

特許・非特許文献同時検索における理想DBの研究

知的財産情報検索委員会
第 1 小委員会*

抄 録 特許文献情報と、科学技術文献等の非特許文献情報との同時検索DBのあり方について、知財担当の視点のみならず、技術者・研究者視点及び経営者視点での検討を行った。特に技術者・研究者視点においては、特許情報は技術情報でもあり、科学技術文献等と同様に検索を行いたいというニーズがあり、そのための理想のDBとしては、検索インターフェースが共通であり、しかもその検索タームが複雑ではなく、適合性の高い結果が簡単にかつリーズナブルなコストにて得られるようなものが挙げられる。

目 次

1. はじめに
2. 特許文献DBと非特許文献DBの現状及びその検索手法
3. 特許文献DBと非特許文献DBの現状の問題点
4. 特許・非特許文献の理想DBに関する考察
 4. 1 理想DBの概要
 4. 2 技術者・研究者向けDB
 4. 3 知財担当者向けDB
 4. 4 経営者向けDB
5. まとめ

1. はじめに

本小委員会では、知的財産戦略大綱に記されている「特許情報とそれに関連した技術情報を、研究者が容易に検索できる環境の整備」の検討を行っており、2002年度は、特許調査における非特許文献調査の重要性を確認した¹⁾。本稿においては、特許・非特許文献の同時検索DBのあり方について、現状の問題点を踏まえた上で検討した。検討に当たっては、知財担当の視点のみならず、技術者・研究者視点及び経営者視点も考慮し、各ユーザー層の固有のニーズに着目した上で議論を進めた。その結果として、あ

くまでユーザー側から見た理想的なDBの提言という形態となったが、今後のDB構築時に少しでもお役に立てばと考える。なお、本稿においてデータベース (DB) は広義の検索システムを指すものとする。

2. 特許文献DBと非特許文献DBの現状及びその検索手法

現状では、特許情報を検索できるデータベース (DB) は有料・無料を問わず多種類に亘る。以下に特許文献DB、非特許文献DBの現状と検索手法を示す。

(1) 特許文献DBとその検索手法

日本の特許情報を検索する場合において、有料 (商用) DBでは、PATOLIS (パトリス)、NRI (野村総研)、ATMS (富士通)、JP-NET (日本パテントデータサービス)、HYPAT (発明通信社) 等が挙げられる。

また、海外の特許情報を検索する場合、有料 (商用) DBでは、DWPI、INPADOC、

* 2003年度 The First Subcommittee, Intellectual Property Information Search Committee

※本文の複製、転載、改変、再配布を禁止します。

USPATFULL, CLAIMS, EPPATFULL, PCTPATFULL (Dialog/STN), Chinese Pat Abs, FRENCH PAT (Dialog), PATDPA (STN), Plus-PAT (Questel-Orbit), Patent-Web (MicroPatent), Delphion (トムソン) 等が挙げられる。

無料DBでは、日本国特許庁HP内のIPDLを始めとして、各国特許庁HP (USPTO, esp@cenet等) が多数存在する。

実務においては、これらのDBを用いて、分類 (IPC, FI, Fターム, US Class, ECLA等) とキーワード等を用いて絞り込み検索を行い、抄録を確認して必要に応じて公報全文を入手し、内容を把握するのが一般的である。

(2) 非特許文献DBとその検索手法

現状では、非特許文献を検索できるDBは、有料 (商用) DBでは、JSTPlus (JOIS/STN), CA (STN), Scisearch (Dialog/STN), INSPEC (Dialog/STN) 等が挙げられる。DialogやSTNで検索可能な非特許文献DBは、前記以外にも、分野ごとに相当数のファイルが存在する。

また、新聞記事 (日経テレコン21, G-Search), 書籍・雑誌記事 (BOOKPLUS, MAGAZINEPLUS), 公開技報 (発明協会) 及び各社技報も非特許文献の一部としてみなされるとともに、特許法29条第1項第3号において、「電気通信回線を通じて公衆に利用可能となった発明」も新規性喪失の理由になることから、インターネット上で公開された情報も非特許文献として考慮せねばならない。これに相当するDBとして、IP.com (Manning&Napier) や、ホームページ登録サービス (発明協会) が挙げられる。

