

オンラインセミナー

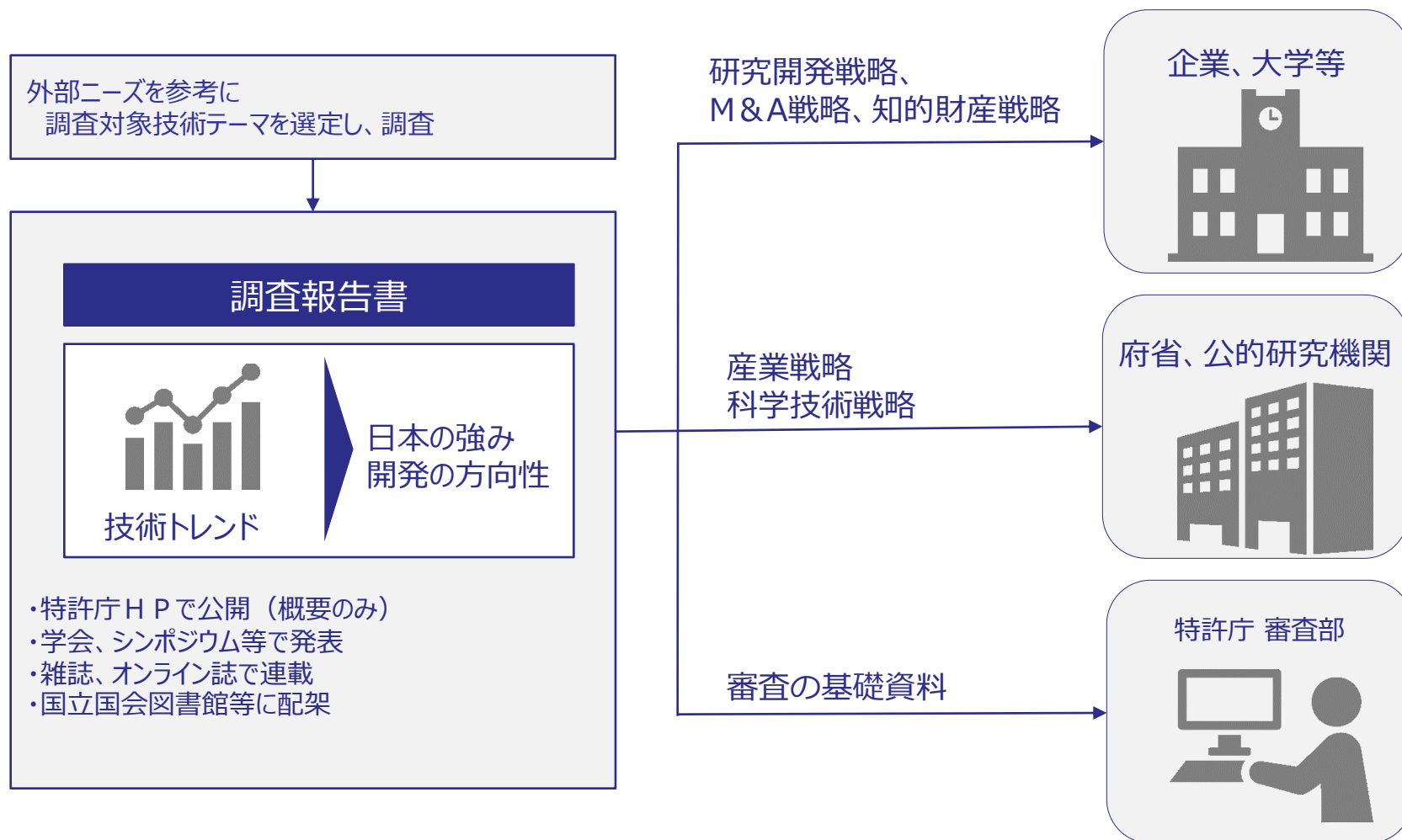
令和6年度 特許出願技術動向調査 —メタバース時代に向けた音声・音楽処理—

令和7年7月
特許庁審査第四部電気機器（音響システム）

1. 特許出願技術動向調査について.	P. 2
2. 委員名簿.	P. 3
3. 調査概要.	P. 4
4. 市場動向.	P. 7
5. 政策動向.	P. 8
6. 特許出願動向.	P. 10
7. 研究開発動向.	P. 20
8. 総合分析（まとめ）.	P. 24

1. 特許出願技術動向調査について－特許出願技術動向調査とは－

- 「特許出願技術動向調査」※¹は、注目度の高い技術テーマを対象に、その特許出願動向等を調査して技術トレンドをつかみ、日本の研究開発の方向性を見定めるもの。
- 調査結果は報告書としてまとめられ、迅速・的確な審査の基礎資料として、また、企業・大学・研究機関等が研究開発戦略等を策定するための基礎資料として活用される。



2. 委員名簿

(敬称略、所属・役職等は令和7年1月現在)

委員長

【学術関係】

猿渡 洋 様 (東京大学大学院 情報理工学系研究科 教授)

委員

【学術関係】

尾本 章 様 (九州大学大学院 芸術工学研究院 教授)

高道 慎之介 様 (慶應義塾大学 理工学部 情報工学科 准教授)

【産業関係】

鎌土 記良 様 (日本電信電話株式会社 コンピュータ&データサイエンス研究所)

須見 康平 様 (ヤマハ株式会社 研究開発統括部 先進技術開発部 音楽情報グループ)

※委員は五十音順に記載

3. 調査概要

－調査背景－

- 音を記録する技術や音を再現する技術は現代社会において重要であり、かつ人々の生活に密接に関わっている。音に関する製品やサービスの態様も大きく変化している。例えば、音楽を聴くために以前はCDを利用していたが、現在ではインターネットを利用した配信サービスが主流になっている。また、AI技術の発展に支えられた音声合成や自動作曲に基づくサービスを一般にも利用できる環境になっている。
- また、こうした技術の発展は音声や音楽が活用される場면을多様化させることに繋がる。その一例としてメタバースが挙げられる。
- 他方、技術の発展に伴い、新たな課題も生じている。AIで著名人の声を模倣するなど、好ましくない利用が見られる問題や、AIが他人の著作物に似た楽曲を生成するといった著作権に関する問題が挙げられる。また、AIそのものに対するリスクも認識され、その対応について議論されている。
- 音声や音楽を取り扱う際、このような課題にも向き合いつつ、さらに人々にとって有用である音声や音楽を届けるための技術を提供していくことが必要である。そこで本調査では、音声・音楽処理技術に関する研究開発、技術開発状況を把握し、将来に向けてどのような方向性を定めることが必要であるかを検討するために、特許出願や論文発表の動向を調査することとした。

