化学、計算化学
- ケモメトリックス
- 構造活性相関、構造物性相関
- 知識情報学およびそのためのソフトウェア
- コンビナトリアルケミストリー
- 通信・システム
- ラボラトリーオートメーション
- 上記分野に関する教育

## 化学情報学

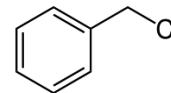
ケモインフォマティ(ツ)クス

- 化学情報をどのように扱うか
  - 情報検索
  - データベース
    - 構造情報
    - ◁完全構造検索、部分構造検索、類似検索▷
  - グラフ理論
- データベースの応用領域
  - 合成経路設計・反応設計
  - 逆合成

### 事例: 化学構造表現における ビットストリングの方法の比較

- 各種ビットストリングの化合物データベースに対する特性
  - ビットストリング(化学構造情報表記法の一つ)
    - 構造要素の有無を1ビット(1/0)で表現
    - ビット列の類似性(Tanimoto係数など)
    - 高速な検索が可能(大規模データベース)
  - 構造要素により複数のビットストリングが提案
    - 特性を比較する

### 事例: ビットストリングの方法の特性



- Bitstring X  
C O N H F Cl Br  
1 0 0 1 0 1 0
- Bitstring Y  
C H 3 C H 2 C H C = C C = N  
0 1 1 1 1 0

### 情報化学

- 化学情報学
- 理論化学、計算化学
- ケモトリックス
- 構造活性相関、構造物性相関
- 知識情報学およびそのためのソフトウェア
- コンビナトリアルケミストリー
- 通信・システム
- ラボラトリーオートメーション
- 上記分野に関する教育

### 計算化学

- 計算機を用いた化学
  - 量子化学
  - 分子軌道法
  - 分子力学
  - 分子動力学
  - 各種シミュレーション

### 情報化学

- 化学情報学
- 理論化学、計算化学
- ケモトリックス
- 構造活性相関、構造物性相関
- 知識情報学およびそのためのソフトウェア
- コンビナトリアルケミストリー
- 通信・システム
- ラボラトリーオートメーション
- 上記分野に関する教育

### ケモトリックス 計量化学

- 多様な分析特徴と物性値との相関
- 多変量解析<分析の方法論>
  - 主成分分析
  - 回帰分析
  - パターン分類

## 事例:タンパク質二次構造予測

- ニューラルネットワーク
  - パターン認識(機械学習)
    - 入力パターン:アミノ酸の種類×前後を含めた配列
    - 出力パターン:ヘリックス、シート、(ターン)、ランダム

## 情報化学

- 化学情報学
- 理論化学、計算化学
- ケモトリックス
- 構造活性相関、構造物性相関
- 知識情報学およびそのためのソフトウェア
- コンビナトリアルケミストリー
- 通信・システム
- ラボラトリーオートメーション
- 上記分野に関する教育

## 構造物性相関

- 構造と物理的性質との関係の発見
  - 材料の開発
  - 生理活性→構造活性相関
    - 薬の開発
- 構造記述子
  - 三次元構造
    - 活性部位とのドッキング

## 情報化学研究

- 研究テーマの設定(問題点の明確化)
- 関連研究の論文講読
- 仮説(モデル)の設定
- 研究テーマ・仮説の新規性のチェック
- 実験(仮説検証)計画の立案
  - 実験方法の選択、開発
  - 分析方法の選択、開発
- 実験(計算機実験)
- 実験結果の考察(仮説の妥当性の検証)
- 研究成果の新規性の再チェック

## 情報化学研究

- 研究テーマの設定(問題点の明確化)
- 関連研究の論文講読
- 仮説(モデル)の設定
- 研究テーマ・仮説の新規性のチェック
- 実験(仮説検証)計画の立案
  - 実験方法の選択、開発
  - 分析方法の選択、開発
- 実験(計算機実験)
- 実験結果の考察(仮説の妥当性の検証)
- 研究成果の新規性の再チェック

## 研究テーマの設定(問題点の明確化)

- 情報化学研究
  - 化学情報学領域
    - 化合物検索システム
  - 計算化学領域
    - 反応メカニズム解析
  - ケモトリックス領域
    - 化合物の分類(構造-物性)



### 事例:タンパク質活性部位の自動検出

- 酵素の活性部位
  - 基質との相互作用
- 基質の無いタンパク質の活性部位
  - どのように検出すれば良いか

### 情報化学研究

- 研究テーマの設定(問題点の明確化)
- 関連研究の論文講読
- 仮説(モデル)の設定
- 研究テーマ・仮説の新規性のチェック
- 実験(仮説検証)計画の立案
  - 実験方法の選択、開発
  - 分析方法の選択、開発
- 実験(計算機実験)
- 実験結果の考察(仮説の妥当性の検証)
- 研究成果の新規性の再チェック

### 仮説(モデル)の設定

- 化学
  - 構造表記
  - 類似度表現
  - 分析手法(スペクトル)
- 情報学/情報工学
  - パターン認識
    - 階層型ニューラルネットワーク
  - 自動分類
    - 自己組織化マップ

### 事例:機能とアミノ酸配置

- 機能発現
  - アミノ酸
- 同じ機能
  - 同じアミノ酸配置
- 同じ機能を持つタンパク質
  - 同じアミノ酸配置の探索 → 機能部位の発見

### 情報化学研究

- 研究テーマの設定(問題点の明確化)
- 関連研究の論文講読
- 仮説(モデル)の設定
- 研究テーマ・仮説の新規性のチェック
- 実験(仮説検証)計画の立案
  - 実験方法の選択、開発
  - 分析方法の選択、開発
- 実験(計算機実験)
- 実験結果の考察(仮説の妥当性の検証)
- 研究成果の新規性の再チェック

