

# 平成 2 0 年度 特許出願技術動向調査報告書

## インターネット社会における検索技術 (要約版)

### < 目次 >

第 1 部 調査概要および情報収集方法 .....	1
第 2 部 特許出願動向分析 .....	14
第 3 部 基本特許・重要特許分析 .....	29
第 4 部 研究開発動向分析 .....	37
第 5 部 市場・政策動向分析 .....	45
第 6 部 総合分析 .....	52

平成 2 1 年 4 月

特 許 庁

問い合わせ先  
特許庁総務部企画調査課 技術動向班  
電話：03 - 3581 - 1101 (内線 2155)

## 第1部 調査概要および情報収集方法

### 第1章 検索技術の概要

#### 第1節 調査の目的と背景

近年のインフラ整備によるブロードバンド化がインターネットの普及を推し進める一方で、大量の情報の中から必要な情報を見つけ出す情報検索技術がインターネットの活用を促している。特に、インターネット利用者が情報にアクセスするための手段として、検索サービスが一般的に使われている。

これまでの動向を簡単にまとめると、1990年前半に、HTTPとHTMLの標準化や表示するためのWebブラウザの開発されたことや、インターネット対応するOSが一般向けに提供されたことでWeb利用者が爆発的に増加した。また、Webの利用とインターネットの接続サービスの拡大による相乗効果により、インターネットの利用の大部分はWWW(World Wide Web)の利用と認識されるようになってきている。

また、一般的な情報の流通は、Webサイトからインターネットを経由して、利用者のブラウザで表示される、という経路となっている。インターネットには、一般的な情報の所在を知るための検索機能がないため、検索サービスに対するニーズが高い。そのため、検索サービスは、独自に情報を収集・集積し、収集した情報を解析することにより情報検索機能を実現している。実現されている検索機能において、扱う情報の種別や情報収集の方法、検索結果など様々な点で技術的な差異が存在し、プレイヤー間の競争につながっている。例えば、現在検索サービスの市場シェアの高い米国グーグル社の検索機能における中心的な技術とされるPageRankの特許<sup>1</sup>が1997年に出願されていることから分かるように、WWWの普及に合わせて本格的に開発が行なわれていると考えられる。

このような背景のもと、情報検索技術に関する特許出願件数についても著しい増加が見られる。本調査では、「インターネット社会における検索技術」の分野について、特許出願技術動向調査として、出願された特許と公開されている論文、各種文献等を対象に調査分析することにより、技術発展状況や研究開発状況を明らかにし、日本および外国の技術競争力や産業競争力を明らかにすることが目的である。加えて、日本が取り組むべき課題や今後の展望として目指すべき研究開発、技術開発の方向性について検討するものである。

#### 第2節 本調査で対象とした検索技術の定義

本調査において、「インターネット社会における検索技術」を以下のように定義する。

「インターネットをベースとする検索技術」を対象とするものであり、つまり、

・インターネットを利用し、検索式を与えることにより、情報を抽出する仕組み

と定義する。

いわゆる検索サービスの中で使われる検索エンジンを構成する技術を中心に詳細な分析を行なうものとする。より一般的な技術範囲としては、インターネット上で流通している情報、

---

<sup>1</sup> スタンフォード大学: “Computer-implemented linked document scoring method applicable for analyzing linked databases involves processing linked documents based on linked document scores assigned according to linking document score” (US6285999B1, 1997年出願)

あるいは、インターネットを経由してアクセスできる情報、を対象として、検索機能を実現する上で必要な技術要素を扱う。

「検索技術」を構成する主な技術要素としては、以下の技術が挙げられる。

- ・ 情報収集・更新技術：流通する情報を収集する
- ・ インデキシング技術：高速に検索するためにインデクス情報を付与する
- ・ 一致検出技術：クエリとして与えられた情報に合致する情報を出力する

また、情報そのものを解析する技術に加えて、検索機能を利用するための入出力に関連する技術と基盤となるシステム技術が存在する。

なお、検索技術には様々な基礎技術が関与しており、応用範囲も広いため、例えば、

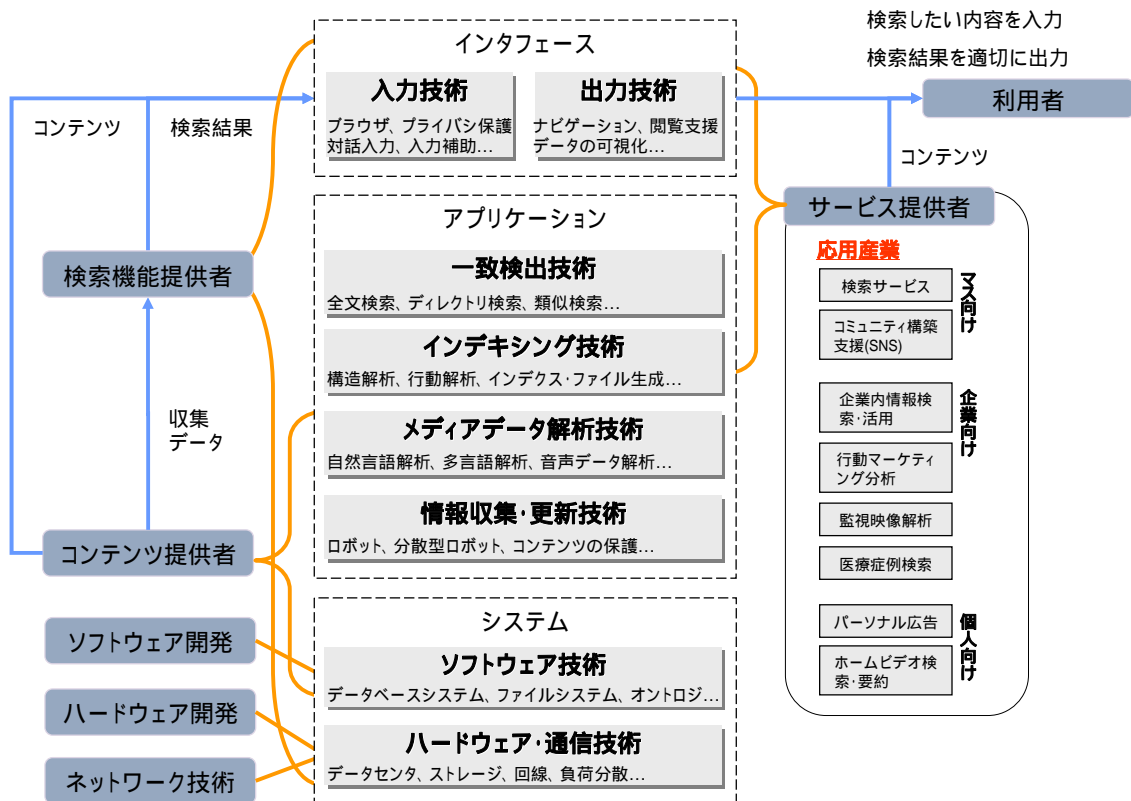
- ・ インターネットに関係しない検索技術（書籍の内容検索技術など）
- ・ 特定メディアの汎用的な解析技術（画像特徴量の抽出技術など）
- ・ 検索技術の応用分野に関する基盤技術（カーナビのハードウェア技術など）

といった技術も調査対象として考えられる。こうした技術についても、後述する調査対象範囲については技術分類に従って分類するが、その詳細な分析については行なわないものとする。

### 第3節 検索技術分野の技術俯瞰図

先に示したように、検索技術分野における要素技術の構成、各機能を提供するプレイヤー、および、情報の流通の関連性を技術俯瞰図として、図 1-1-1 に示す。

図 1-1-1 技術俯瞰図



- ・各要素技術は、インタフェース、アプリケーション、システムの3分野に大別できる。
- ・プレイヤーとしては、検索機能の提供者と利用者だけでなく、検索対象となるコンテンツの提供者、ソフトウェア/ハードウェア/ネットワークといったシステム共通基盤の開発者、さらに、検索技術を応用したサービス提供者が考えられる。
- ・マス向け、企業向け、個人向けにサービスを提供する様々な応用産業が存在する。

#### 第4節 本調査における技術の分析軸

調査対象となる特許、および、論文について、それぞれがどのような特徴を持つのか、分析するための軸として、表 1-1-1 に示すように「要素技術」「メディア」「課題」の3軸を設定した。それぞれの分析軸の関係を図 1-1-2 に示す。

要素技術としては、図 1-1-1 の技術俯瞰図に示したように検索技術に関連する技術として、「システム」「アプリケーション」「インタフェース」の3つの分野に分類することができ、その下にそれぞれ大区分、小区分が存在する。詳細については、「1 技術軸の詳細」に示す。

また、メディア軸としては、検索あるいは処理対象として、どのような情報を扱うものかを示す軸を設定する。文字情報、マルチメディア情報を中心として、14 の区分に分類する。詳細については、「2 メディア軸の詳細」に示す。

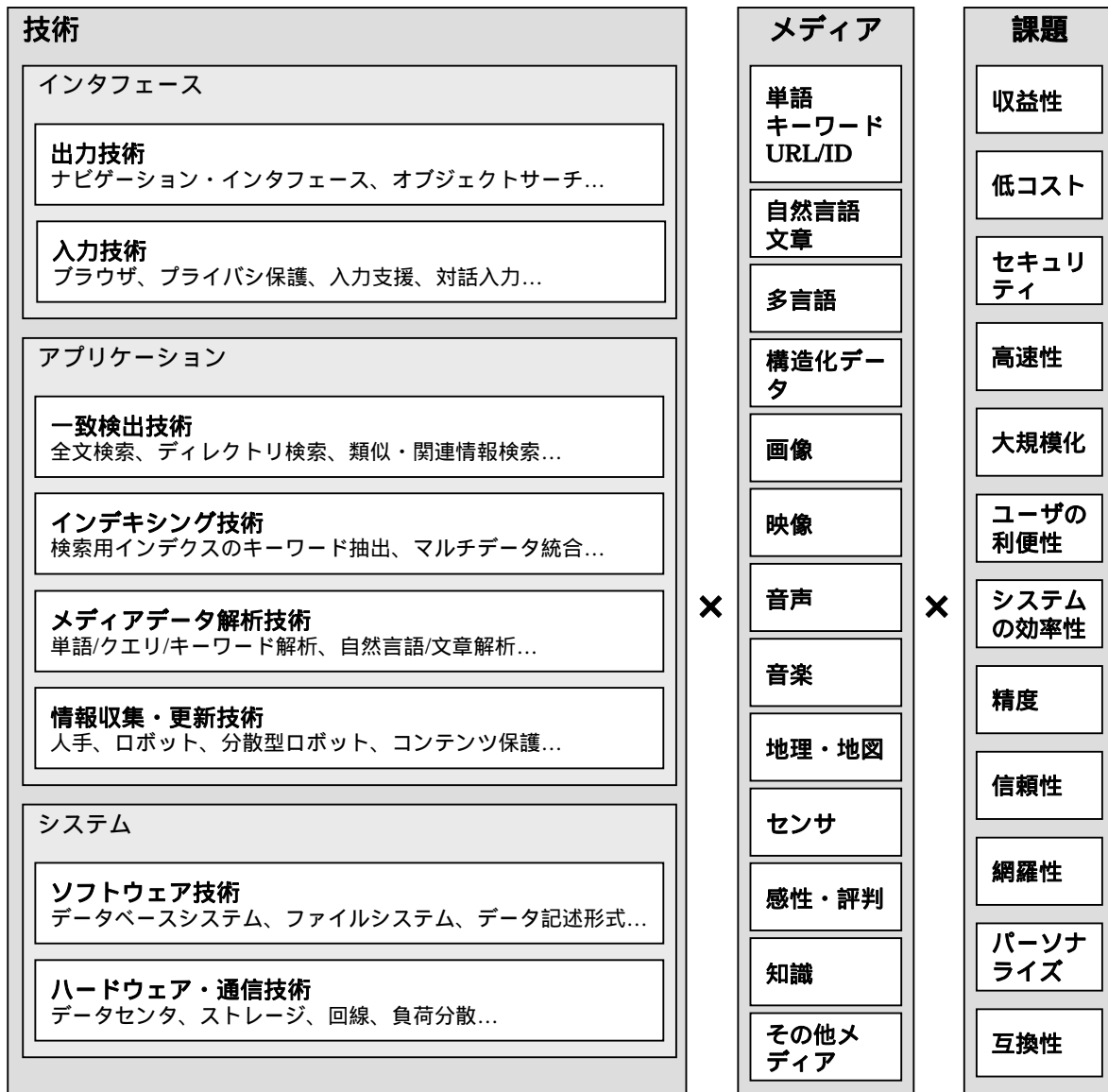
さらに、課題に関する分析軸として、特許および論文が解決しようとする課題に着目して分類したものであり、収益性、低コスト、高速性、利便性向上等の12の区分を与えている。詳細については、「3 課題軸の詳細」に示す。

特許、および、論文はそれぞれ、技術、メディア、課題の各軸において一つ以上に分類されると考えられる。例えば、課題として「ユーザビリティ」を向上させる「音声」メディアに対応した「技術：ブラウザ」などというように、各軸において1つ以上の分類が行なわれる。特許、論文によっては、複数の技術、複数のメディア、複数の課題に関連する場合も多い。なお、いずれにも分類されないものについては、今回の調査対象外の特許および論文として扱う。

表 1-1-1 本調査における分類軸

分析の観点	内容	詳細
要素技術	どのような技術要素を利用しているか、という観点から技術俯瞰図に示した要素技術に基づき分類・分析を行なう。3つの技術分野に対して、8つの大区分を用意している。	表 1-1-2
メディア	どのようなメディアを対象にしているか、という観点から分類・分析を行なう。	表 1-1-3
課題	どのような課題を解決するものか、という観点から分類・分析を行なう。	表 1-1-4

図 1-1-2 特許および論文の分析軸



1. 技術軸の詳細

まず、各特許、論文について、どのような技術を用いているかで分類する。これにより、各国の特許でそれぞれどのような技術が近年注目を集めているか、といったことが読み取れることができる。

技術としては、まずシステム/アプリケーション/インタフェースの各適応分野で大別し、さらに以下に示す大分類、小分類に分類する。

各分類の概要説明、想定される技術例と共に、表 1-1-2 に技術軸として一覧を示す。

表 1-1-2 技術軸

分野	大分類	小分類	概要	対象技術例
システム	ハードウェア・通信技術	データセンタ	特別なデータセンタの構築、活用	データセンタ監視
		ストレージ	特別なストレージ、ハードディスクの構築、活用	ハードディスク、RAID、分散ストレージ
		回線	回線の高速化、ネットワーク構築	ブロードバンド回線、無線回線
		負荷分散	複数のコンピュータを用いて負荷を分散させる方法	ロードバランサー
		並列処理	複数のコンピュータを用いて演算の大規模化に対応、処理の高速化させる方法	グリッド、クラスタ
		その他ハードウェア・通信技術		
	ソフトウェア技術	データベースシステム	特別なデータベースシステムの構築、活用	RMDB、XMLDB
		ファイルシステム	特別なファイルシステムの構築、活用	ext、ReiserFS、XFS
		プログラミング言語	特別なプログラミング言語の構築、活用	Java、perl、PHP、Python、Ruby
		データ記述形式	データの内容や構造、意味を記述するためのフォーマット	XML、JSON、RDF、OWL、SPARQL、セマンティック情報の記述
		キャッシュ	キャッシュの特別な利用方法	キャッシュ、中間ファイル、プロキシ
その他ソフトウェア技術				
アプリケーション	情報収集・更新技術	人手	人手による情報収集、更新の方法	ディレクトリ型サーチエンジンの構成
		ロボット	一般的なロボットを用いた情報収集、更新の方法	探索エージェント、フォーカストクローラ
		分散型ロボット	多数のロボットを用いて行う情報収集、更新の方法	分散型ロボットの基本形
		P2P型ロボット	複数のロボットが自律的に通信しながら行う情報収集、更新の方法	P2P型情報収集
		コンテンツ保護、不正利用防止	収集した情報の著作権を保護するなど、コンテンツの不正利用を防止する方法	DRM、電子すかし
		その他情報収集・更新技術		
		メディアデータ解析技術	単語/クエリ/キーワード解析	それぞれの単語の重要度などの解析
	自然言語/文章解析		自然文の構造や意味を解析する技術	分かち書き、形態素解析、構文解析、意味解析
	構造化データ解析		XMLなどの構造化されたデータを解析する技術	セマンティックウェブ
	多言語解析		日本語、英語以外の言語の処理や、複数言語に対応した処理に関する技術	オントロジー、文字コード、機械翻訳、多言語横断検索
	画像データ解析		画像データを対象に解析して、インデックスとして利用可能な情報を抽出する技術	画像特徴量抽出
	映像データ解析		映像データを対象に解析して、インデックスとして利用可能な情報を抽出する技術	映像特徴量抽出、シーン検出、ハイライト検出
	音声データ解析		音声データの重要度などの解析	音声認識
	音楽データ解析	音楽データの重要度などの解析	音楽特徴量抽出	

分野	大分類	小分類	概要	対象技術例	
		地図、地理データ解析	地図、地理など実世界の位置情報を用いた解析	GPS、3Dモデル処理	
		センサデータ解析	多様なセンサから得られた情報を用いた解析	体感センサ、加速度センサ、人感センサ	
		感性・評判データ解析	「きれい」「かわいい」「よい」といった印象、嗜好、評判など感性データを用いた解析、真理ではない人の感情や感覚のデータの解析	感性特徴量抽出、評判の数値化	
		知識データ解析	Wikipediaにあるような知識や、イントラネットのQ&Aといったデータの解析、真理となるようなデータの解析	知識体系の生成	
		その他メディアデータの解析技術	その他のデータの解析		
	インデキシング技術	検索用インデクス・キーワードの抽出、作成、更新	インデクス作成のためのキーワード抽出、記述方法	キーワード抽出、記述言語、ハッシュ生成	
		検索用インデクス・ファイルの生成、更新、維持	インデクス用のファイルの作成、更新、管理方法	転置ファイル(インバーテッドファイル)、索引更新方式、リアルタイムインデキシング	
		検索用インデクスのマルチデータ統合	画像とテキスト、テキストと位置情報など、複数の種類のメディアのデータを統合し、インデクス化する方法	メディア対応検索、センサフュージョン	
		検索用インデクスのための構造解析	データソースがどれだけリンクされているか、いつ更新されたものかななどを解析する技術	ページタイプ解析、リンク構造解析、PageRank	
		その他インデキシング技術			
	一致検出技術	ディレクトリ検索	一致するカテゴリ、ディレクトリを返す方法	カテゴリ付与	
		全文検索	キーワードでテキスト全文を検索する方法	分かち書き処理(単語ごと)、インデックス作成・検索、文字列照合検索、n-gram	
		類似、関連情報検索	入力や検索結果と類似する情報、関連する広告などを提案する方法	協調フィルタリング、リコメンデーション、潜在的意味インデキシング、検索連動型広告	
		コンテキストマッチ	時刻、場所などの状況やユーザの好みを反映して結果を返す方法	パーソナルエージェント	
		検索結果の評価	スコア、ランキングなどにより検索結果をソートする方法	適合度、新鮮度、人気度、引用度、ページタイプ、信頼性評価、適合性フィードバック	
		その他一致検出技術			
	インタフェース	入力技術	ブラウザ	検索結果を入力、表示するブラウザ	ズーム技術を用いた階層構造のブラウジング、対話機能、ユーザ適応
			プライバシー保護	個人のプライバシーを保護して検索が行えるよう配慮したもの	セキュア計算
			入力支援	想定される選択肢を列挙するなど、検索の入力を手助けする方法	オートコンプリート
対話入力			システムと複数回のやりとり(対話、絞り込み)により検索が行えるもの	ラダリング	
自然文入力			「日本で一番高い山は?」といった自然文で入力を行うもの	質疑応答	

分野	大分類	小分類	概要	対象技術例
		画像入力	画像を検索キーとして入力するもの	カメラ画像による検索
		音声入力	音声を検索キーとして入力するもの	マイク入力
		位置情報入力	位置情報を検索キーとして入力するもの	GPS
		その他センサ入力	その他のセンサ情報を検索キーとして入力するもの	体感センサからの入力
		メタサーチ・タグサーチ	ソーシャルブックマークのようにタグなど、データそのものではなくメタデータを対象として行う検索	協調化、適応化、ソーシャルブックマーク
		携帯端末用入力	携帯電話などに特化した入力方法	PO-Box、T9、予測変換
		ブール型論理演算式入力	ANDやORなどの演算子を用いて効率的に検索する方法	AND、ORなど論理演算による検索条件設定
		API入力	他のシステムで容易に利用できるようなAPIを提供するもの	API
		その他入力技術		
	出力技術	ナビゲーション・インタフェース	出力結果を効率的に眺めるためのインタフェース	コンテンツの先読み、ナビゲーションウィンドウ
		オブジェクトサーチ	ただ検索結果データを返すのではなく、与えられた入力の意味について、ユーザの意図に沿ってデータの内容をまとめ、提示するもの／複数の検索結果を意味解釈し、情報を統合するもの	横断検索
		クラスタリング、ソート、カテゴリ化	出力結果を自動的に分類、ソートなどしてまとめるもの	順序付出力、カテゴリ分類表示
		可視化、縮約	大量の検索結果を図示、グラフ化、縮約するなどして、可視化するもの／検索結果の統計的な分析	ネットワーク・グラフ表示、縮約表示
		マルチメディアデータ表示	画像、音声などテキスト以外の出力を効果的に行うもの	並列画像収集システム
		携帯端末用出力	携帯電話などに特化した出力方法	モバイル環境における対話的情報提示
		情報フィルタリング	出力結果からユーザに不適切な情報を排除するもの	ホワイトリスト、ブラックリスト、確率モデル、ベイジアンネットワーク
		他機器、システム連携	家電など他の電子機器や、他の検索システムと連携し、検索結果を応用するもの	IT家電、マッシュアップ
		データ配信	ユーザの入力を待つのではなく、新しい情報が見つかるなど状況に応じて検索結果を他に自動的に配信するもの	プッシュ型配信、一斉配信、地震情報配信、EPG

## 2. メディア軸の詳細

メディア軸では、各特許、論文がどのようなメディアを入出力の対象としているのかを表1-1-3に示すデータの種類に従って分類する。近年、マルチメディアや実環境のデータを対象とした検索技術が注目されているが、これにより、実際に特にどのようなメディアが伸びているかを読み取ることができる。

また、メディア軸では、対象とする特許や論文が想定している入出力メディアを広く分類する。そうした技術の中でも、映像データからハイライトの検出を行うなど、システム内部で特に特定のメディアの特徴量などを解析するものについては、さらにメディア解析技術に



て分類 / 分析を行う。同様に、検索キーの入力として、特定メディアをとるものは同様に「画像入力」「音声入力」などの入力技術において分類 / 分析を行う。

表 1-1-3 メディア軸

テキストデータ	マルチメディアデータ	各種データ
単語/キーワード/URL/ID	画像	地図、地理
自然言語/文章	映像	センサ
構造化データ	音声	感性・評判
多言語	音楽	知識
その他メディア		
メディア不特定		

### 3. 課題軸の詳細

課題軸では各特許、論文がどのような課題を解決するかを分析する。検索技術の課題は技術の発展により様々な移り変わりがあったと考えられるため、技術分類だけでは追いつけない、そうした社会的なトレンドを把握することを目指す。

各課題の概要と事例は以下の通り。

表 1-1-4 課題分析軸

課題分類	概要	事例
収益性	収益に貢献する技術	広告、課金システム
低コスト	運用・保守などのコストを削減する技術	メンテナンスフリー、低電力、自動処理
セキュリティ	セキュリティを向上させる技術	暗号化、安全性、改竄・ウィルス対策
高速性	処理を高速化させる技術	高速処理
大規模化	大量のデータなどに対し、大規模な処理を可能にする技術	スケールフリー
ユーザの利便性	使いやすい、便利、分かりやすいなど、ユーザの利便性を向上させる技術	操作性、認知性、マーキング、しおり、履歴活用
システムの効率性	通信量が少ない、メモリ消費が少ないなどシステムの効率性を向上させる技術	効率性
精度	検索結果に正しい情報が含まれ、間違った情報が入らないようにして質を向上させる技術	一致度、適合率
信頼性	精度向上の中で、特にスパムやノイズなど、信頼性を下げるような情報を排除し、信頼性の高い情報だけに絞り込む技術	信頼性ネットワーク、スパム・ノイズ除去、フィルタリング
網羅性	検索結果に幅広い情報が含まれるよう、網羅性を向上させる技術	再現率
パーソナライズ	個々人に合った検索結果を出力する技術	コンテキストマッチ、ユーザマッチ
互換性	他のシステムと連携できるよう変換するなど、互換性を高める技術	互換性