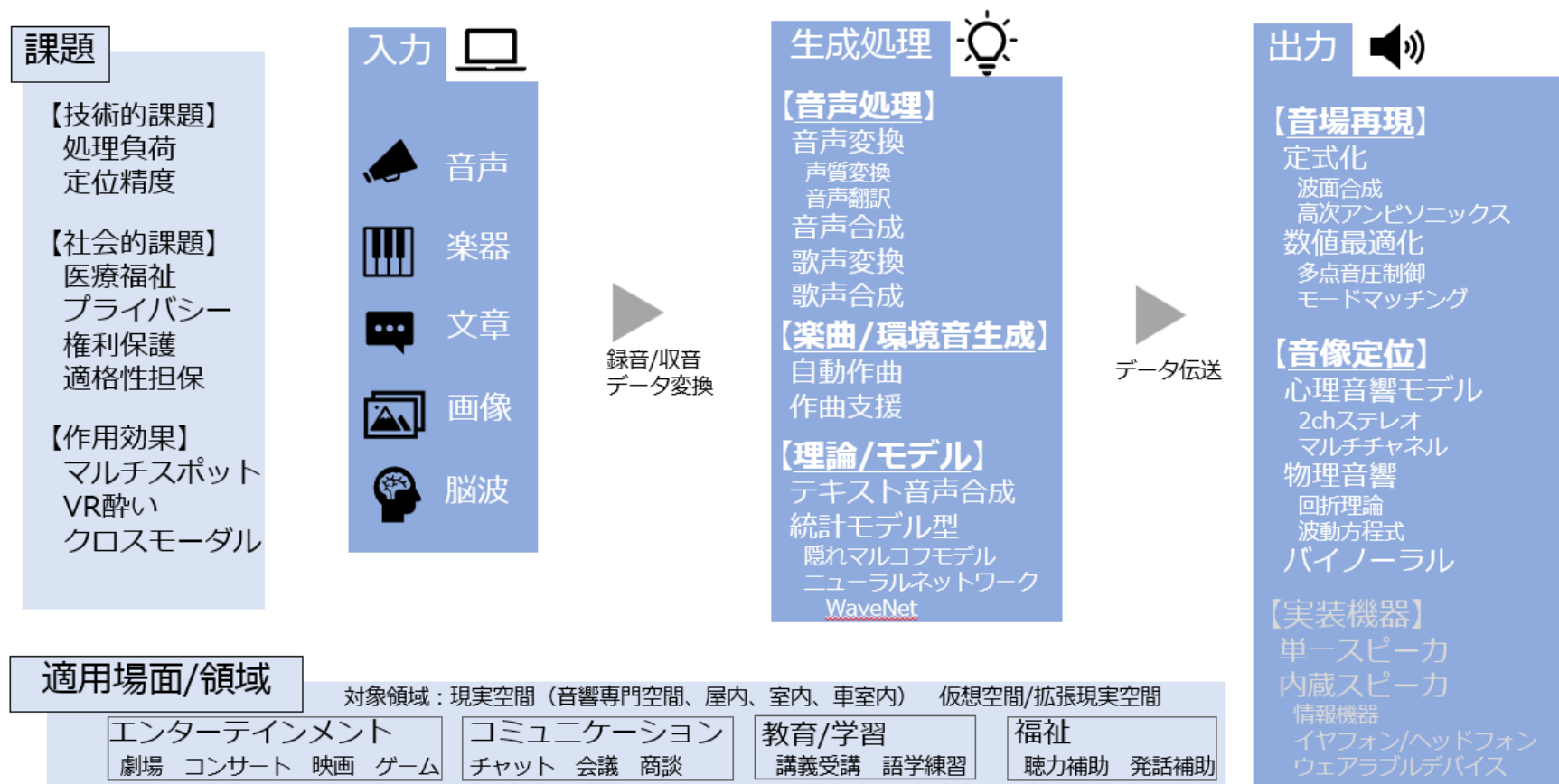
3. 調査概要

ー調査対象技術ー

■ 本テーマの調査対象範囲を図1に示す。

■ 本調査では、音場再現・音像定位、音声合成、音声変換、歌声合成、歌声変換、自動作曲・作曲支援に関する技術を対象とした。また、これらの技術を活用する領域や場面、これらに付随して発生する課題といった側面も対象とした。

図1 【技術俯瞰図】



※入力時あるいは生成処理時の符号化、実装機器そのものについては調査対象外

出典：トヨタテクニカルディベロップメント株式会社にて作成

3. 調査概要

ー調査対象文献、データベースー

- 音場再現・音像定位、音声合成、音声変換、歌声合成、歌声変換、自動作曲/作曲支援に関する技術について特許出願動向、研究開発動向を調査。
- 調査対象の文献は、読込み解析により技術区分に分類。

□ 特許出願動向

調査対象文献：特許文献 … 約26,000件

使用した検索データベース：Orbit Intelligence (Questel社)

調査対象期間：2013年-2022年（優先権主張年）

□ 研究開発動向

調査対象文献：主要な国際会議論文 … 約16,000件

※データベースより論文、発表内容に限定

使用した検索データベース：Scopus (Elsevier社)

調査対象期間：2013年-2023年（発行年）

4. 市場動向

ー音声・音楽の市場動向ー

- 日本のコンテンツ市場は2022年に12兆4,418億円であった。音楽配信売上は年々増加し、2023年には1,164億9,800万円となり、ストリーミングが約90%を占めている。
- メタバースの世界市場は2021年の388.5億ドルから2030年には6,788億ドルにまで拡大すると予測されている。
- AIを巡る議論が行われている。米国では音楽生成AIサービスに対する著作権侵害訴訟が起きている。欧州ではAI法案に関して、政治合意に達した。この中には、音声コンテンツを生成または操作するためにAIシステムを利用する場合、そのコンテンツが人工的に生成または操作されたものであることを明確かつ区別可能に開示する必要があるとの内容が盛り込まれている。

5. 政策動向

－各国・地域における施策－

- 日本では国立研究開発法人を中心に音声・音楽に関するプロジェクトが推進されている。
 - 国立研究開発法人日本医療研究開発機構では、「喉頭摘出者の音声再生プロジェクト」が実施されている。本プロジェクトは、頭頸部がんなどの病気により喉頭を摘出する患者を対象とした臨床研究で、手術で音声を失ってしまう前に音声を保存し、手術後に代替発声による音声と音声変換技術を用いて、自分の音声を取り戻すことを目標としている。
 - 国立研究開発法人情報通信研究機構のユニバーサルコミュニケーション研究所では、言語の壁を越えた音声コミュニケーションを実現するために多言語音声翻訳技術の研究開発に取り組んでいる。31言語に対応した多言語音声翻訳アプリ「VoiceTra」、テキストを音声に変換し、聴覚障害者と健聴者とのスムーズなコミュニケーションを支援するアプリ「こえとら」や「SpeechCanvas」を提供している。また、翻訳結果等を円滑に伝達するための肉声レベルの音声を合成する技術及び自然性劣化を極力抑えつつ、話速等の声質を制御する技術の開発、日常会話の音声認識及び音声合成に関して対象とする言語の拡大に取り組んでいる。
- 独立行政法人日本学術振興会は音声合成や自動作曲に関する研究を支援している。
 - 自動作曲については、「音声認識手法を応用した自動作曲・自動作詞・自動伴奏の研究」をテーマとする研究が開始され、自動作曲システム「Orpheus」を開発した嵯峨山茂樹氏を中心に、Orpheus のversion4 に位置づけられる自動作曲・自動作詞の次段階の原理の探究と実現、自動伴奏システム「Eurydice」のversion2 の実現と社会応用などを主要目標として研究開発が進められた。さらに、「日常生活におけるルーティンワークを愉しく適切に遂行するための楽曲の自動生成」をテーマとする研究が開始され、同研究では、ユーザの感性に即し、理想的な作業を誘導するような作業用楽曲を自動生成するシステムの構築に取り組んでいる。

5. 政策動向

－各国・地域における施策－

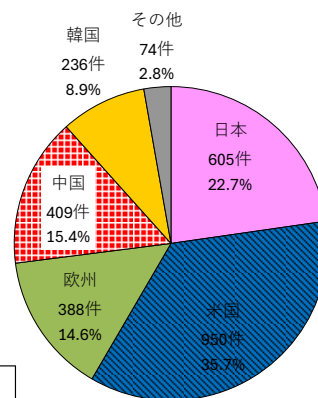
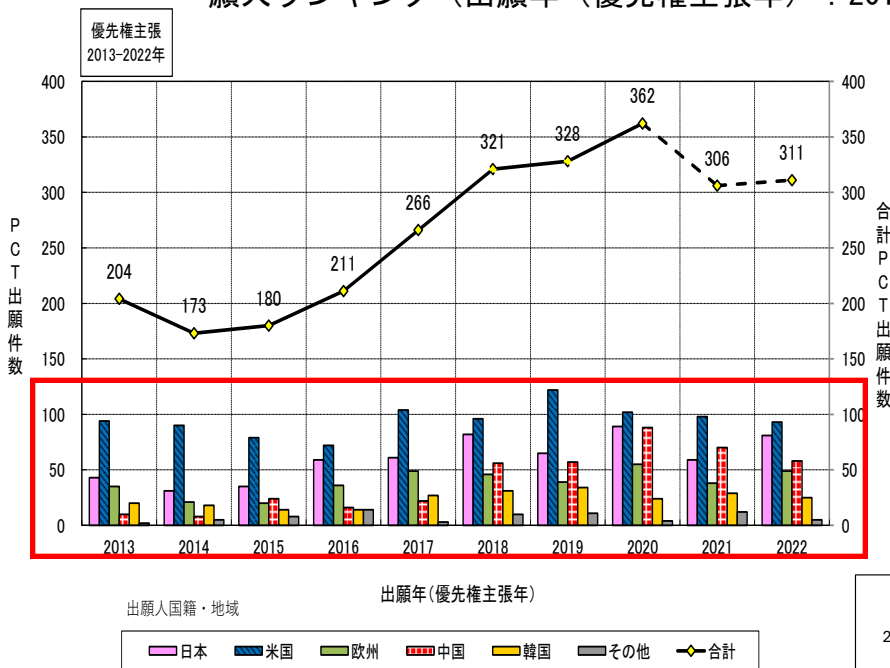
- 米国のIEEEは音響学、音声および信号処理、およびそれらの応用に関する世界最大かつ最も包括的な国際会議であるICASSP 2024を開催した。米国の連邦取引委員会では音声クローンの不正使用防止のための技術開発を促進する取り組みが実施された。優秀作品として、AIアルゴリズムを使用した音声パターンの区別技術や音声クローン検出技術が選出された。
- 欧州の欧州研究執行機関は生成AIによるディープフェイクの研究プロジェクト「SOLARIS」を実施している。フランスの国立音響音楽研究所は音声・音楽処理技術に関する研究を行っている。
- 中国の音声・言語情報処理国家工学研究センター（National Engineering Research Center of Speech and Language Information Processing）は、音声合成、自然言語処理、機械学習などの分野に関する研究開発を実施している。自己教師あり表現に基づく音声合成モデルの性能についての比較、歌詞と楽譜情報を人間のような表現力と高忠実度を持つ歌声に変換する歌声合成システムの開発が行われた。
- 韓国では「地域大学優秀科学者支援事業」を通じて、高品質な韓国語音声合成技術、ディープラーニングベースのエンドツーエンド方式音声合成技術、リアルタイム処理音声合成技術、少量の音声データを使用した適応型話者変換および多話者音声合成技術の開発が進められた。韓国科学技術情報通信部はボイスフィッシング防止技術やエンドツーエンド音声合成システムの開発を支援し、韓国科学技術院は音声合成や自動作曲に関する研究を支援している。

