

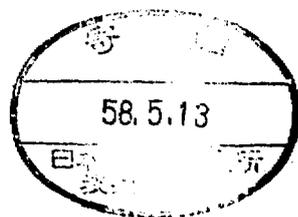
原子過程データ検索のための
データベース

市川行和

1982年7月3日受理

IPPJ-DT-104

1983年3月



名古屋大学プラズマ研究所

原子過程データ検索のための
データベース

市川行和
(1982年7月3日受理)

IPPJ-DT-104

1983年3月

Bibliographic Data Base on Atomic and Molecular Data

Yukikazu Itikawa

Resea Information Center, Institute of Plasma Physics, Nagoya University

A comparative study is made on three bibliographic data bases: INSPEC, ORNL, AMPIC, GAPHYOR. An on-line retrieval is carried out for searching a number of specific atomic and molecular data. Characteristics of each data base are clarified and suggestions are given for use of those data bases.

原子過程データ検索のためのデータベース

§ 1. はじめに

核融合研究にとって原子過程の知識は今や不可欠のものとなっている。しかし、関係する原子過程の種類は非常に多く、適確な情報を必要な時に即座に得ることは困難である。また、核融合研究に必要な原子過程の研究は原子分子物理学の分野にとっても新しいものが多く、既存のデータ集等では間に合わず直接原論文からデータを得る必要がある場合が多い。そのような文献検索に有用なのが文献リストを計算機可読な形にした文献情報データベースおよびそれをオンライン検索するための検索システムである。現在、組織的に作成が行われ、一般の研究者が利用できるデータベースがこの分野で3種類ある。本レポートはそれらのデータベースの相互比較を行い、それぞれの特長を明らかにするとともに、今後の有効利用のための指針を与える。さらに、文献検索の具体例を提供することにより、各データベースを利用する際の参考とする。

ここでとりあげるデータベースは次の3種類である。

(1) INSPEC (AM) : INSPECは英国IEEが作成する物理学全般を対象とする文献情報データベースで、その内容はPhys. Abstracts誌と同じである。その中から原子分子物理に関係した部分のみをとり出してプラズマ研企画情報センターで編集したものをINSPEC* (AM) とする。

(2) ORNL-AMPIC: 米国オークリッジ国立研究所データセンターで作成している原子過程専門のデータベース。

(3) GAPHYOR: バリ大学オルセー校のプラズマ物理研究室において、J. L. Delcroix教授が中心となって作成しているプラズマ・気体中の原子過程を対象とした文献情報データベース。この内、INSPECとORNLはプラズマ研企画情報センターにおいてオンライン検索サービスが行われており、GAPHYORについても今後サービスが可能になるよう検討中である。

以下、§ 2ではこれらのデータベースについて簡単に説明する。次に、いくつかの具体例についてオンライン検索の結果を§ 3で示す。§ 4では検索結果にもとづいて各データベースの特長が述べられ、最後に§ 5ではそれらのデータベースの相互比較と今後の有効利用についてまとめがなされる。

§ 2. 原子分子関係文献情報データベース

ここでは原子分子関係の比較的広い範囲を対象としている3種類のデータベースをとりあげる。それらの扱う分野 (category) および含まれている情報の種類が第I表、第II表にそれぞれまとめられている。以下、簡単に各データベースを紹介する。それぞれの詳細および利用の仕方については参考文献を参照してほしい。

(1) INSPEC (AM) 1,2)

INSPEC (International Information Services for the Physics and Engineering Communities) は英国のIEE (The Institution of Electrical Engineers) が作成しているもので、それぞれの物理学、電気工学、計算機・制御工学の文献情報を集めた3部 (A, B, C) から成っている。物理

学の部分はPhysics Abstracts誌と同じ内容のものがデータベース化されている。1969年から発行されており、現在約2,300種類の雑誌から関連する文献情報を収集している。収録レコード数は1981年末で約150万件に達し、年に18万件の割合で増加している。このデータベースの特長の一つはアブストラクトが含まれていることで、原論文を見なくても内容がある程度分かるようになっている。物理学全体をカバーするために分類方法やキーワードのつけ方はかなり大まかにならざるを得ず、後述するように特定の専門分野の文献をきめ細かく探すにはあまり向いていない。INSPEC全体はぼう大なものなので、プラズマ研企画情報センターではその中から原子分子関係およびプラズマ物理関係のものを抽出してそれぞれINSPEC(AM)、INSPEC(PLASMA)として計算機に格納している。ここではINSPEC(AM)のみを検索の対象とする。

(2) ORNL-AMPIC³⁾

米国オークリッジ国立研究所には核融合のための原子分子データセンターがあり、古くから活動を行っている。その主な活動の一つが原子分子関係の文献情報データベースの作成である。これは1950年以降の文献について情報が集められており、そのカバーする分野もかなり広く(必ずしも核融合に直接関係のあるものばかりではない)、粒子と固体との相互作用や物質中の粒子の通過に伴う現象(阻止能など)などプラズマ-壁相互作用として最近注目をあびている分野までも含んでいる。また、原子分子専門のデータベースなので分類もきめ細かく関係する原子分子についての情報も明記してあるなど、この分野の文献情報検索には大変使い易い。このデータベースはオークリッジ・センターにおいて部内での使用のためにのみ作られているもので一般公開はしていない。同センターの好意によりプラズマ研企画情報センターに提供され、企画情報センターではそれをオンライン検索が可能ないように整備した。(同データベースの一部は文献リストとして印刷・出版されている⁴⁾)なお、1978年版よりレコードの収録形式が変更になったため、それ以降のものをORNL NEWとして別のデータベースとして扱う。

