

第7回特許情報シンポジウム

論文資料集

2023年2月20日(月)

オンライン開催

一般社団法人アジア太平洋機械翻訳協会

一般財団法人日本特許情報機構

特許庁における機械翻訳に関する取組

令和5年2月20日

特許庁 総務部 総務課

情報技術統括室 室長補佐（特許情報利用推進班長）

名和 大輔



1

特許情報提供の意義・目的

2

特許情報提供における機械翻訳の活用

3

機械翻訳に関する調査事業

4

特許庁が提供する機械翻訳文の利用方法

1

特許情報提供の意義・目的

2

特許情報提供における機械翻訳の活用

3

機械翻訳に関する調査事業

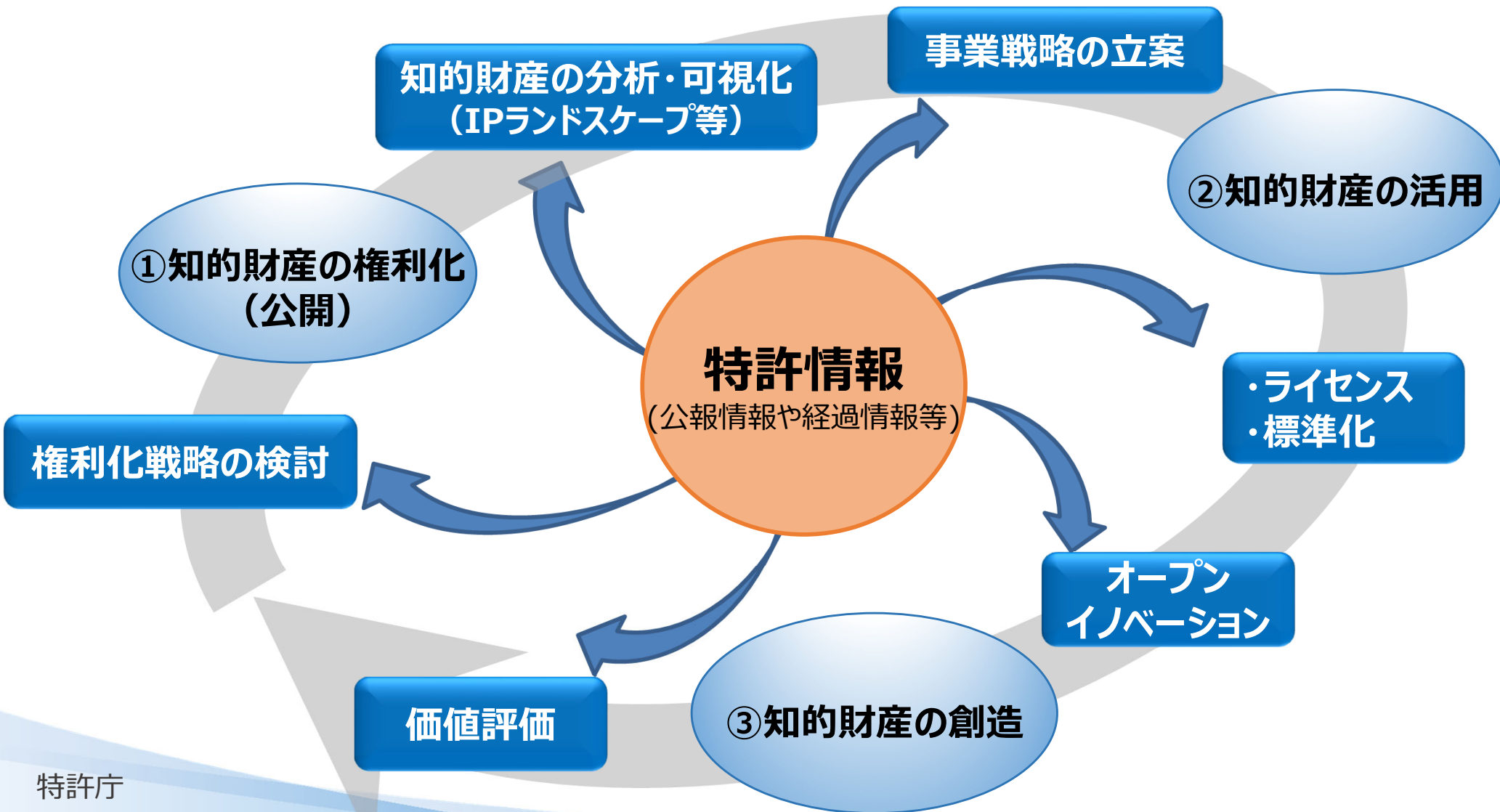
4

特許庁が提供する機械翻訳文の利用方法

知財エコシステムと特許情報

- 特許情報は知財エコシステムにおける促進剤。

特許情報…特許・実用新案・意匠・商標の出願・権利化に伴って生み出される情報(公報情報や経過情報等)