## 第5節 注目研究開発テーマ

本調査が対象とする検索技術分野の技術のうち、前述した分析軸の中から、特に注目して動向を調査・分析すべき分野として、注目研究開発テーマを設定する。

注目研究開発テーマとしては、近年の技術動向、および、市場動向を踏まえて、特許技術の全体動向や委員会での議論に基づき特に注目すべきものとして、以下の7テーマを設定した。

該当するテーマに分類された特許や論文を分析対象として、それぞれの出願人/著者や注目特許/注目論文、出願特許数/論文数の推移など詳細を分析し、動向をまとめるものである。

以下に注目研究開発テーマと選定理由を示す。いずれのテーマに対応する軸が存在しており、分析する際にはその軸に分類された特許を対象にする。

表 1-1-5 注目研究開発テーマおよび選定理由

番号	テーマ名	選定理由
1	地理・地図データ解析 (技術軸：地理・地図データ解析より)	モバイルデバイスを中心に、地図・地理データと連携した実世界情報の検索サービスが注目を集めている。実世界ならではの課題も多いため、技術手法が模索されている。
2	検索結果の抽出・統合 (技術軸：オブジェクトサーチより)	情報量の増加に伴い、全文検索での結果を一覧表示するような従来型の検索サービスだけでなく、検索結果をユーザの意図に沿って絞り込む、複数の検索結果を統合する、といったサービスが登場している。
3	類似・関連情報検索 (技術軸：類似・関連情報検索より)	検索サービスの主要収入源である広告は、検索クエリに類似・関連する情報を提示している。広告以外でも様々なメディアが一元的に検索できるようになっており、類似情報や関連情報の検索技術が重要視されている。
4	テキストデータ解析 (技術軸：自然言語/文章解析、感性・評判データ解析、知識データ解析より)	検索システムの高度化に伴い、クエリ解析とデータソース解析の両面において、自然言語処理技術が求められている。特に感性や知識といった言語表現を解釈していくことはサービスの進化のために必須である。
5	検索用インデクス・キーワードの抽出、作成、更新 (技術軸：検索用インデクス・キーワードの抽出、作成、更新より)	取り扱うデータが大量になるにつれ、そこからどのようにインデクスを作成するかは重要な技術的課題になっている。また、インターネットでは常に情報の量が変化し、データがなくなってしまうことも多々あるため、インデクスの保守・更新が重要な問題である。
6	他機器連携/データ配信 (技術軸：他機器連携、データ配信より)	検索システムがPCだけでなく携帯電話や情報家電などにも搭載されるようになったため、複数の機器でどのように検索結果を共有するか、ある機器で得た検索結果を他機器でどのように生かすかというのがサービスのポイントになっており、他機器連携やデータ配信技術が注目されている。
7	画像・映像解析 (技術軸：画像データ解析、映像データ解析より)	インターネットの高速化に伴いメディアデータが容易に利用できるようになったため、こうしたデータの検索技術が次なる課題となっている。特に画像、映像はデータ量も多く、検索ニーズも高いため注目する必要がある。

## 第2章 調査対象とする特許母集団の設定について

### 第1節 母集団の設定

本調査が対象とする文献は、「日本、米国、欧州、中国、韓国、オーストラリア、ロシア」において出願あるいは登録された、「優先権主張年ベースで1990年から2006年」の、インターネット社会における検索技術に関する特許文献である。なお欧州への出願については、EPOへの広域出願および2008年5月30日現在のヨーロッパ特許条約(EPC)加盟国である34ヶ国への出願と定義し、そのうち特許データベースWPIにて取得可能な20ヶ国を調査対象として設定した。対象とした出願先国・地域をまとめると以下の通りである。

技術領域については、国際特許分類(IPC)としてG06F17/30(情報検索; そのためのデータベース構造)が付与されている特許文献のうち、検索技術に特徴があるもの(キーワードとして「検索」あるいは「サーチ」含むもの)かつインターネットをベースとするもの(インターネット関連のキーワードを含むもの)と定義した。

表 1-2-1 調査対象国・地域

No	コード	名称	区分
1	JP	日本	日本
2	US	米国	米国
3	AT	オーストリア	欧州
4	BE	ベルギー	欧州
5	CZ	チェコ共和国	欧州
6	DE	ドイツ	欧州
7	DK	デンマーク	欧州
8	EP	欧州特許庁	欧州
9	ES	スペイン	欧州
10	FI	フィンランド	欧州
11	FR	フランス	欧州
12	GB	イギリス	欧州
13	HU	ハンガリー	欧州
14	IE	アイルランド	欧州
15	IT	イタリア	欧州
16	LU	ルクセンブルク	欧州
17	NL	オランダ	欧州
18	PT	ポルトガル	欧州
19	RO	ルーマニア	欧州
20	SE	スウェーデン	欧州
21	SK	スロバキア	欧州
22	CH	スイス	欧州
23	NO	ノルウェー	欧州
24	CN	中国	中国
25	KR	韓国	韓国
26	AU	オーストラリア	オーストラリア
27	RU	ロシア	ロシア

## 第2節 利用データベースおよび検索結果

本調査において特許文献の検索・抽出に利用したデータベースは、日本国特許庁に出願されている特許文献については PATOLIS- 、日本国特許庁に出願されていない特許(海外特許)については Derwent World Patent Index(WPI)である。

国内特許母集団の検索結果、海外特許母集団の検索結果をそれぞれ表 1-2-2、表 1-2-3 に示す。データベースの項目の違いにより、多少条件設定の文言が異なっているが、趣旨は同様である。

国内特許については、海外特許の場合と抽出条件を揃えた場合(キーワード検索の対象を「抄録」とした場合)は 2,991 件が抽出された。一方、キーワード検索の対象を「全文」とした場合には 17,808 件が抽出された。本調査においては、基本的には前者の 2,991 件にて国際比較等の大半の分析を実施するが、日本国内の技術動向の分析においては、より範囲を拡張した後者の 17,808 件を対象とすることとする。

一方の海外特許については、6,628 件が抽出された。ただし本検索結果は、ファミリー単位の件数であることに注意が必要である。一方で以降の分析においては、特に断りのない限りファミリーを構成する個々の特許出願単位の集計を行っている。

以上の検索結果を公報単位で整理し直すと、調査対象国・地域全体(日米欧中韓、オーストラリア、ロシア)で 15,123 件(国内特許の検索範囲を拡張した場合は 29,941 件)、日米欧中韓で 13,807 件(同、28,625 件)である。

表 1-2-2 国内特許母集団の検索結果

検索条件	時期範囲	優先権主張日(あるいは出願日) = 19900101:20061231
		公開・公表・再公表日 <= 20080630
	技術範囲	IPC = G06F17/30
		全文 = (検索 or サーチ)
検索結果 ( and and and and )		17,809 件

検索日=2008/7/30

キーワード検索の対象を「抄録」とした場合は、2,991 件

表 1-2-3 海外特許母集団の検索結果

検索条件	時期範囲	優先権主張日 = 19900101:20061231
		発行日 <= 20080630
	国・地域	AT, BE, US, CN, CZ, DE, DK, EP, ES, FI, FR, GB, HU, IE, IT, KR, LU, NL, NO, PT, RO, SE, SK, EP, CH, AU, RU
		IPC = G06F17/30
技術範囲	抄録 = (SEARCH or RETRIEVAL or RETRIEVE)	
	抄録 = (INTERNET or WWW or WEB)	
検索結果 ( and and and and and )		6,628 件

検索日=2008/8/20

### 第3章 分析にあたっての留意点

#### 第1節 出願人国籍別データの集計について

各特許の出願人国籍は、筆頭出願人の国籍とした。その際、「欧州国籍」は2008年5月30日現在のヨーロッパ特許条約（EPC）加盟国である34ヶ国（Austria, Belgium, Bulgaria, Switzerland, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, United Kingdom, Greece, Croatia, Hungary, Ireland, Iceland, Italy, Liechtenstein, Lithuania, Luxembourg, Latvia, Monaco, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, Turkey）と定義した。

#### 第2節 技術区分への分類方法

本調査ではPATOLISおよびWPIを用いた検索により得られた特許文献の抄録を実際に読み込むことで、各技術区分への分類を行った。その際、ある1件の特許文献が複数の技術区分にまたがって属するような分類を許している。

なお要素技術の分析軸においてどの区分にも分類されなかった特許文献については「該当なし」として検索ノイズと見なし、以降の集計結果からは除外している。

#### 第3節 年推移について

調査対象とする母集団は2008年6月30日までに公開・公表・再公表がなされた特許文献である。出願から公開までの期間、あるいはPCT出願後の国内以降までの期間、データベースへの収録の遅れの影響等から、年次変化を追う際に、直近の出願件数については必ずしも実数を反映していない可能性がある点には注意が必要である。

登録件数については、データ収録の問題に加えて、審査中あるいは審査請求判断前の特許が存在することから、近年のデータについては今後増加する可能性がある点には注意が必要である。

また、米国の出願件数について、出願早期公開制度の採用以前である2000年11月29日以前については、登録件数を出願件数としてカウントし集計している点にも注意が必要である。

#### 第4節 出願人について

米国においては、本来は出願人は発明者本人であるが、本調査では他地域との比較のために、譲受人を出願人として集計している。データベース収録の都合上、譲受人が必ずしも記載されていない場合があるが、その際にもファミリー情報等に基づいて出来る限り企業を特定している。また、企業買収等により出願人が変更になっているものについてもできる限り、名寄せを実施して、2008年12月末時点における企業の名称を採用している。

#### 第5節 登録商標について

本報告書では、原則的に商品名、及び、サービス名の表記については、登録商標に従った。なお、その他本報告書の図、表等に記載する会社名、商品名、サービス名は一般に各社の商標、または、登録商標である。

表 1-3-1 本報告書に登場する登録商標<sup>2</sup>

商標名	権利者
GOOGLE グーグル	グーグル・インコーポレイテッド
YAHOO! ヤフー	ヤフー！ インコーポレイテッド
Geocities	
Baidu	百度株式会社
NAVER	NHN CORPORATION
MSN	マイクロソフト コーポレーション
ASK	アイエーシー サーチ アンド メディア, インコーポレイテッド
goo	エヌ・ティ・ティ レゾナント株式会社
EBAY ebay	イーベイ インコーポレイテッド
Marsflag	株式会社マーズフラッグ
shooti	株式会社リクルート
KIZASI	株式会社シーエーシー
BIGLOBE	NEC ビッグロブ株式会社
AOL	エーオーエル エルエルシー
INFOSEEK	楽天株式会社
DAUM	ダウム コミュニケーションズ コーポレーション
CONCEPTBASE	株式会社ジャストシステム
BizSearch	アクセラテクノロジー株式会社
HAKONE	リアルコム株式会社
fast	Fast Search Transfer ASA (現在、マイクロソフト)
WIKIA	Wikia, Inc
HAKIA	Hakia, Inc
Excite	伊藤忠商事株式会社
Lycos	ライコス インコーポレイテッド
Altavista	アルタビスタ カンパニー

<sup>2</sup> 登録商標名の大文字小文字は区別しない。

## 第2部 特許出願動向分析

### 第1章 全体動向分析

#### 第1節 出願先国別件数推移

全ての出願人国籍（日米欧中韓、その他国籍）による出願について、出願先国別の出願件数推移を図 2-1-1 に示す。累積出願件数が最も多いのは米国への出願であり、次いで日本への出願が多い。また、件数の推移を見ると、2000年が大きなピークとなっており、その後の2001年、2002年と大きな減少が見られる。近年、米国への出願は件数を再び伸ばしているが、日本への出願は減少傾向が伺える。また、出願先国別の登録件数推移を図 2-1-2 に示す。米国での登録の比率が最も高く、66%を占めている。次いで、日本での登録と、欧州での登録が多い。

なお2000年に大きなピークがある要因としては、いわゆるITバブルの影響が考えられる。

図 2-1-1 出願先国別出願件数推移

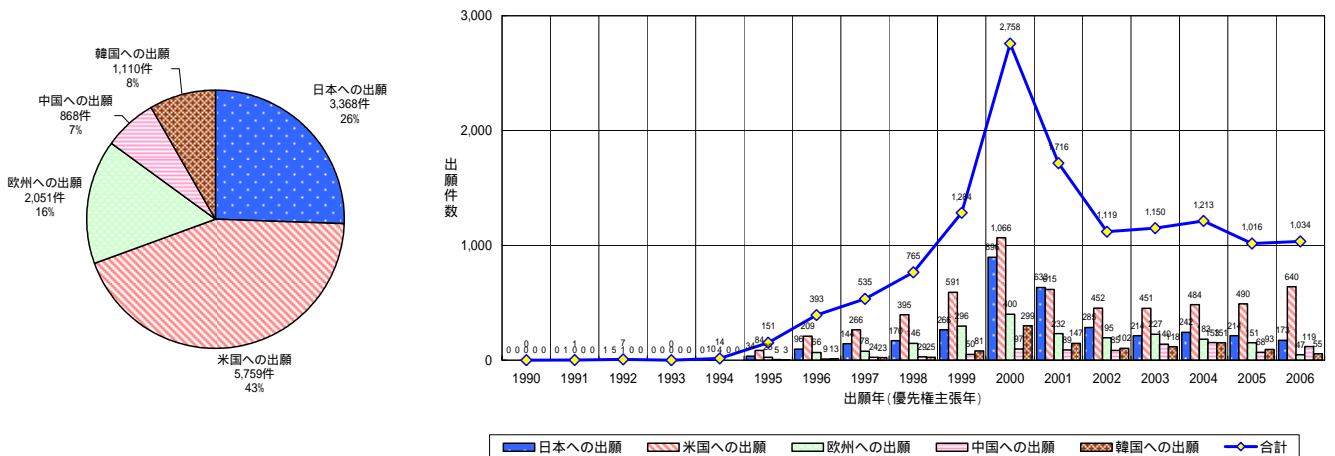
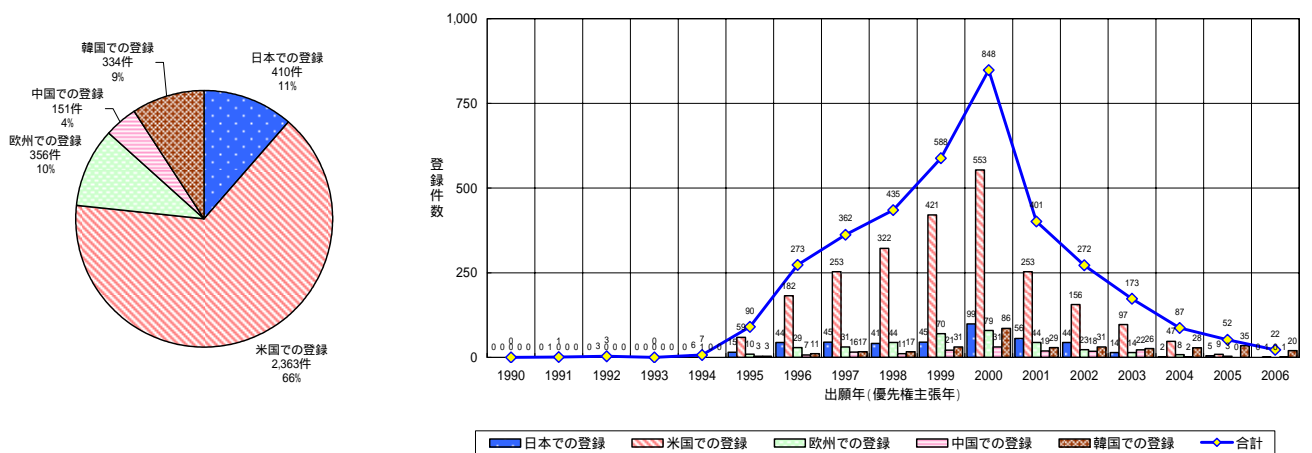


図 2-1-2 出願先国別登録件数推移



## 第2節 出願人国籍別件数推移

日米欧中韓への出願について、出願人国籍別の出願件数推移を図 2-1-3 に示す。出願件数が最も多いのは米国国籍であり 48%と全出願のおよそ半数を占めている。次いで、日本国籍、欧州国籍の順に多い。出願件数推移を見ると、日本国籍、米国国籍ともに、2000 年が出願件数のピークとなっている。近年では、米国国籍はほぼ横ばいの傾向が見られるが、日本国籍は継続的な減少傾向が見える。

同様に、出願人国籍別の登録件数推移を図 2-1-4 に示す。出願件数に比べてさらに、米国国籍の比率が高くなっており、登録件数全体の 63%を占めている。次いで多いのは日本国籍であり、全体の 19%を占めている。

図 2-1-3 出願人国籍別出願件数推移（出願先：日米欧中韓）

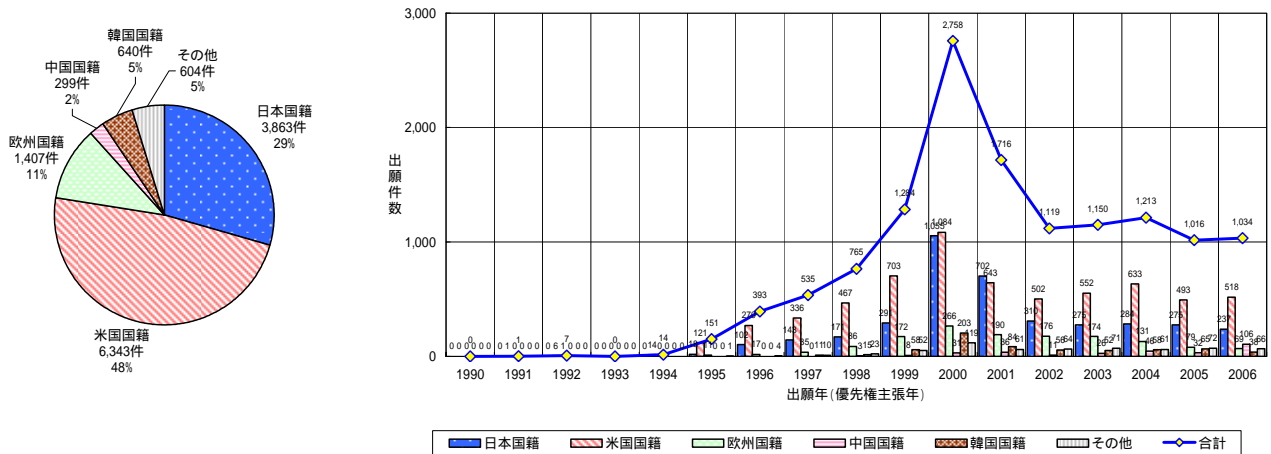
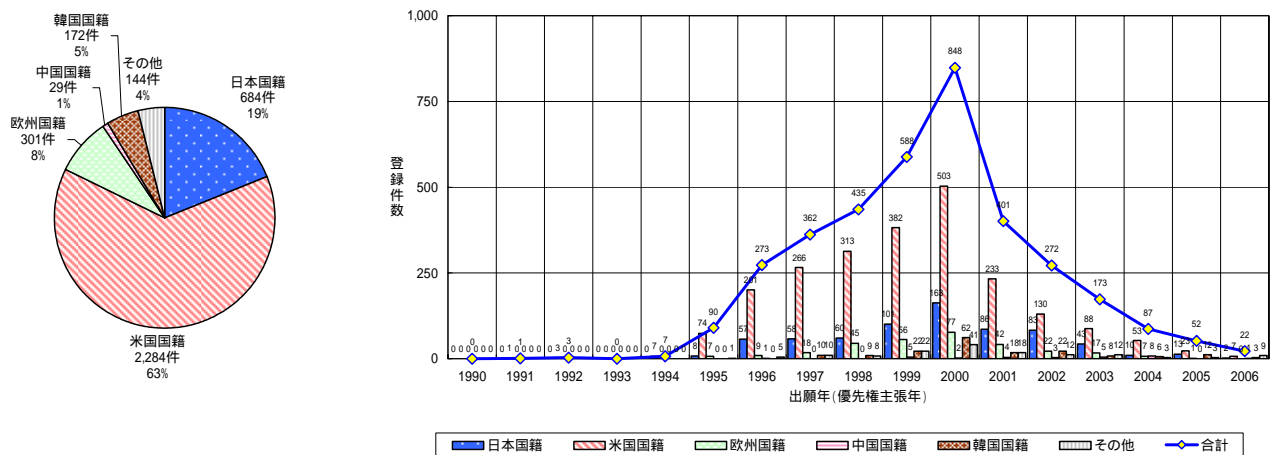


図 2-1-4 出願人国籍別登録件数推移（出願先：日米欧中韓）

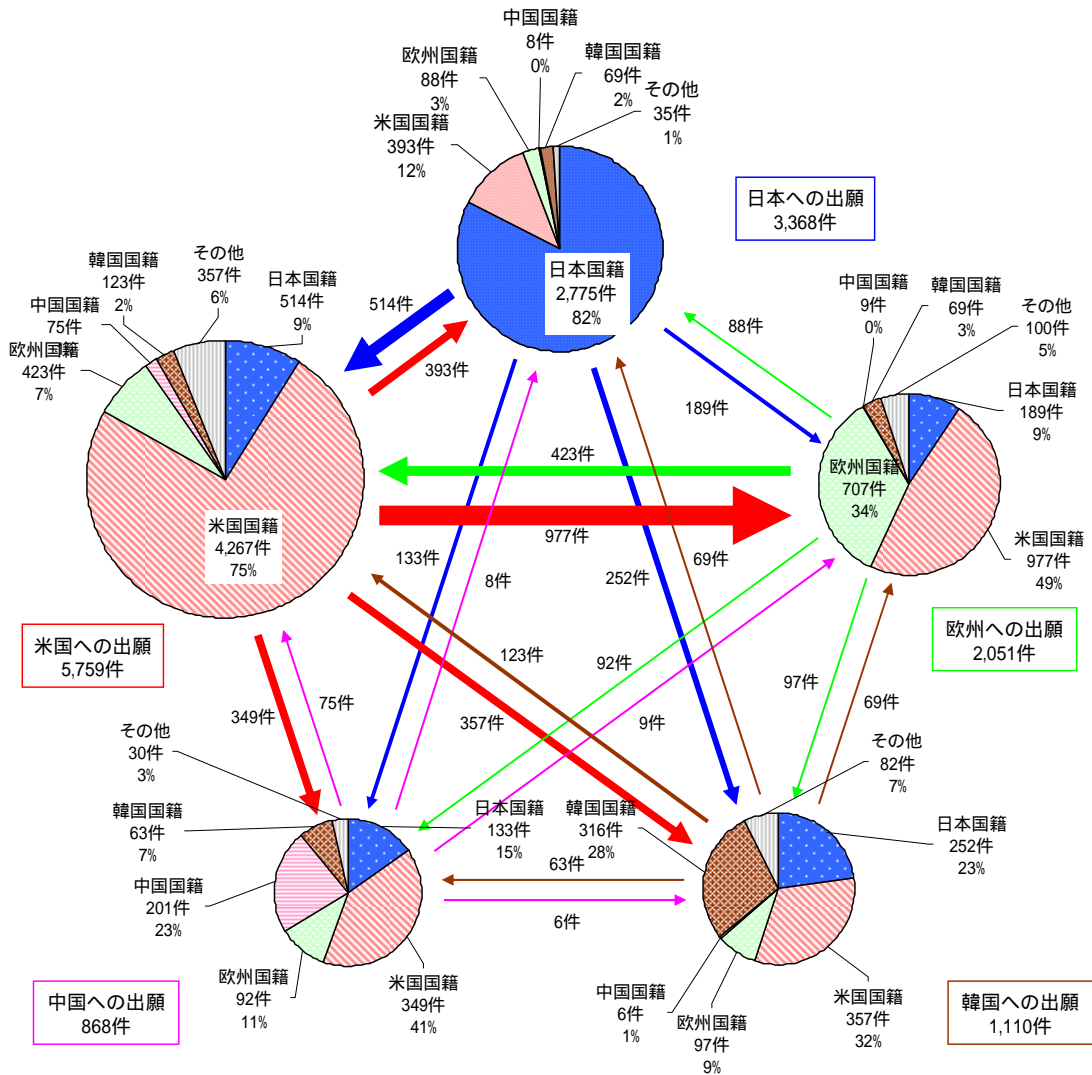




### 第3節 出願先国別出願人国籍別出願件数収支

日米欧中韓への出願について、出願先国別かつ出願人国籍別の出願件数収支を図 2-1-5 に示す。米国への出願が最も多く、次いで日本への出願が多い。日本国籍による出願は、自国に対する出願が最も多く、次いで米国、韓国、欧州への出願が多い。米国国籍についても自国への出願が最も多いが、それ以外の国・地域に対しても出願が多い。