(3) GAPHYOR⁵⁾

パリ大学オルセー校・気体・プラズマ物理研究室のJ. L. Delcroix教授は10年程前から気体あるいはプラズマ中の原子過程について文献情報データベースを作成しているが、現在ではそれが組織化されGAPHYOR(Gaz Physique Orsay)という名前で公開されている。原子衝突や光と原子分子の相互作用だけでなく、原子分子そのものの諸性質(エネルギー単位など)についても文献情報が集められているのが特長である。また、関与する原子分子の状態に関する詳しい情報も含まれており、非常にきめ細かく文献を探ることができる。このデータベースの内容は文献リストとして定期的に刊行されているが⁶⁾オルセーではオンライン検索が可能となっている。我が国ではプラズマ研企画情報センターにその一部が導入され、オンライン検索のテストが行われている。

§ 3. 検索の実例

ここでは前章で紹介したデータベースを実際に利用し、次の8種類の原子過程データあるいは関連する物理量についての情報が得られる文献を探す。

- (1) 電子衝突によるリチウム原子の電離の断面積
- (2) 電子衝突による酸素分子の振動励起の断面積
- (3) 2価の酸素イオン (O^{++})の遷移確率
- (4) 準安定状態にあるアルゴンと窒素分子の衝突において励起状態の移動が起る過程の断面積
- (5) ヘリウム原子による陽子の弾性散乱、特にその微分断面積
- (6) 鉄原子の光電離断面積
- (7) 1価の炭素イオン (C^+)の電子衝突による励起の断面積
- (8) 1価のリチウム・イオンの分極率

これらの原子過程は必ずしもその全てが核融合研究に深い関わりがあるわけではないが、なるべく広い範囲から例題を選ぶように心がけた。なお3種類のデータベース間の比較を容易にするために、1970年以降に出版された論文のみを対象とする。

一般にオンライン検索の際の質問の仕方はデータベースにより、また与えられた問題(どのような情報がほしいか)により千差万別であり、標準的な方法というものはない。データベースによっては効率良く検索するためにはかなりの熟練を要するものもある。ここでは、①問題がかなり単純であること、および②初めての人にも分るような検索例を示すことも目的の一つであること、から出来るだけ単純な質問を使って検索することとした。特にINSPECの場合には、アブストラクトとキーワード(一定の辞書(シソーラス)にのっている語の中から選んでつける標準語(SI)と索引作成者が勝手につける自由語(FI)とがある)が検索の手がかりとして使えるのでかなり多様な質問の仕方が可能である。しかしここでは単純に、キーワード(SIとFIの両方を使う)を手がかりとして検索することを標準とした。各問題に対して実際に使った質問式を第Ⅲ表に示す。各質問式の意味についてはそれぞれのデータベースの「利用の手引き」を参照してほしい。^{2, 3, 5)}

第Ⅲ表にはまた、検索の結果が量的に示されている。ここで各記号は次のような意味をもつ。

n …… 検索の結果得られた全文献数

a …… nの中で目的のデータが与えられている文献の数

b …… nの中で目的のデータそのものは得られないが何らかの関連する情報が得られる文献の数

q …… 1970年以降に出版された論文の中で、目的のデータを与えていることが分かっている文献の総数(関連情報のみのものは除く)

s …… qの中で今回の検索では見つからなかった(従ってnには含まれない)が、当該データベースに含まれていることが明らか(たとえば著者名で検索してみると分る)な文献の数
bには予備的な結果しか与えない国際会議報告や研究所レポートなども含まれる。qを正確に求めるのは一般に容易ではない。電子衝突や光衝突あるいは原子(イオンも含む)の遷移確率のようにある特定の分野に限ると、ほとんど完璧な文献リストが出版されているのでそれを参考にする。^{7, 8)} それ以外の場合には、ここで利用した三つのデータベースの少くともどれか一つで見つけられたものを集めてその総数をqとする。参考のためにそのような文献のリストを附録として掲げ検索結果を記入してある。

最後に、各データベース間の比較を定量的に行うために次の量を導入する。

検索の精度 (precision) : $P = a/n$

(関連文献まで含めた場合の精度) : $P' = (a+b)/n$

必要な文献の検出率 (recall) : $R = a/q$

必要な文献の収録率 (coverage) : $C = (a+s)/q$

検索の結果、不必要な情報 (ノイズ) が現われる程度を示すノイズ率は $1 - P$ で与えられる。P および R が共に 1 に近ければ必要にして十分な情報が得られるわけである。各問題毎にこれらの量を計算した結果を第IV表に示す。これをもとに各データベースの比較を次章で行う。