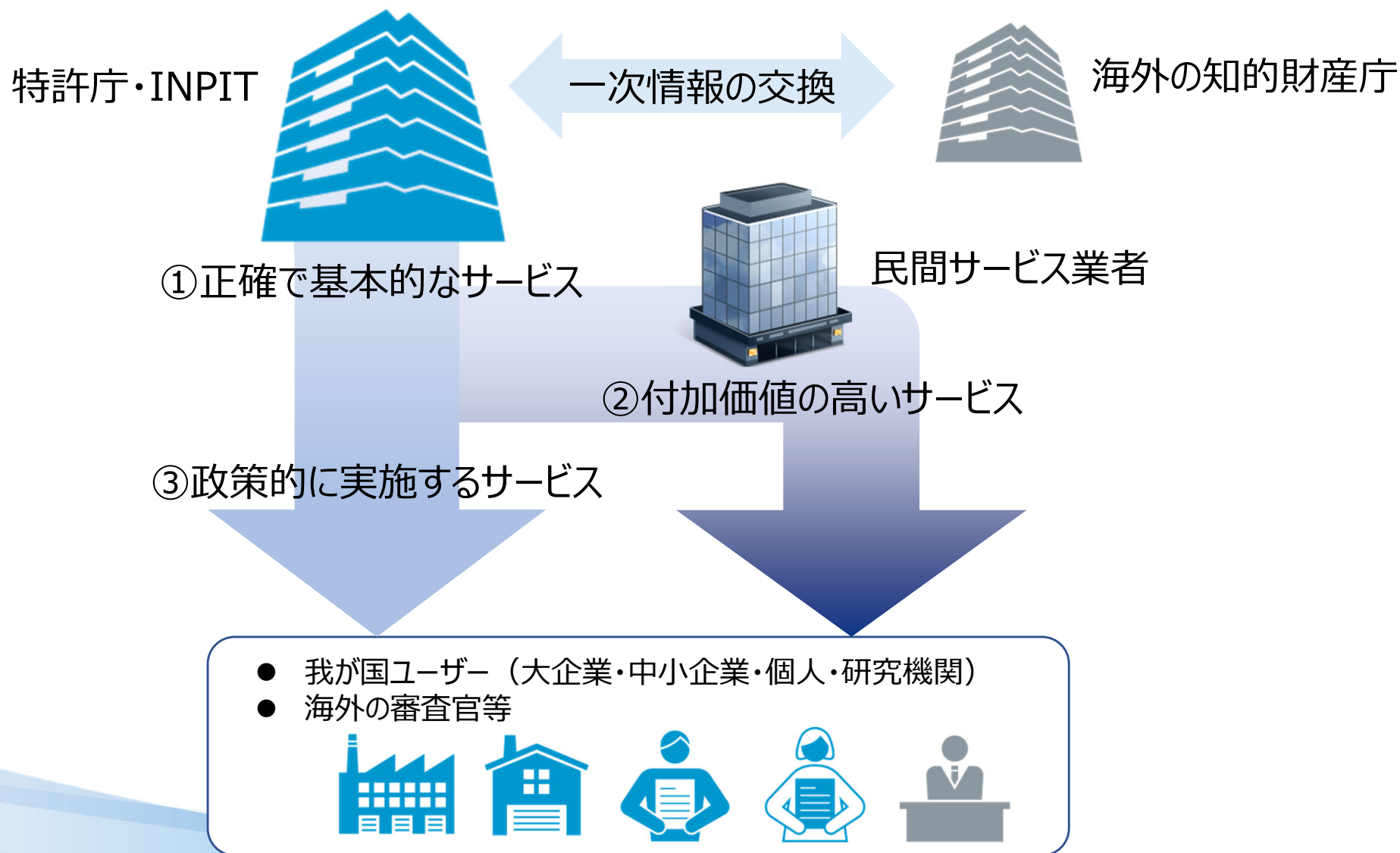


特許情報の提供における国と民間事業者の役割

- 国が正確で基本的な一次情報を提供し、民間事業者が付加価値の高いサービスを提供。

参考：産業財産権情報利用推進委員会（長官私的懇談会）報告（平成15年3月）

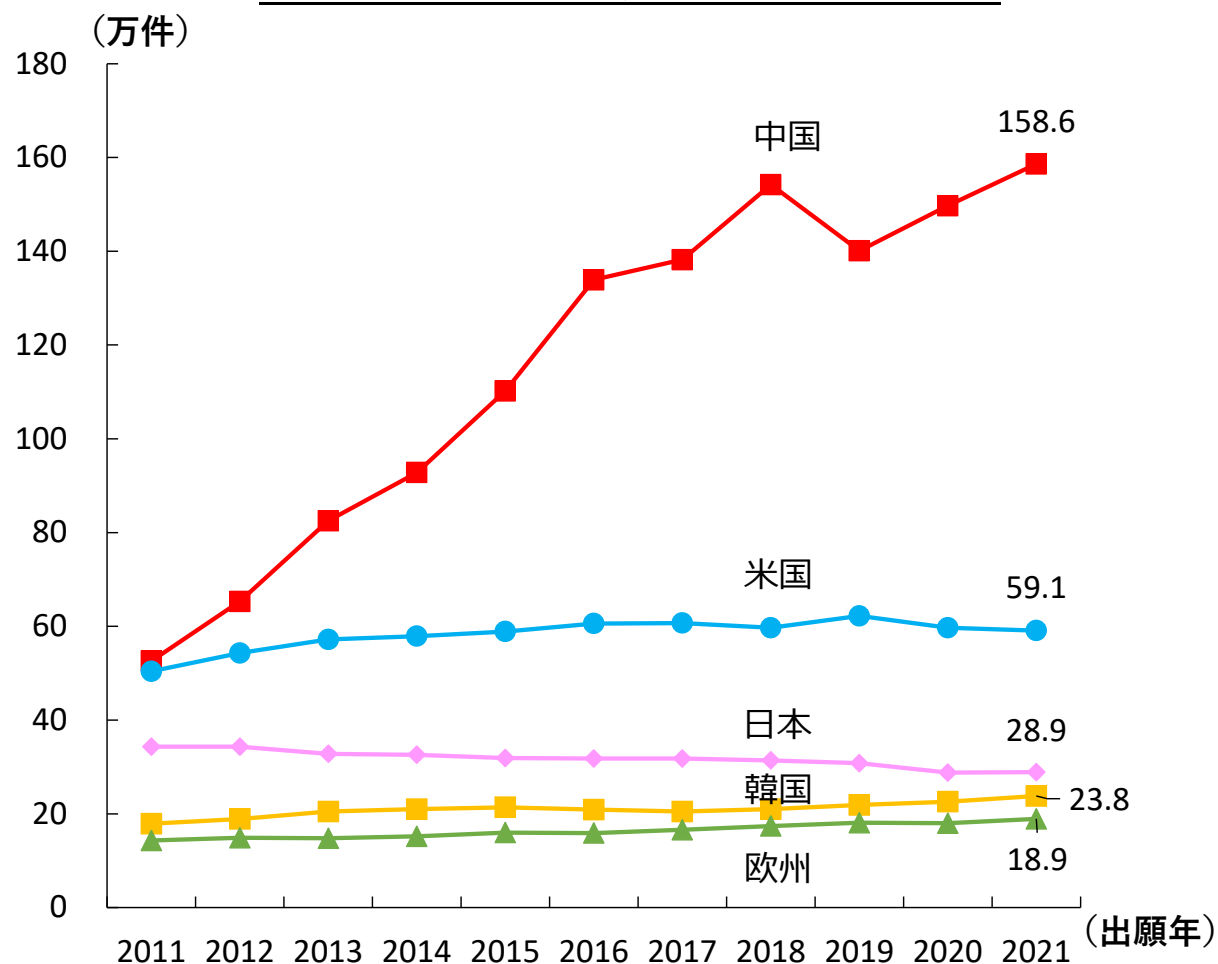
産業構造審議会 知的財産分科会 情報普及活用小委員会（平成28年5月）



各国の特許出願動向

- 中国における特許出願件数は増加傾向を維持する一方、他国は横ばい傾向。
- 日本では、AI関連発明の特許出願件数が増加。（データ非表示）

日米欧中韓における特許出願件数

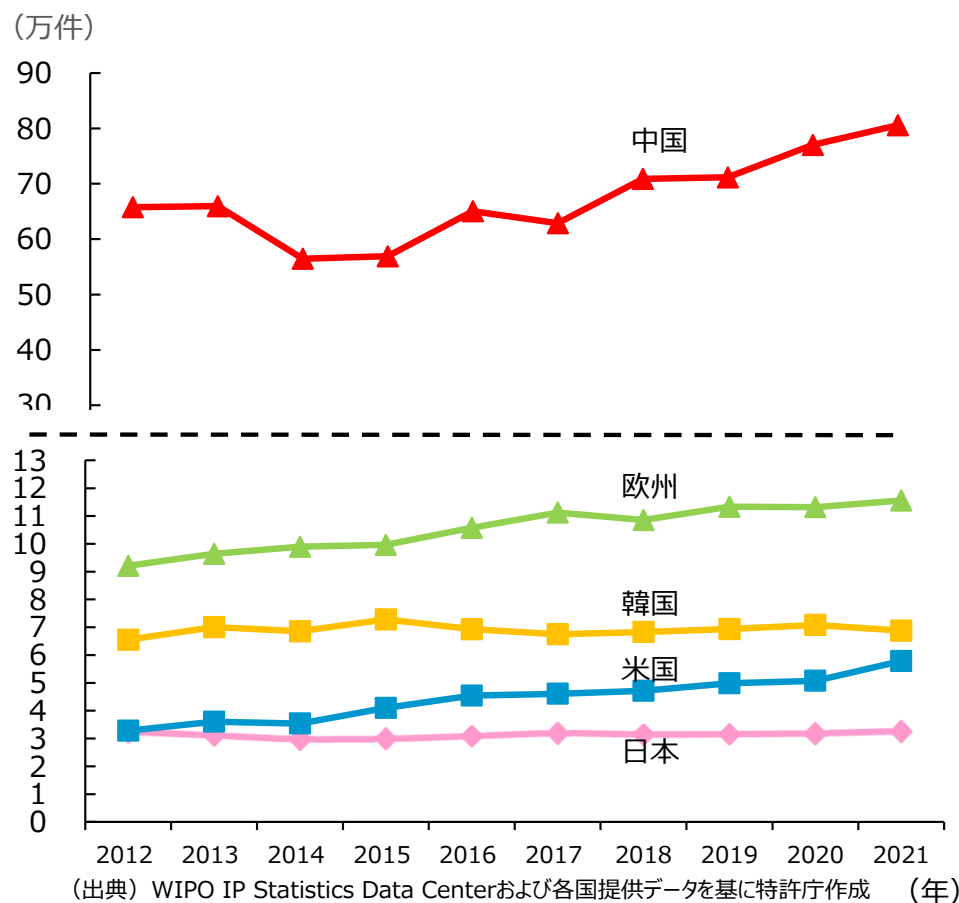