無料DBでは、JOPAL (WIPO) があり、非特許文献にもIPCが付与されているのが最大の特徴である。

これらの非特許文献DBにおいては、一般的にキーワード検索が主体となっている。

また、各々の非特許文献DBには、独自の分類 (JICST分類等) がそれぞれ付与されていることがあるが、特許のような統一性がなく、上位・下位の展開もばらついていることがある。

3. 特許文献DBと非特許文献DBの現状の問題点

前記のように、特許及び非特許文献DBでは、文献別にDBが作成されており、個々のDBにて独自の検索インターフェース (検索画面, 出力画面等) が設けられ、独自の検索ターム (特に分類) にて検索するようになっている。よって、特許文献・非特許文献を検索するための共通的な検索インターフェースや、特許・非特許を通じて共通する検索タームが提供されていないのが現状であり、現時点では、特許・非特許の両者を同一環境下で横断的又は同時に検索することは困難である。

以下、知財担当者や、技術者、研究者等のユーザーが特許・非特許の両DBを使用する際の共通の問題点を考察するとともに、個々のユーザーにおける個別の問題点を考察する。なお、代表的な個々のユーザーとして、①技術者、研究者、②知財担当者、③経営者 (ないしはその代行者) の3者を探り上げることとした。

(1) ユーザーに共通する検索上の問題点

特許・非特許文献の両者を検索しようとするユーザーにとってみれば、前記のように、特許・非特許の両者を同一環境下で横断的又は同時に検索することは困難である。したがって、例えば他社の懸念特許を無効化しようとする調査を、特許文献だけでなく非特許文献に対して実施しようとしても、特許文献DBと同等の精度に基づいた検索を行うことができないといった問題点がある。

※本文の複製、転載、改変、再配布を禁止します。

また、両者を個別に検索しようとするユーザーにとっては、各DB毎に異なる検索タームや検索手法を確実に習得する必要がある。

更には、特許文献DB・非特許文献DBの双方に言えることとして、DBの種類が多岐にわたるため、その中から目的に応じたDBを選択することが困難な場合もある。

これらの①同一検索式での検索が不可能、②各々のDB検索手法の習得が煩わしい、③DBの選択が困難、といった事項が、特許文献DB・非特許文献DBの両者を検索し、両文献の調査結果を活用しようとする際の問題点である。

その結果として、両者を調査する必然性を充分認識していない限り、どちらか一方の検索だけで調査を終えてしまい、本来得られるべき情報を逸してしまう場合もある。

(2) 技術者・研究者の検索ニーズに対する問題点

技術者・研究者にとっては、特許情報は権利情報であるとともに技術情報でもあり、科学技術文献と同様に検索を行いたいというニーズがある。そして、一般に業務多忙な技術者・研究者が業務の合間を縫って検索を行う際のニーズは、明らかに、検索が簡単に行えることである。

しかしながら、現状では特許文献DB・非特許文献DBが別々の検索インターフェースや検索タームを備えており、各DBに対する検索スキルを別々に習得する必要があるため、技術者・研究者に対して、ユーザーフレンドリーな検索環境が整っているとはいえない。

換言すれば、特許文献・非特許文献の各々を検索するDBが別々に組み立てられていることが、技術者・研究者における検索意欲を著しく減退させている可能性が高い。

技術者・研究者にとって、特許文献DB・非特許文献DBは、検索インターフェースが共通であり、しかもその検索タームが複雑ではなく、

適合性の高い結果が簡単にかつリーズナブルなコストにて得られるような環境が必要である。

(3) 知財担当者の検索ニーズに対する問題点

知財担当者が特許に関する様々な調査を行う際には、非特許文献の検索も特許文献と同等の検索環境にて行いたいというニーズがある。

この知財担当者のニーズに対する問題点は、前記の(1)(2)項の問題と同様であるが、更に、無効資料調査等において、特許文献DB・非特許文献DBにて詳細かつ複雑な検索を行いたいというニーズも少なくなく、両DBにてかなり詳細な検索ができるよう、複雑な検索式が立てられ（必要ならばコマンド方式での検索環境の併設）、様々な文献出力方式が選択できることが望まれる。

