

コンピュータソフトウェアデータベース 電子化情報作成事業について

財団法人ソフトウェア情報センター
常務理事
ソフトウェア特許情報センター長
永田 雅博

PROFILE

1995年特許庁審査第四部審査長、1999年審判部書記課長、2001年審査業務部情報システム課長、2002年工業所有権研修所長を経て、現職



1

はじめに

コンピュータソフトウェアデータベース（CSDB）はソフトウェア関連発明の審査の運用指針が改定されたことにより、1997年からCSDBの構築が開始されており、2006年度まで38万6千件のデータが蓄積されている。「知的財産推進計画2005」の「特許情報の利用環境を整備する」項においては、CSDBに関し、引き続き、新たな文献の充実や検索機能の向上、著作物利用許諾が得られた一次文献・抄録はIPDLを通じて公開する施策が挙げられている。

以下、CSDB電子化情報作成事業についての概説である。

2

ソフトウェア特許情報センター 発足の経緯

(1) 審査の運用指針及び審査基準の改訂

特許庁は、ソフトウェア技術の著しい進展に対応するため、コンピュータソフトウェア関連発明の審査の運用指針を改訂し、1997年4月より新しい指針の運用を開始し、これにより、「プログラムを記録した記録媒体」又は「構造を有するデータを記録した記録媒体」が特許の対象となることとなった。

さらに、2000年12月には、コンピュータソフトウェア関連発明の審査基準を改訂し、「媒体に記録されていないコンピュータプログラム」を「物の発明」として

取り扱うこととした。また、併せて、ビジネス関連発明に関する記述も明確にされた。

そして、その後、ソフトウェア等情報財の特許保護強化とネットワーク取引の促進等を図るため、特許法等の一部が改正（プログラム等が特許法上の「物」に含まれることを明記）され、2002年9月1日に施行された。

(2) データベース構築の必要性

コンピュータソフトウェア関連分野は技術開発のスピードが速いうえに、前述の審査の運用指針改訂に伴い、それまで特許と無関係と考えられていた、例えば、電子マネー、電子商取引等のビジネス関連分野に関する特許出願が増加することが予想された。

このような事情から、ソフトウェア関連発明（「ビジネス関連発明」を含む）に係る出願の審査については、特許文献を中心とした既存の先行技術文献調査システムでは十分に対応することができない面があった。

このため、特許庁では、業界等からの強い要望もあり、コンピュータソフトウェア関連分野の先行技術調査を充実させるため、1997年度から、コンピュータソフトウェアマニュアル（頒布されているがその入手が比較的難しい）、単行本、雑誌、学会論文誌、企業技報等のビジネス関連文献を含む非特許文献を収録するコンピュータソフトウェアデータベース（以下、「CSDB」という。）を構築することとなった。

(3) ソフトウェア特許情報センターの設置

特許庁が構築するCSDBに蓄積する電子化情報を作成するために、1997年7月、(財)ソフトウェア情報センター（略称、「SOFTIC」）にソフトウェア特許情報セン

ター（以下「センター」という）がその附属機関として設置された。

センターでは、CSDBに有用な非特許文献の収集と、収集した文献を解析して、検索キーであるコンピュータソフトウェアターム（CSターム）の付与、フリーワードの抽出、抄録の作成を行い、それら文献の一次文献情報や解析情報等を電子化情報として作成し、特許庁へ納品する業務を行なっている。

3 CSDBについて

(1)CSDB電子化情報作成

①CSDB構築に必要な非特許文献の収集

収録される文献は、マニュアル（コンピュータソフトウェアマニュアル、ビジネスマニュアル、ゲームソフトウェアマニュアル）、単行本、技術雑誌、非技術雑誌、学会論文誌、企業技報、団体機関誌及び学会予稿集であり、コンピュータソフトウェアデータベース検討委員会（委員長：相澤英孝一橋大学教授 SOFTICに設置）及び特許庁の承認を経て収集している。