図 2-1-5 出願先国別出願人国籍別出願件数収支（出願先：日米欧中韓）



## 第2章 技術区分別動向分析

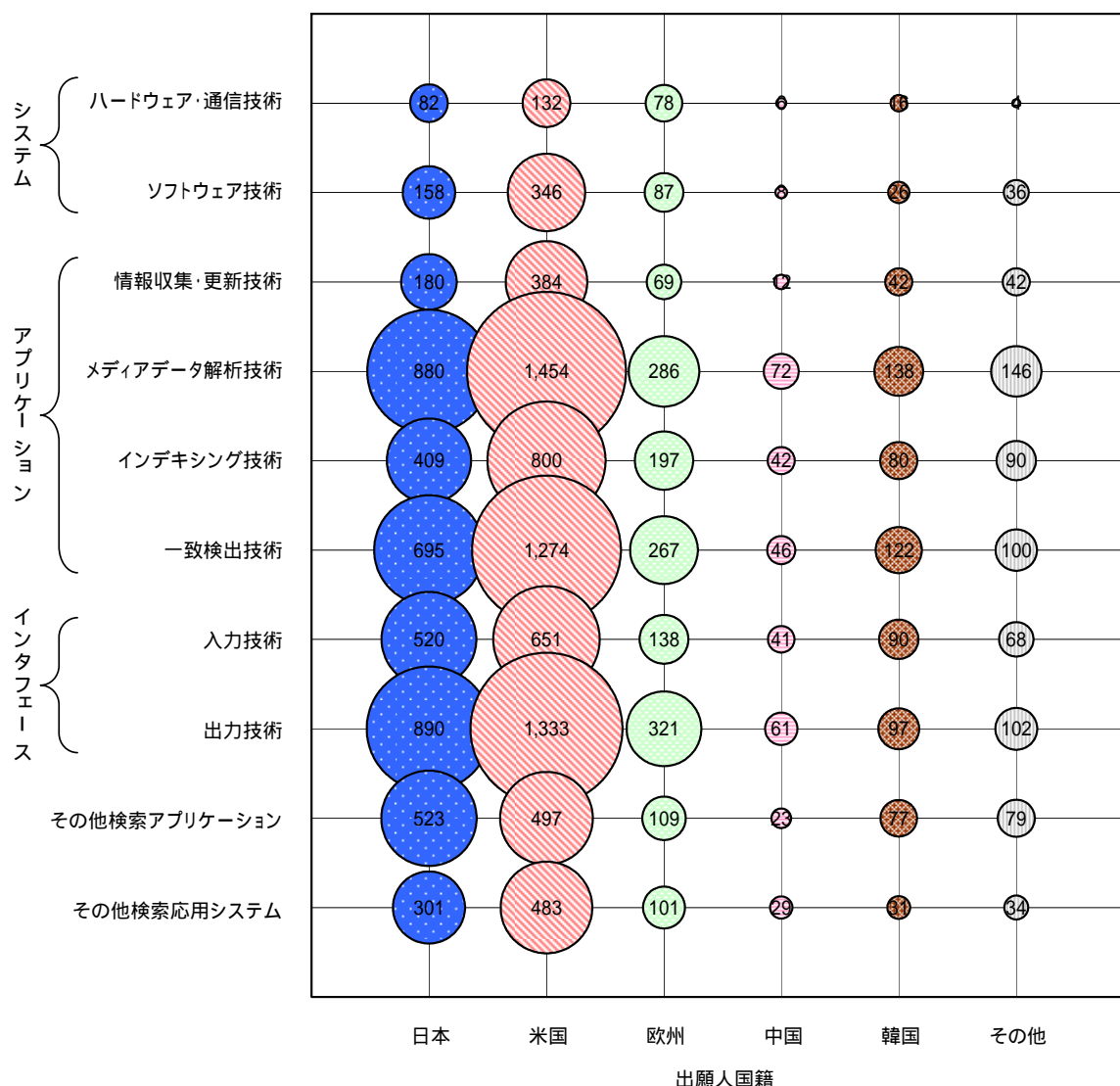
### 第1節 要素技術別

#### 1. 技術区分別 - 出願人国籍別出願件数

日米欧中韓への調査対象期間（優先権主張年ベースで1990年から2006年）における出願について、要素技術の中区分別かつ出願人国籍別の累積出願件数を図2-2-1に示す。

多くの区分について、米国国籍による出願が最も多く、特に件数が多いのは「メディアデータ解析技術」「一致検索技術」「出力技術」である。日本国籍による出願でも同様の3つの区分が多い。また、「入力技術」や「その他検索アプリケーション」については、他の区分に比べて日本国籍による出願が多い傾向が伺える。

図2-2-1 要素技術中区分別 - 出願人国籍別出願件数（出願先：日米欧中韓）



## 2. 技術区分別 - 出願人国籍別件数推移

全ての出願人国籍（日米欧中韓、その他国籍）による出願について、要素技術別の出願件数推移を図 2-2-2 から図 2-2-9 に示す。いずれの要素技術においても、累積出願件数が最も多いのは米国国籍の出願人によるものであり、次いで日本国籍、欧州国籍の順となっている。いずれの技術もこれら三者で全体の 90% 程度を占めている。

件数の推移では 2000 年が大きなピークとなっており、「情報収集・更新技術」を除いては、その後の 2003 年から 2004 年に申請数が持ち直し、再び小さいピークが現れている。

図 2-2-2 出願人国籍別出願件数推移（ハードウェア・通信技術）

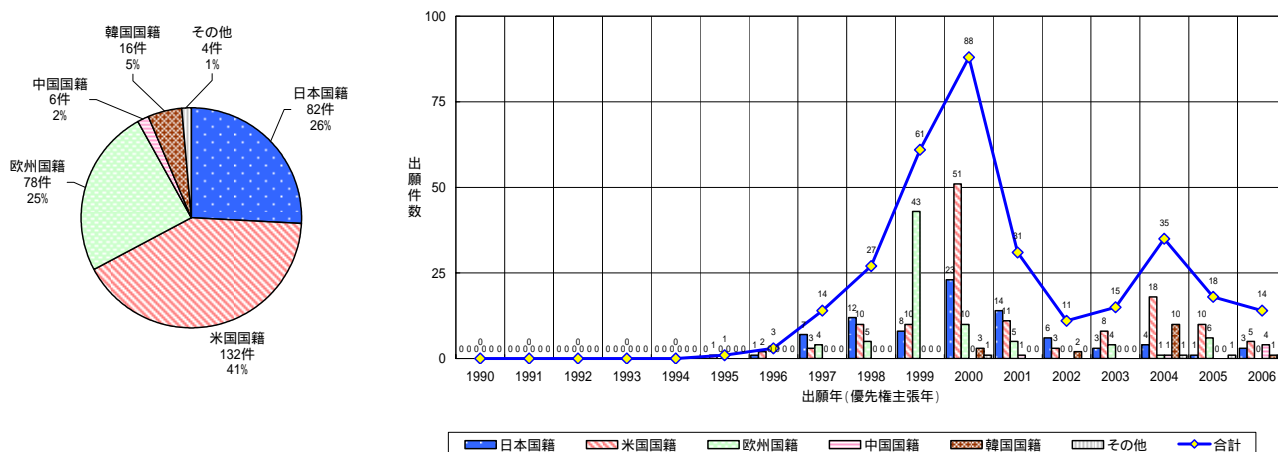


図 2-2-3 出願人国籍別出願件数推移（ソフトウェア技術）

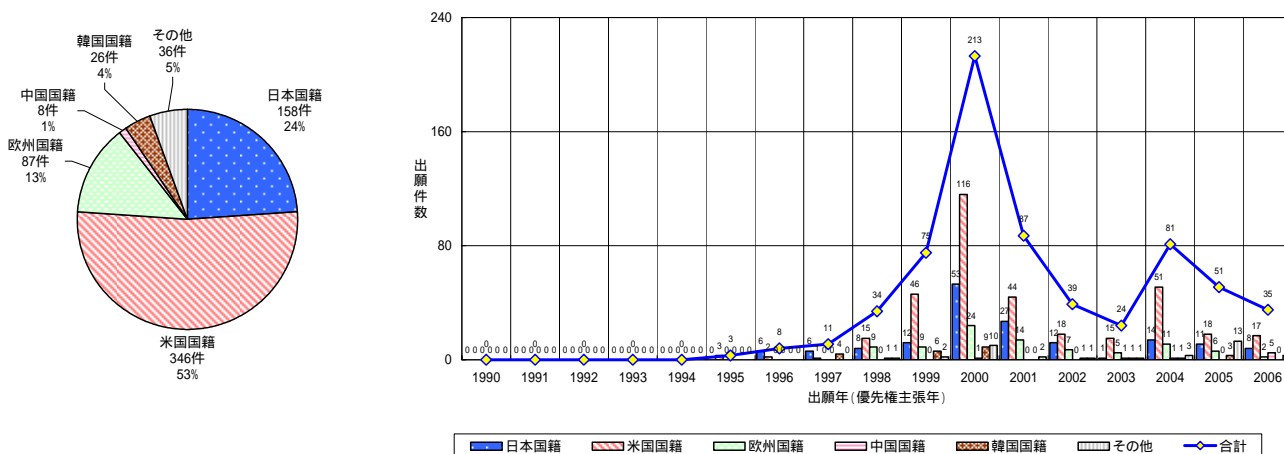


図 2-2-4 出願人国籍別出願件数推移 (情報収集・更新技術)

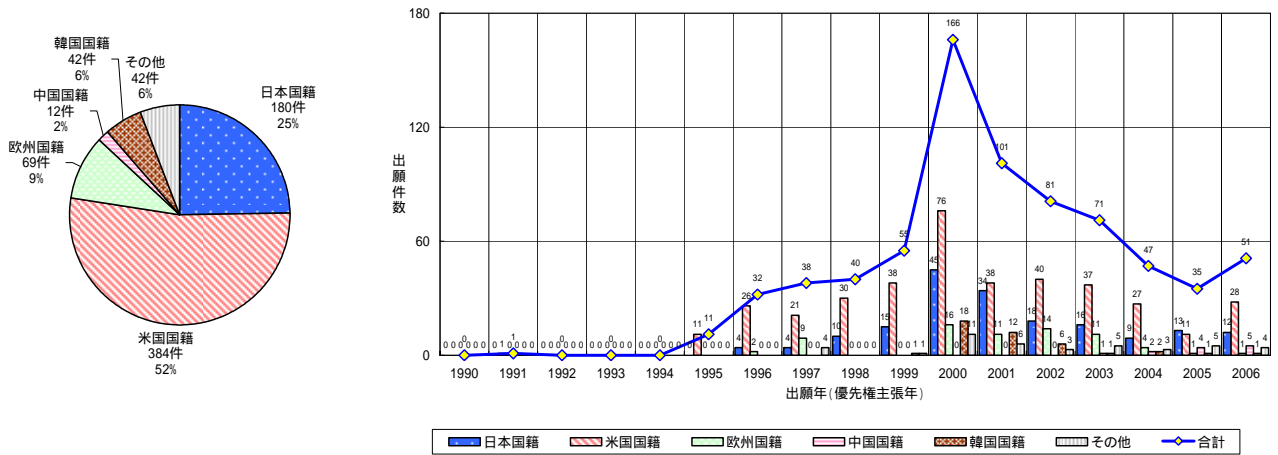


図 2-2-5 出願人国籍別出願件数推移 (メディアデータ解析技術)

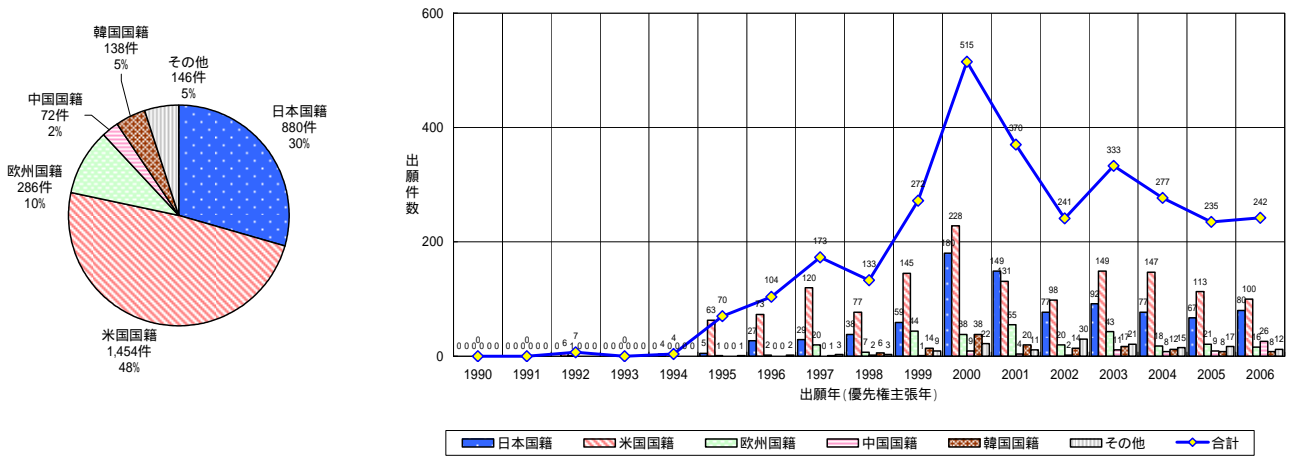


図 2-2-6 出願人国籍別出願件数推移 (インデキシング技術)

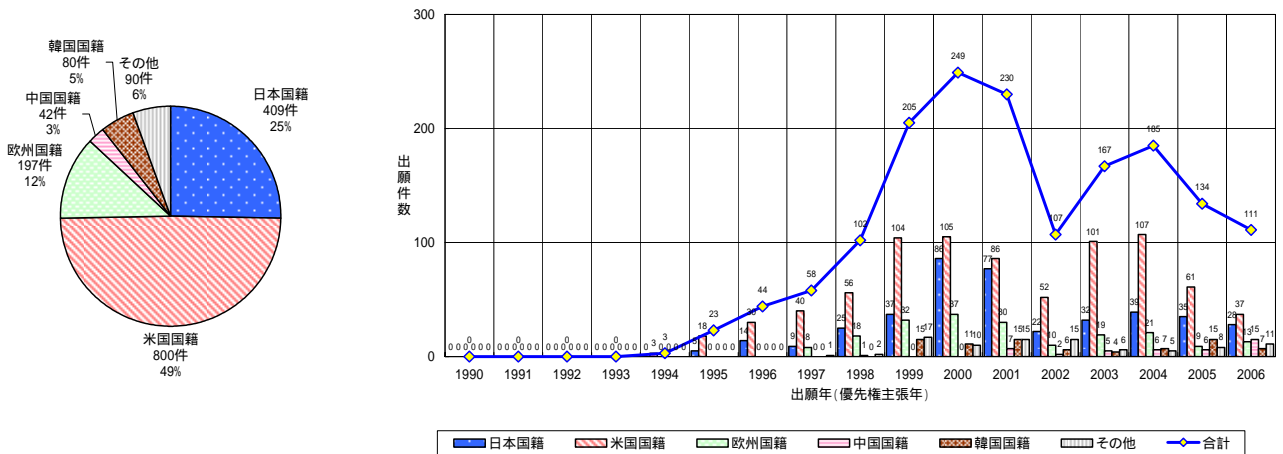


図 2-2-7 出願人国籍別出願件数推移（一致検出技術）

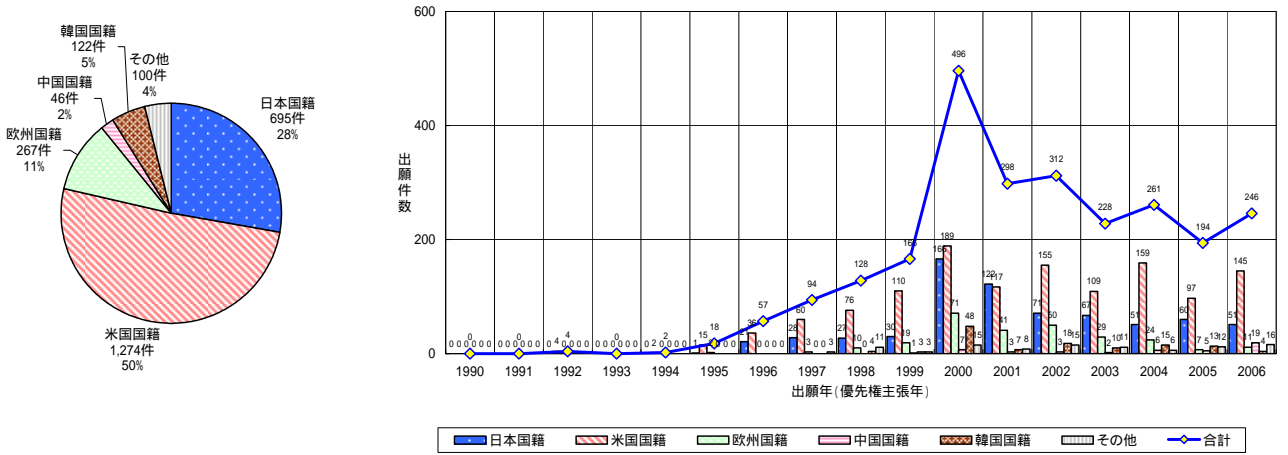


図 2-2-8 出願人国籍別出願件数推移（入力技術）

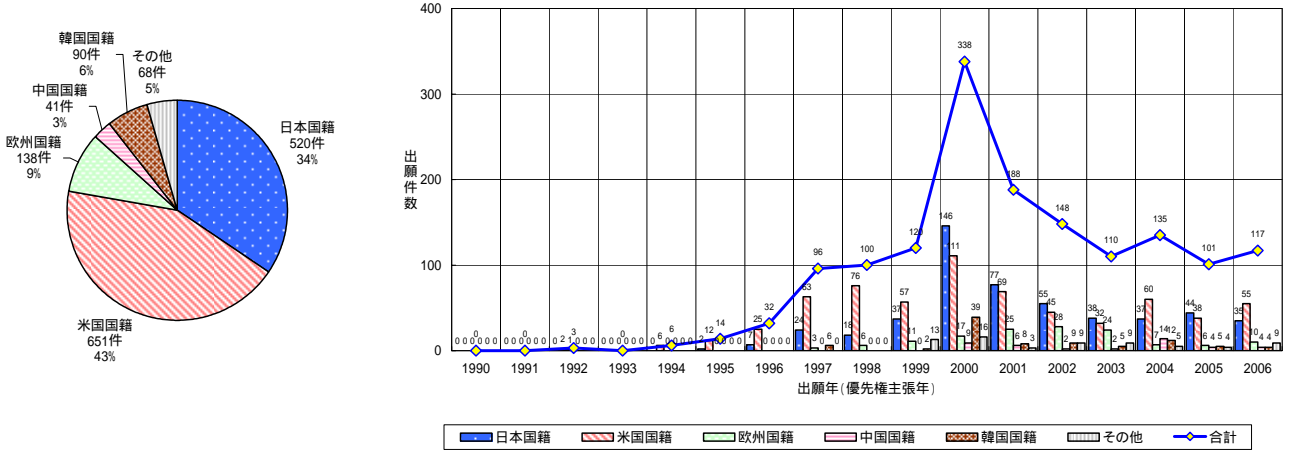
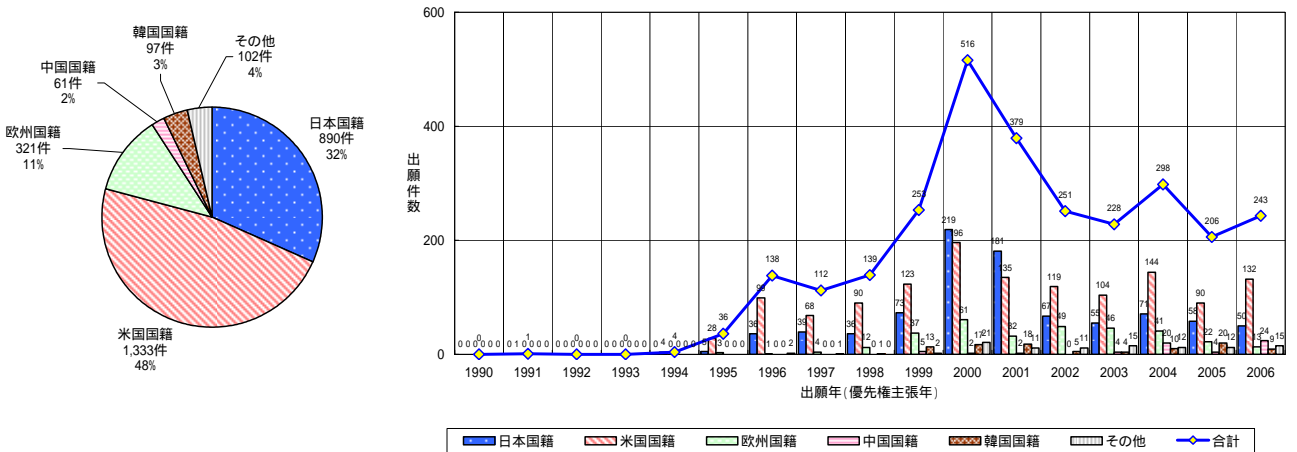


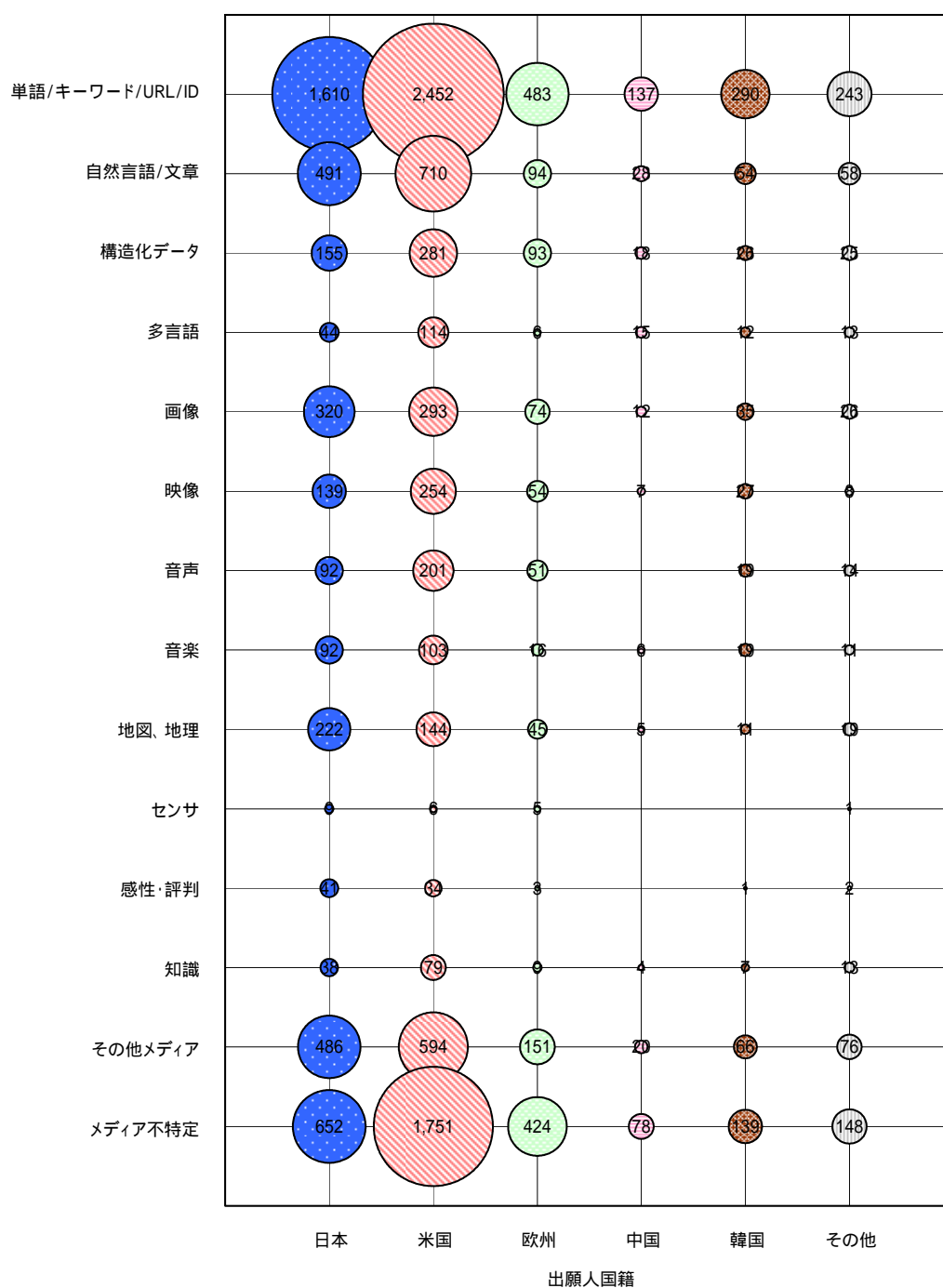
図 2-2-9 出願人国籍別出願件数推移（出力技術）



## 第2節 メディア別

日米欧中韓への調査対象期間（優先権主張年ベースで1990年から2006年）における出願について、取扱いメディア別かつ出願人国籍別の累積出願件数を図2-2-10に示す。全体的な傾向として、「単語/キーワード/URL/ID」および特定のメディアに限定しない「メディア不特定」の件数が多いことが伺える。また、多くの区分で米国国籍の件数が最も多いが、「画像」「地図、地理」「感性・評判」については日本国籍の件数が最も多い。