(4) 経営者向けの検索ニーズに対する問題点

近年、自社が保有する特許等の知的財産における価値評価をし、経営戦略等に活用しようとする動きが散見されることから、経営者（及び投資家）を代表的なユーザーとして設定し、彼らに対する考察も行った。

実際には、経営者本人が検索を実行することは稀と思われるため、その検索代行者が実検索を行うと思われる。その際、具体的な検索に関するニーズとしては、前記の技術者・研究者あるいは知財担当者における検索ニーズと同様であるが、検索結果に関しては、技術者及び知財担当者のニーズと異なる場合がある。

すなわち、経営者にとってみれば、ヒットした個々の文献の内容よりも、トータルとして得られたデータによってマクロ的に解析される結果に興味があり、更に、特許・非特許の両文献が同じDB環境に蓄積されていれば、多種多様の情報解析が可能となる。

しかしながら、現状では、特許文献と非特許文献とを融合させて、例えばある企業の注目技

※本文の複製、転載、改変、再配布を禁止します。

術分野における具体的な開発状況をマクロ的に推察・解析する、といった解析機能を有するシステムが存在しない。

4. 特許・非特許文献の理想DBに関する考察

理想的なDBの考察に先立ち、DBの本来の目的と現状を今一度整理してみる。

DBとは、基本的に、検索結果から検索者が欲する文献を発見するためのツールである。すなわち、検索者が最終的に欲しいものは一次文献である。一次文献としては、特許公報、科学技術文献、一般の雑誌・図書、新聞・雑誌記事、webサイト上のカタログ文書、あるいは個人がwebサイトで公開する文書など、様々なものが考えられる。しかし、現状では文献の種類毎に扱うDBベンダーがまちまちで、各々個別に契約しなければならず、更にはDB利用料金の高額なものが多い。そして、各々ユーザーインターフェースが異なっており、更には検索手法も異なるものが多い。

ところで、今日のインターネット上の検索サイトでは、既に一部の出版社の書籍紹介ページ等が検索可能となっている。すなわち、非特許文献情報の一部については、エンドユーザーレベルでの検索がある程度実現し得ている。もし、これに科学技術文献の抄録文情報と各国特許庁IPDLが発行する電子特許公報の情報が加われば、全てのジャンルの文献を対象とする横断的なキーワード検索が、既存のインターネット検索サイトにて実現できる。

そこで、前記の横断的キーワード検索を実現するための仕組みを考察してみる。

(1) 多くの出版社が、文献の抄録文情報をインターネット上に無償解放すると共に、文献の電子文書等をオンライン販売可能な仕組みを構築する。

(2) 特許の抄録情報を既存の検索エンジン

で検索可能とするとともに、その一次文献である電子特許公報を各国IPDLから検索エンジン経由で容易に入手できるようにする。

例えば、科学技術文献を取り扱う出版社の場合、その出版社の最終目的は、取り扱う文献を販売することである。それならば、最終的に検索者が欲する一次文献がスムーズにオンライン販売にて入手できれば、検索者にとっても、出版社にとってもプラスになるものと考えられる。そこで、出版社が文献の抄録文情報をインターネット上に無償解放すれば、インターネット上の検索エンジンを通じて目的の文献がキーワード検索にて入手可能になる。つまり、従来はDBベンダーが負担していた、検索DBの基となるデータを提供する作業の一部を、文献提供者側が負担する考え方である。もちろん、ユーザーによってはすべての種別の文献を必要としないこともあるので、各抄録文情報には文献種別情報を付加し、必要な情報のみを検索できるようにすればよい。

次に、以上の考え方を踏まえた上で、前章の特許文献・非特許文献におけるDBの様々なニーズ及び問題点に対して、理想的なDBのあるべき姿を、各ユーザーの視点に立って考察することとする。

4. 1 理想DBの概要

理想のDBとして搭載すべき機能・特徴として、①ユーザーに応じた検索環境を選択できる、②過不足のない文献が収録される、③検索コストが妥当なものである、といった点が挙げられる。