### 実験(仮説検証)計画の立案

- 化合物検索システム
  - システム開発→評価(適合率、再現率、速度)
- 反応メカニズム解析
  - 理論の選択→得られたメカニズムの妥当性
- 化合物の分類
  - 多様な分類法→方法論の評価



## 事例: 三次元的に同じ部位の探索

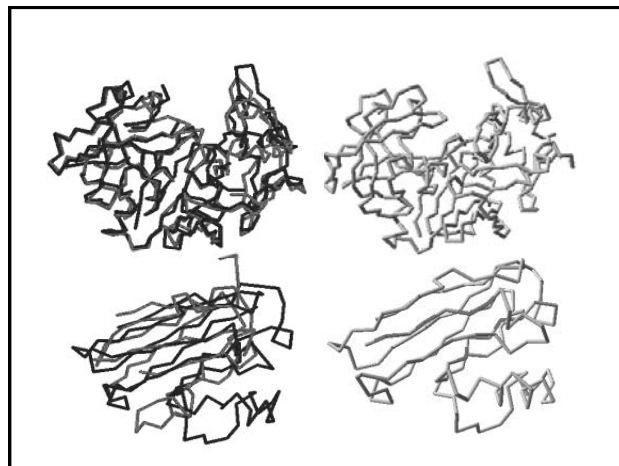
- アミノ酸の配置
- グラフ理論
  - ノードをアミノ酸の対応
  - 対応するアミノ酸間の距離が一致 → エッジ
  - 最大完全部分グラフ
    - 三次元的に同じ最大アミノ酸部分の発見

## 情報化学研究

- 研究テーマの設定(問題点の明確化)
- 関連研究の論文講読
- 仮説(モデル)の設定
- 研究テーマ・仮説の新規性のチェック
- 実験(仮説検証)計画の立案
  - 実験方法の選択、開発
  - 分析方法の選択、開発
- 実験(計算機実験)
- 実験結果の考察(仮説の妥当性の検証)
- 研究成果の新規性の再チェック

## 実験(計算機実験)

- プログラム
  - 新規開発
  - 既存
- データ
  - 新規作成
  - 既存



## 情報化学研究者が求める資料

- 基本的な情報
  - 総説
  - 単行書
  - シリーズ
- 最新の情報
  - 雑誌論文
    - 検索システム

## 情報化学関係の論文誌(学会)

- Journal of Computer Aided Chemistry (2000-)  
(日本化学会情報化学部会)
- Journal of Computer Chemistry, Japan (2002-)
- Journal of Chemical Software (1992-2002)
- JCPE Journal <JCPE Newsletter(1989-1988)> (1999-2002)  
(日本コンピュータ化学会)
- Chem-Bio Informatics Journal  
(情報計算化学生物学会)
- Journal of Chemical Documentation (1961-1974)
- Journal of Chemical Information and Computer Sciences (1975-2004)
- Journal of Chemical Information and Modeling (2005- )  
(The American Chemical Society)
- Journal of Chemical Theory and Computation (2005- ) (The American Chemical Society)

## 情報化学関係の論文誌(商業誌)

- Journal of Computational Chemistry (Wiley)
- Molecular Informatics (Wiley)
- Journal of Chemometrics (Wiley)
- Journal of Molecular Modeling (Springer)
- Journal of Computer-Aided Molecular Design<Perspectives in Drug Discovery and Design> (Springer)
- Journal of Molecular Graphics and Modelling <Journal of Molecular Graphics> (Elsevier)COMP MGMS
- Computational Biology and Chemistry <Computers & Chemistry> (Elsevier)
- Tetrahedron Computer Methodology (Elsevier) -1991
- Chemometrics and Intelligent Laboratory Systems (Elsevier)
- Bioinformatics (Oxford)
- Journal of Chemical Theory and Computation (ACS/COMP)
- Journal of Chemical Information and Modeling (ACS/COMP)

## 情報化学関係論文が掲載されている論文誌

- Journal of the American Chemical Society (ACS)
- Angewandte Chemie(Wiley)<ドイツ化学会誌>
- Journal of Organic Chemistry(ACS)
- Tetrahedron (Elsevier)
- Acta Crystallographica (Wiley)
- Analytical Chemistry (ACS)
- Journal of Analytical Chemistry (Springer)
- Analytica Chimica Acta (Elsevier)
- Journal of Chemical Physics(AIP)
- International Journal of Quantum Chemistry(Wiley)
- Biochemistry (ACS)
- Journal of Molecular Biology (Elsevier)
- Nucleic Acids Research (Oxford)
- Journal of Medicinal Chemistry (ACS)
- Journal of Pharmaceutical Sciences (Wiley)
- Drug Discovery Today ( Elsevier)

## 学生にどのように資料を探させるか

- 図書
  - 数がさほど多くない
    - 個人的に提示
- 論文
  - ほとんどが電子ジャーナル
    - 研究者自身で検索

## 情報化学研究者が求める図書館員

- 周辺領域の知識を求める
  - 専門知識は研究者が持つ
    - アイデアの拡張のために周辺の情報を探る
      - どのような検索システムを使えば良いのか
      - どのような基本的教科書があれば良いのか
- 図書館員が知識を持つ
- 図書館員が研究者の知識を持つ

## おわりに

- 利用者の分析
- 提供できる情報の拡大
  - 専門資料知識
  - 研究者の知識
- 情報の発信