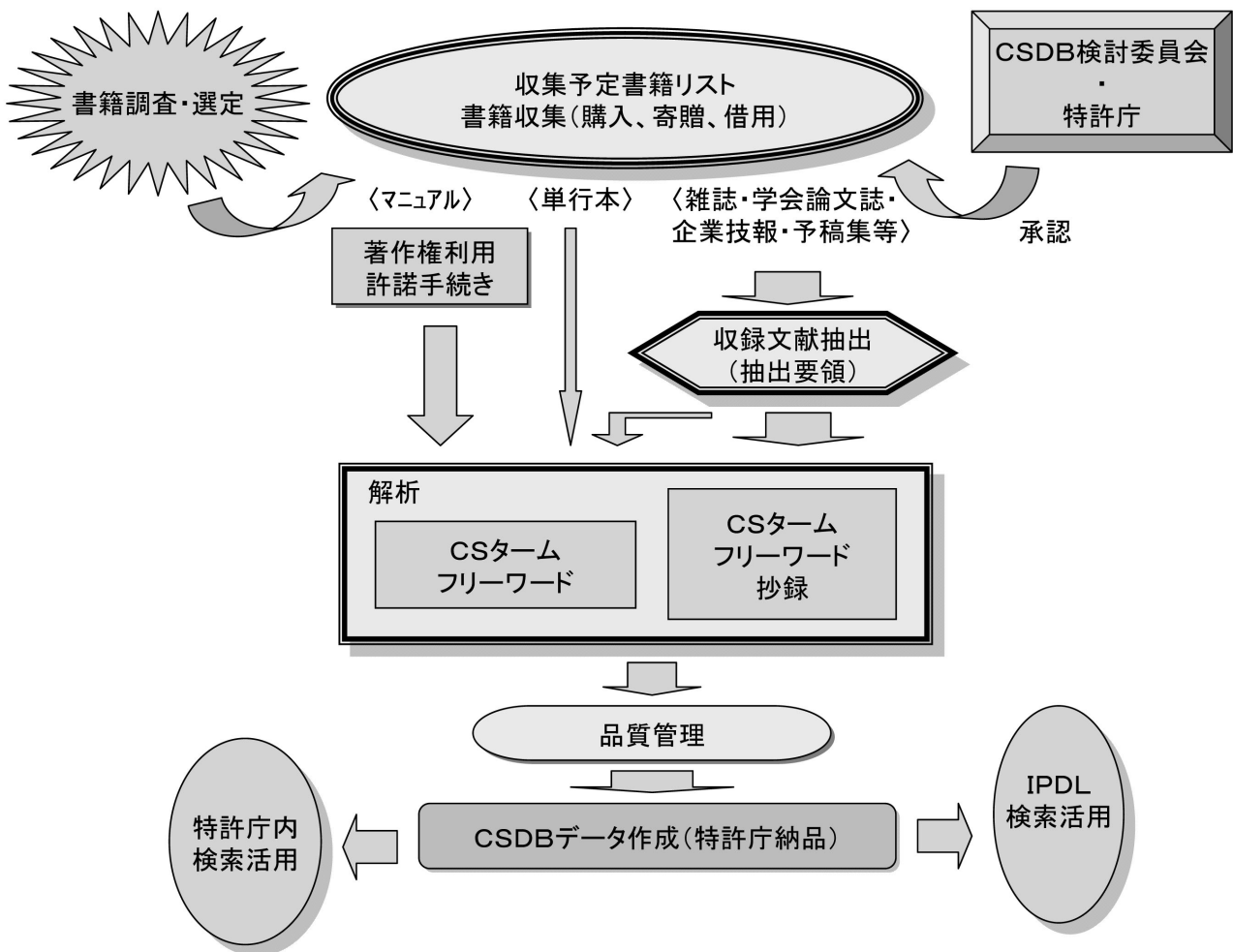


図1 概念検索で参照した特許を検索式で除く利用法

電子媒体（CD-ROM）については、2005年度から電子媒体のみで発行されているコンピュータソフトウェアマニュアルの収集を開始した。

②CSDBのIPDLによる外部公開

特許庁は、2003年5月から、それまでに蓄積したCSDBの電子データのうち、著作物利用許諾を要しない書誌事項等に係る電子データは、IPDLを通じて外部公開を行っている。さらに、2004年9月からは、他の商用データベース等で入手することが困難なマニュアル類を中心に、特許庁が著作物利用許諾を得られたものは、それら一次文献及び抄録も公開している。