以下、図1に基づいて説明する。まず、検索するユーザーは様々であり、基本的にはどのようなユーザーにも対応可能とすることが望まれる。

そして、全ユーザーに対応可能なDB検索画面が設定されるべきであり、当然、ある程度の

※本文の複製、転載、改変、再配布を禁止します。

情報までは無料で検索できる環境の提供が必要と考える。そこで、インターネットによる接続環境のもと、既存の無料検索サイトなどが全ユーザーの基本的な検索画面として提供されるのが望ましい。

この検索サイトにおいては、特許文献については各国特許庁HPに掲載されている抄録情報を、また、非特許文献（科学技術文献のみならず新聞、雑誌、更には製品カタログ等の情報も含めた広範囲の文献情報であることが望ましい）については、それらの文献を取り扱う出版社や書籍販売業者が発行する抄録情報を検索対象とする。そのためには、特許情報の抄録については、現状のように各国特許庁HP内でのみ検索可能とするのではなく、無料検索サイトでの検索もできるよう前記情報をインターネット上に解放する。また、非特許文献抄録については、出版社や書籍販売業者が発行する文献の抄録を、自身が運営するwebサイトに掲載し、無償で検索可能とする。更に、これら抄録情報はXMLにて解放し、文献種別情報が付加された状態とする。ユーザーは、この検索サイトにて、各国特許庁HPや出版社、書籍販売業者webサイトの抄録情報をシームレスに検索・抽出することができ、しかも文献種別情報により検索したいジャンルを自由に選択できるので、絞り込み検索が容易となる。

一方、知財担当者にとっては、簡単な検索を行う場合、このような無料検索サイトで所望の文献ジャンル内で検索を行えば事足りることもある。しかし、より複雑な検索を行う場合は、キーワードだけで検索を行う検索サイトでは要求が満たされないので、既存の各種特許・非特許文献DBや、特許・非特許文献をシームレスに検索できる特許・非特許統合DB（4.3節にて後述）といった複雑な検索ができる環境（有料DBサービス）を選択すればよい。無料検索サイトにて予備検索を行い、有料DBサービスに

て詳細検索を行うといった利用方法も考えられる。

更に、経営者・投資家にとってみれば、特許・非特許文献をシームレスに検索した結果を基に様々な解析が行えるデータ解析サービスへインターネット経由で接続できる環境が整っていればよいと考える。

このような環境がインターネット上に構築されれば、様々なユーザーに応じた検索ニーズを確実に満たせる理想のDBが構築されるものと考えられる。

ただし、当然のことながら、このような理想のDBを構築するに当たっては、無料で抄録情報を提供する出版社や書籍販売業者の利益も確保できるビジネスモデルを構築する必要がある。例えば、図1にもあるように、抄録までは無料提供するが、一次情報（全文PDF等）を有料にて販売することで対応できる可能性がある。

情報を提供する出版社や書籍販売業者にとってみても、抄録情報の無料提供のためのコストがかかるものの、この情報をインターネット上に掲載することで、一次情報の販売促進につながられれば、抄録作成コストを吸収できるはずである。また、多少精度が低下する恐れがあるものの、著者抄録があるものについては、それを利用することも考えられる。勿論、コストとの兼ね合いで、有料DBサービスに提供される抄録文とは、充実度に差異が設けられていても構わない。