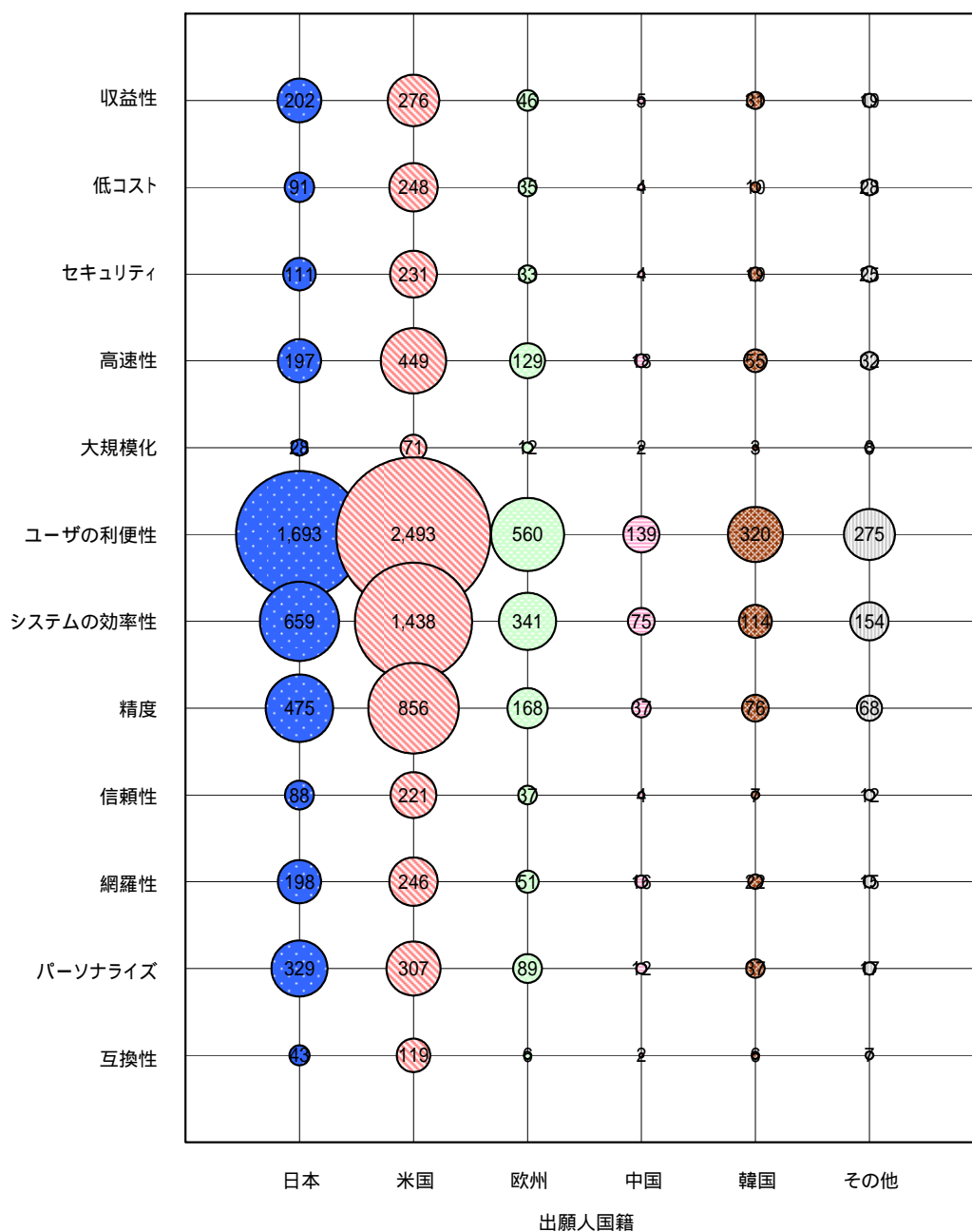
図 2-2-10 メディア別 - 出願人国籍別出願件数（出願先：日米欧中韓）



### 第3節 課題別

日米欧中韓への調査対象期間（優先権主張年ベースで1990年から2006年）における出願について、課題別かつ出願人国籍別の累積出願件数を図2-2-11に示す。全体的な傾向として、「ユーザの利便性」や「システムの効率性」に関する特許が多いことが伺える。これまでと同様に、多くの区分について米国国籍の件数が最も多いが、「パーソナライズ」については日本国籍が最も多く、また「収益性」「ユーザの利便性」「網羅性」についても、他の区分と比べると比較的日本国籍の件数が多い。

図2-2-11 課題別 - 出願人国籍別出願件数（出願先：日米欧中韓）



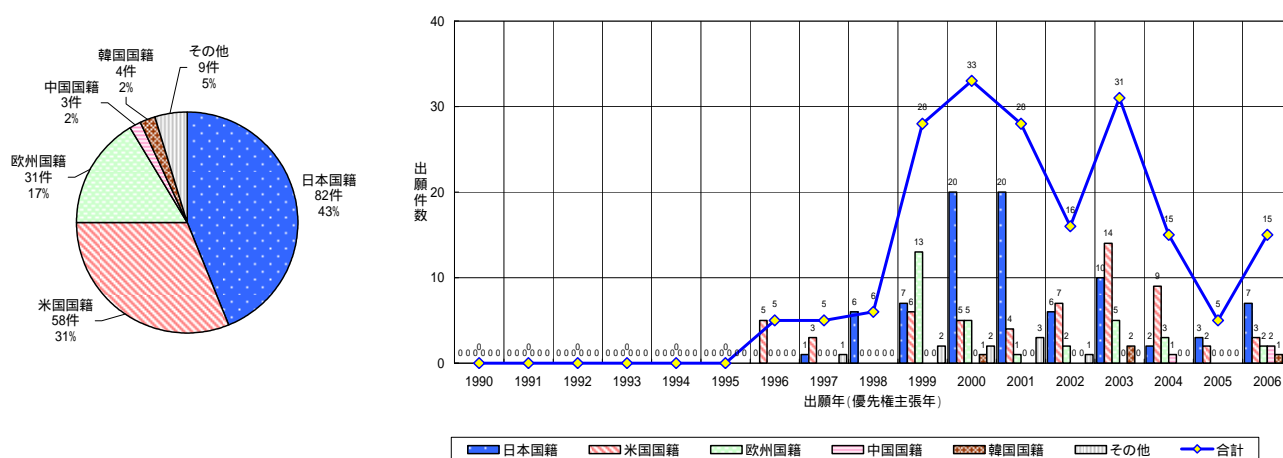
### 第3章 注目研究開発テーマの動向分析

#### 第1節 出願人国籍別出願動向

##### 1. 地理・地図データ解析

注目研究開発テーマ「地理・地図データ解析」について、出願人国籍別の出願件数推移を図 2-3-1 に示す。日本国籍が最も多く、全体の 43% を占めている。日本は従来から地図リテラシが高いと言われており、携帯電話ユーザも多いことから、地図情報サービスのニーズが高いと考えられる。また、他に比べて欧州国籍の件数も比較的多い。推移を見ると 2000 年に一度ピークがあり、その後減少しているが、2003 年に再び増加している。

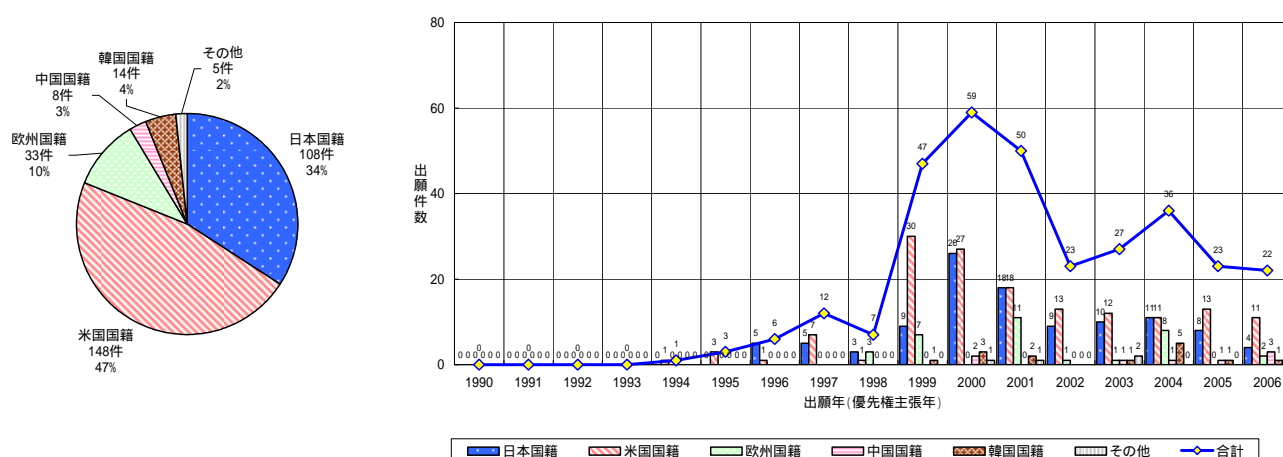
図 2-3-1 「地理・地図データ解析」出願人国籍別出願件数推移（出願先：日米欧中韓）



##### 2. 検索結果の抽出・統合

注目研究開発テーマ「検索結果の抽出・統合」について、出願人国籍別の出願件数推移を図 2-3-2 に示す。累積件数では米国籍が最も多く、推移を見てもほぼ常に他国籍を上回っている。2000 年に件数のピークが見られるが、その後も 2004 年に再度盛り上がりを見せている。2004 年頃は Web 2.0 やマッシュアップという言葉が話題になり始めたころであり、期待感は継続していると考えられる。

図 2-3-2 「検索結果の抽出・統合」出願人国籍別出願件数推移（出願先：日米欧中韓）

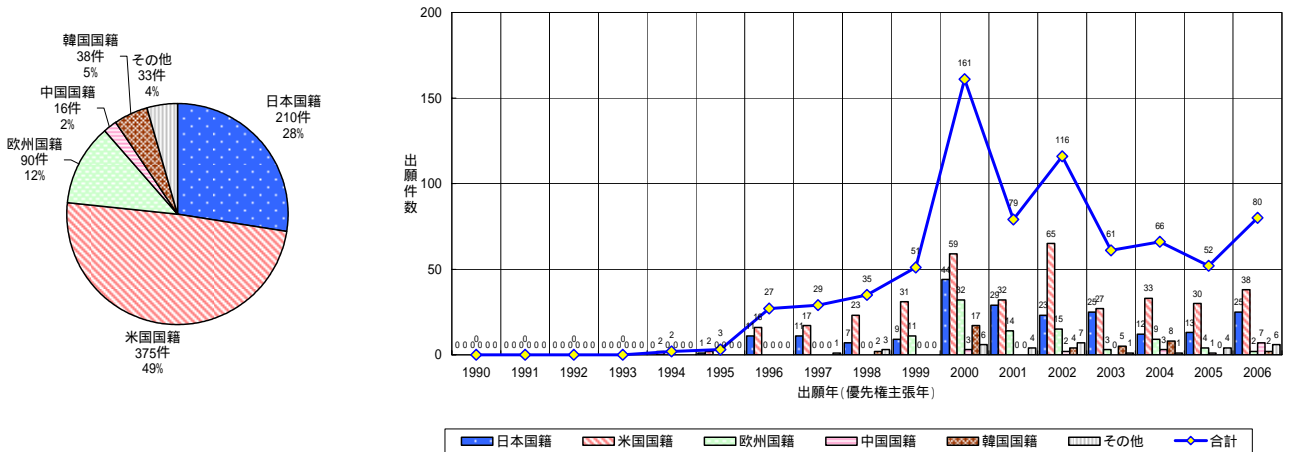




### 3. 類似・関連情報検索

注目研究開発テーマ「類似・関連情報検索」について、出願人国籍別の出願件数推移を図2-3-3に示す。累積件数では米国国籍が最も多く、およそ半数を占めている。また、推移を見ても常に他の国籍を上回っている。他の区分と同様に2000年に件数のピークが見られるが、その後も2002年にピークがあり、2006年にも増加傾向が伺える。2000年のピーク以降も継続的に研究開発されているテーマと考えられる。

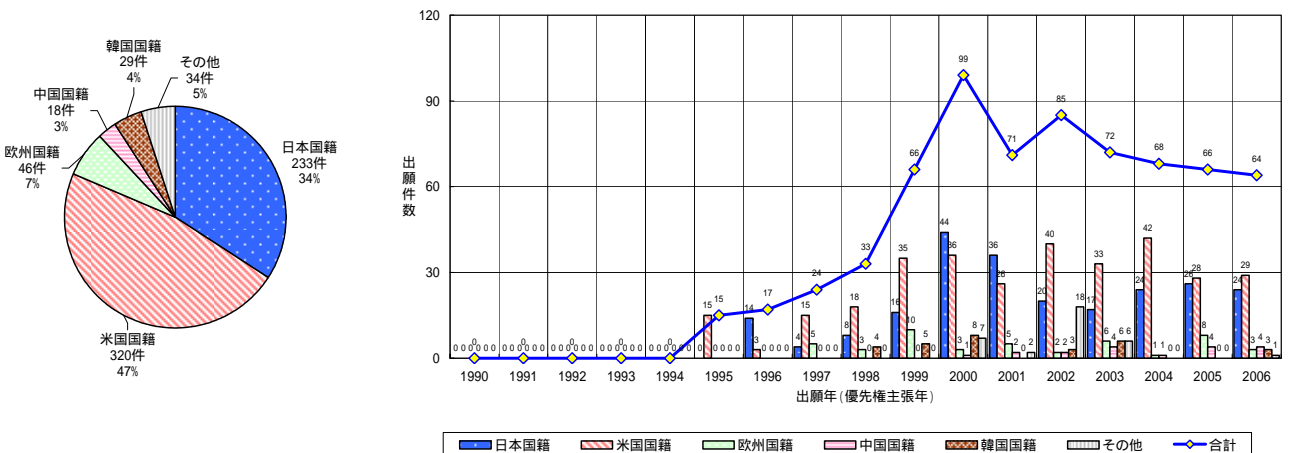
図 2-3-3 「類似・関連情報検索」出願人国籍別出願件数推移（出願先：日米欧中韓）



### 4. テキストデータ解析

注目研究開発テーマ「テキストデータ解析」について、出願人国籍別の出願件数推移を図2-3-4に示す。累積件数では米国国籍が最も多く、およそ半数を占めている。また、推移を見ても概ね他の国籍を上回っているが、1996年、2000年、2001年については日本国籍の件数が最も多い。他の区分と同様に2000年に件数のピークが見られるが、その後の件数の減少がほとんど見られない点が特徴的である。

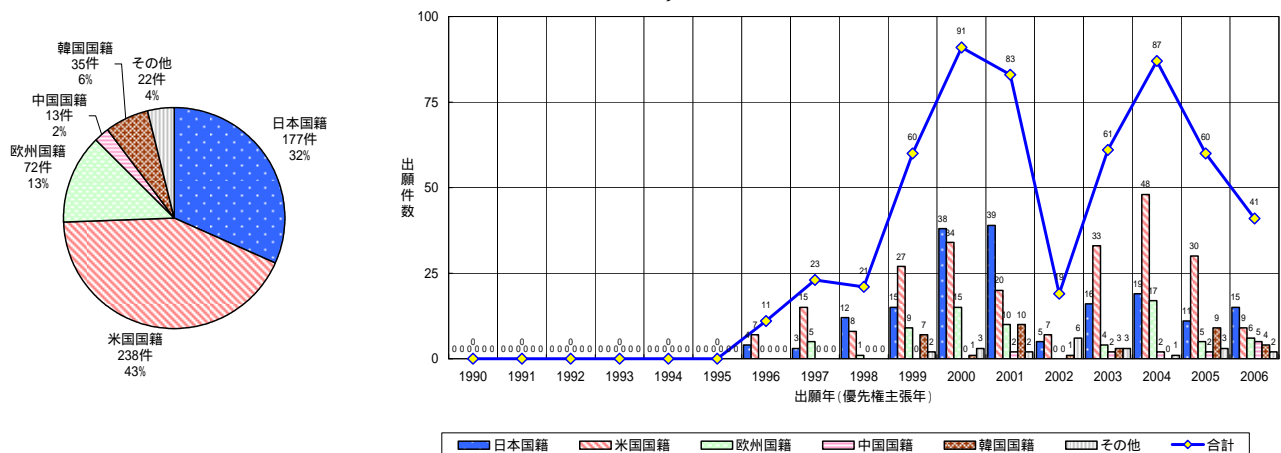
図 2-3-4 「テキストデータ解析」出願人国籍別出願件数推移（出願先：日米欧中韓）



## 5. 検索用インデクス、キーワードの抽出、作成、更新

注目研究開発テーマ「検索用インデクス、キーワードの抽出、作成、更新」について、出願人国籍別の出願件数推移を図 2-3-5 に示す。累積件数では米国国籍が最も多いが、日本国籍も全体の 32% を占めている。2000 年前後と 2004 年前後の 2 度、出願が多くなされている。しかし出願者は、前者では日本国籍が中心である一方、後者は米国国籍が大半である。国産検索エンジンが淘汰された一方、米国では Wikia や Hakia、Powerset、Cuil といった検索サービスが次々に登場しており、基盤技術の研究開発が続けられていると考えられる。

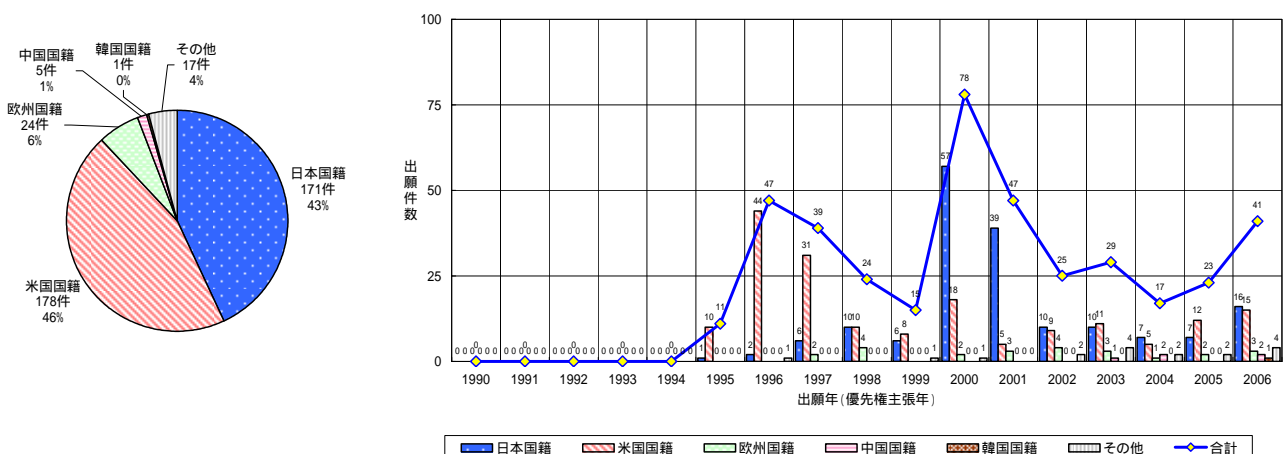
図 2-3-5 「検索用インデクス等の抽出、作成、更新」出願人国籍別出願件数推移（出願先：日米欧中韓）



## 6. 他機器連携 / データ配信

注目研究開発テーマ「他機器連携 / データ配信」について、出願人国籍別の出願件数推移を図 2-3-6 に示す。累積件数では日本国籍と米国国籍の件数がほぼ同等となっている。推移を見ると、1996 年と 2000 年に大きなピークがあり、前者は米国国籍、後者は日本国籍による出願がピークの要因となっている。近年は日本国籍、米国国籍ともに件数を再び伸ばしている。

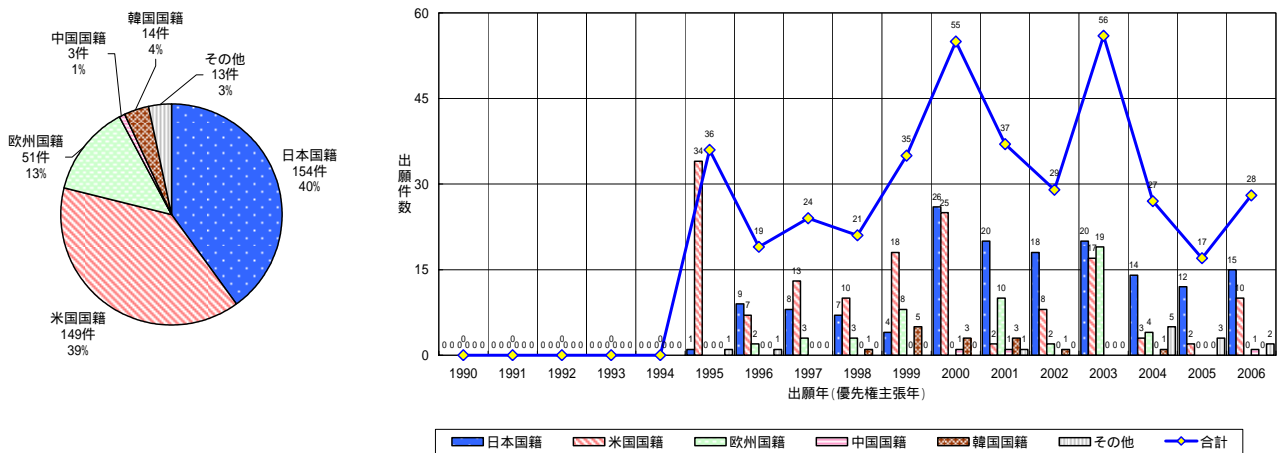
図 2-3-6 「他機器連携 / データ配信」出願人国籍別出願件数推移（出願先：日米欧中韓）



## 7. 画像・映像解析

注目研究開発テーマ「画像・映像解析」について、出願人国籍別の出願件数推移を図 2-3-7 に示す。累積件数では日本国籍と米国国籍の件数がほぼ同等となっている。推移を見ると、1995 年、2000 年、2003 年にそれぞれ大きなピークがあり、はじめのピークは米国国籍、2 回目のピークは日本国籍、米国国籍、3 回目のピークは日本国籍、米国国籍に加えて欧州国籍によるものである。

図 2-3-7 「画像・映像解析」出願人国籍別出願件数推移（出願先：日米欧中韓）



## 第4章 出願人別動向分析

### 第1節 出願人別 - 出願先国別出願件数上位ランキング

日米欧中韓への調査対象期間（優先権主張年ベースで1990年から2006年）における出願について、出願人別かつ出願先国別の出願件数上位ランキングを表2-4-1に示す。

日本への出願については、日本電気、日本電信電話、ソニーの順である。上位はIBMを除き日本国籍の出願人が占めている。米国への出願については、IBM、マイクロソフトが特に多い。また、欧州への出願については、マイクロソフト、IBM、グーグルの順である。中国への出願では、IBM、マイクロソフトが特に多く、次いでグーグルの順である。韓国への出願については、NHN、マイクロソフト、IBMと続いている。

表2-4-1 出願人別出願件数上位ランキング（出願先：日米欧中韓）

日本への出願			米国への出願			欧州への出願			中国への出願			韓国への出願		
順位	出願人	件数	順位	出願人	件数	順位	出願人	件数	順位	出願人	件数	順位	出願人	件数
1	日本電気(日本)	194	1	IBM(米国)	484	1	マイクロソフト(米国)	87	1	IBM(米国)	77	1	NHN(韓国)	57
2	日本電信電話(日本)	180	2	マイクロソフト(米国)	373	2	IBM(米国)	73	2	マイクロソフト(米国)	63	2	マイクロソフト(米国)	56
3	ソニー(日本)	123	3	ヤフー(米国)	124	3	グーグル(米国)	66	3	グーグル(米国)	35	3	IBM(米国)	43
4	日立製作所(日本)	121	4	ソニー(日本)	100	4	ソニー(日本)	57	4	フィリップス(欧州)	35	4	サムスン(韓国)	37
5	富士通(日本)	108	5	グーグル(米国)	99	5	ヒューレット・パッカード(米国)	56	5	ソニー(日本)	34	5	フィリップス(欧州)	35
6	東芝(日本)	105	6	富士通(日本)	78	6	オーバーチュア(米国)	51	6	サムスン(韓国)	21	6	ソニー(日本)	34
7	キヤノン(日本)	73	7	ヒューレット・パッカード(米国)	77	7	フィリップス(欧州)	48	7	ZHANG W(中国)	17	7	LG(韓国)	31
8	リコー(日本)	71	8	AT&T(米国)	56	8	フランステレコム(欧州)	31	8	オーバーチュア(米国)	16	8	グーグル(米国)	24
9	IBM(米国)	64	9	アメリカオンライン(米国)	53	9	ブリティッシュ・テレコミュニケーションズ(欧州)	28	9	パナソニック(日本)	14	9	オーバーチュア(米国)	22
10	三菱電機(日本)	64	10	サン・マイクロシステムズ(米国)	53	10	サムスン(韓国)	26	10	無錫甬鐘科技(中国)	13	10	韓国電子通信研究院(韓国)	20