§ 4. 各データベースの特長

前章で述べられた検索結果をもとに各データベースの特長をまとめると次のようになる。

(1) INSPEC

必要な文献はほぼ 100% 収録されている。しかし、検索によって実際に見つけ出される確率 (R) はかなり低い。今回は主としてキーワードを手がかりに検索を行ったので、このことはキーワードのつけ方が良くないことを示す。さらに、キーワードはほとんど全ての場合、論文のタイトルとアブストラクトを基にしてつけるので、原論文のタイトルやアブストラクトの書き方が良くない (つまり、十分な情報を与えていない) ことがそもそもの原因になっている場合が多い。たとえば問題 (1) の例で、リチウムをはじめ一連のアルカリ金属について研究を行った場合にキーワードには単に ALKALI METAL としか書かず、各元素名を書いていないものがある。もちろんこの場合に、ALKALI METAL を使って検索すれば良いが、そのときは余分な文献が多数現われる (ノイズが多くなる) 恐れがある。このように、関係のありそうなキーワードをいくつか用いて検索を行えばそれだけ検出率は良くなるが無駄が非常に多くなる。また、キーワードでなくアブストラクトを対象として検索を行えば検出率を増すことが出来るがやはりノイズが格段に増加する。質問に使っている単語が主題に関係なく使用されることがあり全く関係のない論文が検出されたりするからである。INSPEC では検索精度 (P) もかなり低い。すなわち、一般にノイズが多い。これは、このデータベースが原子過程専用ではないので衝突過程や関与する原子分子に関する情報をキーワード (もっと広くはタイトルやアブストラクト) の中から探さなければならず、ズバリそのものを特定することが困難なことによる。たとえば問題 (4) で準安定状態にあるアルゴン原子を検出するのに METASTABLE というキーワードを手がかりとしたが、アルゴンではなくその衝突相手の粒子が準安定状態にある場合を扱った論文も少数ではあるが得られた。以上を要するに、INSPEC ではほしい情報を規定する単語 (群) をうまく見つけられればかなり精度良く文献を探することができる。ノイズが多くなることを構わなければ、キーワード、アブストラクト、タイトルを対象にいろいろな単語を組み合わせれば検索を行えばかなりの程度はほしい論文を見つけることが可能である。いったん情報が得られると、アブストラクトを見るだけでその文献の内容がある程度判明するし、著者の住所が必ずついているので原論文の別刷を請求することも出来る。これらのことは他のデータベースにはない特長である。

(2) ORNL

文献の収録率はかなり良いがINSPECほどではない。これは作成組織が小さいからであろう。検出率は平均して約60%で、収録されている論文のみに限れば80%にもなる。このデータベースでは原子過程の種類を指定するのにカテゴリー分類が使える。しかしその分類では粗すぎて特定できない場合(たとえば、分子の振動励起というカテゴリーはなく励起に含まれてしまう)には、論文のタイトルを手がかりに検索しなければならない。タイトルのつけ方が不十分だとモレが出てくる。検索精度は40%位であるが、関連情報を与える文献まで含めると(P'), 80%位に増加する。これは、次のGAPHYORの場合と同様に、関係のある論文は出来るだけ全て含めて利用者の方で選択するという方針でデータベースを作成しているためである。1977年版以前のものである。一つの論文でいく組もの衝突系を扱っている場合それらを一括して表示して(例、(He, Ne, Ar) + (H₂, N₂))。個々の組み合わせを一つずつ示していない(例、He+H₂, Ne+N₂, Ar+H₂, Ar+N₂)。そのために関与する原子分子の名前で検索すると実際にはその衝突系を扱っていない論文も出てきてしまう(上記の例では、He+N₂は扱っていない)。1978年以降の版(ORNL, NEW)では個々の組み合わせが示されており、特定の衝突系を指定して検索することができる(第Ⅲ表の質問式参照)。従って、ORNL, NEWだけをとると検出精度はほぼ100%(P'でみれば)となる。以上のことから分るように、ORNLは原子過程専用のデータベースであるので、原子過程に関しては細かく情報を探ることができる。特に、多少質の低い情報(関連情報)が混ざることが構わなければ、その検索効率はかなり良い。ただし、アブストラクトがないので原論文を見ないと本当に必要な情報かどうかの判断が出来ず、大量の結果が出たときは選別に困ることがある(もちろん、タイトルである程度の判断は出来る)。

(3) GAPHYOR

文献の収録率はかなり悪い。平均すると50%である。特に1975年以前の論文は収録されていない割合が高い。つまり、GAPHYORの作成は1970年から始まっているがはじめはDelcroix教授が個人でやっており組織的になったのが1975年頃であることが収録率に反映している。したがって文献の検出率もかなり悪い。ただし1975年以後のものに限ればかなり良くなる。検索精度はaのみでは平均して40%弱である。bまで含めると90%を超える。すなわち、全く無関係な文献が出てくる確率はほとんどないが、質の低い関連情報が多く目的とするデータが見つかる確率はそれほど高くはない。このデータベースの作成方針は少しでも関係のある論文は全て収録することである。これはどんな場合でも何らかの手がかりが得られるという利点はある反面、あまり役に立たない難な情報が多過ぎて困るという欠点もある。GAPHYORの特長は、論文のタイトルやアブストラクトが含まれていない代わりに、その論文で扱っている原子過程に関して詳しい情報が記号化されて付与されていることである。それを利用することにより非常にキメ細かい検索が行える。すなわち、全くのノイズが出てくるのは分類を間違えるなど誤った情報が含まれる場合にほとんど限られる。結論として、GAPHYORは1975年以後に発表された論文に限れば情報もこまかく検索の効率も大変良い。しかし、かなり多数の関連情報が含まれることがあり、無駄が多い場合がある。