(出典) 【日本】特許行政年次報告書2022年版 第2部第1章1. 【米国】USPTO提供資料 (2017-2020年、2021年暫定値)
【欧州】EPOウェブサイトPatent Index 2021 (European patent applications)
【中国】CNIPA Annual Report 【韓国】KIPOウェブサイト及びKIPO提供資料 (2021年暫定値)

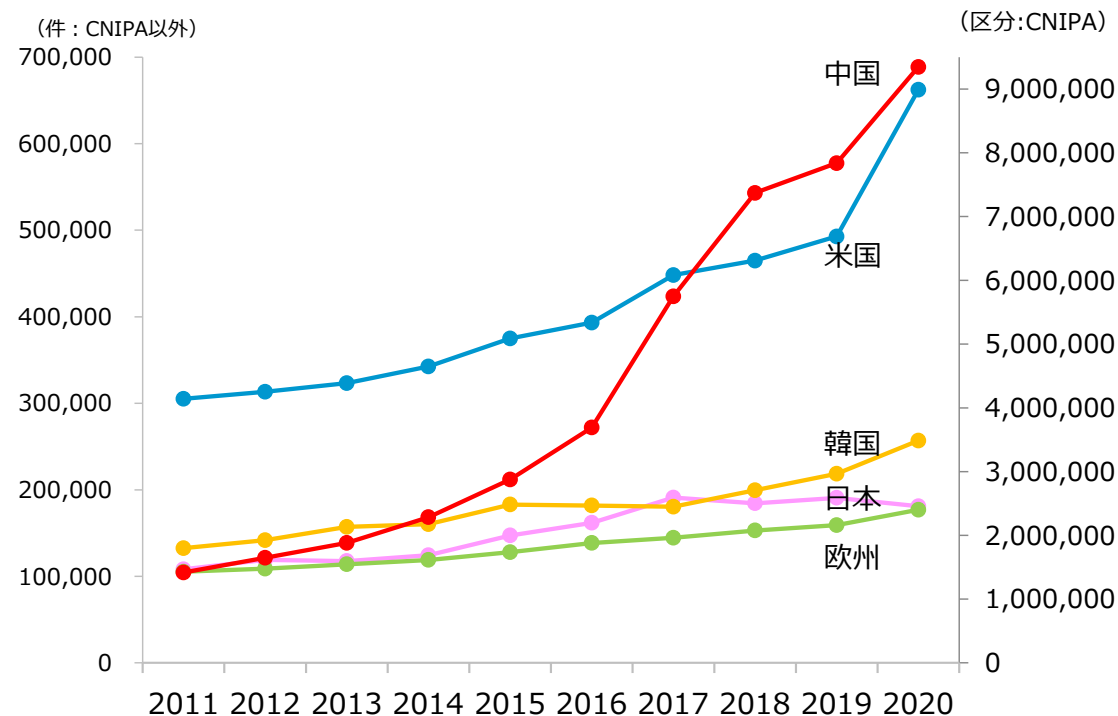
各国の意匠／商標出願動向

- 意匠／商標でも中国における出願件数は増加。特に商標で著しく増加。
- 意匠について、2020年以降、コロナ関連物品が含まれる生活用品等や住宅設備用品が増加。（データ非表示）
- 商標について、IT関連を含む機械分野の出願（区分数）が顕著に増加。（データ非表示）