6. 特許出願動向

－音声・音楽処理に関するPCT出願－

- 調査対象期間（出願年（優先権主張年）2013～2022年）における音声・音楽処理に関するPCT出願の出願人国籍・地域別PCT出願件数推移、出願件数比率及びPCT出願件数上位出願人ランキングを示す。PCT出願件数合計は2,662件である。
- 件数比率を見ると、米国籍が35.7%と最多で、これに日本国籍22.7%、中国籍15.4%、欧州籍14.6%、韓国籍8.9%と続いている。どの年においても米国籍の出願が最も多い。
- 2018年以降、中国籍からのPCT出願が大きく増加している。

図2 【出願人国籍・地域別PCT出願件数推移、出願件数比率及びPCT出願件数上位出願人ランキング（出願年（優先権主張年）：2013～2022年）】



順位	出願人	国籍	属性	件数
1	ソニー	日本	企業	234
2	ドルビー	米国	企業	147
3	マイクロソフト	米国	企業	108
4	グーグル	米国	企業	102
5	ノキア	欧州	企業	99
6	クアルコム	米国	企業	93
7	サムスン	韓国	企業	82
8	ファーウェイ	中国	企業	78
9	フラウンホーファー研究機構	欧州	研究機関	68
10	ヤマハ	日本	企業	62

注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

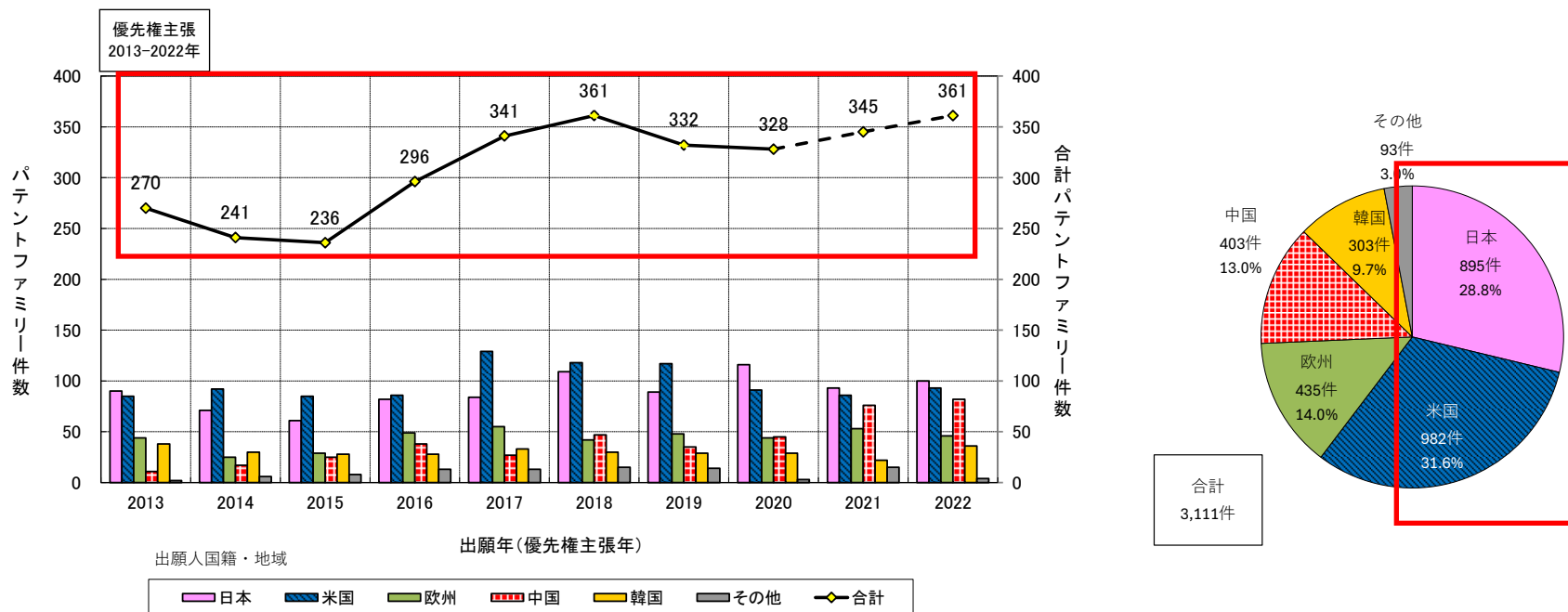
6. 特許出願動向

－音場再現・音像定位の出願動向－

＜技術区分：音場再現・音像定位＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分「音場再現・音像定位」の出願人国籍・地域別のパテントファミリー件数推移を示す。
- 米国籍、日本国籍のパテントファミリーが多い。
- 他の技術区分との比較になるが、特段大きな増加は見られず、常に研究開発が一定の特許出願に結び付いている様子が見られる。

図3 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「音場再現・音像定位」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

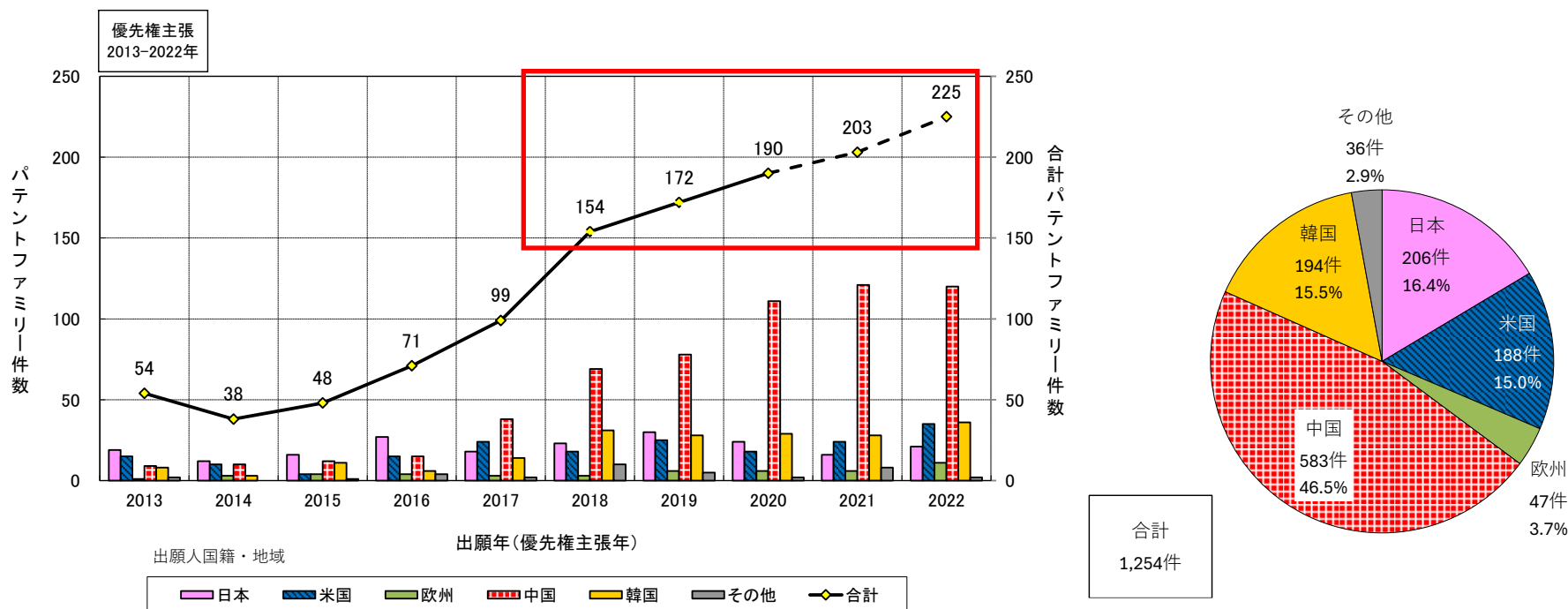
6. 特許出願動向

－音声変換の出願動向－

<技術区分: 音声変換>

- 調査対象期間(出願年(優先権主張年)2013～2022年)における大区分「音声変換」の出願人国籍・地域別のパテントファミリー件数推移を示す。
- 音声変換に関するパテントファミリー件数は年々増加の一途を辿っている。