以上、ここでとりあげた三つのデータベースについてそれぞれの特長を述べた。利用者はそれぞれ

の特長を生かしたデータベースの使い方をすることが望まれる（次章参照）。

最後に、検索結果を評価する際に気をつけなくてはならないのは検索システム固有の欠陥による問題点である。特にINSPECとORNLの検索に用いたシステム、FAIRS, は汎用のものであり、それらのデータベースに合わせて作られたものではない。たとえば、INSPECではO₂は'O SUB 2'と表わされているがFAIRSでは'O SUB 2'の形のままで検索することは出来ない。ここでは'O SUB'を用いて検索した。当然のことながら、このようにするとO₂ばかりではなくO₃などもひっかかってしまう。検索精度が悪い原因にはこのようなものも含まれている。また、ORNLとGAPHYOR はもともと印刷された文献リストを作るためのものでオンライン検索をすることを念頭においてその記録形式が考えられていない。従って、データベースと検索システムとが必ずしもマッチしないところがある。関連情報まで入れても検索精度が異常に小さい（すなわち、ノイズが極めて多い）場合があるのはこのような事情が支配している例である。

§ 5. まとめ

前章の各論をもとにデータベース間の相互比較を行うと次のようにまとめられる。

- (1) 特定の原子分子の関与する特定の原子過程についてデータの所在を調べるにはGAPHYOR が最もキメ細かく検索ができ効率が良い。
- (2) しかしGAPHYOR は、1975年以前の論文については収録率が悪い、多数の質の低い関連情報が含まれる、得られる情報には文献のタイトルモアブストラクトも含まれていないので論文の送別が即座には出来ない、などの欠点がある。これらのことが問題となる場合にはORNLを用いるのが良い。
- (3) 何らかの手がかりが得られれば良いというのであればINSPECも役に立つ。しかし一般に、INSPECはキメ細かい検索には不向きであり無駄が多い。INSPECはむしろある特定の主題（“原子の自動電離”とか“トカマク中の不純物”など）について情報を集めるのに適している。ただしその場合でも質問式を工夫しないと効率が悪い。
- (4) 以上を総合すると、単に手がかりが得られれば良いというのではなく、ある程度十分な情報を得たいときには、二つないし三つのデータベースを併用する必要があることがわかる。附録の表を見れば分るように、三つのデータベースを検索した結果全てに共通して現われた文献というものは意外に少ない。これは一見奇妙なことであるが、各データベースの作成方針や内容の表わし方（フォーマット）の違いが原因であろう。とにかく、三つのデータベースを併用しても重複による無駄はそれ程多くないようである。

以上原子過程という狭い分野であるがデータベースをオンライン検索した結果を分析しまとめてみた。今後ますます情報過多になる社会ではいかに上手にそれらの情報を使っていくかが問題となってくるであろう。そのためには情報を効率良く利用するための研究が必要になると思われる。⁹⁾ 本報告がそのような研究の一部として役立てば幸いである。

本研究における文献検索は、INSPECとORNLについてはプラズマ研究所核融合研究企画情報セン

ターにおいて検索システムFAIRSを用いて行われ、GAPHYORについてはバリ大学オルセー校において検索システムSYGALを用いて行われた。データベースORNLはオークリッジ国立研究データセンターの好意で企画情報センターに提供されているものであり、同センター関係者、特にC. F. Barnett, D. H. Crandallの両氏に感謝したい。また筆者をオルセーに招いてGAPHYORの検索の機会を作って下さったJ. L. Delcroix教授にも感謝する次第である。

参 考 文 献

- 1) INSPECの紹介については、たとえば山崎 昶：情報管理23 536 (1980)。
- 2) プラズマ核融合研究企画情報センター：「データベース利用の手引き」第1版(1979)
- 3) 「ORNLデータベース利用の手引き」(プラズマ研企画情報センターにて準備中)
- 4) Bibliography of Atomic and Molecular Processes, DOE/ER 0044(1978版), 0074(1979版), 0118(1980版)
- 5) GAPHYOR:Manual of operation for on line retrieval, Laboratoire de Physique des Gaz et Plasmas, Universite de Paris-Sud 発行
- 6) Bulletin Signalétique 166 GAPHYOR, Centre National de la Recherche Scientifique 発行(年4回)
- 7) Bibliography of Low Energy Electron and Photon Cross Section Data, NBS Special Publ. 426(1976); Suppl. 1(1979)およびJILA Information Center Rept. No. 18(1980); No. 21(1981).
- 8) Bibliography on Atomic Transition Probabilities, NBS Special Publ. 505(1978); Suppl. 1(1980).
- 9) データベースのオンライン検索について問題点を述べた文献として、たとえば、T. E. Doszkoecs, B. A. Rapp and H. M. Schoolman : Science 208 25(1980).