## 第2節 出願人別 - 出願先国別登録件数上位ランキング

日米欧中韓での調査対象期間（優先権主張年ベースで1990年から2006年）における登録について、出願人別の登録件数上位ランキングを表2-4-2に示す。日本での登録では、IBMが最上位に来ている。米国での登録、中国での登録についても同様である。また、欧州での登録ではヒューレット・パッカーが、韓国での登録ではNHNがそれぞれ最上位に来ている。

表2-4-2 出願人別登録件数上位ランキング（出願先：日米欧中韓）

日本での登録			米国での登録			欧州での登録			中国での登録			韓国での登録		
順位	出願人	件数	順位	出願人	件数	順位	出願人	件数	順位	出願人	件数	順位	出願人	件数
1	IBM(米国)	51	1	IBM(米国)	332	1	ヒューレット・パッカー(米国)	28	1	IBM(米国)	33	1	NHN(韓国)	28
2	日本電気(日本)	39	2	マイクロソフト(米国)	150	2	IBM(米国)	27	2	マイクロソフト(米国)	6	2	IBM(米国)	26
3	日本電信電話(日本)	32	3	ソニー(日本)	53	3	ブリティッシュ・テレコミュニケーションズ(欧州)	9	3	ソニー(日本)	6	3	サムスン(韓国)	20
4	富士通(日本)	31	4	ヒューレット・パッカー(米国)	52	4	ソニー(日本)	8	4	サムスン(韓国)	5	4	韓国電子通信研究院(韓国)	15
5	東芝(日本)	19	5	サン・マイクロシステムズ(米国)	51	5	テリア(欧州)	8	5	日本電気(日本)	5	5	LG(韓国)	11
6	日立製作所(日本)	18	6	AT&T(米国)	38	6	サムスン(韓国)	7	6	パナソニック(日本)	5	6	オーバーチュア(米国)	8
7	ソニー(日本)	14	7	アメリカオンライン(米国)	38	7	クリアスピード(欧州)	7	7	東芝(日本)	4	7	CHUTNOON CO LTD(韓国)	7
8	パイオニア(日本)	10	8	ゼロックス(米国)	38	8	エリクソン(欧州)	7	8	フィリップス(欧州)	4	8	KIM S H(韓国)	7
9	オーバーチュア(米国)	9	9	富士通(日本)	32	9	フィリップス(欧州)	6	9	ノキア(欧州)	3	9	KT(韓国)	5
10	インクリメント・ピー(日本)	9	10	オーバーチュア(米国)	28	10	オーバーチュア(米国)	6	-	-	-	10	ネットピア(韓国)	5
-	-	-	-	-	-	10	ファストサーチ&トランスファ(欧州)	6	-	-	-	-	-	-
-	-	-	-	-	-	10	マイクロソフト(米国)	6	-	-	-	-	-	-
-	-	-	-	-	-	10	サン・マイクロシステムズ(米国)	6	-	-	-	-	-	-

ただし中国での出願については、3件以上の出願人のみ掲載

### 第3部 基本特許・重要特許分析

#### 第1章 基本・重要特許の変遷に関する分析

##### 第1節 対象テーマ

本調査における基本・重要特許分析の対象として、表 1-1-5 に示した注目研究開発テーマに設定した7テーマを対象とする。

##### 第2節 分析の流れ

本調査における基本・重要特許の分析の流れを以下に示す。

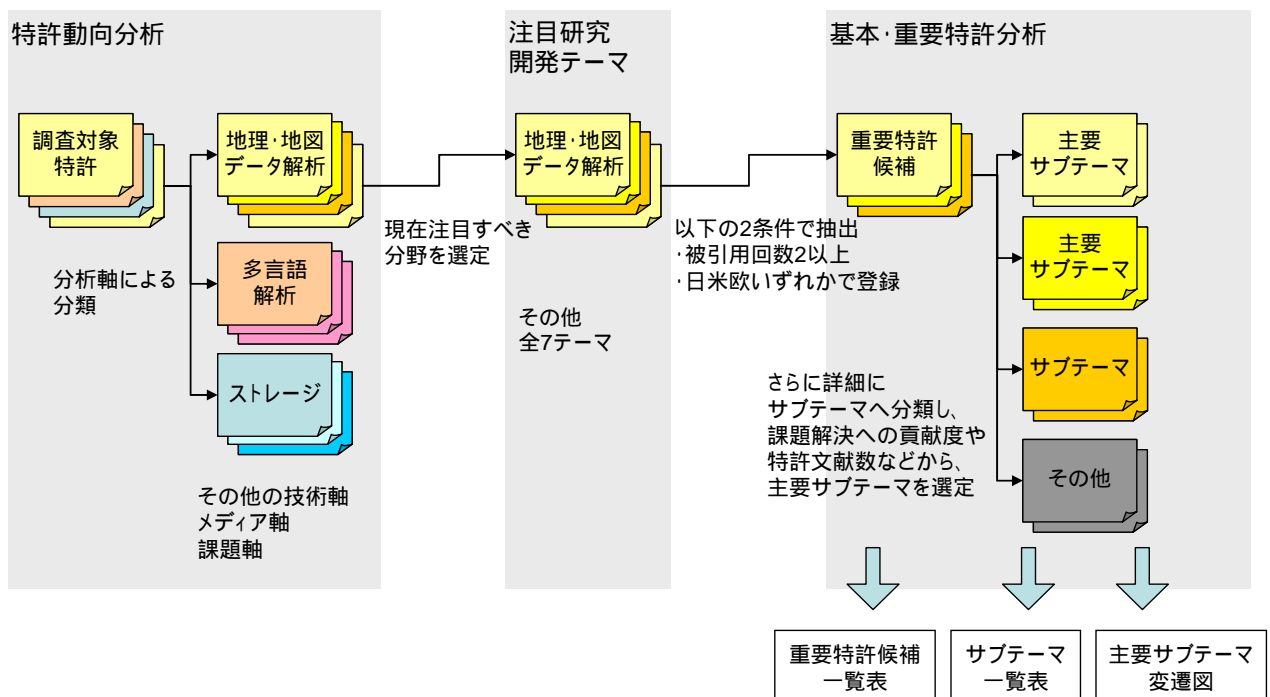
まず、注目研究開発テーマそれぞれに分類された特許文献を、ファミリー単位で抽出した。特許ファミリーは、同様の発明について優先権を利用して出願を成した結果構成される特許群を指し、ファミリーの定義は特許検索に利用した WPI ( World Patent Index ) の定義に従っている。ここから本調査では、重要特許候補として「被引用回数が2回以上」「日米欧のいずれかで登録」の両条件を満たすものに絞り込んだ。同時に重要特許候補については今一度内容を精査し、サブテーマとして詳細に分類した。

次に、テーマ毎に得られたサブテーマをサブテーマ一覧表としてまとめた。サブテーマの中から特に「技術の課題解決への貢献度」が高く、「関連する特許文献数」が多いものは主要サブテーマとして抽出し、サブテーマ一覧表にて付記している。

最後に、主要サブテーマとそこに含まれる特許について、主要サブテーマ変遷図としてまとめた。

なお被引用回数は、国内特許母集団については PATOLIS にて、海外特許母集団については PCI ( Patents Citation Index ) にて取得した。

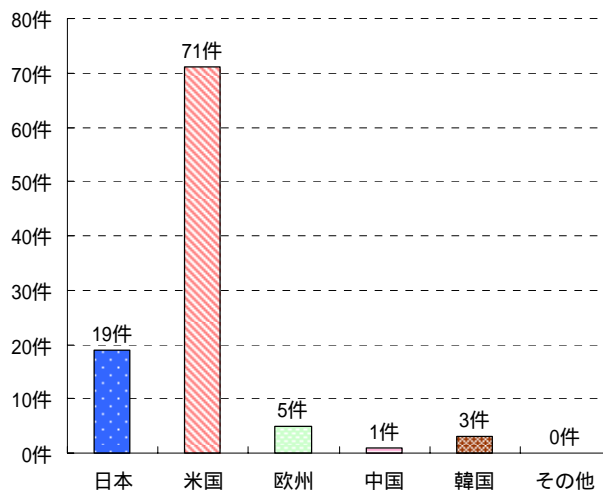
図 3-1-1 基本重要特許分析の流れ



## 第2章 分析結果のまとめ

基本重要特許分析により抽出したインターネット社会における検索技術分野の本調査における基本・重要特許の件数について、出願件数との比較により出願人国籍別に整理した結果を図 3-2-1 に示す。米国国籍が 71 件と最多であり、日本国籍はそれに続く 19 件となっている。なお、被引用回数で絞りこんだため、近年の特許はわずかであり、特許動向と同様に 2003 年以降に限定して見た場合「類似・関連情報検索」で米国国籍が 1 つ、「テキストデータ解析」で米国国籍と欧州国籍が 1 つずつに留まっている。

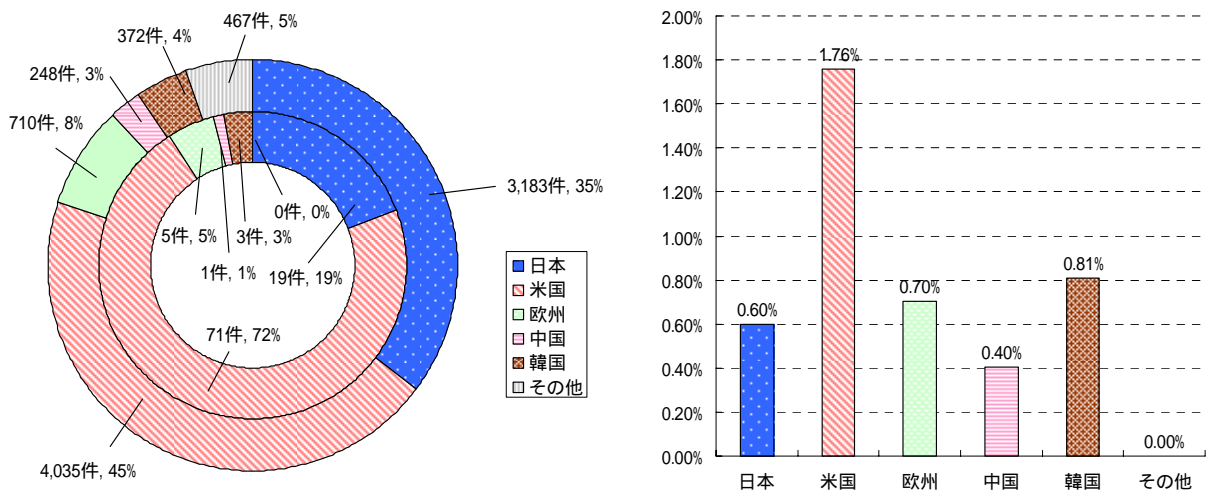
図 3-2-1 基本・重要特許件数（ファミリー単位）



また、図 3-2-2 の円グラフは出願件数と基本・重要特許件数との対比を、右側は出願件数に占める基本・重要特許件数の割合を示している。なお、基本重要特許分析はファミリー単位に実施しているため、件数についてもファミリー単位の集計となっている。出願件数全体では日本国籍が 35%、米国国籍が 45%であるが、基本・重要特許件数では日本国籍が 19%、米国国籍が 72%と差が広がっている。出願件数に占める基本・重要特許の割合でも、米国国籍の 1.76%が最大となっている。

図 2-2-2 基本・重要特許割合（ファミリー単位）

（左：出願件数と基本・重要特許件数の対比、右：出願件数に占める基本・重要特許の割合）



### 第3章 テーマ別分析結果

「類似・関連情報検索」をテーマ別分析結果の例として示す。

#### 第1節 類似・関連情報検索

地理・地図データ解析今日ではさまざまなウェブサービスが存在しているが、主要収入源はウェブページの内容や検索クエリといったコンテキストに応じた広告の配信である。そのため広告配信技術は収益性という課題に最も関係する技術として注目が必要である。また、ユーザのコンテキストを把握して求める情報を提示するという意味では、広告以外においても同様の課題が存在しており、ここでは類似・関連情報検索として広く分析を行う。

本テーマの重要特許と、その変遷図を図 3-3-1 に示す。重要特許は重要特許候補 102 件をサブテーマに分類した結果、主要サブテーマに分類されたものである。サブテーマは表 3-3-1 に示したものであり、主要サブテーマとして、「パーソナライズド広告」「情報の関連度評価」「類似情報検索システム」を選定した。以下、主要サブテーマ別に、重要特許の概要をまとめる。

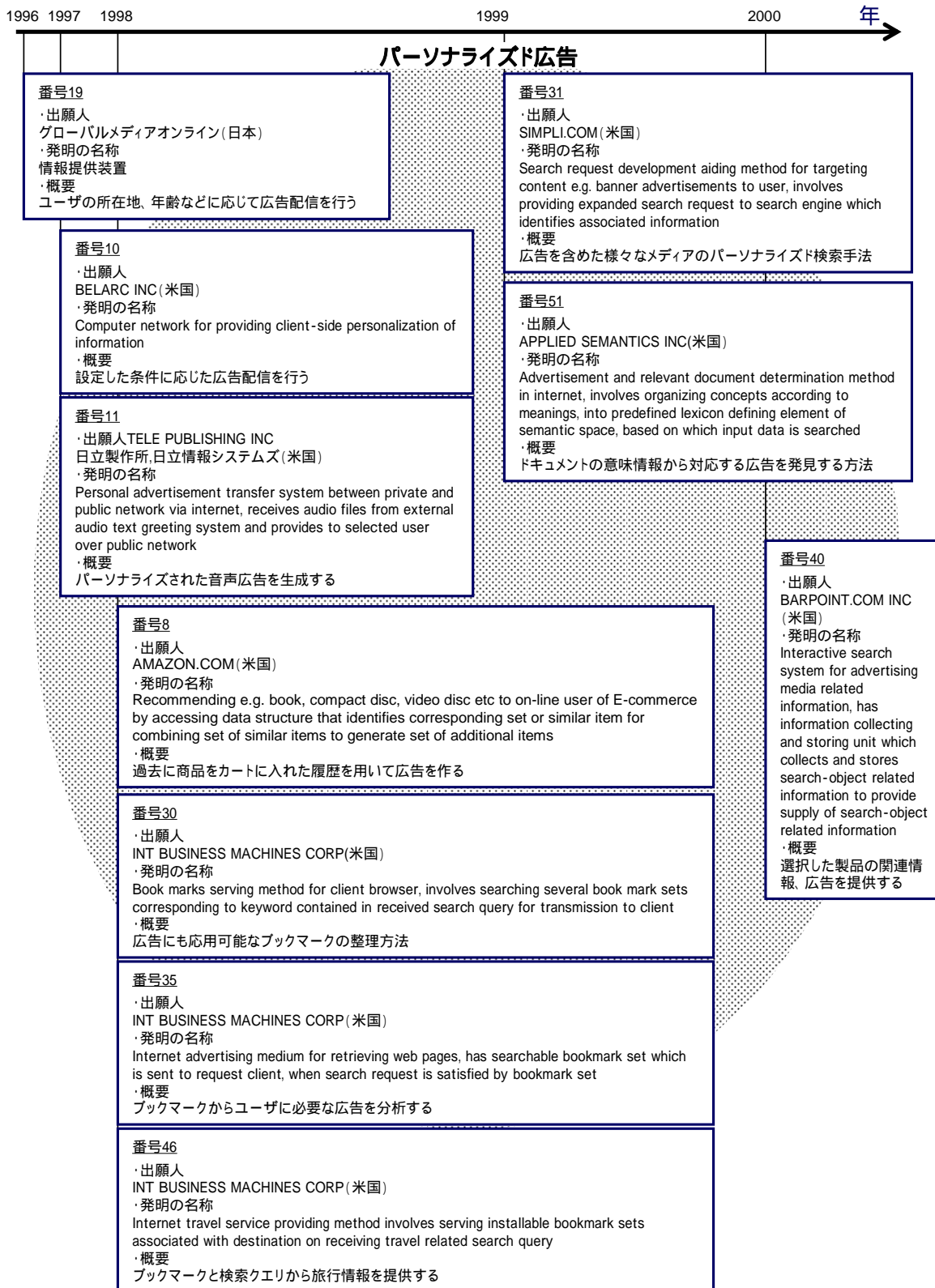
「パーソナライズド広告」は類似情報検索を広告に用いることにより、検索条件やユーザの嗜好に応じた広告を配信するものである。特許文献として表中番号 8、10、11、19、30、31、35、40、40、46、51 が挙げられる。8 は過去に商品をカートに入れた履歴を用いて広告を作るものである。10 は設定した条件に応じた広告配信を行うもの、11 はパーソナライズされた音声広告を生成するものである。19 はユーザの所在地、年齢などに応じて広告配信を行うものである。30 は広告にも応用可能なブックマークの整理方法である。31 は広告を含めた様々なメディアのパーソナライズド検索手法である。35 はブックマークからユーザに必要な広告を分析するもの、40 は選択した製品の関連情報、広告を提供するもの、46 はブックマークと検索クエリから旅行情報を提供するものである。また、51 はドキュメントの意味情報から対応する広告を発見する方法である。

「情報の関連度評価」は類似・関連情報検索を行うための情報の評価方法に関するものである。特許文献として表中番号、12、21、27、41、42、53 が挙げられる。12 は情報をクラスタとして分類し、最も類似のクラスタを発見する手法である。21 はユーザのスケジュールから自動的に関連情報を収集し、関連度を計測するものである。27 は他の情報からの参照により関連度を得る方法である。41 は検索クエリのテンプレートを用いることで関連度の低い情報を排除する方法であり、42 は様々な種類の履歴データを用いて関連度を評価する方法である。53 はリンク構造から関連度を評価する方法である。

「類似情報検索システム」は類似検索をシステムとして取り入れた例である。特許文献として表中番号 20、28、54、58、66、69 が挙げられる。20 は知識データベースで利用し、質問に最適な回答を選ぶものである。28 は色表現から類似パターンを発見するシステム、54 は質問回答システムにおいて、回答をクラスタ化することで類似情報を得るものである。58 は類似する関心を持つユーザを発見するシステムである。66 はカラーヒストグラムから類似画像を発見するシステム、69 は検索クエリの部分文字列関連情報を発見するシステムである。



図 3-3-1 主要サブテーマの特許変遷 (類似・関連情報検索)



### 情報の関連度評価

**番号12**  
 ・出願人  
 SIEMENS CORP RES INC(米国)  
 ・発明の名称  
 WWW search facilitating method in computer - involves processing incoming query by selecting cluster centroid that is most similar to incoming query, and returning weight associated with selected cluster as weight of current search engine  
 ・概要  
 情報をクラスタとして分類し、最も類似のクラスタを発見する手法

**番号21**  
 ・出願人  
 APPLE COMPUTER INC(米国)  
 ・発明の名称  
 Automatic retrieval method of information relevant to user's scheduled event - involves retrieving identified relevant information resulting from searching information source using formulated query  
 ・概要  
 コーザのスケジュールから自動的に関連情報を収集し、関連度を計測する

**番号27**  
 ・出願人  
 GOOGLE INC  
 日立製作所,日立情報システムズ(米国)  
 ・発明の名称  
 Document identification method for web-based applications, involves refining relevance sources for documents in initial set based on local score value that quantifies amount of reference of at least two documents  
 ・概要  
 他の情報からの参照により関連度を得る方法

**番号41**  
 ・出願人  
 ASK JEEVES INC(米国)  
 ・発明の名称  
 Natural language information retrieval method for website searching, involves receiving results based on initiated information retrieval commands associated with selected query template  
 ・概要  
 検索クエリのテンプレートを用いることで関連度の低い情報を排除する方法

**番号53**  
 ・出願人  
 ALPHA S A(米国)  
 ・発明の名称  
 Relevance rank determination method for web page, involves adjusting content-based relevance rank, based on link structure of pages including link rank values from in-coming links  
 ・概要  
 リンク構造から関連度を評価する方法

**番号42**  
 ・出願人  
 GOOGLE INC(米国)  
 ・発明の名称  
 Document scoring method in information retrieval system, involves obtaining multiple types of history data associated with identified document  
 ・概要  
 様々な種類の履歴データを用いて関連度を評価する方法

### 類似情報検索システム

**番号20**  
 ・出願人  
 アイエーシー サーチ アンド メディア(米国)  
 ・発明の名称  
 グラマーテンプレートクエリシステム  
 ・概要  
 知識データベースで利用し、質問に最適な回答を選ぶもの

**番号28**  
 ・出願人  
 イーストマン コダック(米国)  
 ・発明の名称  
 類似のパターンの検索と取り出しのための方法及びシステム  
 ・概要  
 色表現から類似パターンを発見するシステム

**番号54**  
 ・出願人  
 松下電器産業(日本)  
 ・発明の名称  
 情報検索システム  
 ・概要  
 質問回答システムにおいて、回答をクラスタ化することで類似情報を得るもの

**番号58**  
 ・出願人  
 インターナショナル ビジネス マシーンズ(米国)  
 ・発明の名称  
 類似する関心を有する人物の間のネットワ - キングを容易にし、情報に関する共同作業検索を容易にする、インタ - ネット・ベースの方法  
 ・概要  
 類似する関心を持つユーザを発見するシステム

**番号66**  
 ・出願人  
 エル ジー エレクトロニクス(韓国)  
 ・発明の名称  
 HMM Dカラースペースに基づいたカラ - 量子化方法を用いたマルチメディアの検索方法  
 ・概要  
 カラーヒストグラムから類似画像を発見するシステム

**番号69**  
 ・出願人  
 AT & T CORP(米国)  
 ・発明の名称  
 Query substring indexing method for database search system involves matching query substring to large collections of string data in database using positional q-grams, maintained in index structure  
 ・概要  
 検索クエリの部分文字列関連情報を見出すシステム

表 3-3-1 サブテーマ一覧 (類似・関連情報検索)

サブテーマ ID	サブテーマ名	特許件数	主要サブテーマ
A	パーソナライズド広告	10	
B	情報のカテゴライズ	10	
C	関連情報の発見	8	
D	情報のパーソナライズ	8	
E	情報の関連度評価	6	
F	類似情報検索システム	6	
G	複数データソースの利用	3	
H	広告仲介システム	3	
I	ハイパーリンク操作	3	
J	トレンド情報配信	2	
K	商品の関連情報分析	2	
L	分野別検索	2	
M	検索の効率化	3	
N	ガイドシステム	2	
O	情報クローリング	2	

表 3-3-2 重要特許一覧 (類似・関連情報検索)

番号	特許情報ファミリー単位	最先の優先権主張日	被引用回数	発明の名称	出願人/権利者	サブテーマ
8	WO2000017793A1, AU9961447A, US20010021914A1, US6317722B1, US20060195362A1, US7113917B2, US20080033746A1, US20080033821A1, US20080040239A1	1998/9/18	114	Recommending e.g. book, compact disc, video disc etc to on-line user of E-commerce by accessing data structure that identifies corresponding set or similar item for combining set of similar items to generate set of additional items	AMAZON.COM	パーソナライズド広告
10	WO9916003A1, AU9894941A, US6085229A	1997/9/22	71	Computer network for providing client-side personalization of information	BELARC INC	パーソナライズド広告
11	US6073105A	1997/6/13	67	Personal advertisement transfer system between private and public network via internet, receives audio files from external audio text greeting system and provides to selected user over public network	TELE PUBLISHING INC	パーソナライズド広告
12	US5864845A	1996/6/28	56	WWW search facilitating method in computer - involves processing incoming query by selecting cluster centroid that is most similar to incoming query, and returning weight associated with selected cluster as weight of current search engine	SIEMENS CORP RES INC	情報の関連度評価