情報提供側の確実な利益確保を達成しつつ、ユーザーにとっても過不足ない情報を得られる環境が構築されるようにすれば、理想のDBへの実現性が向上すると確信する。

4. 2 技術者・研究者向けDB

技術者・研究者向けの検索ニーズは、先に述べたように、検索が簡単に行えること、すなわちインターフェースが単一で、キーワードで簡

※本文の複製、転載、改変、再配布を禁止します。

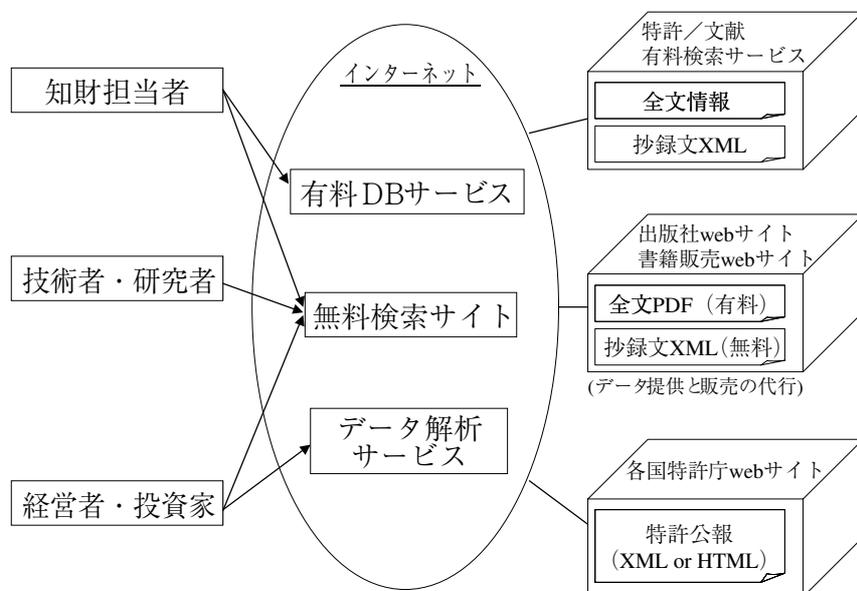


図1 理想DBの概要

単に検索でき、特許・非特許文献に関わらずシームレスに結果が引き出せることである。そして試行錯誤を許容するため、検索料金はできれば無料が望ましい。

そして4.1節で述べた理想DBの概要のうち、無料検索サイトの部分がこれらのニーズを満足させるものに相当する。

特許情報は、所謂サーチャーだけのものではなく、一般ユーザーも一情報として利用する時代となっている。特許・非特許として分かれて存在しているDBを同時検索するという意識ではなく、一般ユーザーが何も考えずに検索をしても、シームレスに特許・非特許の両方の情報が得られるという環境になるのが理想的と考える。

4.3 知財担当者向けDB

次に、知財担当者向けの理想の特許・非特許統合DBについて考察することとする。

知財担当者は、近年多種多様化する特許情報DBの中でも、パトリスのようにコマンド形式のDBを利用することが多く、論理演算を繰り返して試行錯誤することで大量の特許情報の中からの確で適量の検索結果を導き出している。

また、知財担当者にとって非特許文献調査は、特許と同様の技術情報としての価値があるものの、権利としての機能を持たない非特許文献を必須の調査対象とせず、主に特許調査を補完するものとして利用され、特に無効資料調査の対象として位置づけられている。

しかしながら、技術開発が一刻を競う今日、知財担当者にとっても、公開特許公報よりも早期に公開される可能性の高い非特許文献を調査することは、分野によっては出願前の先行技術調査の情報源となるものである。

以下、知財担当者向けの特許・非特許統合DBについて説明する。

特許と非特許文献のシームレスな検索の最大の利点は、技術者を対象とした場合と同様に検索の二度手間を省くという点にある。前記のように知財担当者は、論理演算を繰り返すなどの複雑な検索を行う場合が多く、このような場合に検索の二度手間が省けるということは大変な効果がある。この意味において、前記の無料検索サイトと同様に、既存の特許DBあるいは非特許文献DBもインターネット上に公開される抄録文情報を収集してインデキシングすることに

※本文の複製、転載、改変、再配布を禁止します。

より、文献種別を超えた横断的なキーワード検索が低コストで実現できるだろう。

そして、知財担当者の利用を促す上で一番大きなウェイトを占めるのが、多様な検索タームと、それぞれの検索タームの特許・非特許の統合性であり、特に技術分類の統合性は不可欠である。

すべての特許文献には世界で統一された技術分類（国際特許分類）が付与されている。したがって、非特許文献情報にも、この特許と統一した国際特許分類を付与することができれば、知財担当者も絞込みが容易になり適合率が高い検索を行えるようになる。インターネット上で無償公開されている非特許文献の抄録文に対して、概念検索等の手法で特許公報との類似性を算出し、国際特許分類を機械的に付与することにより、無料検索サイトでは為し得ない、国際特許分類を使った非特許文献及びwebサイトの検索が、特許文献と共にシームレスに実現できるだろう。国際特許分類を付与する場合、サブグループまで付与されていた方が好ましいが、メイングループ、サブクラスまでの付与であっても、ある程度の検索精度向上は見込めるものとする。