附 録

検索の結果、それぞれのデータベースからどのような文献が見つかったかを一覧表にして示す。第III表でa（検索の結果見つかったもの）またはS（検索では見つからなかったが当該データベースに含まれていることが明らかなもの）を出版年および第一著者名のアルファベット順に示してある。なお、以下の略号を使用する。

A U……………第一著者名

D P……………出版年

J N……………雑記名（簡単のためにGAPHYORで使用しているコードで示す。実際の雑記名との対応は本附録の最後に掲げてある）

V I I P……………巻 ページ

Table I. Categories covered by the data bases

Category	INSPEC(AM)	ORNL [*]	GAPHYOR
Properties of atoms and molecules	yes		yes
Photon collisions	yes	yes	yes
Electron collisions	yes	yes	yes
Heavy particle collisions	yes	yes	yes
Particle penetrations in gases and solids	yes	yes	
Particle interactions with solids		yes	
Transport phenomena and other macroscopic properties in gases	yes**	yes	yes

* The scope of ORNL data base may change after 1982.

** Most of this category is covered by INSPEC(PLASMA).

Table II. Information included in the data bases

	INSPEC	ORNL	GAPHYOR
Title	yes	yes	
Authors' names	yes	yes	yes ^{a)}
Authors' affiliation	yes	yes ^{b)c)}	yes ^{c)}
Journal name	yes	yes	yes
Volume	yes	yes	yes
Pages	yes	yes	yes
Year of publication	yes	yes	yes
Abstract	yes		
Key words	yes		
Collision systems	(yes) ^{e)}	yes	yes ^{d)}
Processes	(yes) ^{e)}	yes	yes
Collision energy	(yes) ^{e)}	yes	yes

a) First author only.

b) Only after 1978.

c) Name of the country.

d) Details of the initial/final states of the colliding particles are also specified.

e) These informations are not given separately, but included in the key words as well as in the titles and abstracts.

Table III. Sample problems and results of retrieval

- n: total number of papers found
a: number of papers giving the relevant data (called 'relevant papers')
b: number of papers giving only related information
q: number of relevant papers published after 1970
s: number of relevant papers not found by the present retrieval but stored in the data base

All the relevant papers (either included in a or s) are listed in the Appendix.

(1) e + Li ionization		q = 7
<u>INSPEC</u>	<u>ORNL</u>	<u>GAPHYOR</u>
SI/FI= LI or LITHIUM SI/FI= ELECTRON IMPACT SI/FI= IONIZATION or IONISATION not TAGALL= INNER n= 36 a= 2 b= 7 s= 5	CA= E05 RS= LI YE= 197@ n= 12 a= 6 b= 4 s= 0 <u>ORNL NEW</u> CA= E05 RI= @LI n= 2 a= 0 b= 2 s= 0	FILE 1 XX= LI PP= LII CP= 3 DS= IN n= 21 a= 2 b= 14 s= 0
(2) e + O ₂ vibrational excitation		q = 5
<u>INSPEC</u>	<u>ORNL</u>	<u>GAPHYOR</u>
FI= 0 SUB SI/FI= ELECTRON IMPACT SI/FI= VIBRATION@ n= 20 a= 3 b= 7 s= 2	CA= E03 RS= 02 TI= VIBRATION@ YE= 197@ n= 7 a= 2 b= 5 s= 2 <u>ORNL NEW</u> CA= E03 RI= 'e + O ₂ ' TI= VIBRATION@ n= 1 a= 0 b= 1 s= 0	FILE 1 XX= 0 PP= 02 CP= 3 DS= EX SS= 02//v n= 11 a= 1 b= 10 s= 0

(3) 0^{++} transition probabilities

q = 29

INSPEC

FI= 0 SUP
SI/FI= TRANSITION PROBAB@
or OSCILLATOR STRENGTH@

n= 23 a= 5 b= 0
s= 21

GAPHYOR

FILE 1
XX= 0
PP= 01/2+/*
CP= 1
DS= TR

n= 21 a=14 b=5
s= 4

(4) $Ar^* + N_2$ excitation transfer

q = 33

INSPEC

FI= N SUB
SI/FI= AR or ARGON
SI/FI= METASTABLE
SI/FI= EXCITATION TRANSFER
or ENERGY TRANSFER

n= 26 a= 14 b= 0
s= 17

ORNL

CA= A03 or A05
or A09
RS= N2
RS= 'AR SUP*'
YE= 197@

n= 30 a= 19 b= 3
s= 3

ORNL NEW

CA= A03 or A10
or A11

RI= 'Ar* + N₂'

n= 4 a= 3 b= 0
s= 0

GAPHYOR

FILE 2
XX=AR
YY= N
PP= AR1/0+/*
QQ= N2
CP= 4
DS= XX or DX
DQ= F

n= 23 a= 16 b= 7
s= 3

(5) $H^+ + He$ elastic scattering (OCS)