番号	特許情報 ファミリー単位	最先の優先 権主張日	被引用 回数	発明の名称	出願人/ 権利者	サブテ ーマ
19	WO9727531A1, AU9714563A, EP877314A1, JP09526727X, CN1214132A, KR99081831A, KR305902B, US20020095407A1, US6639608B1, US20040201616A1, JP3602030B2, CN1146781C, JP2000348065, JP2000155741, JP3602030B	1996/1/23	39	情報提供装置	グローバルメ ディアオンライン	パーソ ナライ ズド広 告
20	WO2000057302A1, AU2000033782A, EP1163609A1, US6584464B1, JP2002540506	1999/3/19	37	グラマ - テンプレ - トクエリシステ ム	アイエーシー サーチ アンド メディア	類似情 報検索 システ ム
21	US5842009A	1996/9/30	37	Automatic retrieval method of information relevant to user's scheduled event - involves retrieving identified relevant information resulting from searching information source using formulated query	APPLE COMPUTER INC	情報 の 関 連 度 評 価
27	US6526440B1	2001/1/30	25	Document identification method for web-based applications, involves refining relevance sources for documents in initial set based on local score value that quantifies amount of reference of at least two documents	GOOGLE INC	情報 の 関 連 度 評 価
28	EP1128282A2, US6584465B1, JP2001282848	2000/2/25	25	類似のパタ - ンの検索と取り出しの ための方法及びシステム	イーストマン コダック	類似情 報検索 システ ム
30	US6212522B1	1998/5/15	19	Book marks serving method for client browser, involves searching several book mark sets corresponding to keyword contained in received search query for transmission to client	IBM	パーソ ナライ ズド広 告
31	AU2000062003A, WO2000079436A2, US7089236B1	1999/6/24	19	Search request development aiding method for targeting content e.g. banner advertisements to user, involves providing expanded search request to search engine which identifies associated information	SIMPLI.COM	パーソ ナライ ズド広 告
35	US6247021B1	1998/5/15	17	Internet advertising medium for retrieving web pages, has searchable bookmark set which is sent to request client, when search request is satisfied by bookmark set	IBM	パーソ ナライ ズド広 告
40	US6675165B1	2000/2/28	13	Interactive search system for advertising media related information, has information collecting and storing unit which collects and stores search-object related information to provide supply of search-object related information	BARPOINT.COM INC	パーソ ナライ ズド広 告

番号	特許情報 ファミリー単位	最先の優先 権主張日	被引用 回数	発明の名称	出願人/ 権利者	サブテ ーマ
41	US20030069880A1, US7403938B2	2001/9/24	13	Natural language information retrieval method for website searching, involves receiving results based on initiated information retrieval commands associated with selected query template	ASK JEEVES INC	情報の 関連度 評価
42	US20050071741A1, WO2005033978A1, IN2006CN01476P4, US7346839B2	2003/9/30	12	Document scoring method in information retrieval system, involves obtaining multiple types of history data associated with identified document	GOOGLE INC	情報の 関連度 評価
46	US6256639B1	1998/5/15	11	Internet travel service providing method involves serving installable bookmark sets associated with destination on receiving travel related search query	IBM	パーソ ナライ ズド広 告
51	US6816857B1	1999/11/1	9	Advertisement and relevant document determination method in internet, involves organizing concepts according to meanings, into predefined lexicon defining element of semantic space, based on which input data is searched	APPLIED SEMANTICS INC	パーソ ナライ ズド広 告
53	US20030061214A1, US7010527B2	2001/8/13	9	Relevance rank determination method for web page, involves adjusting content-based relevance rank, based on link structure of pages including link rank values from in-coming links	ALPHA S A	情報の 関連度 評価
54	EP1156430A2, CN1324046A, KR2001105241A, US20020065845A1, JP3654850B2, US6993517B2, JP2002041573, JP3654850B	2000/5/17	9	情報検索システム	松下電器産業	類似情 報検索 システ ム
58	TW548558A, US6745178B1, JP3629438B2, JP2001357068, JP3629438B	2000/4/28	8	類似する関心を有する人物の間のネットワ - キングを容易にし、情報に関する共同作業検索を容易にする、インタ - ネット・ベ - スの方法	IBM	類似情 報検索 システ ム
66	BR2001003093A, CN1326160A, EP1176804A2, JP2002027277A, US20010055128A1, KR2001107506A, TW505893A, KR374063B, EP1517540A2, US6919971B2, JP3676259B2, JP2005216295A, CN1619543A, CN1227612C, JP2005216296	2000/5/26	6	HMM Dカラ - スペ - スに基づいたカラ - 量子化方法を用いたマルチメディアの検索方法	エル ジー エ レクトロニクス	類似情 報検索 システ ム
69	US7010522B1	2002/6/17	5	Query substring indexing method for database search system involves matching query substring to large collections of string data in database using positional q-grams, maintained in index structure	AT & T CORP	類似情 報検索 システ ム

## 第4部 研究開発動向分析

### 第1章 調査概要および情報収集方法

#### 第1節 調査対象とする論文母集団の設定について

##### 1. 母集団の設定

本調査が対象とする文献は1990年から2007年12月までに特定の論文誌、研究会、国際会議で発表された、インターネット社会における検索技術に関する学术论文である。

##### 2. 利用データベースおよび検索結果

国内外の論文について調査を行ったが、概要では海外論文について紹介する。表4-1-1に示す関連する主要な国際会議を調査対象とした。SIGIR(International Conference on Research and Development in Information Retrieval)は情報検索全般を取り扱っている大きな国際会議である。WWW(International Conference on World Wide Web)は、Webにおける情報システムやコンテンツなどを対象とした国際会議である。また、今日の検索サービスは大規模データベースの設計に大きく関係していることから、VLDB (International Conference on Very Large Data Bases)も対象とした。いずれも注目度の高い国際会議である。

海外特許同様に以下の条件で絞り込みを行っている。

(WEB or WWW or INTERNET) and (SEARCH or RETRIEVAL or RETRIEVE)

表 4-1-1 海外の対象論文誌 / 研究会報告名

論文誌 / 研究会報告名	論文数
ACM SIGIR (情報検索に関する国際会議)	1400
ACM WWW (ウェブ全般に関する国際会議)	1062
VLDB (大規模データベースに関する国際会議)	155
合計	2617

ただし、VLDBは1992年、WWWは1994年から開催されている。

以上、国内外を合わせて3228件を調査対象論文とした。

#### 第2節 分析にあたっての留意点

特許との比較のため、各論文と研究者所属機関国籍の対応は一論文一国籍とし、筆頭著者の所属機関から導いている。所属機関名とその国籍は特許同様に名寄せを行ったものである。また所属機関別集計においては、一著者一所属とし、各著者の筆頭所属機関から導いている。国籍の不明なものは「その他」としている。

対象論文のうち1990年のものは0件であったため、グラフは1991年から2007年で表示している。ただし、特許との比較については、特許側に合わせ1990年から2006年で表示している。

欧州国籍の内訳、技術区分への分類方法などは特許分析に準じる。

## 第2章 全体動向分析

### 第1節 研究者所属機関国籍別論文献数推移

本章より、海外での論文発表について分析結果を示す。

研究者所属機関国籍別の発表件数を以下に示す。最も多いのは米国国籍の発表であり、次いで欧州国籍である。日本国籍は米国国籍に次いでいた特許と異なり、論文では中国国籍よりも少ない。件数の推移は右肩上がりであるが、2004年に大きく伸び、そこからしばらく横這いで、2007年に再び大きく伸びている。こうした変化は米国国籍に顕著である。特許が近年横這いであることと比較しても、右肩上がりであることは目立つが、対象となる学会そのものが拡大したことも大きな要因である。

図 4-2-1 研究者所属機関国籍別論文献数推移

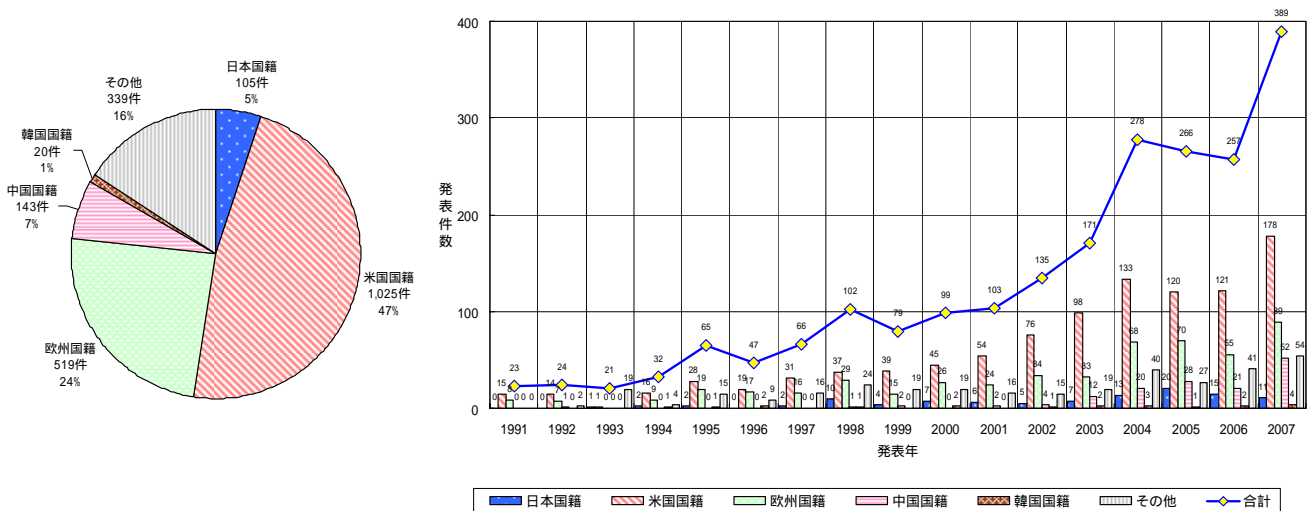
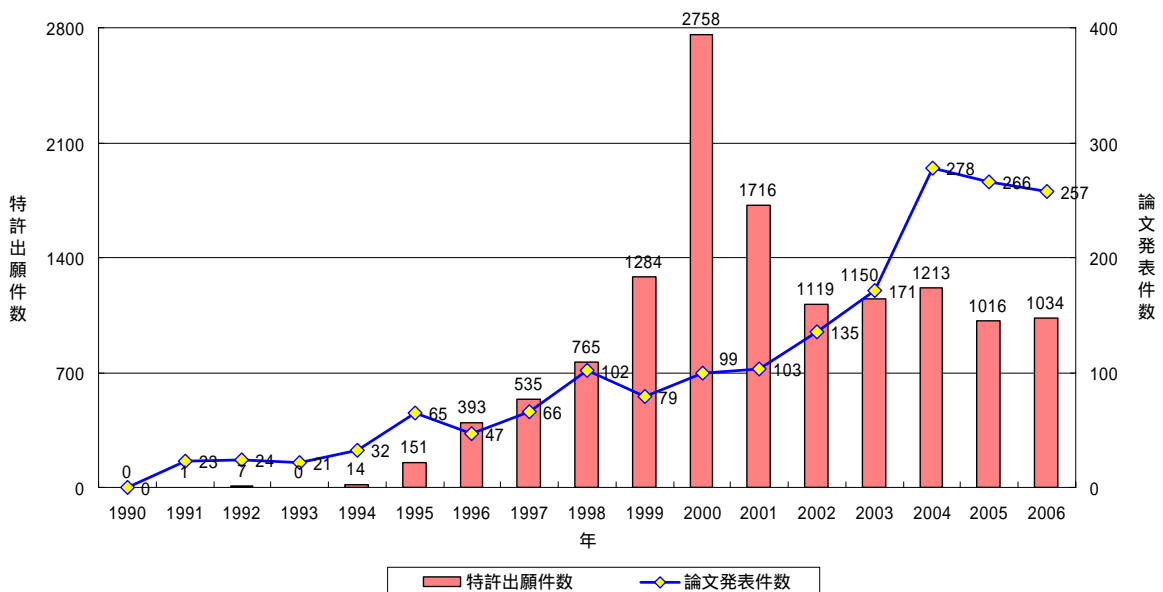
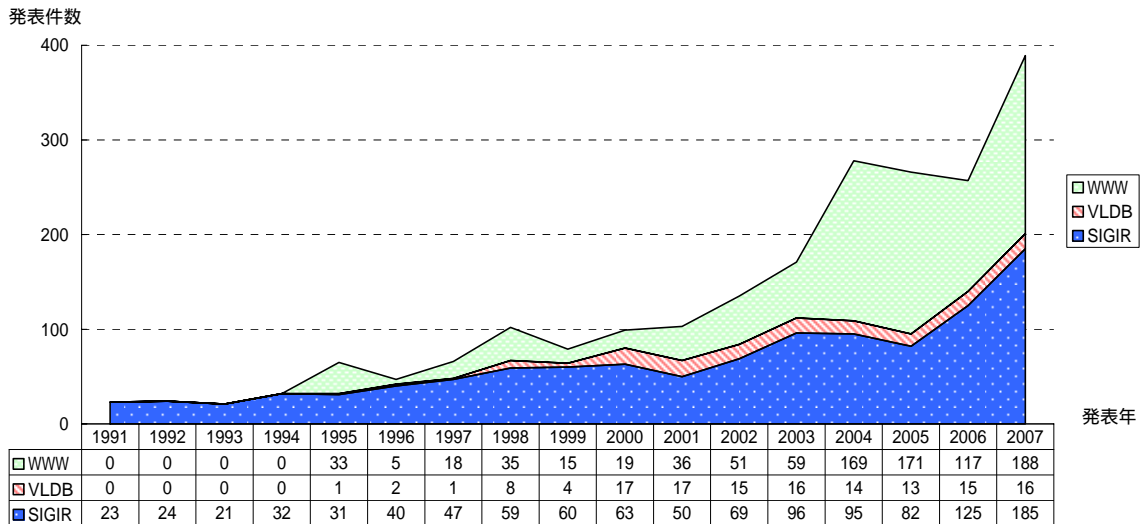


図 4-2-2 論文件数と特許動向との比較推移



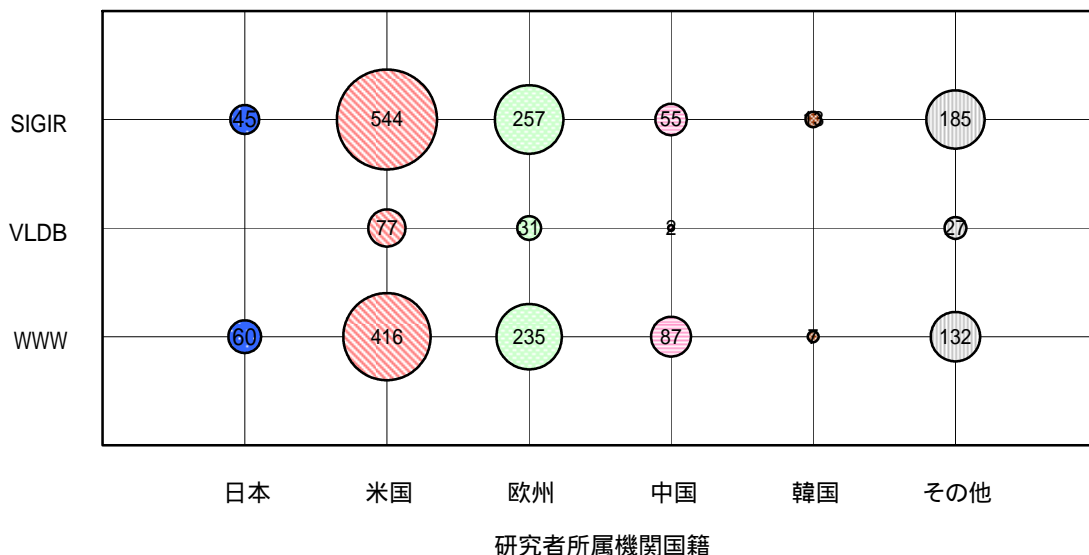
論文誌別の内訳推移についても以下に示す。WWW は 1998 年と 2004 年以降にポスターセッションが開催されたため、急増している。SIGIR が 2000 年を境に急増しているが、VLDB は近年あまり変化がない。

図 4-2-3 論文誌別の件数内訳推移



各論文誌の研究者所属機関国籍内訳は以下の通りである。米国国籍がいずれにおいても中心であるが、SIGIR よりも WWW のほうが比較的多国籍である。

図 4-2-4 研究者所属機関国籍別 - 論文誌別件数





### 第3章 技術区分別動向分析

#### 第1節 技術区分別 - 研究者所属機関国籍別論文件数

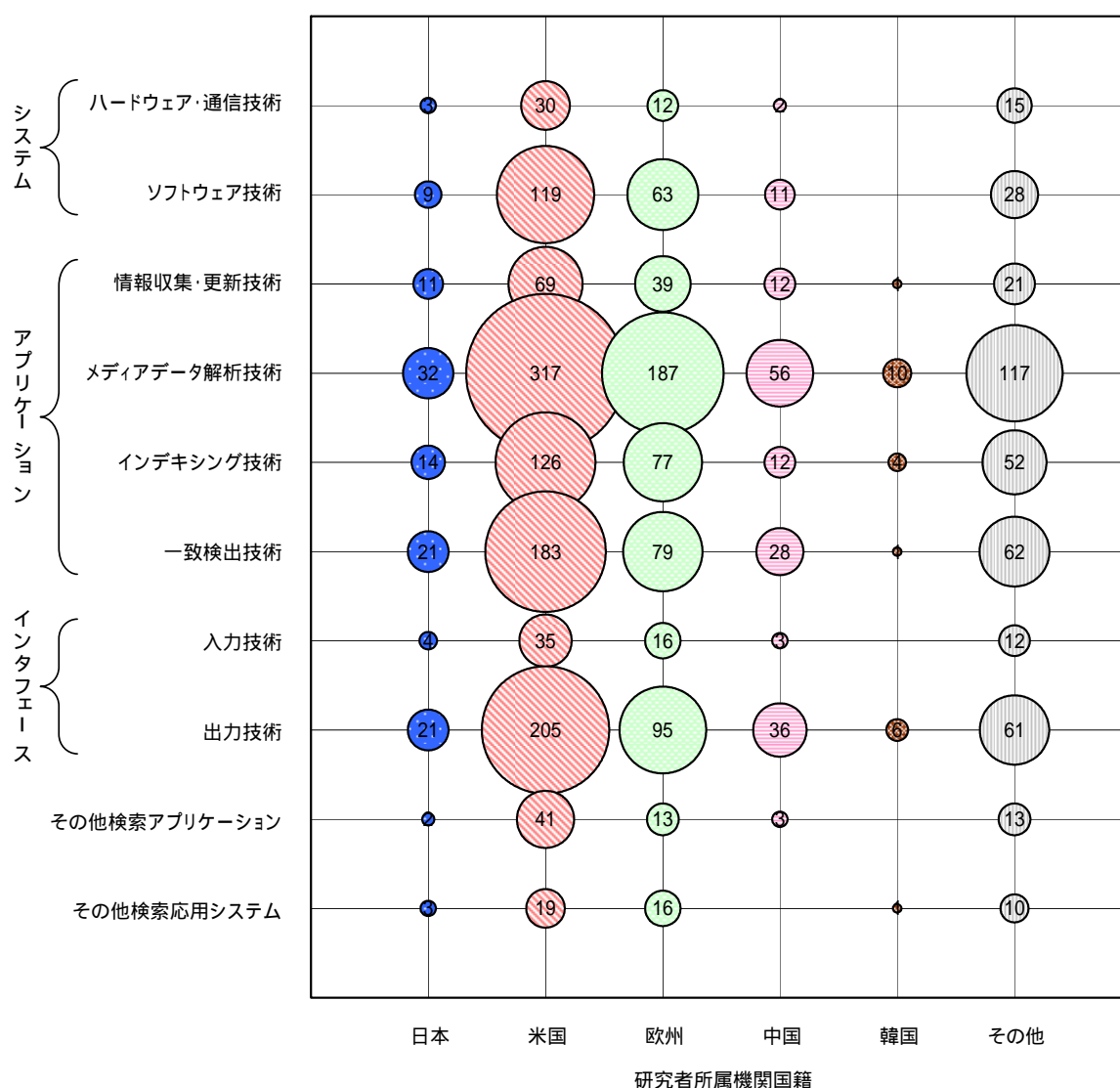
##### 1. 要素技術別

##### (1) 要素技術中区分別

要素技術の中区分別かつ研究者所属機関国籍別の累積論文件数を以下に示す。

多くの区分について、米国国籍による発表が最も多く、特に件数が多いのは特許同様に「メディアデータ解析技術」「一致検索技術」「出力技術」である。日本国籍による発表でも同様の3つの区分が多い。

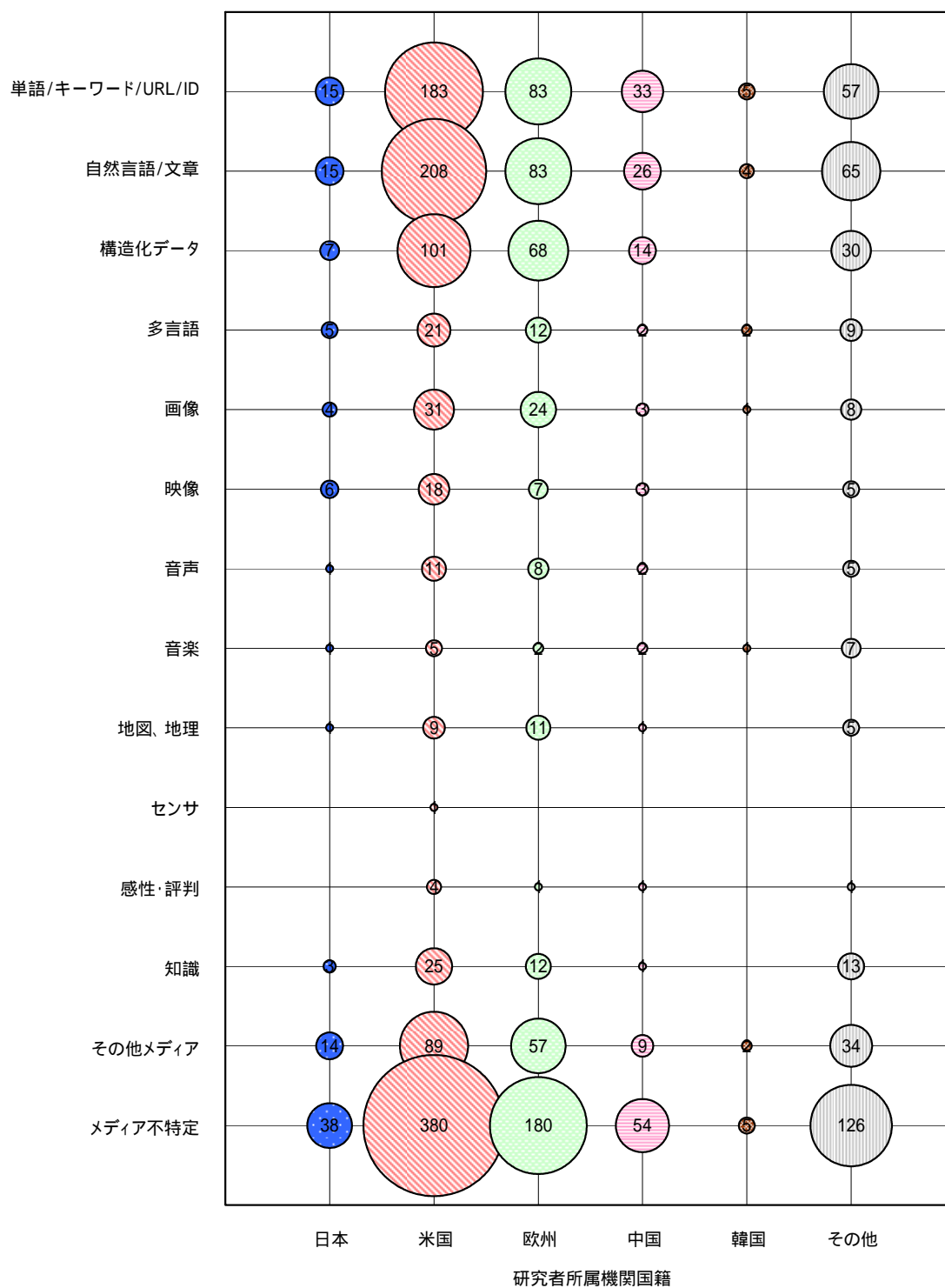
図 4-3-1 技術区分別 - 研究者所属機関国籍別論文件数



## 2. メディア別

取扱いメディア別かつ研究者所属機関国籍別の累積論文件数を以下に示す。米国国籍では「自然言語/文章」が「単語/キーワード/URL/ID」を上回っている。さらに「メディア不特定」が他のメディアより多いのは特許とは異なる点である。