また、特許文献と非特許文献との書誌事項の異なり、すなわち、出願人と所属機関、発明者と著者、公開日と発行日とを共通して検索できる機能は必須である。更に非特許文献には、所属機関名として、略称名が記載される場合もあり、正式社名に置き換えるなどのデータの統一を図る必要がある。

また、キーワード検索を可能にするために、最低限、抄録の収録は不可欠であり、更に抄録に対する概念検索機能が付加されれば、多様なニーズにも対応できる。また、無料検索サイトよりも精緻な検索を実現するために、より詳細な抄録をDB業者が自前で作成したり、出版社等の一次文献提供元から有償で導入する等の方

策も有効であろう。そしてより好ましくは、文献の全文テキスト化が可能になれば、より詳細且つ精緻な検索が可能となり、知財担当者のみならず全てのユーザーが高精度で検索できるようになる。

更に、引用文献からの特許もしくは非特許文献への相互リンクは、利用頻度が高いと思われる、利便性の点から付加が望まれる機能である。

このような項目や機能が付加され、収録内容も豊富であるならば、利用料が有料であっても利用頻度は高いものと思われる。また一次情報提供者との関係で全文収録に問題が生じるのであれば、全文は検索用に収録し、その入手は技術者・研究者向けのDB同様に一次情報提供者の各webサイトで直接入手する形態も有効であると考える。

4. 4 経営者向けDB

経営者（及び投資家）をユーザーと捉えた場合、前記のように、検索利便性等のニーズよりも、結果出力としてどのような情報を得ることができるかが重要となる。知的財産情報を、経営戦略や企業価値評価等に有効活用させるためには、マクロ的な解析とそのビジュアル化が要求されるからである。

例えば、特許情報をマクロ解析しビジュアル化するツールとして、パテントマップという手法が存在し、そのシステムも数多く市販されている。しかしながら、作成されたマップがどれほど経営戦略に役立っているかには若干の疑問がある。その原因の一つとして、実施予定のないアイデア出願の存在が挙げられる。一般的にマクロ解析においては、個々の出願が実施され得るものか否かまでは解析できないため、全ての出願が同格として扱われ、情報の不確かさを露呈する恐れがある。ところが、仮に特許情報と科学技術文献情報とが同じDB環境に蓄積されていれば、科学技術文献に発表された発明は

※本文の複製、転載、改変、再配布を禁止します。

実施される可能性が高いと想定できるため、相互の相関関係を解析することにより、ある程度の確率で実施され得る出願を抽出することが可能となる。更にそのDBに、新聞・雑誌情報等やカタログ情報が含まれていれば、製品紹介などの記事・記載から、前記の実施発明の抽出確率を高めることが可能となる。このような情報をマップ化することにより、より精度の高い情報解析が可能となり、経営戦略や企業価値評価等に有効に活用できるものと考えられる（例えば、図2参照）。

更に、科学技術文献情報や、企業情報（電子カタログを含む）、新聞雑誌情報、経営情報などにより、特許情報の非速報性（公開まで1年半）を補うことが可能となり、技術トレンドを解析するのに効果的である。

このように、特許情報と非特許情報とが同じDB環境に蓄積されていれば、多様なデータ解析を行うことが可能となり、更にそれらをビジュアル化することにより、経営者（及び投資家）のユーザーを満足させるDBが実現できるもの

と考える。

5. まとめ

以上説明した理想のDBが実現すれば、特許・非特許文献の垣根を越えたシームレスな検索が可能となり、それぞれのユーザー層のニーズに応じた情報を得ることが容易となる。

ところで、現在の特許文献及び非特許文献の検索環境は、未だ保護に手厚く傾いているものと感じている。これは先行技術調査の精度を阻害し、出願件数の増大を招き、特許庁においては実体審査処理を遅延させてしまう。「車輪の再発明」と言われるような、重複発明の無駄な特許出願を省くことは、出願人、特許庁だけでなく、社会全体にプラスに働く。そのためには、できるだけ短時間に、低コストで目的の技術文献を見つけられる、超高精度な情報検索環境の構築が急務である。そのためには、情報が備える価値を正しく評価し、適切な形で利用できるよう環境が整っていることが望まれる。