図4 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「音声変換」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

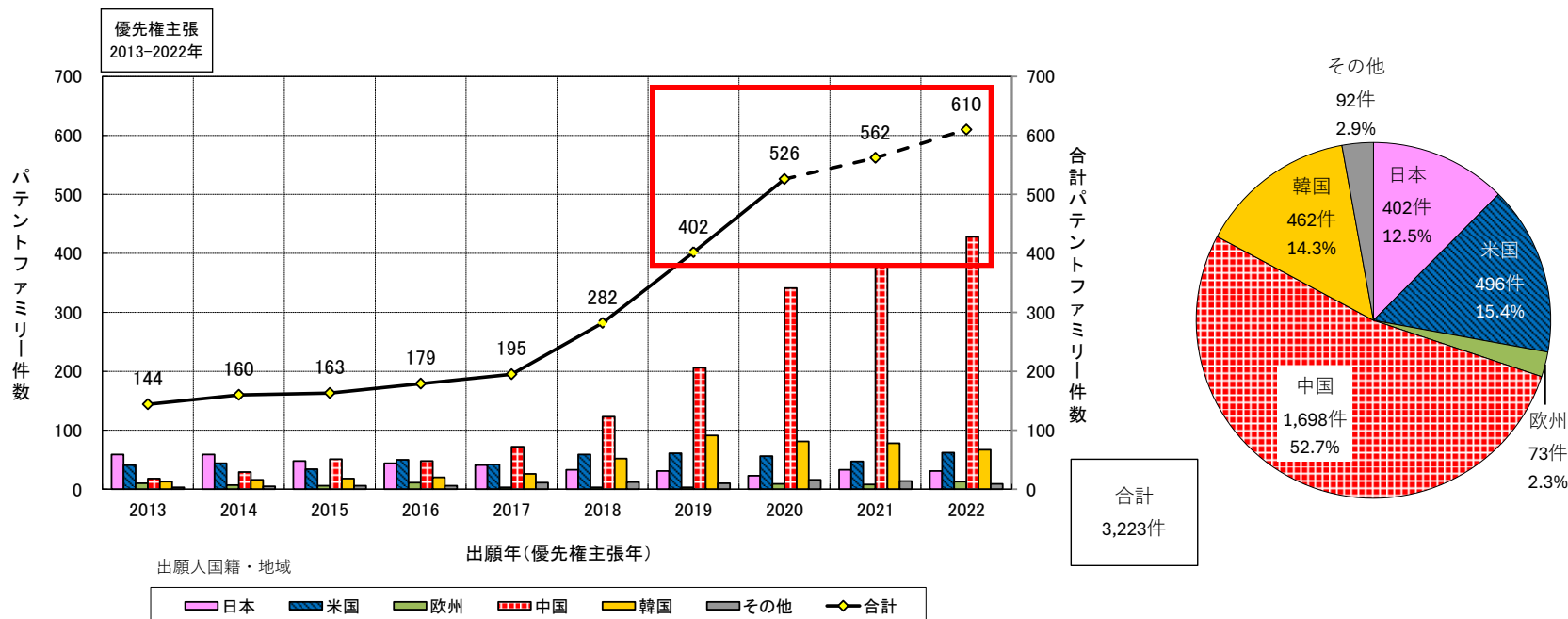
6. 特許出願動向

－音声合成の出願動向－

＜技術区分：音声合成＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分「音声合成」の出願人国籍・地域別のパテントファミリー件数推移を示す。
- 音声合成に関するパテントファミリー件数は年々増加の一途を辿っている。
- 音声変換、音声合成ともに、特に2018年以降の件数増加が顕著に表れている。これについては2016年に発表されたWaveNetの論文が大きく影響しているものと考えられる。
- WaveNetは音波形を直接生成するディープニューラルネットワークであり、より自然な発音に近い音声を合成することに成功したアルゴリズムである。

図5 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「音声合成」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

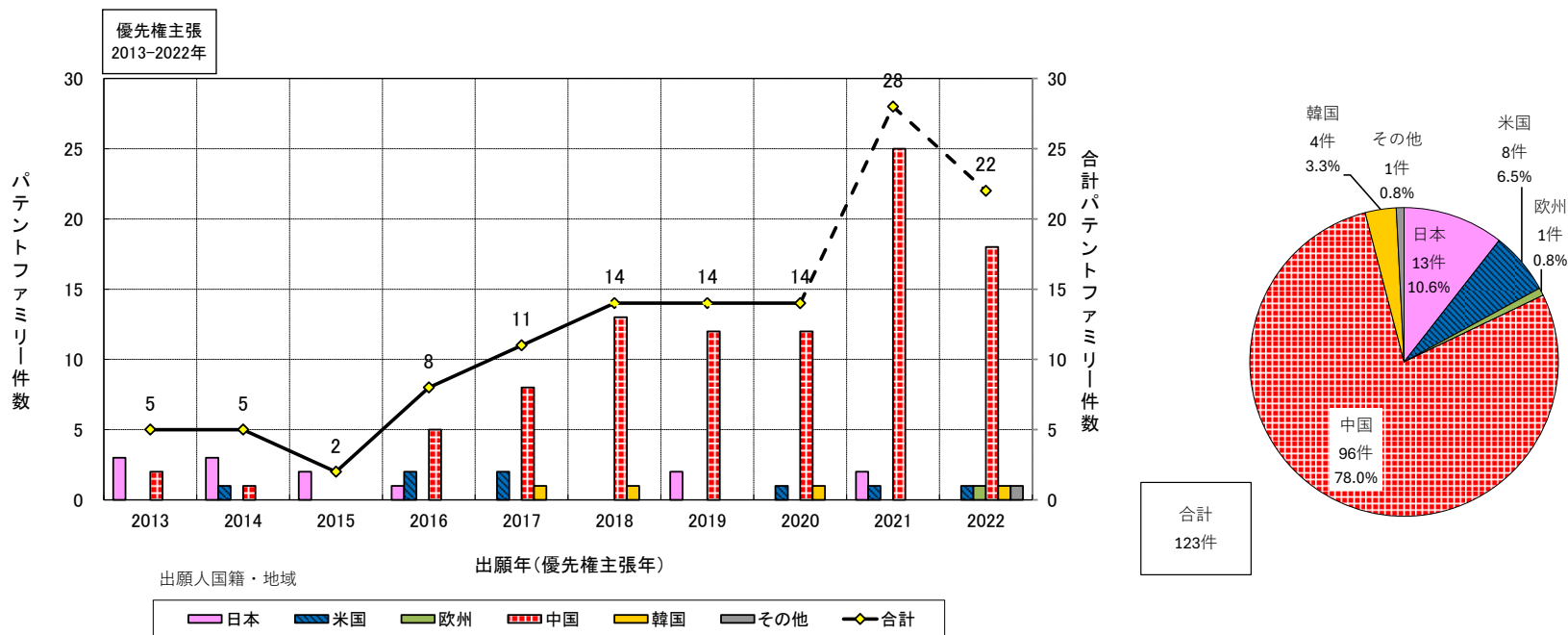
6. 特許出願動向

－歌声変換の出願動向－

＜技術区分：歌声変換＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分「歌声変換」の出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率を示す。
- 中国籍の出願が圧倒的に多い。
- 他の技術区分と比較するとパテントファミリー件数は少数である。

図6 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「歌声変換」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