センターとして、これにあわせて、著作物利用許諾手続きに係る協力と、その許諾情報を電子化情報として取り込みを行っている。

③収録文献の抽出・解析

収集した文献のうち、当センターの主席部員がCSDB構築に必要な文献の抽出を行い、次いで、外部解析者がCSターム解析マニュアルに従ってCSタームの付与、フリーワードの抽出、抄録の作成を行い（マニュアル・単行本等については主席部員が解析を行っている）、これら解析結果については、主席部員による品質管理（内容校閲・指導）が行われる。

④電子化情報の作成

解析済みの文献については、予め、イメージ情報を作

成すると共に、付与されたCSターム、抽出されたフリーワード、作成された抄録並びに各文献の書誌事項、目次、索引等をテキスト情報に変換し、電子化情報を作成する。（[図1] 参照）

⑤電子化情報の作成状況

CSDB電子化情報の年度別作成状況は表1のとおりであり、2006年度末における累積（実績）は38万6千件に至っている。

(2)CSDB構成の概要

CSDBは、検索用DBと文献照会用DBの2つのデータベースから構成されている。

検索用DBには、各文献の一次文献、二次文献（抄録）のテキスト情報、書誌事項、目次・索引情報、各文献に付与されたCSターム、各文献から抽出されたフリーワード等が蓄積され、また、文献照会用DBには各文献の全頁がイメージ情報として蓄積されている。

検索は、CSタームを用いたインデックス検索と、フリーワード検索、書誌事項、目次・索引、抄録、一次文献のテキスト検索の組み合わせで行う。なお、IPDLでの検索結果は、著作物利用許諾が得られている文献を除いては、書誌事項、CSターム、フリーワードが公開されている。

(3)CSタームの概要

CSタームは、特許文献の検索用データベースのため

表1：CSDB電子化情報の年度別作成状況（実績）

（単位：件数）

文献種別／年度	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
マニュアル	2	4,602	4,513	3,211	3,911	4,362	4,374	4,457	4,435	3,950
単行本	305	208	387	1,689	989	538	526	443	465	350
雑誌	4,513	18,758	18,656	19,260	19,764	23,584	26,656	26,764	20,121	24,656
学会論文誌	1,336	11,370	12,133	11,910	11,420	8,218	7,243	13,249	14,374	10,623
団体機関誌	0	0	0	0	578	36	104	61	43	31
企業技報	1,319	2,562	2,212	1,830	1,238	1,162	1,597	1,926	2,064	1,403
学会予稿集	1	52	29	22	22	22	24	26	8,127	11,295
合計	7,476	37,552	37,930	37,922	37,922	37,922	40,524	46,926	49,629	52,308
累計（実績）	7,476	45,028	82,958	120,880	158,802	196,724	237,248	284,174	333,803	386,111

（注）学会予稿集：2005年度より論文単位に作成

に開発されたFターム (File Forming Term) と同様に、ある概念を表すインデックスである。CSタームはコンピュータソフトウェア関連技術において、国際特許分類等の既存のインデックスでは検索しにくいものも検索可能となるように開発されている。

また、CSタームは252のタームから構成され（表2参照）、多くのソフトウェア関連の文献に共通する観点を上位レベルに抽出して構成しており、タームの数が膨大になるのを防ぐため、一分野に特有な技術までは細展開されていない構成となっている。

表2：CSタームの大分類とターム数

	大分類	タームの数
BB	ハードウェア	6
CC	コンピュータアーキテクチャー	9
DD	ソフトウェア	15
EE	データ	7
FF	計算機関係理論	10
GG	計算機数学	21
HH	コンピュータ方法論	35
JJ	情報システム	78
KK	コンピュータ応用分野	62
LL	その他	9