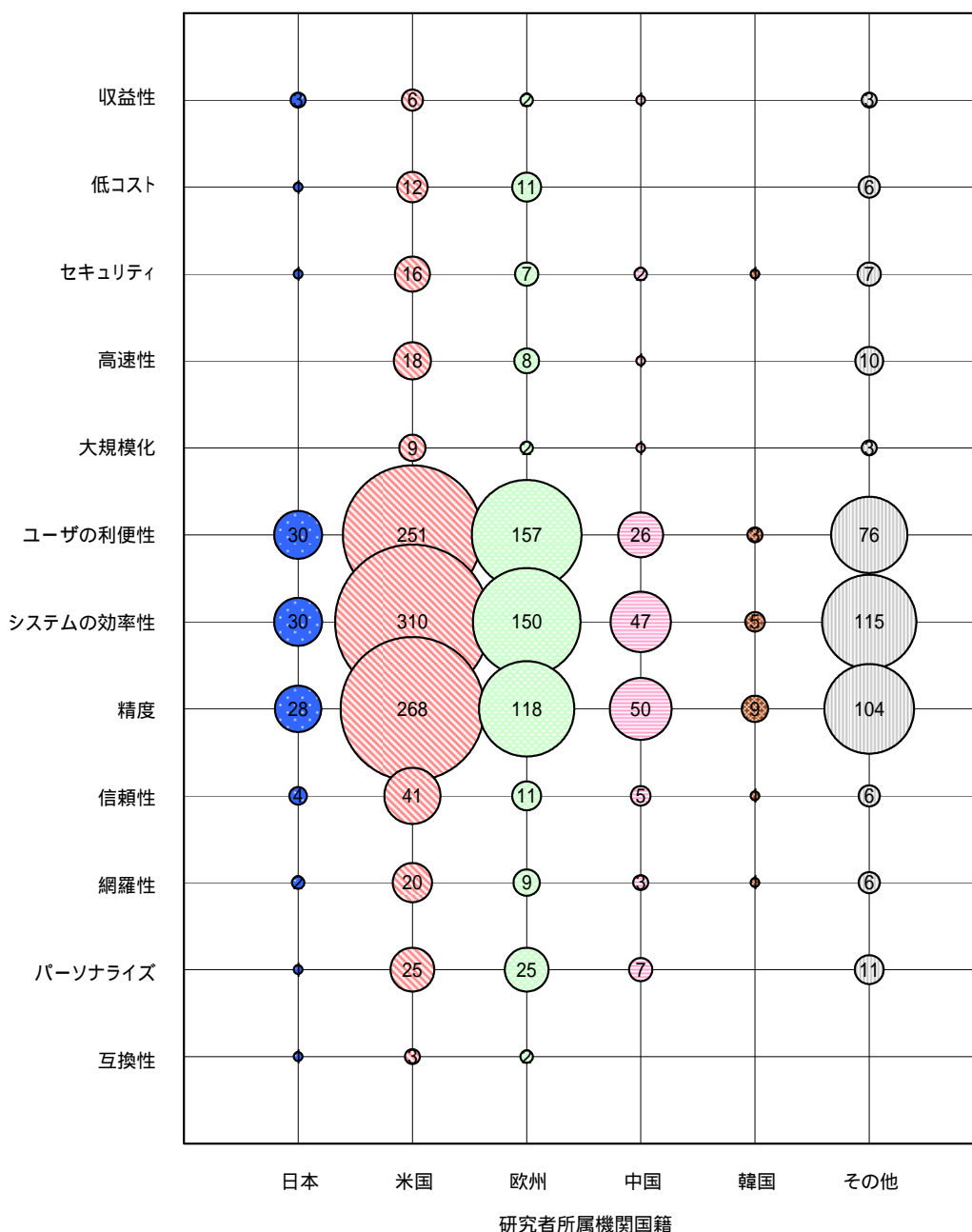
図 4-3-2 メディア別 - 研究者所属機関国籍別論文件数



### 3. 課題別

課題別かつ研究者所属機関国籍別の累積論文件数を以下に示す。特許では「ユーザの利便性」が他を圧倒しているが、論文では「システムの効率性」「精度」がそれを上回っており、欧州国籍だけが「ユーザの利便性」が筆頭になっている。

図 4-3-3 課題別 - 研究者所属機関国籍別論文件数



## 第4章 研究者所属機関・研究者別動向分析

### 第1節 研究者所属機関別発表件数上位ランキング

研究者所属機関別の発表件数上位を表 4-4-1 に示す。米国は企業としてマイクロソフト、IBM、ヒューレット・パッカード、ヤフーなど、大学としてカリフォルニア大学、マサチューセッツ大学、カーネギーメロン大学、イリノイ大学、メリーランド大学など、大半を占めている。

アメリカ以外では中国の香港中文大学、シンガポールのシンガポール大学が登場しているが、日本の最高位は日本電気の 21 位である。

表 4-4-1 研究者所属機関別発表件数上位ランキング

順位	研究者所属機関	件数
1	マイクロソフト（米国）	126
2	IBM（米国）	109
3	カリフォルニア大学（米国）	82
4	マサチューセッツ大学（米国）	78
5	カーネギーメロン大学（米国）	67
6	イリノイ大学（米国）	47
7	ヤフー（米国）	42
8	ヒューレット・パッカード（米国）	40
9	香港中文大学（中国）	38
10	シンガポール国立大学（シンガポール）	36
11	メリーランド大学（米国）	35
12	AT&T（米国）	32
	グラスゴー大学（英国）	32
14	スタンフォード大学（米国）	30
	ロイヤルメルボルン工科大学（オーストラリア）	30
16	ワシントン大学（米国）	29
17	コーネル大学（米国）	28
18	マサチューセッツ工科大学（米国）	27
19	清華大学（中国）	23
	グーグル（アメリカ）	23

## 第2節 研究者別発表件数上位ランキング

研究者の発表件数上位を表 4-4-2 に示す。研究機関別上位の米国大学が上位を占め、同組織での共著関係も多い。

表 4-4-2 研究者別発表件数上位ランキング

順位	研究者	所属	件数
1	W. Bruce Croft	マサチューセッツ大学 (米国)	28
2	Zheng Chen	マイクロソフト (米国)	26
	Wei-Ying Ma	マイクロソフト (米国)	
4	James Allan	マサチューセッツ大学 (米国)	20
5	Cheng Xiang Zhai	イリノイ大学 (米国) カーネギーメロン大学 (米国)	19
6	Ji-Rong Wen	マイクロソフト (米国)	18
7	Susan T. Dumais	マイクロソフト (米国)	17
	James P. Callan	カーネギーメロン大学 (米国)	
9	Justin Zobel	ロイヤルメルボルン工科大学 (オーストラリア)	15
10	Mark Sanderson	シェフィールド大学 (英国)	14
	C. Lee Giles	ペンシルバニア州立大学 (米国)	
	Hua-Jun Zeng	マイクロソフト (米国)	
13	Ophir Frieder	イリノイ大学 (米国)	13
	Jian-Yun Nie	モンリオール大学 (フランス)	
15	Charles L. A. Clarke	ウォータールー大学 (カナダ)	12
16	Ravi Kumar	ヤフー (米国)	11
	酒井哲也	東芝 (日本) ニュースウォッチ (日本)	
	Jan O. Pedersen	ゼロックス (米国)	
19	Soumen Chakrabarti	IBM (米国)	10
	田中克己	京都大学 (日本)	
	Javed A. Aslam	ダートマス大学 (米国)	
	Douglas W. Oard	メリーランド大学 (米国)	
	Edward Allan Fox	バージニア工科大学 (米国)	
	Ryen W. White	グラスゴー大学 (英国)	
	Joemon M. Jose	グラスゴー大学 (英国)	
	Norbert Fuhr	ドルトムント大学 (ドイツ) ドウイスブルグ - エッセン大学 (ドイツ)	
	Alistair Moffat	メルボルン大学 (オーストラリア)	
	Mounia Lalmas	グラスゴー大学 (英国)	
	Krishna Bharat	デジタル・イクイップメント・コーポレーション (米国)	
	Andrew Tomkins	IBM (米国)	

## 第5部 市場・政策動向分析

### 第1章 政策動向

#### 第1節 日米欧中韓における政策

図 5-1-1 に日米欧中韓の情報検索技術関連政策、および、図 5-1-2 に出願先国別 - 出願人国籍別出願件数を示す。

日本については、米国と比較して政策面での遅れはあるものの、近年、文部科学省主導の情報爆発や経済産業省主導の情報大航海等、次世代検索技術開発のための政策が講じられており、今後、情報検索技術関連の研究・サービスの増加に伴い、特許出願数も増加すると予想される。近年、取扱いメディアをセンサとした特許出願件数も増えており、さらに情報大航海等が研究の盛り上がりを促進させるであろう。

米国では、1990年代初頭から国防総省の機関である DARPA や NIST が主導となり、情報技術関連の政策が講じられ、検索エンジンの誕生に貢献、サービスも盛んになった結果、今回の調査対象領域における出願件数が他地域と比較して多くなったと考えられる。

欧州については、欧州産検索エンジン開発のためのプロジェクトが行われている。中でも、セマンティック Web への対応を目指したドイツの Theseus や、次世代のマルチメディア検索プラットフォームの開発を目的とした EU の Pharos 等の政策により、今後関連の特許出願件数が増えると想定される。

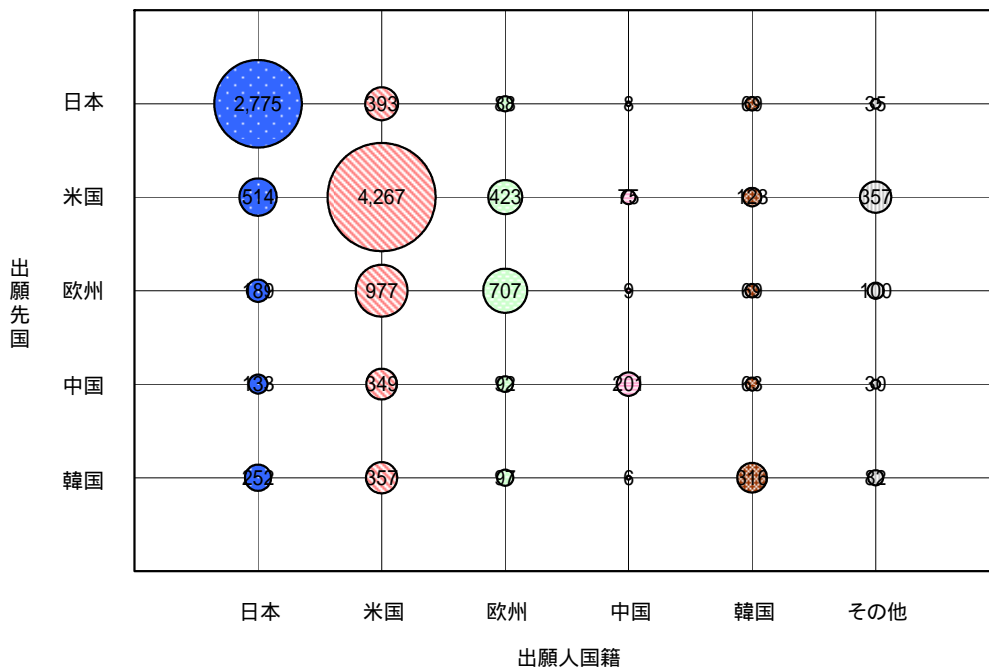
中国では、検索技術を銘打った政策は講じられていないものの、ハイテク産業技術の開発を目的とした科学技術部主導の 863 計画や 973 計画において、情報検索技術分野の研究への助成が行われている。今回の調査対象領域への出願はあまり行われていない。

韓国については、情報検索技術研究に対する主だった政策的な支援は行われておらず、同分野の研究は、民間企業や学術機関に委ねられている。同国は、ブロードバンド普及等、情報化インフラ整備のための政策は進んでおり、その結果、画像・映像・音楽といったコンテンツの利用や、ユーザが作成するコンテンツ（UCC：User Created Contents、CGM、UGC とも呼ばれる）の構築が進んだ。特許出願数への影響はそれ程大きくないものの、近年、若干画像、映像に関する特許出願件数が増えている。

図 5-1-1 日米欧中韓の情報検索技術関連政策

国	プロジェクト名	予算	取扱メディア	~ 1996年	1997年	1998年	1999年	2000年	2001年	2002年	2003年	2004年	2005年	2006年	2007年	2008年	
日本	情報学	総額2億円	自然言語等						←								
	e-society	不明	テキスト								←						
	情報爆発	総額30数億円	自然言語等									←					
	情報大航海	総額150億円	センサ、行動履歴												←		
	情報信憑性	総額16億円	文字、音声、映像												←		
	Sensing Web	総額2.5億円	センサ												←		
	NTCIR	不明	自然言語/文章			←											
米国	TIPSTER	不明	自然言語/文章	←													
	TIDES	不明	自然言語/文章 多言語				←									→	
	TREC	不明	自然言語/文章	←													
	DLI	フェーズで年間 2,400万ドル	自然言語/文章	←		←											
	DAML	総額7,000万ド ル	構造化言語					←						→			
	UIMA	不明	非構造化データ									←				→	
	手書き画による3次元検索	総額80万ドル (推定)	3D画像										←	→			
	概念の関連性を検索する技術	総額50万ドル (推定)	自然言語/文章						←								
	GALE	総額4,630万ド ル	多言語・音声・テキ スト													←	
	Quaero	2億5,000万 ユーロ	音声、画像、映像等 マルチメディア											←			
欧州	Theseus	不明	文書、音声、画像等												←		
	Pharos	1,425万ユーロ	マルチメディア													←	
	FP7	1億100万ユー ロ	セマンティクス													←	
	VITALAS	469万ユーロ	マルチメディア													←	
	欧州電子図書館	不明	単語、地図、写真、 画像、映像										←				
	CLEF	不明	自然言語/文章						←								
	中国	863計画	不明		←												
973計画		不明			←												
第11次5ヵ年計画		不明													←		
インターネット情報サービス管理 方法に関する法規		不明							←								
金盾プロジェクト		不明														←	
韓国	IT839政策	2兆ウォン										←					

図 5-1-2 出願先国別 - 出願人国籍別出願件数



## 第2章 市場動向分析

### 第1節 日米欧中韓の検索サイトシェアまとめ

表 5-2-1 に、上述した日米欧中韓の検索サイトシェアを地域別に纏めたものを表す。表から分かる通り、Google はアルファベット言語圏比率の高い米国・欧州でシェアが高く、アルファベット言語圏ではない日本・中国・韓国ではシェアを伸ばしていない。特に、中国・韓国では自国の検索サイトである Baidu、Naver のシェアが高くなっている。これらの背景には、アルファベット言語圏以外の特有の言語処理技術や、各国の文化的な差異による検索サイトの使い方の違い、インターネットが普及した背景の違い等、複数の要素があると考えられる。

表 5-2-1 日米欧中韓の検索サイトシェア比較

検索サイト <sup>3</sup>	地域				
	日本	米国	欧州	中国	韓国
Google (米国)	31.3%	63.0%	79.2%	27.0%	1.7%
Yahoo! (米国)	56.2%	19.6%	2.0%	2.4%	4.6%
MSN (米国)	2.8%	8.3%	1.9%	0.0%	0.0%
Baidu (中国)	0.0%	0.0%	0.0%	60.9%	0.0%
Naver (韓国)	0.0%	0.0%	0.0%	0.0%	72.1%
ebay (米国)	0.0%	0.0%	3.1%	0.0%	0.0%
Daum (韓国)	0.0%	0.0%	0.0%	0.0%	15.4%
Ask (米国)	0.0%	4.8%	0.0%	0.0%	0.0%
Yandex (ロシア)	0.0%	0.0%	2.2%	0.0%	0.0%
AOL (米国)	0.0%	4.3%	0.0%	0.0%	0.0%
BIGLOBE (日本)	3.8%	0.0%	0.0%	0.0%	0.0%
goo (日本)	2.0%	0.0%	0.0%	0.0%	0.0%
Sogou (中国)	0.0%	0.0%	0.0%	3.1%	0.0%
Empas (韓国)	0.0%	0.0%	0.0%	0.0%	1.1%

### 第2節 各国の市場と特許出願動向・重要特許との関連分析

#### 1. 日本

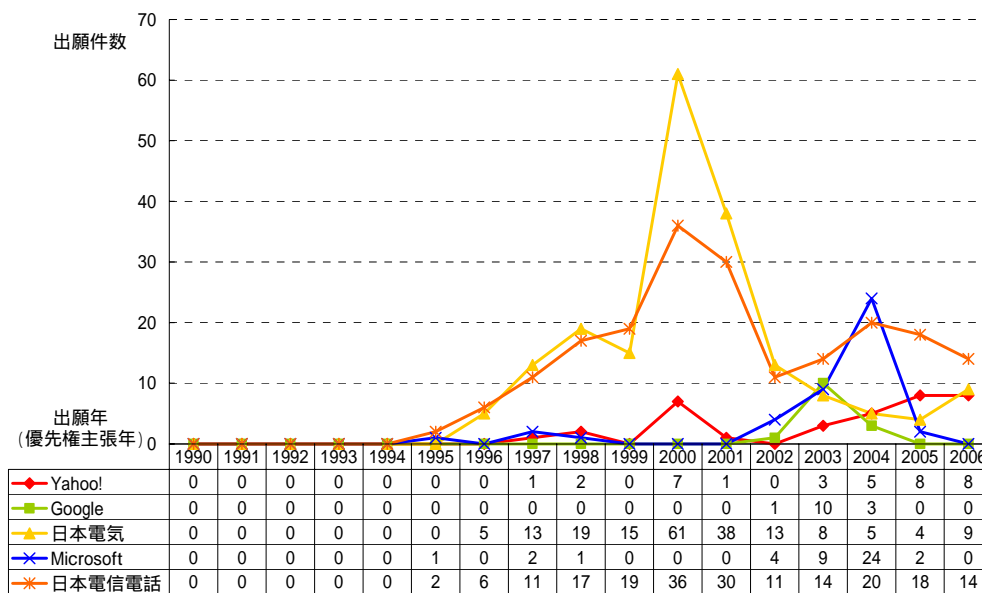
##### (1) 特許出願動向との関連分析

日本における検索サイトの市場シェア上位5社に関する今回の調査対象領域での特許出願件数の推移（日本への出願）を、図 5-2-1 に示す。図 5-2-2 から読み取れるように、今回の調査対象領域へは、BIGLOBE（NECビッグロブ株式会社の商標登録）を展開するNECビッグロブの親会社である日本電気は、2000年には60件程の特許を出願しているものの、その後減少傾向にある。gooを展開するNTTレゾナントの親会社である日本電信電話では、2000年以降も他社と比較して日本への出願を積極的に行っていると言える。MSNを展開するMicrosoftは、2002年以降日本への出願を増やしている。一方、日本でシェアの高いYahoo!とGoogleについては、今回の調査対象領域では、日本への出願はあまり行われていない。

<sup>3</sup> 検索サイトの名称に関しては、各社の商標、または、登録商標である。



図 5-2-1 日本検索サイト上位 5 社の検索技術関連特許出願件数推移（日本への出願）<sup>4</sup>



## (2) 重要特許との関連分析

日本国籍の重要特許は表 5-2-2 の通りである。家電メーカーを中心に「他機器連携/データ配信」に関する特許が目立つ一方、検索エンジンの根幹「検索用インデクス・キーワードの抽出、作成、更新」や「検索結果の抽出・統合」は 0 件であり、市場動向と同様に、海外に押されている様子が見てとれる。

表 5-2-2 日本国籍の重要特許件数

テーマ	特許件数
地理・地図データ解析	2
検索結果の抽出・統合	0
類似・関連情報検索	2
テキストデータ解析	4
検索用インデクス・キーワードの抽出、作成、更新	0
他機器連携/データ配信	7
画像・映像解析	4

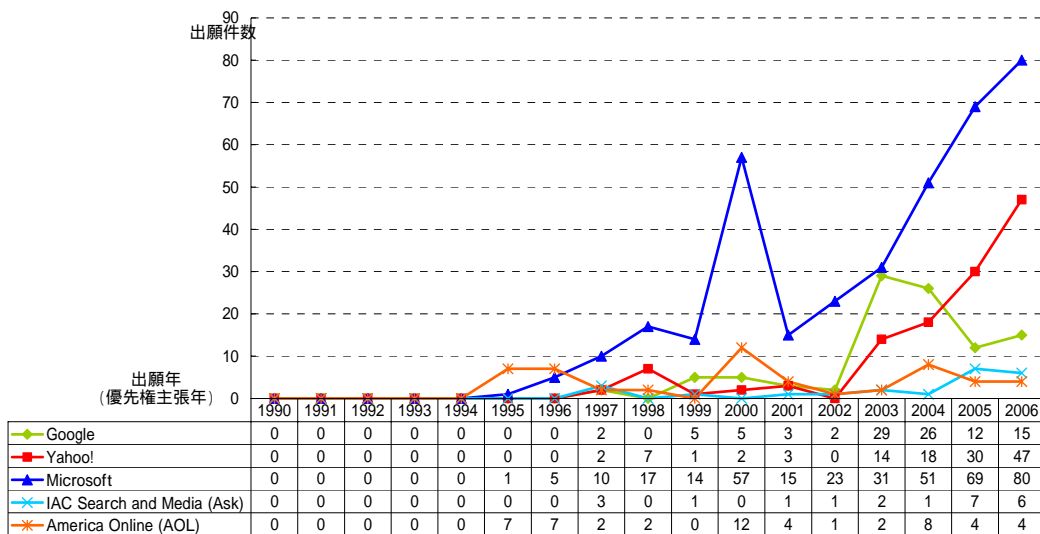
## 2. 米国

### (1) 特許出願動向との関連分析

図 5-2-2 に、米国における検索サイトシェア上位 5 社の今回の調査対象領域での特許出願件数の推移(米国への出願)を示す。今回の調査対象領域への出願は、5 社の中では Microsoft による出願が最も多く、IT バブルであった 2000 年頃に同社の出願件数は一旦ピークを迎え、その後出願数は減少したものの、2002 年頃から再び増加傾向にある。Yahoo! も 2003 年頃から出願数を増やしている。Google は 6 割を超える市場シェアを有するものの、今回の調査対象領域への特許出願件数は少ない。

<sup>4</sup> 検索サイトの名称に関しては、各社の商標、または、登録商標である。

図 5-2-2 米国検索サイト上位 5 社の検索技術関連特許出願件数推移（米国への出願）<sup>5</sup>



## (2) 重要特許との関連分析

米国国籍の重要特許は表 5-2-3 の通り。ほとんどのテーマで他国を圧倒しているが、特に「検索結果の抽出・統合」「類似・関連情報検索」「検索用インデクス・キーワードの抽出、作成、更新」など検索サービスの基本技術に強みを持っている。

表 5-2-3 米国国籍の重要特許件数

テーマ	特許件数
地理・地図データ解析	5
検索結果の抽出・統合	10
類似・関連情報検索	19
テキストデータ解析	8
検索用インデクス・キーワードの抽出、作成、更新	18
他機器連携 / データ配信	3
画像・映像解析	8

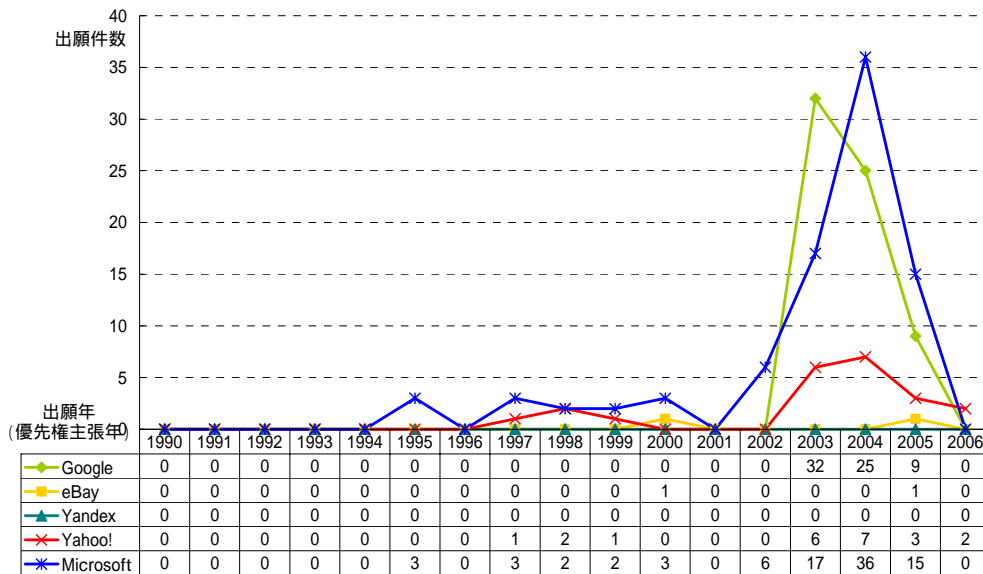
## 3. 欧州

### (1) 特許出願動向との関連分析

図 5-2-3 に、欧州における検索サイト上位 5 社の今回の調査対象領域での特許出願件数の推移（欧州への出願）を示す。今回の調査対象領域への出願は、5 社の中では Microsoft による出願が最も多く、同社は 1995 年頃から今回の調査対象領域へ特許を出願していることが分かる。市場シェアがもっとも高い Google も、欧州では 2003～2004 年頃に目立って多く出願していることが分かる。Yahoo! も 2003～2004 年頃に出願数を増やしている。

<sup>5</sup> 検索サイトの名称に関しては、各社の商標、または、登録商標である。

図 5-2-3 欧州検索サイト上位 5 社の検索技術関連特許出願件数推移（欧州への出願）<sup>6</sup>



(2) 重要特許との関連分析

欧州国籍の重要特許は表 5-2-4 の通り。「地理・地図データ解析」「テキストデータ解析」「画像・映像解析」とデータ解析分野に留まっており、市場動向と同様に米国企業がシェアを握っている様子が見てとれる。