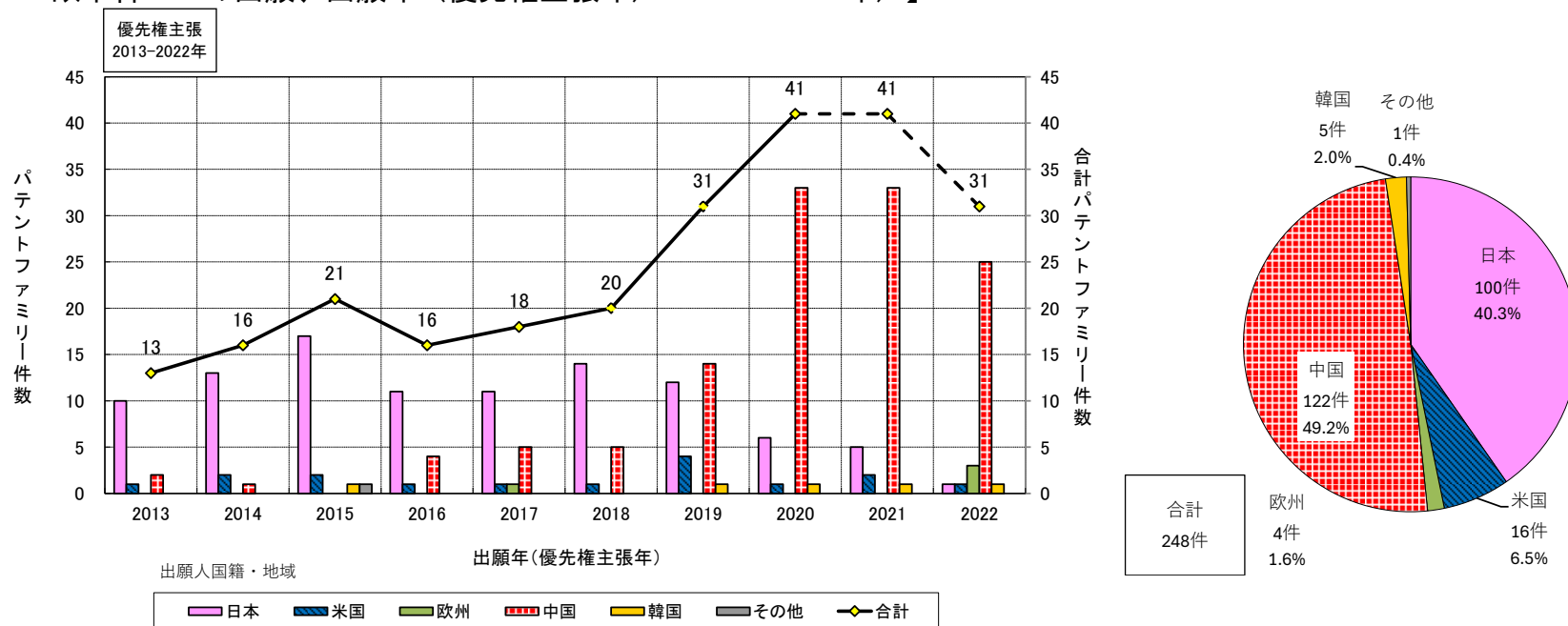
6. 特許出願動向

－歌声合成の出願動向－

＜技術区分：歌声合成＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分「歌声合成」の出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率を示す。
- 調査期間の前半では日本国籍のパテントファミリー件数が優勢であった。
- 2019年以降は中国籍のパテントファミリー件数が急激に伸びている。

図7 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「歌声合成」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

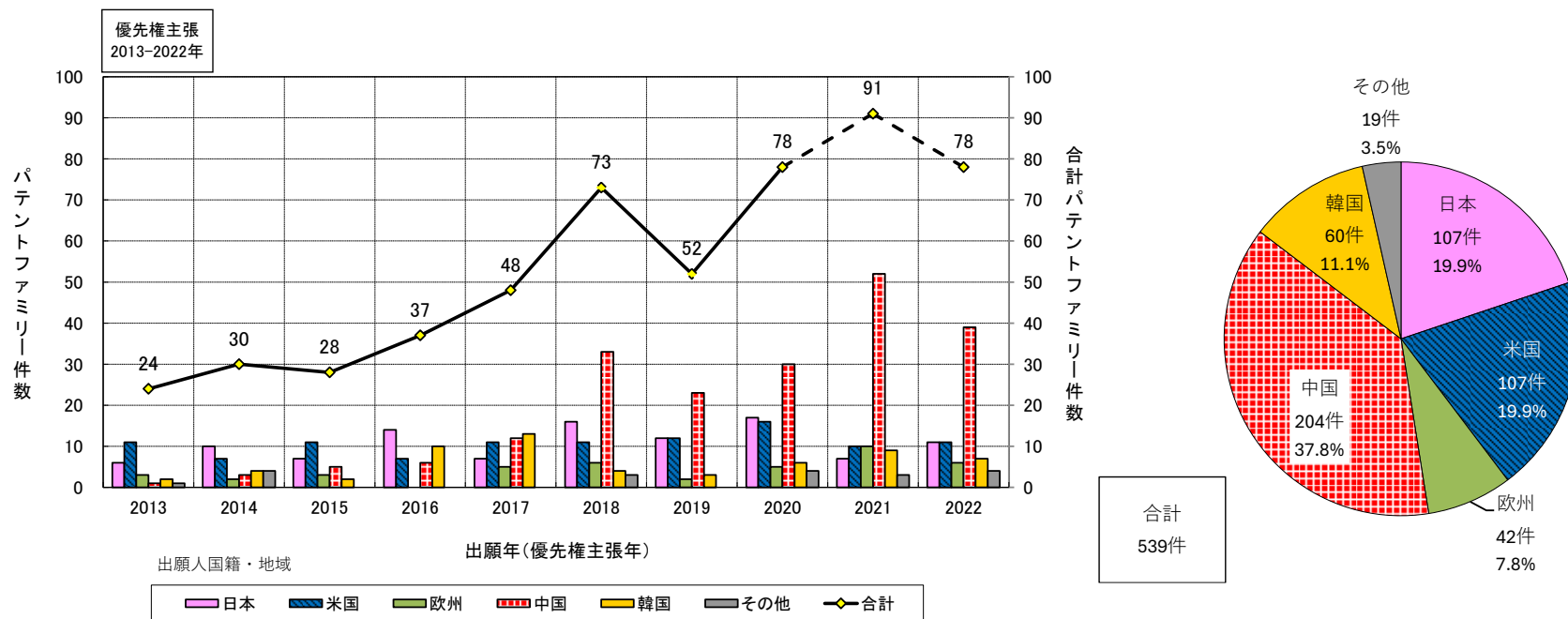
6. 特許出願動向

－自動作曲・作曲支援の出願動向－

＜技術区分：自動作曲・作曲支援＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分「自動作曲・作曲支援」の出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率を示す。
- パテントファミリーは増加傾向にある。
- 2018年以降、中国籍のパテントファミリー件数が急激に増加している。

図8 【出願人国籍・地域別パテントファミリー件数推移及びパテントファミリー件数比率（「自動作曲・作曲支援」、日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】