例えば、情報には鮮度がある。鮮度の高い情

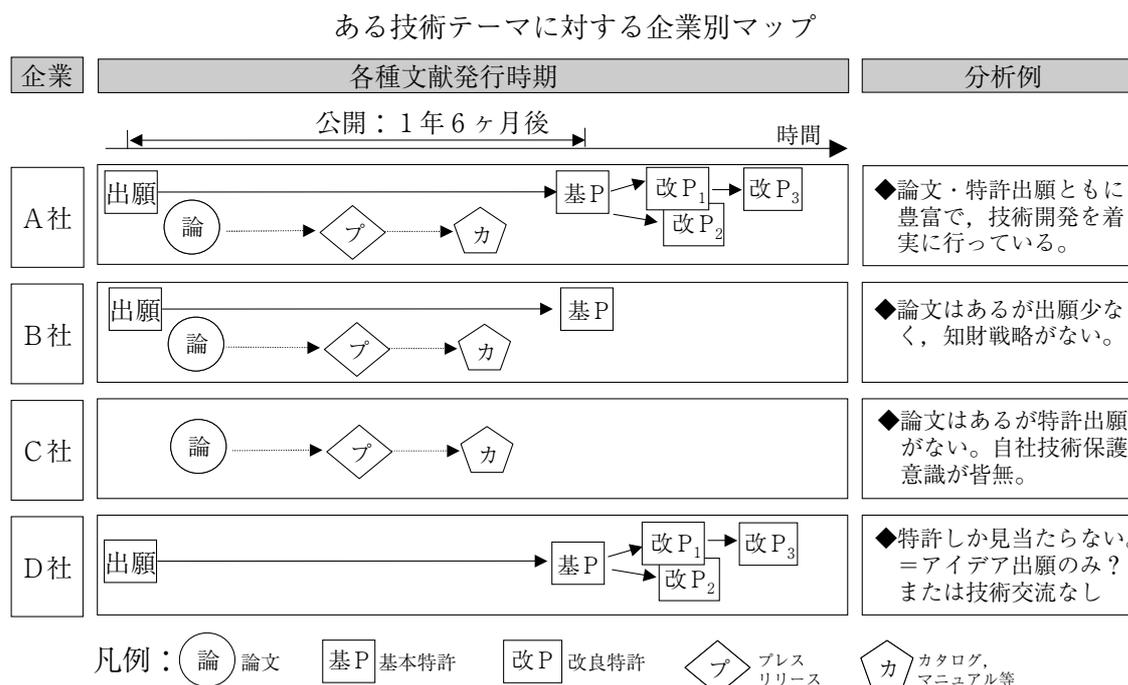


図2 特許・非特許文献の融合マップ例

※本文の複製、転載、改変、再配布を禁止します。

報は付加価値が高い。そのような情報は有料DBでのみ利用可能になっていて良いだろう。一方で、所定の期間を過ぎて鮮度の落ちた情報は公共利用に供するようになれば、特定のDBに依存しない、情報検索環境の強化が図られる。

文献の抄録文情報を適切な形で公共利用に供し、それを無料検索サイトや有料DBが利用することにより、知的財産戦略大綱に記されている「特許情報とそれに関連した技術情報を、研究者が容易に検索できる環境の整備」、ひいては産業の発展に繋がるものと確信する。

なお、本検討に携わった2003年度知的財産情報検索委員会第1小委員会委員は、臼井裕一（副委員長，住友電気工業），佐久間幹雄（小委員長補佐，ブラザー工業），岩本広美（三井化学），小林 敦（トキコ），北條由貴子（住友電装），松田陽一（ティアック），安田真幸（デンソー）である。

注 記

- 1) 知的財産情報検索委員会第1小委員会，知財管理, Vol.53, No.9, pp.1485～1493(2003)

（原稿受領日 2004年7月21日）