日米欧中韓における意匠登録出願件数



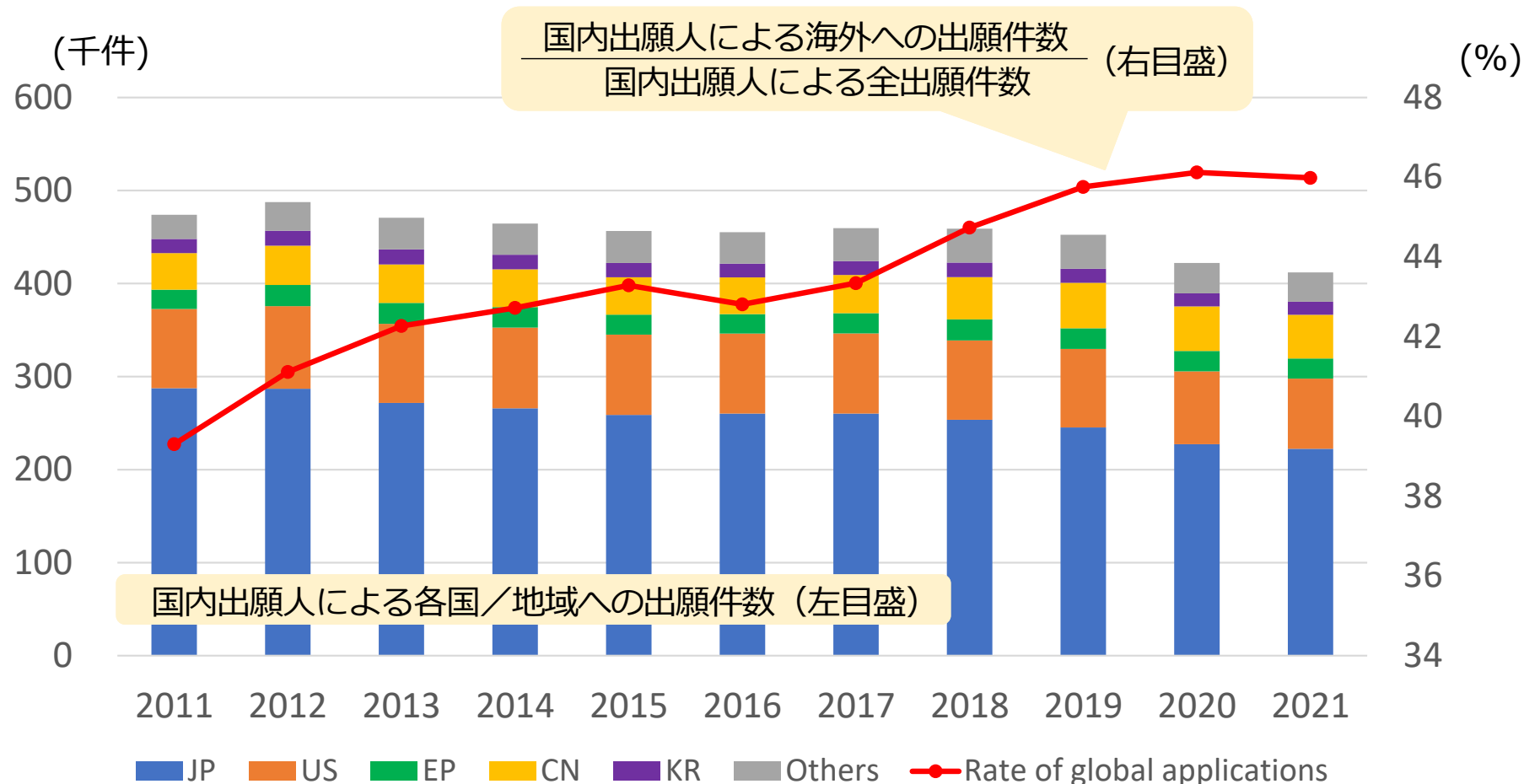
日米欧中韓における商標登録出願件数



(出典)
 【日本】特許庁年次報告書2021年版
 【中国】中国商標戦略年度発展報告（～2017年）及びTM5における
 Report for Common Statistical Indicators（2018年～）を基に特許庁作成
 【その他】WIPO統計
 ※中国は件数での公表を行っていないため、数値は区分数（右軸）。

国内出願人の特許出願動向（国内／海外）

- 国内出願人による特許出願先は、国内が減少傾向にある中、海外割合は増加傾向。
- それに伴い、日本の特許情報の海外への発信、及び海外庁からの特許情報の国内提供は重要度を増している。



産業構造審議会知的財産分科会情報普及活用小委員会での報告書

- 特許庁又はINPITが運営する公的な特許情報提供サービスのあり方等について審議を行い、「特許情報のさらなる活用に向けて」として報告書を公開。（2016年5月）

□ 世界最速・最高品質の審査結果の発信のためのサービス

- ✓ 審査結果の発信にあたっては、・・・新興国も含めた海外庁の審査官等が日本の審査結果を参照しやすいツールとするために必要な機能を積極的に搭載して、世界最高の利便性を実現して審査結果に関するポータルサイト的位置付けを目指していくべきである。また、我が国の審査結果を利用する際には、通常、機械翻訳を介する必要があるが、海外庁の審査官等が我が国の審査結果（審査書類＋引用文献等）を参照しやすいよう機械翻訳の取組を強化し、高精度な機械翻訳文を提供することにより、我が国の審査結果の利用を促進していくべきである。

□ 出願情報・権利情報の確認のためのサービス

- ✓ 海外庁の文献は、日本語以外の言語で記載されているため、多くのユーザーにとって内容を理解するために時間を要することから、現在も中国語、韓国語の特許・実用新案文献に対して行っているように、機械翻訳を活用しつつ、日本語でアクセスできる環境の整備を進めていくべきである。

産業構造審議会知的財産分科会基本問題小委員会での報告書

- 新型コロナウイルスの感染拡大やデジタル化の流れの加速の中で、イノベーション促進に向けた産業財産権制度の改善の在り方、また、特許庁の事務や審査業務の品質向上などを幅広く審議を行い、「ウィズコロナ／ポストコロナ時代における産業財産権政策の在り方―とりまとめ―」として報告書を公開。（2021年2月）

I .これからの審査制度・業務を支えるシステムの在り方

【ユーザニーズを踏まえ、AI等も活用し、効率化・合理化を進める】

● 特許審査制度

- ・ ユーザへの提供価値（審査の質、利便性等）の見直し
- ・ 先行技術調査の強化（毎年350万件超増加する海外特許技術文献を着実に調査できるようAIを活用）、効率化（類似度順スクリーニング等を活用したメリハリのあるサーチ等）
- ・ 高審査負荷の請求項記載形式（マルチマルチクレーム）制限に向けた検討
- ・ 調査対象文献の急増（中国文献爆発）等への対応のため、審査プロセスを改めて見直し、徹底した効率化を推進。