q = 7

INSPEC

TI/SI/FI= HELIUM
TI/SI/FI= PROTON@ or H SUP
TI/SI/FI= ELASTIC
TI/SI/FI= DIFFERENTIAL

n= 12 a= 2 b= 5
s= 5

ORNL

CA= A02
RS= 'H SUP +'
RS= HE
DA= 1

n= 23 a= 6 b= 12
s= 1

GAPHYOR

FILE 2
XX= H
YY= HE
PP= H1/+
QQ= HE 1
CP= 4
DS= EL

n= 15 a= 1 b= 13

s= 0

ORNL NEW

CA= A02
RI='H⁺ + He'

n= 6 a= 0 b= 6
s= 0

(6) $h\nu + Fe$ ionization

q = 7

INSPEC

SI/FI= FE or IRON
SI/FI= PHOTOIONISATION
or PHOTOIONIZATION
not TAGALL= ASTRO@

n= 29 a= 4 b= 13
s= 3

ORNL

CA= H06
RS= FE
YE= 197@

n= 5 a= 3 b= 1
s= 0

GAPHYOR

FILE 1
XX= FE
PP= FE1
CP= 2
DS= IN

n= 12 a= 5 b= 7
s= 1

ORNL NEW

CA= H06
RI= @Fe

n= 2 a= 2 b= 0
s= 0

(7) e + C⁺ excitation

q = 3

INSPEC

FI= C SUP
SI/FI= ELECTRON IMPACT
SI/FI= EXCITATION@
or TRANSITION@

n= 27 a= 1 b= 0
s= 2

ORNL

CA= E 03
RS= 'C SUP +'
YE= 197@

n= 6 a= 0 b= 2
s= 1

GAPHYOR

FILE 1
XX= C
PP= C1/+/
CP= 3
DS= XX

n= 2 a= 0 b= 2
s= 2

ORNL NEW

CA= E03
RI= 'e + C⁺'

n= 2 a= 0 b= 2
s= 0

(8) Li⁺ polarizability

q= 29

INSPEC

FI= LI SUP
SI/FI= POLARISABILIT@
or POLARIZABILIT@

n= 24 a= 15 b= 0
s= 14

GAPHYOR

FILE 1
XX=LI
PP= LI1/+/
CP= 1
DS= PE

n= 23 a= 19 b= 2
s= 1

Table IV. Quantitative comparison of the retrieval

	data base*	Precision		Recall	Coverage
		$P=a/n$	$P'=(a+b)/n$	$R=a/q$	$C=(a+s)/q$
(1) e+Li ionization	I	0.056	0.25	0.29	1.0
	O	0.43	0.86	0.86	0.86
	G	0.095	0.76	0.29	0.29
(2) e+O ₂ vibrational excitation	I	0.15	0.5	0.6	1.0
	O	0.25	1.0	0.4	0.8
	G	0.091	1.0	0.2	0.2
(3) O ⁺⁺ transition probability	I	0.22	0.22	0.17	0.90
	O	-	-	-	-
	G	0.67	0.91	0.48	0.62
(4) Ar* + N ₂ excitation transfer	I	0.54	0.54	0.42	0.94
	O	0.65	0.74	0.67	0.76
	G	0.70	1.0	0.49	0.58
(5) H ⁺ + He elastic scattering	I	0.17	0.58	0.29	1.0
	O	0.21	0.83	0.86	1.0
	G	0.067	0.93	0.14	0.14
(6) hv + Fe ionization	I	0.14	0.59	0.57	1.0
	O	0.71	0.86	0.71	0.71
	G	0.42	1.0	0.71	0.86
(7) e + C ⁺ excitation	I	0.037	0.037	0.33	1.0
	O	0	0.5	0	0.33
	G	0	1.0	0	0.67
(8) Li ⁺ polarizability	I	0.63	0.63	0.52	1.0
	O	-	-	-	-
	G	0.83	0.91	0.66	0.69
	averaged	I	0.24	0.42	0.40
	O	0.38	0.80	0.58	0.74
	G	0.36	0.94	0.37	0.51

* I:INSPEC(AM), O:ORNL and ORNL NEW, G:GAPHYOR

(1) e + Li : ionization

AU	DP	JN	VI/IP	INSPEC	ORNL	GAPHYOR
I.L.Beigman	1970	827	<u>13</u> 775	s	a	
A.K.Liepinsh	1970	484	2 3	s	a	
B.Stern	1970	22	<u>2708</u> 533	s	a	
E.J.McGuire	1971	PRA 70	<u>3</u> 267	s	a	
K.Omidvar	1972	PRA 70	<u>5</u> 1174	s	a	
R.Jalin	1973	JCP 30	<u>59</u> 952	a	a	a
A.Korotkov	1977	498	<u>20</u> 51	a		a