注：2021年以降はデータベース収録の遅れ、PCT出願の各国移行のずれ等で全データを反映していない可能性があるため、破線にて示す。

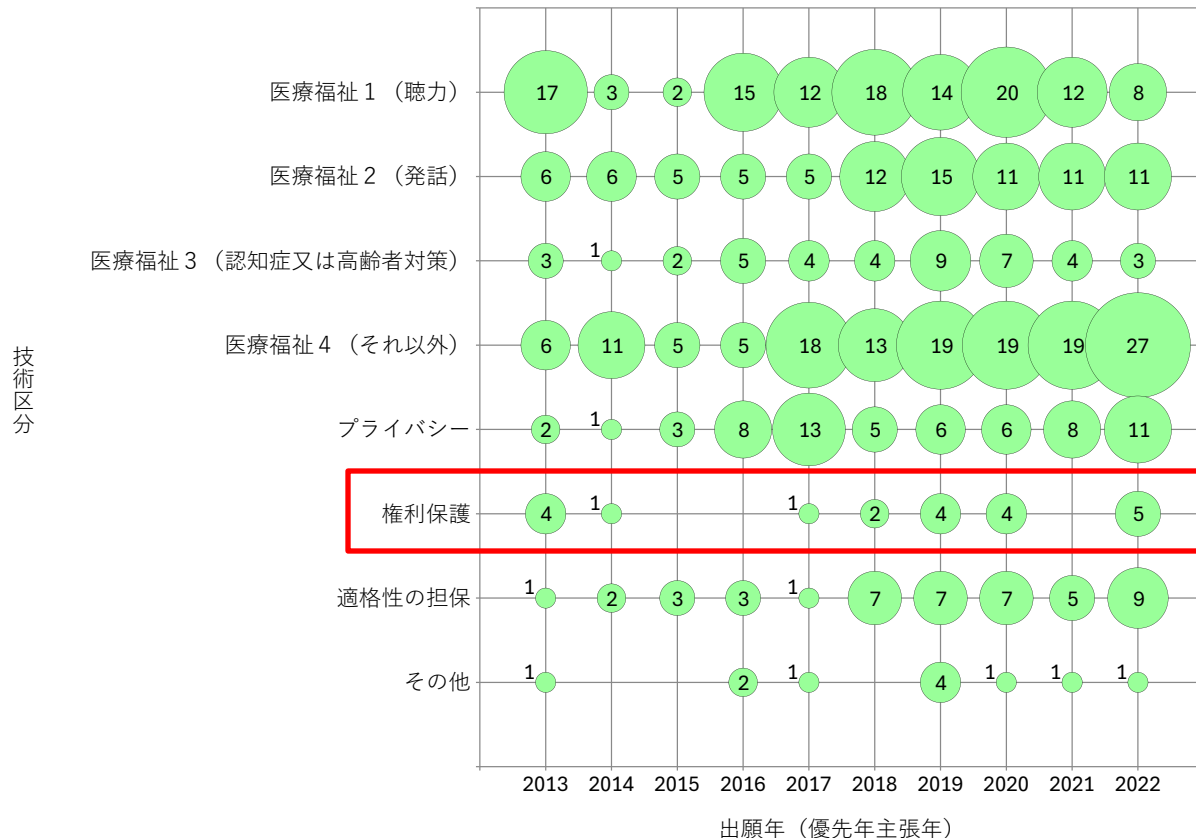
6. 特許出願動向

—社会的課題の出願動向—

<技術区分:社会的課題>

- 調査対象期間(出願年(優先権主張年)2013~2022年)における大区分「社会的課題」別パテントファミリー件数推移を示す。
- 権利保護に関するパテントファミリー件数は他の区分に比べると少ないものとなっている。

図9 【技術区分「社会的課題」別パテントファミリー件数推移（日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】



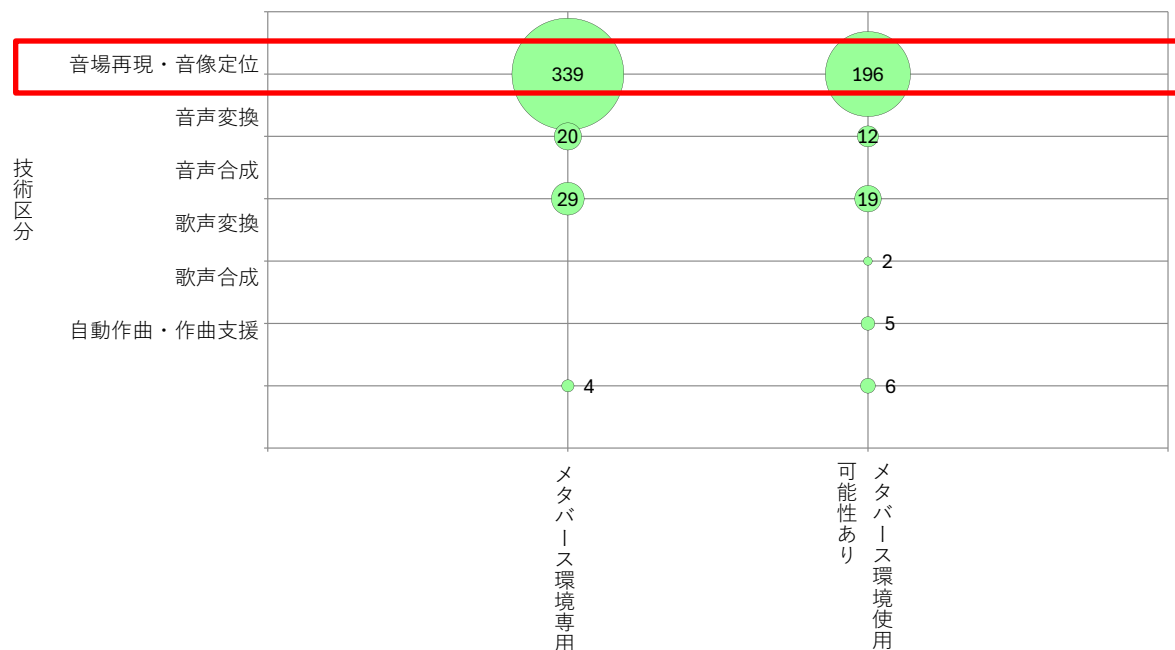
6. 特許出願動向

－大区分×メタバースの出願動向－

＜技術区分：大区分×メタバース＞

- 調査対象期間（出願年（優先権主張年）2013～2022年）における大区分である音場再現・音像定位、音声変換、音声合成、歌声変換、歌声合成、自動作曲・作曲支援の6つの技術区分とメタバースの関係を示す。
- メタバースを意識した特許出願は音場再現・音像定位から多く出願されている。その他の技術区分はこれに比べると少数であり、メタバースを意識した研究開発、技術開発活動は盛んではないものと考えられる。

図10 【技術区分「大区分」別－技術区分「メタバース」別パテントファミリー件数（日米欧中韓W0への出願、出願年（優先権主張年）：2013～2022年）】

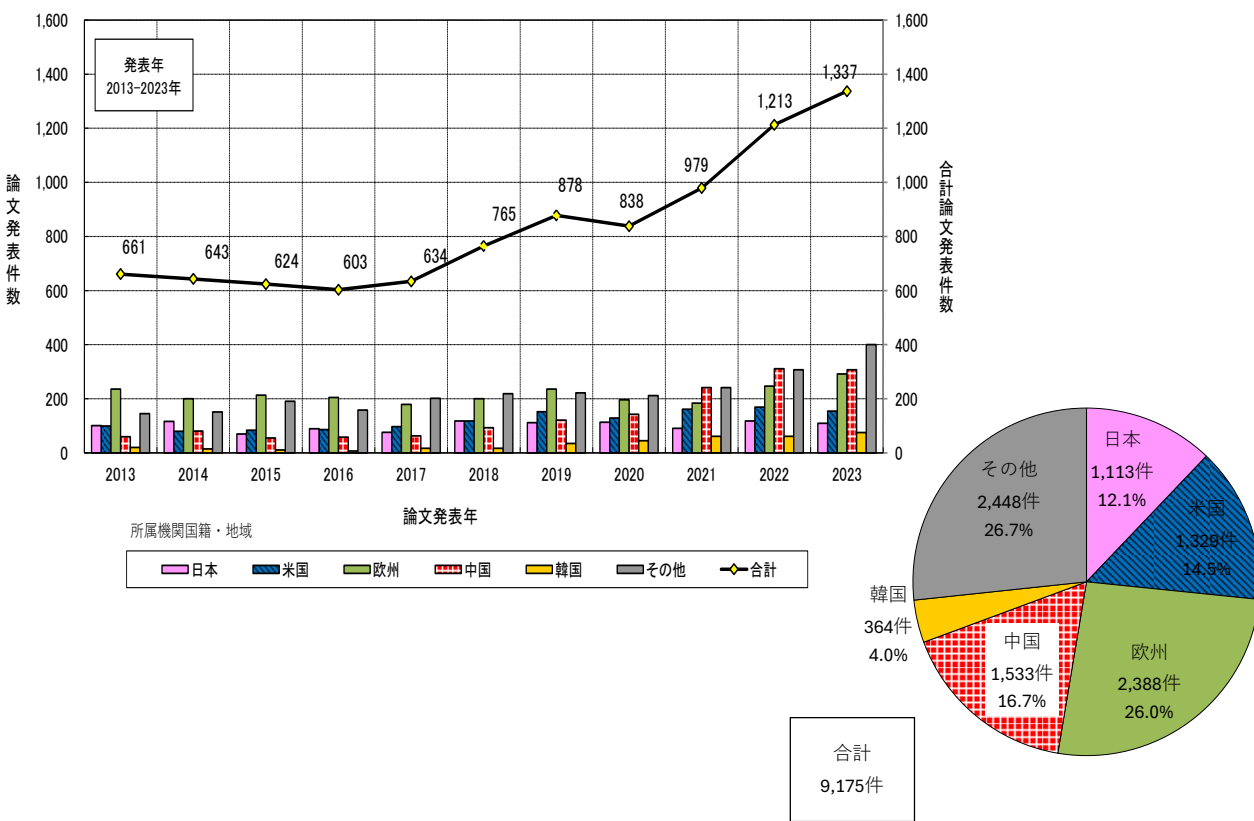


7. 研究開発動向

－全体に関する所属機関国籍・地域別の動向－

- 調査対象期間(論文発表年 2013～2023年)における研究者所属機関国籍・地域別の論文発表件数推移、論文発表件数比率及び論文発表件数上位研究者所属機関ランキングを示す。件数比率を見ると、欧州からの論文発表が26.0%で最も多い。
- 件数推移を見ると、2021年以降は中国籍の論文発表が大きく増加し、年当たりの発表件数は欧州籍と同程度となっている。また、その他の論文発表が増加しており、このうちの多数はインドからの論文発表である。