1

特許情報提供の意義・目的

2

特許情報提供における機械翻訳の活用

3

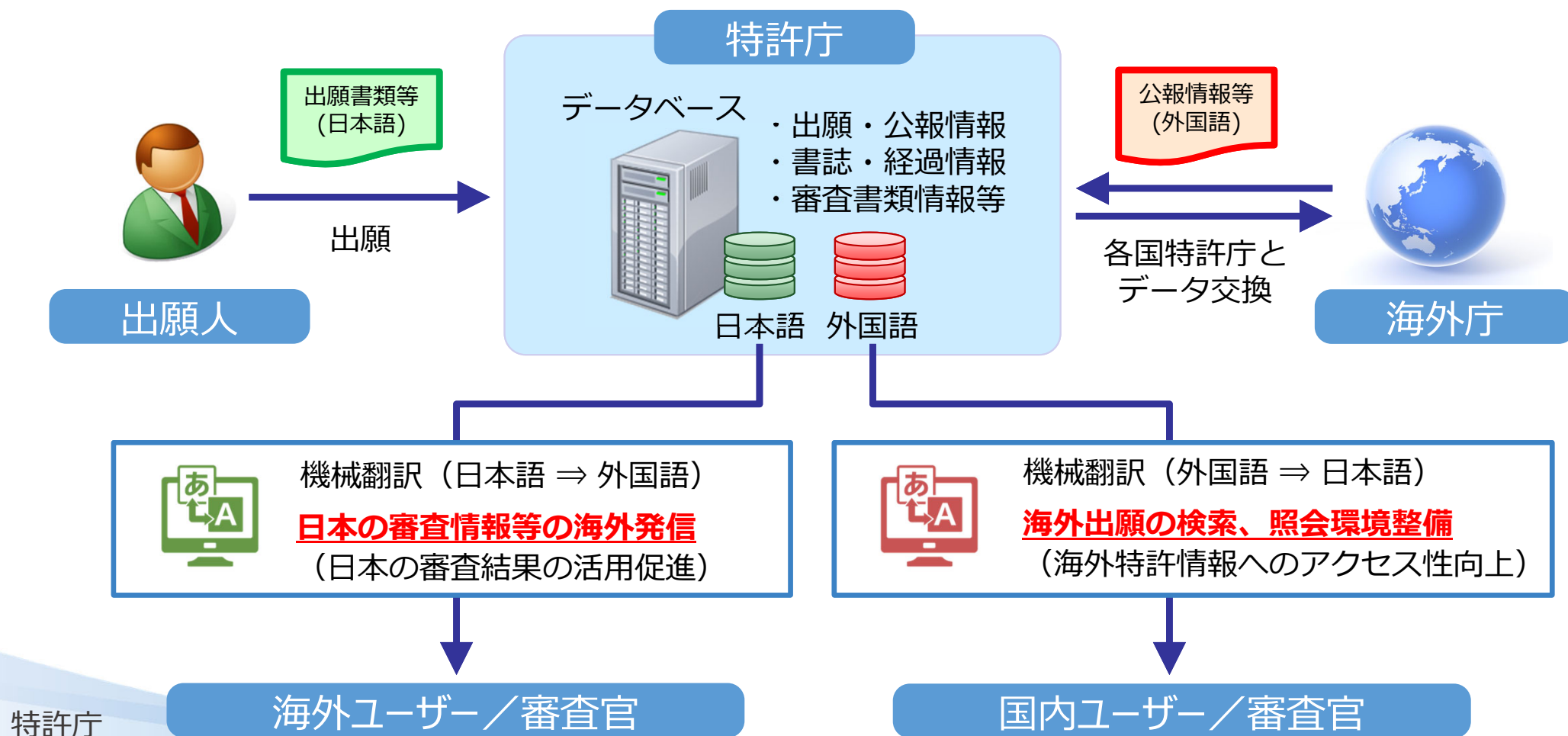
機械翻訳に関する調査事業

4

特許庁が提供する機械翻訳文の利用方法

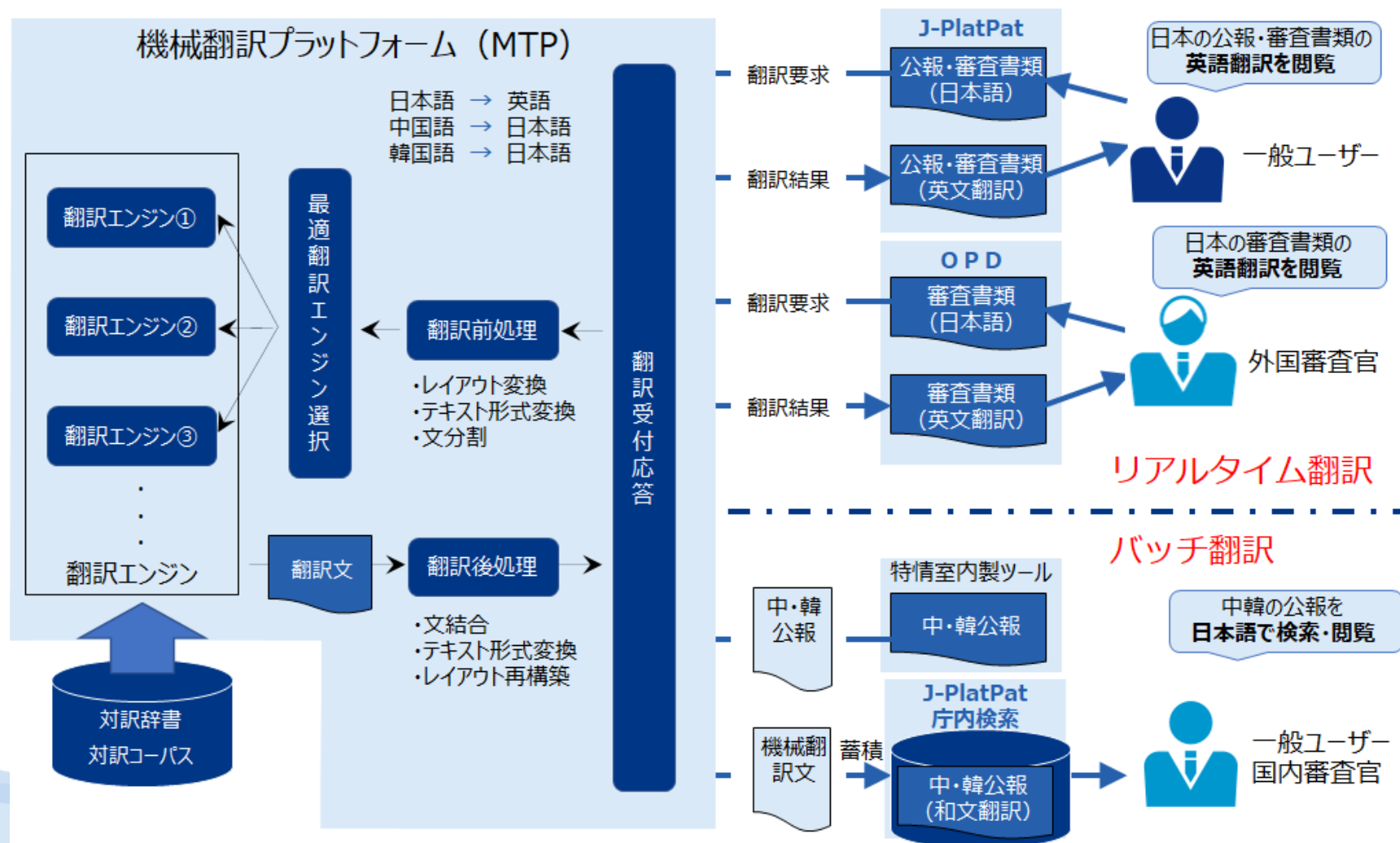
機械翻訳を活用した特許情報提供

- 特許庁では、日本の審査情報等の海外発信／海外出願の検索、照会環境整備の施策における翻訳として、機械翻訳を活用している。
- 機械翻訳した、出願人や海外庁から得られた各種データは、国内外の審査官や一般利用者に提供している。



機械翻訳プラットフォーム（MTP）の構築・運用

- 特許庁では、機械翻訳プラットフォーム（MTP）を通じて、日本の審査情報や公報のリアルタイム英訳、中韓の特許/実用新案のバッチ和訳を提供。
- MTPは、複数の翻訳エンジンを搭載しており、令和元年5月に日英翻訳機能、翌令和2年4月に中日・韓日翻訳機能をリリースし、現在運用中。

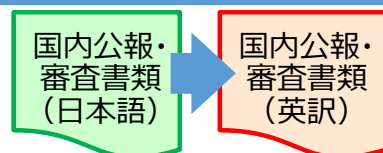


MTPの特徴 ①翻訳エンジンの使い分け

- MTPには、複数の翻訳エンジンが搭載されており、NMTをメインに利用しつつ、SMT及びRBMTの併用や、翻訳前後の文章に対する各種処理等を含む独自システムを通じて、翻訳品質を最適化。