(2) e + O₂ vibrational excitation

AU	DP	JN	VI/IP	I	O	G
D.Spence	1970	70	<u>2</u> 1802	s	a	
F.Linder	1971	90	<u>269</u> 1617	a	s	
F.Koike	1973	52	<u>35</u> 1166	a	s	
S.F.Wong	1973	68	<u>31</u> 969	a	a	
F.Koike	1975	52	<u>39</u> 1590	s		a

(3) 0^{2+} transition probabilities

AU	DP	JN	VI/IP	I	0	G
H.G.Berry	1970	07	<u>5</u> 81	s		
W.Eissner	1970	29	<u>31</u> C4-149			a
J.A.Kernahan	1970	50	<u>60</u> 986	s		a
C.Nicolaides	1970	29	<u>31</u> C4-117			a
C.Nicolaides	1970	675	<u>90</u> 133	s		
C.Nicolaides	1970	62	<u>33A</u> 178	s		
E.H.Pinnington	1970	675	<u>90</u> 93	s		
M.Druetta	1971	50	<u>61</u> 515	s		a
I.Martinson	1971	50	<u>61</u> 519	s		s
C.Nicolaides	1971	70	<u>4</u> 1400	s		a
E.H.Pinnington	1971	62	<u>36</u> 299	s		a
M.W.Smith	1971	234	<u>23</u> 103	s		
C.C.Lin	1972	16	<u>50</u> 2496	s		a
C.A.Nicolaides	1973	20	<u>21</u> 242	s		
H.Nussbaumer	1973	09	<u>166</u> 411	a		
W.Eissner	1974	40	<u>7</u> 2533	a		a
C.Laughlin	1974	30	<u>60</u> 1688	s		a
E.H.Pinnington	1974	16	<u>52</u> 1961	s		a
J.V.Mallow	1976	46	<u>16</u> 409	a		a
O.Bely	1977	231	<u>61</u> 711	s		a
M.W.Chang	1977	09	<u>211</u> 300			a
Z.I.Kuplyauskis	1977	490	<u>41</u> 2626	a		
E.H.Pinnington	1978	16	<u>56</u> 508	s		a
A.K.Bhatia	1979	231	<u>76</u> 359	s		s
J.Mitroy	1979	40	<u>12</u> 1081	s		
H.Nussbaumer	1979	231	<u>72</u> 129	s		
G.Sorensen	1979	29	<u>40</u> C1-157	s		
K.L.Baluja	1980	40	<u>13</u> 829	s		s
H.Nussbaumer	1981	231	<u>99</u> 177	a		s

(4) Ar* + N ₂	excitation		transfer		I	O	G
	AU	DP	JN	VI/IP			
D.W.Setser	1970	30	<u>53</u>	1004	s	a	a
D.H.Stedman	1970	30	<u>52</u>	3957	s	a	
O.P.Bochkova	1971	58	<u>31</u>	359	s	a	
J.M.Calo	1971	30	<u>54</u>	4961	s	a	
T.T.Kassal	1971	30	<u>54</u>	1363	s	a	
P.C.Cosby	1972	30	<u>57</u>	4111	a	a	
J.M.Calo	1972	30	<u>56</u>	682		a	
N.Sokabe	1972	647	<u>13</u>	163	a	a	
J.Le Caloe	1973	30	<u>58</u>	1446	a	a	
L.G.Piper	1973	30	<u>59</u>	3323	s	a	
O.P.Bochkova	1974	58	<u>36</u>	19	a	a	
H.J.De Jong	1974	20	<u>25</u>	129	s	a	
P.H.Winicule	1974	30	<u>61</u>	1548	a	a	
J.Krenos	1976	30	<u>65</u>	5017	s	s	a
R.A.Sanders	1976	30	<u>65</u>	2700	s	a	s
A.N.Schweid	1976	20	<u>42</u>	103	a	a	a
J.Krenos	1977	20	<u>49</u>	447	a	a	a
J.Krenos	1977	30	<u>66</u>	5832			a
V.Puech	1977	30	<u>67</u>	2887	s	a	a
N.Sadeghi	1977	28	<u>38</u>	1283	a	a	
M.Touzeau	1977	28	<u>38</u>	789	a	s	s
A.Yokoyama	1977	315	<u>22</u>	459	a	a	
J.M.Cook	1978	30	<u>69</u>	2562	s		a
W.R.Bennett	1978	70	<u>18</u>	2527	a	a	a
M.Touzeau	1978	20	<u>53</u>	355	a		a
J.E.Velazco	1978	30	<u>69</u>	4357	s	s	a
T.P.Parr	1978	30	<u>69</u>	1613	s	a	s
E.R.Cutshall	1979	30	<u>70</u>	3171	a	a	a
E.A.Gislason	1979	20	<u>67</u>	252	a		a
I.Nishiyama	1979	20	<u>67</u>	258	s		a
V.N.Ochikin	1980	80	<u>48</u>	232	s		a
J.Deronard	1980	30	<u>72</u>	6698	s		a
D.W.Loeb	1981	30	<u>74</u>	3270	s		a

(5) $H^+ + He$ elastic scattering (DCS)