図11 【研究者所属機関国籍・地域別論文発表件数推移、論文発表件数比率及び論文発表件数上位研究者所属機関ランキング
(論文発表年：2013～2023年)】



順位	研究者所属機関	国籍	属性	件数
1	東京大学	日本	大学	120
2	インド工科大学	その他	大学	116
3	Google	米国	企業	114
4	中国科学技術大学	中国	大学	113
5	エディンバラ大学	欧州	大学	108
6	中国科学院	中国	研究機関	97
7	西北工業大学	中国	大学	90
8	清華大学	中国	大学	83
9	日本電信電話	日本	企業	80
10	奈良先端科学技術大学院大学	日本	大学	79

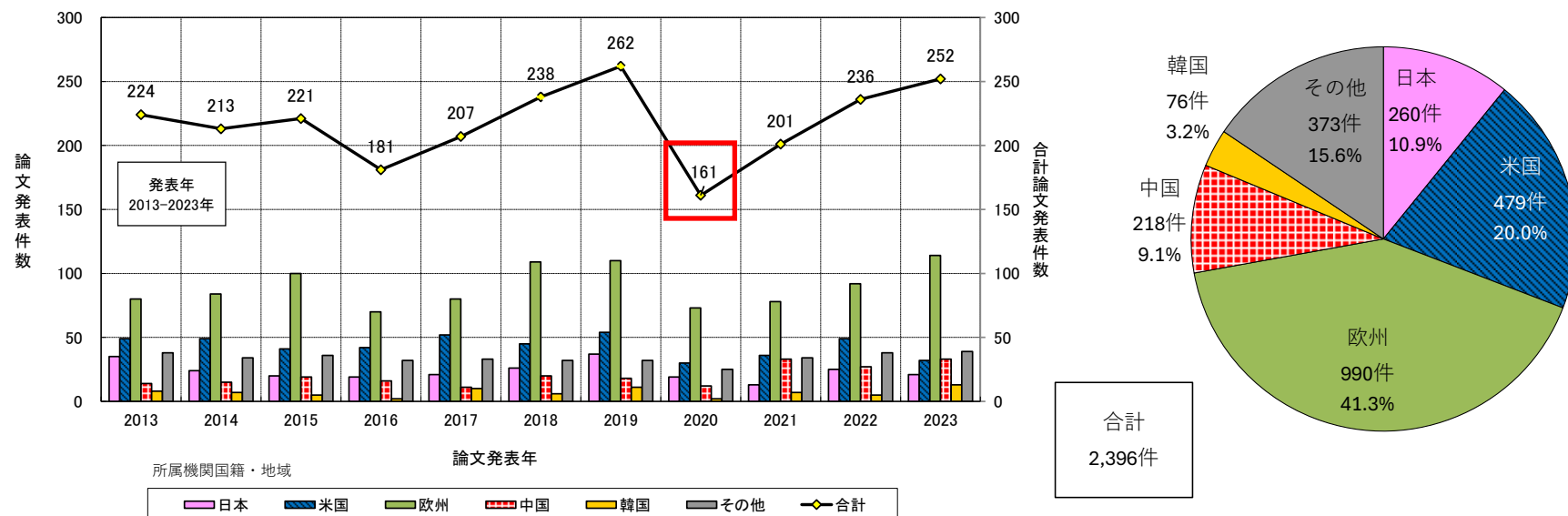
7. 研究開発動向

－音場再現・音像定位の研究開発動向－

＜技術区分：音場再現・音像定位＞

- 調査対象期間(論文発表年 2013～2023年)における大区分「音場再現・音像定位」における研究者所属機関国籍・地域別論文発表件数推移及び論文発表件数比率を示す。
- 欧州籍からの論文発表が突出して多く、次いで米国籍、日本国籍の順である。
- 2020年には論文数が大きく減少している。COVID-19の影響があったものと思われる。今回の調査対象技術のうち、音場再現・音像定位は施設での実験が重要な位置を占める。この時期には研究者や被験者が実験施設に集まることが困難であったという事情が推測される。
- その後に徐々に論文発表件数は増加傾向にあり、2023年時点では2019年とわずかな差しかないまでの件数になっている。音場再現・音像定位に関しての研究開発は一定程度の規模によってなされていると見られる。

図12 【研究者所属機関国籍・地域別論文発表件数推移及び論文発表件数比率（「音場再現・音像定位」、論文発表年：2013～2023年）】



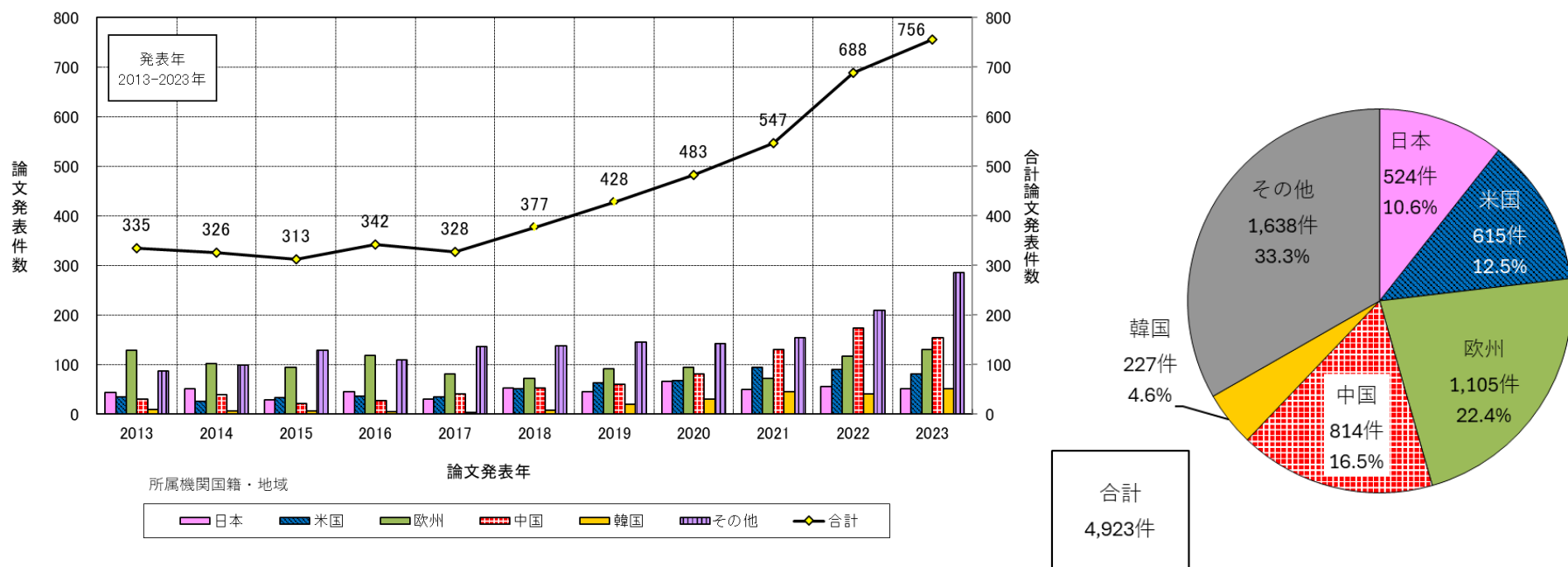
7. 研究開発動向

－音声合成の研究開発動向－

<技術区分：音声合成>

- 調査対象期間(論文発表年 2013～2023年)における大区分「音声合成」における研究者所属機関国籍・地域別論文発表件数推移及び論文発表件数比率を示す。
- 2018 年以降、論文発表件数は増加傾向である。また、その他国籍からの論文発表が多いことがわかる。これの多くはインドからの論文発表である。