表 5-2-4 欧州国籍の重要特許件数

テーマ	特許件数
地理・地図データ解析	2
検索結果の抽出・統合	0
類似・関連情報検索	0
テキストデータ解析	2
検索用インデクス・キーワードの抽出、作成、更新	0
他機器連携 / データ配信	0
画像・映像解析	1

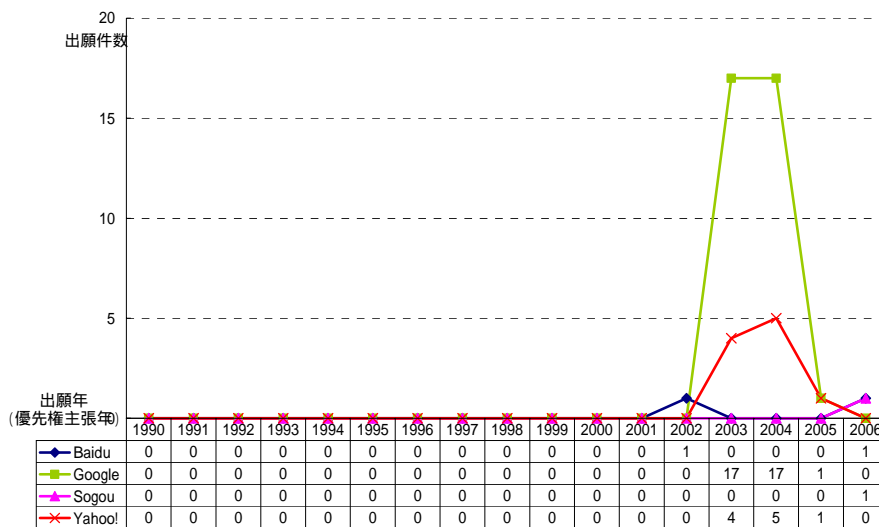
4. 中国

(1) 特許出願動向、重要特許との関連分析

図 5-2-4 に、中国における検索サイトシェア上位 4 社の今回の調査領域への特許出願件数の推移（中国への出願）を示す。今回の調査対象領域への出願件数が最も多いのは Google、続いて Yahoo!となっている。Baidu、Sogou といった中国国籍の検索サイト企業からは、今回の調査対象領域への出願は殆ど行われていない。

<sup>6</sup> 検索サイトの名称に関しては、各社の商標、または、登録商標である。

図 5-2-4 中国検索サイト上位 4 社の検索技術関連特許出願件数推移（中国への出願）<sup>7</sup>



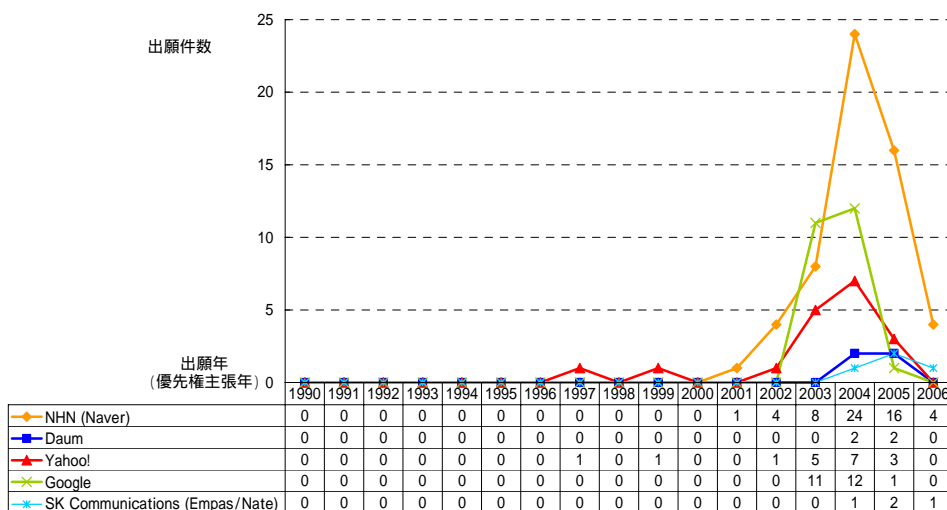
重要特許との関連分析では、中国国籍の重要特許は「検索結果の抽出・統合」の 1 件に留まっている。

## 5. 韓国の検索サイト市場

### (1) 特許出願動向、重要特許との関連分析

図 5-2-5 に、韓国における検索サイト上位 5 社の今回の調査対象領域への特許出願件数の推移（韓国への出願）を示す。2002 年頃から今回の調査対象領域への出願件数が増えているのが分かる。韓国市場で最もシェアの高い検索サイト Naver を提供している NHN の出願件数が最も多い。続いて、Google の出願件数が多い。

図 5-2-5 韓国検索サイト上位 5 社の検索技術関連特許出願件数推移（韓国への出願）<sup>7</sup>



重要特許との関連分析では、韓国国籍の重要特許は 3 件あり、LG エレクトロニクス、メディソン、KT の各社がそれぞれ出願している。

<sup>7</sup> 検索サイトの名称に関しては、各社の商標、または、登録商標である。

## 第6部 総合分析

### 第1章 本調査のまとめ

本報告書における各部の調査結果について、まとめたものを表 6-1-1 に示す。

表 6-1-1 各調査結果のまとめ

部	調査結果のまとめ
第2部 特許出願動向分析	<ul style="list-style-type: none"> <li>日米欧韓では出願件数全体で2000年をピークとしている。検索技術に限らず情報通信分野で多く見られる現象である。各国別に見ると、中国は、2000年以降も増加傾向が見られ、日欧韓では、2000年以降減少傾向であるが、米国は2002年以降、若干の増加傾向を示している。</li> <li>要素技術では、データベース技術である「一致検出」「インデキシング」が、日本からの出願は累積件数で米国の半分程度に止まるが、「メディアデータ解析技術」、「入力技術」などは、近年、日本からの出願が伸びている。特に「地理・地図データ解析技術」は日本からの出願が半数を占めている。</li> <li>メディア解析技術の中でも、特に、画像、音楽、地図・地理情報、感性・評判データを扱う技術では日本からの出願が多い。</li> <li>メディア別では「画像」「地理・地図」で日本からの出願が目立つが、全体として不特定のものも多い。伸び率では日本では「センサ」が高い。</li> <li>課題別では、パーソナライズやユーザの利便性向上に関する日本からの出願が多い。特にパーソナライズは40%を日本からの出願が占め、全ての課題で一番高い割合である。</li> <li>注目研究開発テーマでは「検索結果の抽出・統合」、「類似・関連情報検索」が安定して出願されている。</li> <li>また「テキストデータ解析」「他機器連携・データ配信技術」では日本からの出願が増えている。「画像・映像解析技術」は2000年以降、日本からの出願が最も多い。</li> <li>出願人ランキングでは、IBM、マイクロソフト、ヤフー、グーグル、ヒューレットパッカードなど検索エンジンやデータベースエンジンを持つ企業は米国、欧州、中国、韓国への出願が多い。日本への出願のみ大きく異なっており、日本国籍の出願人が上位を占めている。</li> </ul>
第3部 基本重要特許	<ul style="list-style-type: none"> <li>「地理・地図データ解析」については日本国籍の特許が目立つ。</li> <li>「類似・関連情報検索」は国内外からパーソナライズド広告に関する特許が多数出願されている。</li> <li>「画像・映像解析」では画像分野では特徴量分析技術が目立つのに対し、映像分野ではシーン推定/サマライズに関する技術が多い。</li> </ul>
第4部 研究開発動向	<ul style="list-style-type: none"> <li>論文の発表件数は増加傾向にあり、研究分野として注目されている分野と考えられる。近年は横這いであったが、2007年で増加している。</li> <li>ほぼ全ての技術区分で大半を米国国籍が占めている。特に「メディアデータ解析技術」「一致検出技術」「出力技術」に関する論文が多い。</li> <li>課題の解決では「システムの効率性」「精度」が多く「ユーザの利便性」がそれに次ぐ。</li> <li>発表はカリフォルニア大学、マサチューセッツ大学、カーネギーメロン大学などの米国の大学が中心である。マイクロソフト、IBM、ヤフー等の米国企業も上位に入っているが、米国中心の国際会議ということもあり、日本からの発表は大学、企業共に少ない。</li> </ul>
第5部 市場・政策動向	<ul style="list-style-type: none"> <li>米国では、複数の政府機関であるNSF/DARPA/NASAが横断的に資金提供した電子図書館プロジェクトDLIが始まっており、大規模な検索技術の開発が進められた。</li> <li>市場の寡占化に対応するように、日本、欧州共に大規模な研究開発プロジェクトを実施している。</li> <li>検索エンジンに関する著作権法上の取り扱いについては各国で方針が異なり、民間企業、研究者、国際競争力の観点からもフェアユース規定の導入が提言されている。</li> <li>米国、欧州を中心として世界的に米国の大手検索サービスによる寡占化が進んでいる。</li> <li>対照的に、中国、韓国、ロシアといった国では、自国の企業の市場シェアが高い。米国企業の検索エンジンはアルファベット以外の言語を対象とする自然言語処理や多言語対応が利用者から不十分であると判断されている可能性がある。</li> <li>日本では検索サービスを含めたポータルサービスが好まれており、ブックマークとしても利用されている。またブログ検索などのサービスを提供するベンチャ企業も多い。</li> </ul>

## 第2章 本調査結果の分析

本調査結果を分析すると、大きく以下の三つ背景が読み取れる。

### 第1節 検索対象メディアの拡大

検索技術というと一般的には情報収集・更新技術、インデキシング技術、一致検出技術が中心と考えられるが、本調査では、特許出願件数を技術中区分別に見た通り、米国国籍で一番出願が多いのはメディアデータ解析技術であり、日本国籍でも出力技術に次いで多いという結果になった。ベンチャ企業からの出願も多い。伸び率を見ても、メディアデータ解析技術が日米ともに高い伸び率を示しており、累積件数で見ると、日本国籍の出願が全体の約 1/3 を占めている。

特に、注目研究開発テーマの1つである「画像・映像解析」においては累積件数で日本国籍、米国国籍がそれぞれ全体に対して約 40% を占めており、日米は同程度の出願をしているといえるが、出願件数推移で見ると、出願ピークの時期が異なっている。

また研究開発動向分析からこの分野において論文発表が活発に行われていること、市場動向分析からインターネットにおける検索技術はテキストを対象にしたものが中心であったが、今日では多くの企業が画像、映像といったマルチメディア検索サービスを提供していることも分かった。

つまり、テキスト検索サービスは多くの国で寡占状態にあるが、マルチメディア検索についてはまだ広く普及していないため、現在はベンチャ企業や研究機関を交じて多くの企業による激しい競争環境にあるとも言える。

さらにメディアデータ解析技術に限らず特許出願全体での取扱いメディアを見ると、中心は画像や映像であることが分かる。伸び率で見ると、米国国籍においては映像と音楽、日本国籍では画像と音声と、やや方向性が異なっていることも読み取れる。これらは画像解析や音声認識が日本で長く研究されていること、一方で映像や音楽の配信サービスが米国を中心に多く登場していることとも符合する。

加えて「地理地図データ」を検索対象とする特許出願については、日本からの出願が全体の約半数を占めており、2006 年以降も増加傾向にある。

このように、検索技術は検索対象をテキストからマルチメディアへと拡大、拡散していると分析できる。

### 第2節 利用者のニーズにあった情報の提示とサービス連携化

検索技術が図書館を中心に利用されていた時代から、検索サービスは原則としてどのような利用者に対しても同じ結果を返していた。しかし、利用者毎に提示すべき情報は異なるものであり、利用者のニーズにあった情報を適切に提示する技術が望まれている。そのための技術として注目研究開発テーマに示した「検索結果の抽出・統合」、「類似・関連情報検索」がある。「検索結果の抽出・統合」は、米国国籍の出願を中心に、安定して出願されていることが分かる。「類似・関連情報検索」についても、安定して出願されていることが分かる。

また、ユーザのニーズにあった情報を抽出するためには、テキストデータをどのように抽出するかが問題となる。本調査では、注目開発テーマの一つとして「テキストデータ解析」を選定した。日本国籍の出願は堅調に伸びている。

しかし市場動向を見ると、日本市場においても多くの検索サービスが海外発のものである。日本向けに特化したサービスというのは現状ではそれほど多いとはいえず、現状ではグローバルなサービスを各国へ展開する際に、言語や国に合わせて機能やサービスを調整する程度に留まっていることが多い。

また検索のパーソナライズは、利用者のニーズにあった情報を提示するための技術であり、利便性、効率性、精度といった根本的課題に次ぐ課題となっており、2003年以降も年間100件前後の出願が行われている。

パーソナライズが注目される背景には、広告収益に大きく依存した検索サービスのビジネスモデルがある。すなわち、より良い広告を提供するため、パーソナライズ技術が必要となるという流れである。これは注目研究開発テーマ「類似・関連情報検索」における重要特許の多くがパーソナライズド広告であることから見てとれる。特に日本ではモバイル検索の需要が高いこともあり、パーソナライズを課題とする特許出願の40%を占め、積極的に取り組んでいることが分かる。さらに注目研究開発テーマの1つである「他機器連携・データ配信技術」もまた、パーソナライズ技術を含んでいる。

また市場動向からは、パーソナライズの精度向上にはユーザの属性や行動パターンを多く把握する必要があるため、Google や Yahoo! といった大手企業がベンチャ企業を積極的に買収し、サービスメニューを拡充させていることも読みとれる。

### 第3節 検索サービスのグローバル化と寡占化

本調査を通じて、ある特定技術分野に圧倒的に強い国がいるというような、国籍別の極端な差異はほとんど見られなかった。技術中区分別、メディア別、課題別、いずれにおいても国籍によって規模の大小は異なるが、傾向は似ている。これはインターネットという環境の特性もあり、検索技術/検索サービスがグローバル化していることを示している。言うまでもなく、各分野に多数の出願を行っている米国がその中心にある。

また、市場動向分析で示したように、検索サービスは各国とも寡占市場にあり、特に日本、米国、欧州各国では Google、Yahoo!、Microsoft と市場シェア上位の企業も同じである。欧米が同じアルファベット圏であり、非アルファベット圏では同じ寡占市場であっても中国の Baidu、韓国の Naver といった自国の企業が提供する検索サービスがシェアを確保していることを考えると、日本は非アルファベット語圏にありながら欧米と近いという特徴的な状況にある。特許出願数においても、日本のみランキングが大きく異なり、上位の日本企業は検索サービスをほとんど行っていない。

ただし、日本の独自性として、検索サービスがポータルサイトの一部として利用されており、「Yahoo」「Google」といった単語が検索ランキングの上位に挙がるなど、ブックマークとして活用されている面もある。サービスはグローバル化されても、使われ方は各国で異なっていることが分かる。また、日本では寡占化の対応策として、ブログ検索など用途を絞り込んだ検索サービスでベンチャ企業が登場している。

政策動向に目を向けると、欧州では「欧州産検索エンジン」をアピールした Quaero、日本では非 Web 情報を対象とした情報大航海プロジェクトが進行中であり、米国大手企業による検索サービスの寡占化に対する危機意識と、独自性の模索が見てとれる。

### 第3章 日本が取り組むべき課題、目指すべき研究開発・技術開発の方向性

これまで各動向分析に基づいて、今後日本が取り組むべき課題、目指すべき研究開発、技術開発の方向性についてまとめる。

#### 第1節 日本の研究開発力が強いメディアに関連した技術開発の強化

##### 【提言1】実空間で利用するデータの検索および利用技術の強化

インターネットから入手できるコンテンツは実世界と関連するデータが増えており、文書と同様に検索対象となっている。特に、地理・地図情報がインターネットや携帯電話等のネットワーク経由でどこでも誰でも利用できる環境が一般的になりつつあり、実空間に関連するデータを活用した様々なサービス、および、関連する特許が出願されていることもあり、取り扱うデータに合わせた検索技術の強化が望まれる。

また、検索結果の利用方法に関しても、従来のPCの画面に表示するだけでなく、利用者がその場で利用できるように、表示する大きさ、メモリ、通信速度等のリソース上の制約がある携帯端末や、利用者個人の状況やニーズに合わせて情報を提示する技術の強化が期待される。

「検索対象メディアの拡大」で示した通り、インターネットを利用する検索対象となるメディアは拡大している。特に「地理・地図データ」を検索対象とする特許出願については、日本国籍による出願件数が全体の約半数を占めている。実際に、「地理・地図データ」を扱うサービスとしては、カーナビ等を中心に実空間で利用するものから、近年ではインターネットを利用する連動サービスやGPS機能を持つ携帯電話を通じて利用するサービス、地図と連動する検索サービスも増えている。

このことから、日本国籍による出願が多く、2002年以降も大幅な落ち込みが少なく、2006年以降増加傾向にあることが分かるように、注目されている分野であると考えられる。今後の応用サービス開発が必要である。

一方で、ただメディアに対応するだけでなく「利用者のニーズにあった情報の提示とサービス連携化」で示した通り、検索周辺技術も重要である。注目研究開発テーマ動向で挙げた「他機器連携・データ配信技術」では、携帯端末への配信技術が日本国籍と米国国籍の出願件数がほぼ同程度あり、近年増加している。その中でもパーソナライズド配信技術は、日本国籍の企業から重要特許が挙げられている。

##### 【提言2】映像・画像メディア向けの技術強化

近年、日本に限らず、各国におけるインターネット利用者のブロードバンド化が進んでおり、リッチコンテンツを扱う機会が増えている。特に、CGM（UGC、あるいは、韓国ではUCCと呼んでいる）と呼ばれるメディアの一つとして、利用者自身が撮影した画像・映像を扱うサービスが増えている。

映像や画像の場合には、データ量が大规模であり、文字・テキストと比較すると1000倍以上あることもあり、検索のためのインデックス作成には工夫が必要である。一般にはコンテンツ作成時に付けられるメタデータを使うことが多いが、メディアに合わせた解析技術により内容に基づくインデックスを作成することで、多量のデータを扱えるようになる。そのため、内容に合わせたメディア解析技術を強化する必要がある。

日本では、画像・映像解析技術については、古くから幅広く研究開発が行なわれてきた



め、技術的な蓄積がある。「検索対象メディアの拡大」で示した通り、特許出願動向調査においても、メディアデータ解析技術に分類される特許に関しては、日本国籍による出願が全体の約 1/3 を占めている。また、検索対象メディアについても、画像に関しては、日本国籍による特許出願は米国国籍分とほぼ同じであり、映像を対象とした特許出願に関しても、米国国籍が多いものの、日本国籍による出願も全体の約 1/3 を占めている。

さらに、注目研究開発テーマの一つである「画像・映像解析」においても、累積件数では日本国籍と米国国籍の件数がほぼ同程度となっている。全体の出願数の推移を見ると、1995年、2000年、2003年にそれぞれ大きなピークがあり、はじめのピークは米国国籍、2回目のピークは日本国籍、米国国籍、3回目のピークは日本国籍、米国国籍に加えて欧州国籍によるものであることが明らかになった。

「検索サービスのグローバル化と寡占化」で述べた通り、欧州において欧州委員会が FP6 等を通じて支援を行なっている Quaero, Pharos といった研究開発プロジェクトでも、音声、画像・映像を対象とする検索エンジン向けの技術開発が行なわれている。

以上のように、日米欧それぞれにおいて継続的に研究開発が行われている画像・映像解析技術について、ブロードバンド化が進むにつれてよりデータ量の大きな映像・画像を扱う技術が必要となることから、映像・画像メディア向けの技術を強化するためにさらなる研究開発が期待される。

## 第2節 日本語及び日本向けサービス特有の技術強化

### 【提言 3】日本特有のサービス・文化に合わせた技術強化

市場動向分析で示したように、中国、韓国、ロシアといった非アルファベット言語圏の国における検索サービス市場は、現状米国の大手企業による独占的な市場にはなっていない。主な理由としては、各国企業の検索サービス市場への先行者利益によるものや市場規模やインフラ整備状況等の経済的理由からアジア圏の言語向けへの対応が遅れていることだけでなく、各国内向けポータルサイトとしての機能や各国の言語・文化に合わせた技術の採用により、国内利用者の満足度が高い検索サービスが提供されているためと考えられる。そのため、日本語独自の技術開発や日本国内向けのサービスの充実が必要である。特に、日本語特有の解析処理として、分かち書きや表記のゆれ、CGM を通じて新たに作られる単語（新語）や隠語的な単語、構文解析などの技術的な課題も多い。また、日本人の嗜好に合ったサービスと提供していくことが必要である。

「検索サービスのグローバル化と寡占化」で示した通り、検索サービスは各国で寡占化しており、サービス展開時には、言語や国に合わせて機能やサービスを修正しているのが実情である。しかし各国における利用者のニーズや文化の違いを満たすためには、自然言語処理技術の開発が欠かせない。特に、日本語向けの自然言語処理の一つとして、意味解釈技術の研究を進めることで、より精度の高い検索サービスが提供できると考えられる。図 2-2-26 に示すように特許出願動向でも日本国籍の出願が全体の約 1/3 を占めることから分かるように、力を入れているところである。

また、やはり「検索サービスのグローバル化と寡占化」で述べた通り、日本では検索サービスで検索されているキーワードの多くがサイト名を示すキーワードであって、米国の検索キーワードとは異なり、ブックマーク的に使われている。また韓国では、検索サービスはポータルサイトの機能の一つであると考えられていることから、利用者にとっては、検索サー

ビスよりもポータルサイトとしての認識が強いことがわかる。これらの事例から分かるように、検索サービスは各国特有のサービスに合わせた技術開発が望まれている。

**【提言 4】新たな分野、情報の質や利用者行動等を考慮した検索技術の強化**

オブジェクトサーチと呼ばれる「検索結果の統合」や利用者の行動履歴等に合わせて「パーソナライズした検索結果を出力」など新しい技術が登場しており、検索サービスは再度、競争市場になる可能性がある。特に、日本からの発信が多いブログを対象とする検索などの特定メディアを対象とした検索サービスや、知識データや評判データなど自然言語で表された情報を対象とした検索など情報の質を考慮した検索サービス、利用者の行動を考慮した検索サービスなど、より利便性の高い検索サービスを実現する技術開発に注力することが期待されている。

「検索サービスのグローバル化と寡占化」で述べた通り、日本国内では日本語によるブログのコンテンツが非常に多いこともあり、ブログを対象とした検索技術を応用したサービスを提供する企業がベンチャ企業を中心に複数現れている。評判情報の検索や市場のトレンドの分析など自然言語解析技術をベースとした技術が多い。

また「利用者のニーズにあった情報の提示とサービス連携化」で示した通り、単純なキーワードによる検索結果だけではなく、自然言語による検索や検索結果をまとめる「情報結果の抽出・統合」「類似・関連情報検索」「テキストデータ解析」といった意味的な解釈、統合することを実現する新しい分野がテーマとして注目されている。近年特許出願件数も増加傾向にあり、日本国籍による特許出願も比較的多いことが分かる。

検索技術を応用したサービスは今後さらに普及する可能性があり、また、自然言語処理技術やメディア解析技術は近年特許出願も増加傾向で日本国籍出願人も技術開発を広く行い技術的蓄積もあることから、新たなメディアを対象とした解析技術や検索技術の開発を注力することが期待される。

### 第3節 日本が弱い技術力の強化

**【提言 5】プラットフォーム技術の強化**

2000 年前後には、国内大手企業でも検索技術について、現在グローバル展開している検索サービス提供事業者と競争環境にあったものの、検索エンジンについては、海外の大手企業による市場の寡占化が進んでいる。既に検索エンジンがインフラ化しているため、検索技術だけではなくサービスを安定的に提供するためのプラットフォーム技術が必要である。

「検索サービスのグローバル化と寡占化」で示した通り、検索サービスにおいては Google など少数の企業による寡占化が世界的に広がっている。出願人にも見られるように、従来から日本国内の大手電機メーカーや通信事業者を中心に研究開発していた全文検索技術はあったものの、検索対象となる量、検索結果の精度・質や検索速度といった性能面での差異が市場シェアに影響を及ぼしたと考えられる。

現状の寡占状態にある検索サービスの市場シェアを考慮すると、大手検索サービスが有する検索サービスに対抗するだけのインフラを民間企業が単独に整備するのは難しい。継続的に競争力の高いサービス提供するために、様々な企業が利用できるプラットフォーム技術の技術開発が必要と考えられる。

システム技術である「ハードウェア・通信技術」、および、「ソフトウェア技術」に関して米国へ出願割合が比較的多いことから国内の企業が注力しているとも考えられる。このことは、クラウドコンピューティング等のプラットフォーム技術が強い米国での優位性を考慮したものと考えられる。

欧州の Pharos プロジェクトにおいては、検索機能をサービスに活用するためのアーキテクチャを重視して、検索エンジンをプラットフォーム内で使えるように ESB(Enterprise Service Bus)技術を活用した開発が行なわれている。この事例からも大手検索エンジンが提供する API を使うことだけではなく、検索エンジンを取り込んだアーキテクチャを考慮したサービス開発を行なう必要がある。