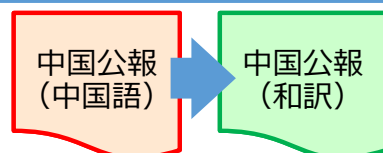
MTPにおける翻訳エンジンの使い分け

日英翻訳



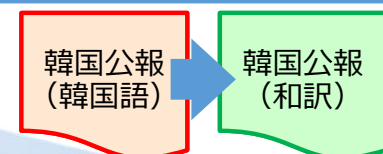
翻訳対象	メイン翻訳エンジン
特許公報（明細書等）	NMT（公報用カスタマイズ）
特許公報（発明の名称）	SMT
審査書類等	NMT（審査書類用カスタマイズ）
書誌情報・特許以外の公報 等	RBMT

中日翻訳



翻訳対象	メイン翻訳エンジン
特許公報（明細書等）	NMT
特許公報（発明の名称）	SMT

韓日翻訳

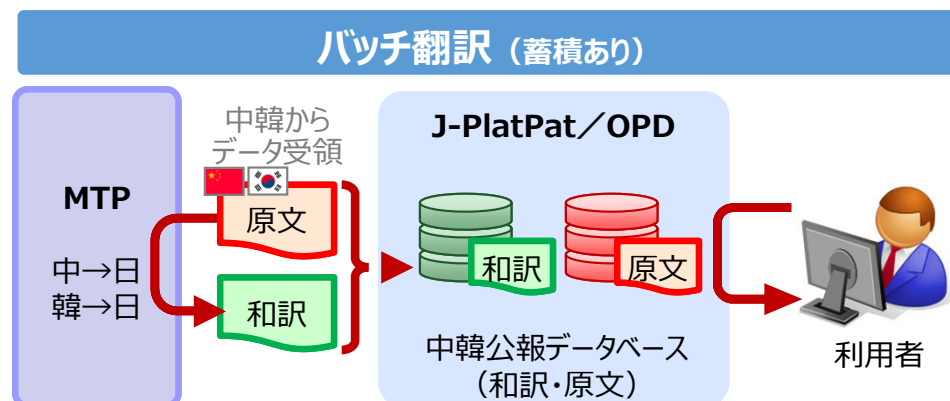
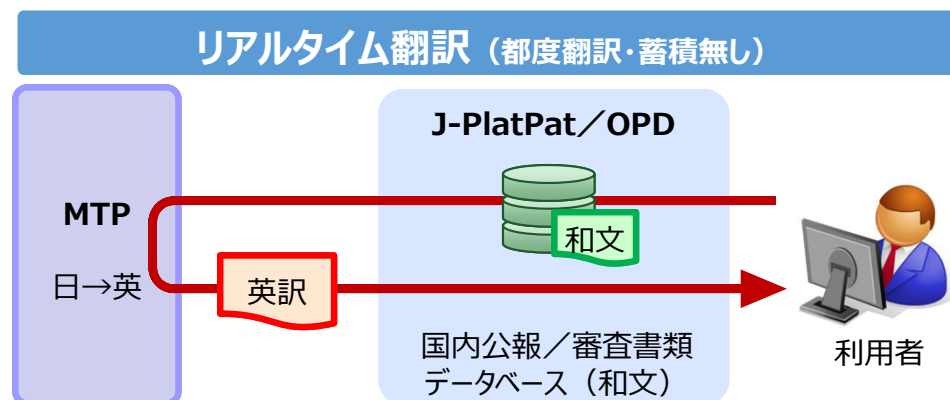


翻訳対象	メイン翻訳エンジン
特許公報	SMT

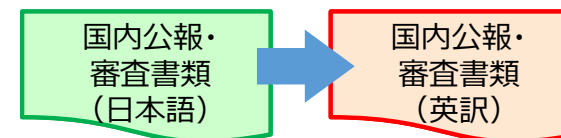
MTPの特徴 ②リアルタイム翻訳とバッチ翻訳

- 日英翻訳では、リアルタイム翻訳（照会の度に都度翻訳）を採用しているのに対し、中日・韓日翻訳では、バッチ翻訳（予め検索システムに翻訳文を蓄積）を採用。

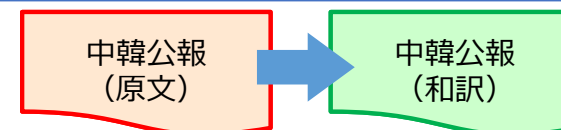
リアルタイム翻訳とバッチ翻訳の比較



メリット	デメリット
最新品質の翻訳文を即時に参照できる	翻訳文を対象としたテキスト検索ができない



メリット	デメリット
翻訳文を対象としたテキスト検索ができる	蓄積時点の翻訳品質（最新の品質でない）



特許情報の翻訳文を提供している主なインターフェース

- 日本国特許庁が保有する特許情報は、J-PlatPat、OPD、FOPISER等を通じて、無償で对外提供。
- これらインターフェースにおいて、特許情報の各種翻訳文を提供。

特許情報プラットフォーム（J-PlatPat）



<https://www.j-platpat.inpit.go.jp>

- 国内外の特許、実用新案、意匠、商標、審決に関する公報、手続、審査経過等を収録
- 「ぷらっと」寄って情報を「ぱっと」見つけれられるような、ユーザーフレンドリーなサービス提供を目指す名称

ワンポータルドシエ（OPD）

<https://www.j-platpat.inpit.go.jp/p0000>
(J-PlatPatインターフェース（一般用）)

- 各庁の特許審査に関連する、手続や審査の状況、各種書類データ等（ドシエ情報）を、原語及び英語で閲覧できるプラットフォーム

外国特許情報サービス（FOPISER）

<https://www.foreignsearch2.jpo.go.jp>

- 露・台・オーストラリア・シンガポール・ベトナム・タイ・EUIPOの特許、実用新案、意匠又は商標の公報情報、及び中国の特許及び実用新案に関する審決の和訳文を収録



WIPO Translate

- 世界知的所有権機関（WIPO）が開発した機械翻訳ツール。PCTの国際公開言語（日本語を含む10言語）について、自由入力文の翻訳が可能。
- WIPOの特許検索（PATENTSCOPE）において、公報／書誌情報等をリアルタイム翻訳する翻訳エンジンとして、Google翻訳と併用している。
- 2016年9月よりNMTを採用。

WIPO Translate

WIPO Translate NMT is a powerful instant translation tool, designed specifically to translate patent texts (now almost all languages are available using Neural Machine Translation technology). Simply cut and paste text from a patent document into the box below and select from the available language pairs, then click on "Translate" (Note that there is a limit of 2000 characters).

NOTE: WIPO Translate not be used for translating undisclosed patent information or other sensitive data as data transmitted via the translation tool is not encrypted)

Text to be translated:

Language pair: English->Japanese (Neural MT)

Domain: ADMIN-Admin, Business, Management & Soc Sci

Show concordances: ☒

Translate

This automatic translation is provided for information only. It may contain discrepancies or mistakes and does not have any juridical value.

- Please hover your mouse over parallel segments of text
- Click to view other proposals
- Select words or phrases on the left to access other translation proposals

WIPO Translate NMT is a powerful instant translation tool, designed specifically to translate patent texts (now almost all languages are available using Neural Machine Translation technology). Simply cut and paste text from a patent document into the box below and select from the available language pairs, then click on "Translate" (Note that there is a limit of 2000 characters).