図13 【研究者所属機関国籍・地域別論文発表件数推移及び論文発表件数比率（「音声合成」、論文発表年：2013～2023年）】



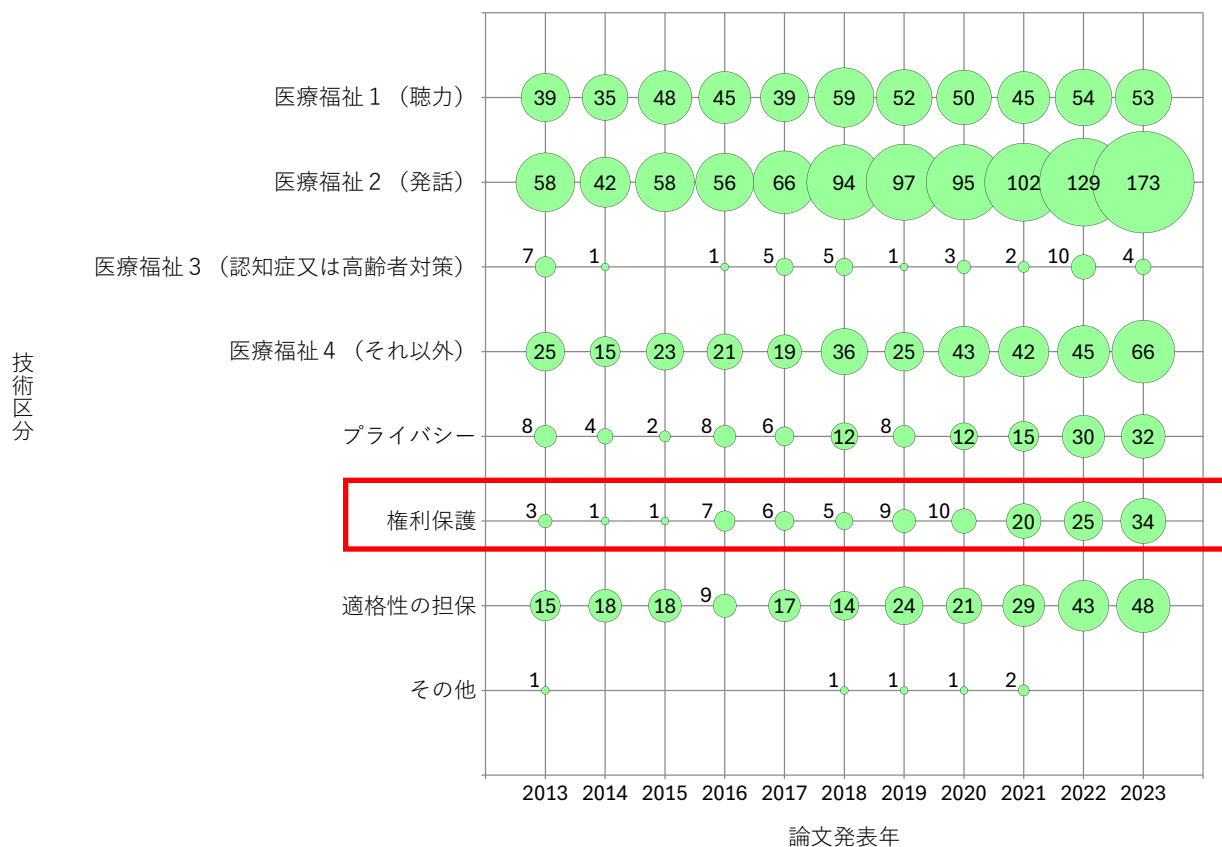
7. 研究開発動向

－社会的課題の研究開発動向－

<技術区分：社会的課題>

- 調査対象期間(論文発表年 2013～2023年)における社会的課題別論文発表件数推移を示す。
これによれば、権利保護に関する論文発表件数は2021年以降、増加が認められる。

図14 【技術区分「社会的課題」別論文発表件数推移（論文発表年：2013～2023年）】



日本の目指すべき方向性に関して検討した結果を提言、示唆としてまとめた。

■ **提言1: 音の伝達に付加価値を与えるための研究開発、技術開発を継続し、日本が今後も存在感、影響力を高めていくことが必要である。**

- 特許出願動向を見ると、**音像定位技術、音場再現技術**ともに毎年一定の量の特許出願が確認できる。また、研究開発動向を見ると、論文発表件数も毎年一定量が確認できる。**音の伝達に対して付加価値を与えるための研究開発活動、技術開発活動が常に行われていると考えられる。**今後、音の伝達の場面はメタバースを始めとして広がっていくと考えられることから、この点の**研究開発活動、技術開発活動**を今後も継続し、日本の存在感、影響力を高めていくべきである。

■ **提言2: 音声合成技術の実用化に際して、社会情勢やそれに基づいた政策意思決定の動向を確認するとともに、悪用防止の研究開発、技術開発へと速やかに移行できるような体制づくりをすることが必要である。**

- 声のなりすましを許さないという社会的な機運が高まっている。**権利保護に関する論文の発表件数は2021年以降に増加傾向を見せており、今後の研究開発活動の活発化が期待される。**欧州では、2024年5月にEU理事会によってAI法案が採択され、音声コンテンツを生成または操作するためにAIシステムを利用する場合、そのコンテンツがAIによるものであることを開示する必要があることが規定された。このように社会的な機運によって政策面での指針が策定されており、これに対応する研究開発活動、技術開発活動が進展していくと考えられる。**日本においても、社会情勢やそれに基づいた政策意思決定の動向を注視し、これらに対応した研究開発活動、技術開発活動の開始に乗り遅れないための準備をしておくことが望ましい。**

- 提言 3 : 音楽知識がなくとも作曲の喜びを提供する研究開発、技術開発を進めることによって、今後の市場において優位性を確保することが必要である。
 - 一般的に市場におけるサービスの展開においては、先行者利益が大きく、後発の参入で先行者に勝つことが難しい場合がある。2018年以降、自動作曲・作曲支援に関する中国の特許出願件数及び論文発表件数の増加が顕著であり、市場導入に向けた研究開発活動、技術開発活動が活発に行われている様子が見られる。そのような事情を鑑み、市場において他国に遅れを取らないよう研究開発活動、技術開発活動を進める必要がある。

- 示唆 1 : 音声技術、音楽技術の進展には技術革新が非常に大きく関わっている。革新的な技術をいち早く製品、サービス開発に取り入れることによって、他に先駆けて消費者に支持される製品、サービスを提供することが必要であると考えられる。
 - 新しい理論の発表後に特許出願、論文発表が増加することは、最新の研究開発、技術開発動向に気を配り、その成果をいち早く製品、サービス開発に取り入れることが産業競争力の確保に必要であることを示している。従って、旧来から存在する技術や理論に対するブレイクスルーの登場に注意し、そのような技術や理論を迅速に取り入れて研究開発活動、技術開発活動を進めて行くことが望ましい。

- 示唆 2 : メタバースへの展開を見据えた場合、音場再現・音像定位技術を活用し、産業競争力を確保することが考えられる。
 - 日本は音場再現・音像定位に強みを持っていることから、メタバースへの展開に関しても、この音場再現・音像定位を基礎として研究開発活動、技術開発活動を継続することが好ましい。

ご清聴ありがとうございました。

※¹本調査の報告書の要約は、特許庁のウェブサイト上（以下URL）で公開しています。

<https://www.jpo.go.jp/resources/report/gidou-houkoku/tokkyo/index.html>

（特許庁 ホーム＞資料・統計＞刊行物・報告書＞出願動向調査等報告書＞特許出願技術動向調査）

※²パテントファミリー件数とは、国内国外を通じて、少なくとも一つの共通の優先権を持ち、技術内容が完全または部分的に一致する関係を有する特許文献群を1件とカウントした件数である。国際パテントファミリー（International Patent Family）件数とは、複数の国・地域への出願を含むパテントファミリー、又は、欧州特許庁への出願やPCT出願（複数の国・地域での権利取得意志に基づくと推定される出願）を含むパテントファミリーを1件とカウントした件数である。国際パテントファミリー件数は、欧州特許庁への出願であればそれだけで1件としてカウントするものであるため、欧州籍の出願件数が多くカウントされ易い特徴があることに留意されたい。

論文だけで終わらせない！ 特許で広がる基礎研究

真理探究タイプの基礎研究者のあなたも
特許の視点を研究にプラスすることで、
研究の幅や視点が広がり、
応用研究の可能性が生まれます！

- ◆あなたはどのタイプ？研究者タイプ別の情報を記載
真理探究タイプ・実用思考タイプ・バランスタイプ
- ◆研究者タイプごとの特許のメリットをストーリーで説明
- ◆特許取得のプロセス(学内フロー～特許取得)を分かりやすく解説
論文投稿と対比させてイメージしやすくしました

特許庁 研究活動 パンフレット で検索



出典:令和6年度特許庁産業財産権制度問題調査「大学の研究者の目線に立った知財の情報提供に関する調査研究」