AU	DP	JN	VI/IP	I	O	G
R.L.Champion	1970	70	<u>2</u> 2327	s	a	
L.D.Doverspike	1970	68	<u>25</u> 909	a	a	
H.F.Helbig	1970	70	<u>2</u> 771	s	a	
J.D.Johnson	1970	762	<u>44</u> 1700	s	s	
W.G.Rich	1971	70	<u>4</u> 2253	s	a	
H.P.Weise	1971	90	<u>26a</u> 1122	a	a	
T.L.Bailey	1975	26	<u>18</u> 339	s	a	a

(6) $h\nu + \text{Fe}$ ionization

AU	DP	JN	VI/IP	I	O	G
H.Kelly	1971	68	<u>26</u> 1359	a	a	
H.P.Kelly	1972	70	<u>5</u> 168	s	a	a
H.P.Kelly	1972	70	<u>6</u> 1048	a	a	a
J.H.Scofield	1976	32	<u>8</u> 129	s		s
J.E.Hansen	1977	40	<u>10</u> 37	a		a
G.G.Lombardi	1978	70	<u>18</u> 2131	s	a	a
R.F.Reilman	1978	70	<u>18</u> 2124	a	a	a

(7)	$e + C^+$	excitation					
	AU	DP	JN	VI/IP	I	O	G
	A.R.G.Jackson	1972	40	<u>5</u> L83	s		
	B.R.Tambe	1977	40	<u>10</u> L249	a	s	s
	L.Brandus	1980	76	<u>25</u> 121	s		s

(8) Li^+ polarizability

AU	DP	JN	VI/IP	I	O	G
J.T. Broussard	1970	30	<u>53</u> 1507	s		
J.M. Schulman	1970	30	<u>53</u> 3662	s		
K.T. Chung	1971	70	<u>4</u> 7	s		
P.K. Mukherjee	1971	27	<u>5</u> 647	a		
M. Natori	1971	52	<u>30</u> 518	a		a
P. Sitz	1971	30	<u>55</u> 1481	s		
A.K. Bhattacharya	1972	27	<u>6</u> 337	a		a
M. Cohen	1972	40	<u>5</u> 184	a		
F. Weinhold	1972	72	<u>327</u> 209	a		
E. Weislinger	1972	22	<u>275</u> 179	s		s
G.P. Arrighini	1973	70	<u>8</u> 577	a		a
K.E. Banyard	1973	30	<u>58</u> 1132	a		
A.K. Bhattacharya	1973	27	<u>7</u> 491	s		a
R.F. Stewart	1973	49	<u>69</u> 1685	s		a
T.J. Venanzi	1973	30	<u>59</u> 523	a		
G.P. Arrighini	1974	40	<u>7</u> 1790	s		a
R. Klingbeil	1974	70	<u>9</u> 1456	s		a
D. Rosenthal	1974	72	<u>337</u> 365	s		a
R.F. Stewart	1974	56	<u>27</u> 779	s		a
N.L. Manakov	1975	58	<u>38</u> 115	s		a
R.M. Glover	1976	30	<u>65</u> 4913	a		a
J.L. Rivail	1976	22	<u>283</u> 111	a		a
W.E. Cook	1977	70	<u>16</u> 1141	s		a
R.M. Glover	1977	30	<u>66</u> 185	a		a
R.M. Glover	1977	30	<u>66</u> 191	a		a
L.J. Bartolotti	1978	56	<u>36</u> 79	a		a
L.J. Bartolotti	1978	56	<u>36</u> 97	a		a
G.D. Mahan	1980	70	<u>22</u> 1780	s		a
M.J. Stott	1980	70	<u>21</u> 12	s		a

Code of Journal Name (JN)

- 07 Appl. phys. lett.
- 09 Astrophys. J.
- 16 Canad. J. phys.
- 20 Chem. phys. lett.
- 22 Comptes Rendus Ac. Sc. B.
- 26 Internat. J. mass spec. ion. phys.
- 27 Int. J. quant. chem.
- 28 J. de phys.
- 29 J. de phys., Colloq.
- 30 J. of chem. phys.
- 32 J. of electron sp. relat. ph.
- 40 J. of phys. B (Proc. phys. soc.)
- 46 J. of quant. spect. rad. transf.
- 49 J. of the chem. soc. Faraday trans.II
- 50 J. of the opt. soc. Amer.
- 52 J. of the phys. soc. jap.
- 56 Molecular phys.
- 58 Opt. and spectroscopy
- 62 Phys. lett. A.
- 68 Phys. rev. letters.
- 70 Phys. rev. A.
- 72 Proc. roy. soc. A.
- 76 Rev. roum. phys.
- 80 Sov. phys. J.E.T.P.
- 90 Z. naturforschung A.
- 231 Astr. and Astrophys.
- 234 Astrophys. J. suppl. ser.
- 315 Chem. phys.
- 484 Izvest. AK. Nauk. latv. SSR ser. fiz. tek.
- 490 Izvest. AK. Nauk. S.S.S.R. Ser. fiz.
- 498 Izvest. vys. uchebn. zavcd. Fis.
- 647 Mem. Fac. Eng. Osaka Univ.
- 675 Nucl. instrum. methods.
- 762 Progr. theor. phys.
- 827 Sov. astronomy.