Edit translation

PATENTSCOPE

1. JOP/1986/1448 - ARTICLES FOR CARRYING CHEMICALS

国内書誌情報

官庁: ヨルダン

出版番号: P/1986/1448

出版日: 05.06.1986

公開番号: P/1986/1448

公開日: 10.03.1986

発明の名称
[AR] 化学物質を運搬するための材料
[EN] 化学薬品運搬用物品

要約
[AR] 化学物質を運搬するための材料
[EN] 化学薬品運搬用物品

WIPO Translate
Google Translate

WIPO Translateで日本語を選択

<https://patentscope2.wipo.int/search/ja/search.jsf>

Patent Translate (EPO)

- 欧州特許庁（EPO）がGoogleの協力により開発した機械翻訳ツール。EPO各国の公用語（28言語）と英・独・仏語間の相互翻訳、及び中・日・韓・露語と英語の相互翻訳。
- EPO提供の特許検索サービス（Espacenet）を通じて提供。
- 2017年5月よりNMTを採用。

The screenshot shows the Patent Translate web interface. On the left is a language selection menu with 28 languages. The main content area displays a patent document in Japanese. A blue highlight box is placed over a paragraph in the document. On the right side, there are options to print the document (PDF only translation or PDF original and translation), a feedback form to improve translation quality, and a list of links including FAQ, Help, Legal notice, and Contact.

Patent Translate
Powered by EPO and Google

French
German

Albanian
Bulgarian
Croatian
Czech
Danish
Dutch
Estonian
Finnish
Greek
Hungarian
Icelandic
Italian
Latvian
Lithuanian
Macedonian
Norwegian
Polish
Portuguese
Romanian
Serbian
Slovak
Slovene
Spanish

注意

この翻訳は元の口頭翻訳と、特定の翻訳の正確性、正確性、完全性、信頼性、または他の品質への適合性は保証されません。申請に提出したものの言語に関する特許などの権利は、機械翻訳の信頼性に基づいて行わないようにしてください。

日本語に関するPatent Translate（リファレンス・トランスレート）はただいま開発中です。翻訳品質の向上に努めております。

明細書 US2021187858A1

優先権の主張

[0001]

このアプリケーションは、35 U.S.C.に基づく継続アプリケーションインパート（CIP）です。5アプリケーションSer.の120。出願特許Ser.。米国特許第16/719,901号。上記で特定された特許出願は、開示の能率性を提供するために、

発明の分野

[0002]

本発明は、一般に、口腔内投与装置のための自己修正方法およびシステムに関する。

背景ノート

[0003]

最近の研究によると、薬物が舌下（舌の下）に投与されると、従来の経口投与よりも速く吸収

Print
PDF (only translation)
PDF (original and translation)

Please help us to improve the translation quality.

Your opinion on this translation:

☐ Human translation
☐ Very good
☐ Good
☐ Acceptable
☐ Rather bad
☐ Very bad

Your reason for this translation:

☐ Overall information
☐ Patent search
☐ Patent examination

Submit

FAQ
Help
Legal notice
Contact

1

特許情報提供の意義・目的

2

特許情報提供における機械翻訳の活用

3

機械翻訳に関する調査事業

4

特許庁が提供する機械翻訳文の利用方法

中日、日英翻訳資源作成事業（平成31～令和3年度事業）

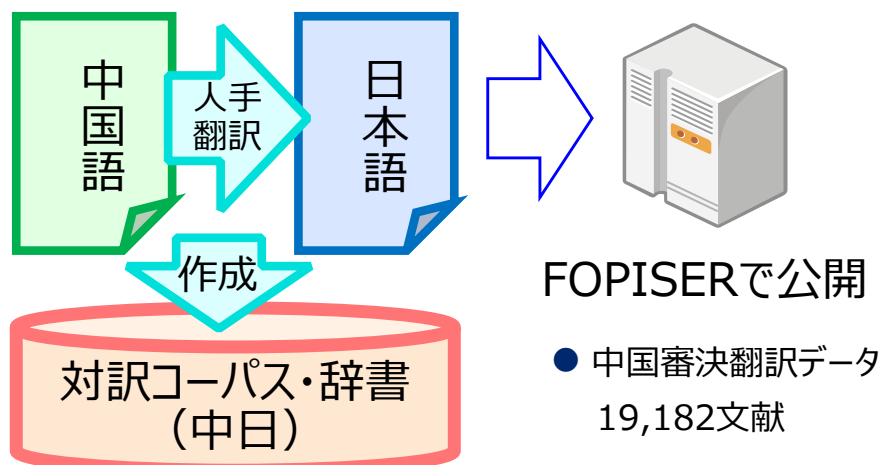
1. 中日機械翻訳に関するデータ作成・解析

- 中国審決の人手和訳、約200万の対訳コーパス作成
- 作成コーパスによる翻訳エンジンの学習効果を解析

2. 日英機械翻訳に関するデータ作成・解析

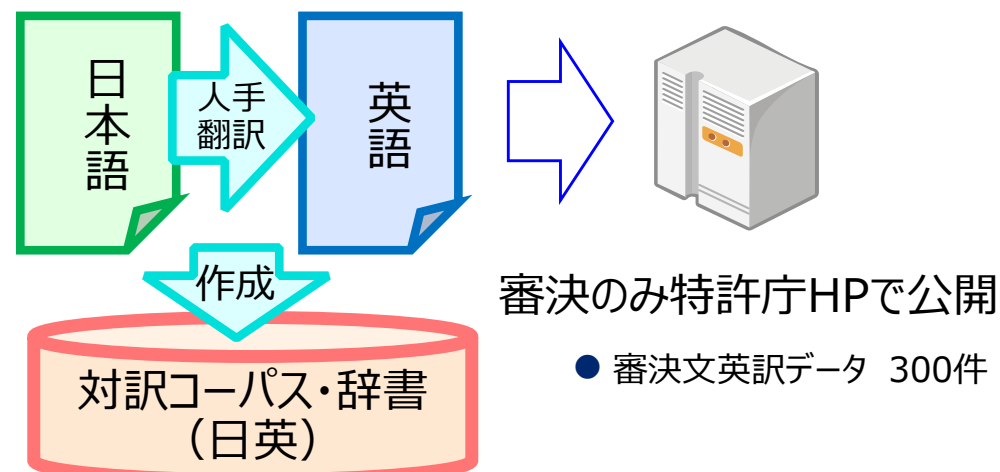
- 審査書類・審決の人手英訳、約200万の対訳コーパス作成
- 作成コーパスによる翻訳エンジンの学習効果を解析

1. 中国審決



- 対訳コーパス：約200万、対訳辞書：10万語

2. 審査書類・審決



- 審査書類対訳コーパス：約200万、対訳辞書：3万語
- 審決対訳コーパス：約10万、対訳辞書：3060語

翻訳品質の自動評価・人手評価

- 自動評価では、評価用コーパスを用いたBLEU値・RIBES値を算出する一方、人手評価では、評価用コーパスの一部を用いて以下の5段階で評価し、平均値を算出。

【内容伝達レベルの評価】

- 5 : すべての重要情報が正確に伝達されている (100%)
- 4 : ほとんどの重要情報は正確に伝達されている (80%~)
- 3 : 半分以上の重要情報は正確に伝達されている (50%~)
- 2 : いくつかの重要情報は正確に伝達されている (20%~)
- 1 : 文意がわからない、もしくは正確に伝達されている重要情報はほとんどない (~20%)

【流暢さの評価】(参考)

- 5 : 文意が明解で、人間が書いた日本語文に近い。
- 4 : 日本語文として不自然な箇所を若干含むが、文意は明解である。
- 3 : 日本語文として不自然な箇所があり、文意がわかりにくい。
- 2 : 日本語文法規則に反する表現をかなり含む。文意がわからない。
- 1 : 日本語文として成立していない。

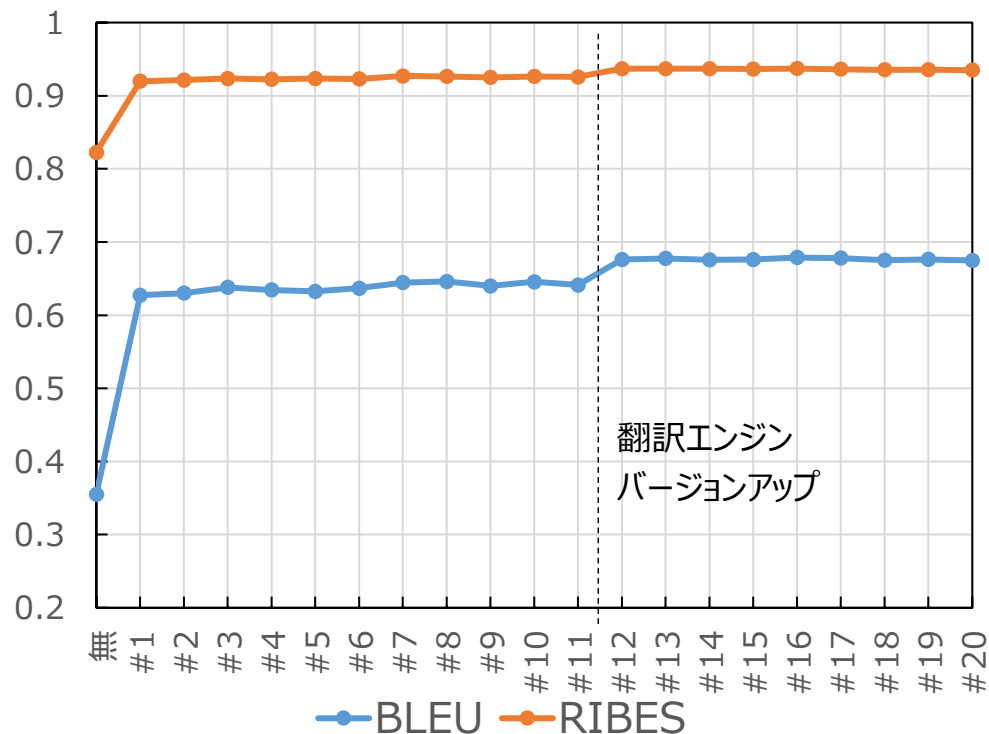
【重要技術用語の翻訳精度の評価】(参考)

- A (適訳語) : 人手翻訳に照らし、技術的に同義かつ一般的に用いられる訳語である。
- B (可訳語) : 技術用語として一般的に用いられる訳語ではないが、意味はおおむね正しい。
- C (誤訳語) : 誤訳である。
- D (不訳語) : 未知語、訳漏れである。

中日コーパス学習による機械翻訳精度（審決）の向上

- 対訳コーパスを10万文対ずつ段階的に学習させ、学習の進度に伴う翻訳精度の変化を調査。
（みんなの自動翻訳@Textra®（NICT）を使用）

1. 自動評価



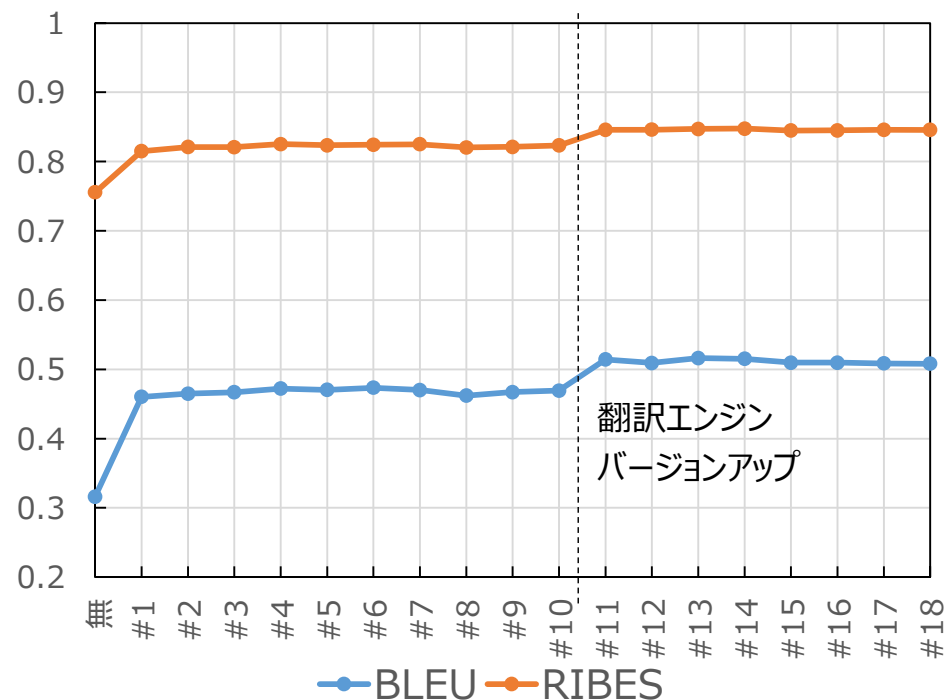
2. 人手評価

エンジン	無学習	第5回	第10回	第15回	第20回	改善度
汎用NT	2.41	4.42	4.32	4.61	4.65	+2.24
特許NT	2.18	4.28	4.30	4.66	4.61	+2.43

日英コーパス学習による機械翻訳精度（審査書類）の向上

- 対訳コーパスを10万文対ずつ段階的に学習させ、学習の進度に伴う翻訳精度の変化を調査。
（みんなの自動翻訳@Textra®（NICT）を使用）

1. 自動評価



2. 人手評価

	無学習	第5回	第10回	第15回	第18回	向上度
全種類	3.14	4.09	4.20	4.67	4.69	+1.55
拒絶査定	3.08	4.13	4.16	4.61	4.69	+1.61
拒絶理由通知書	3.44	4.28	4.45	4.85	4.86	+1.42
補正却下の決定	3.20	4.09	4.18	4.63	4.62	+1.42
前置報告書	2.84	3.85	4.01	4.58	4.59	+1.75

中日・日英コーパス学習によるエラー発生文数の減少

- 対訳コーパスを10万文対ずつ段階的に学習させ、典型的エラーの発生文数を集計。
(みんなの自動翻訳@Textra® (NICT) を使用)

1. 中日 (審決)

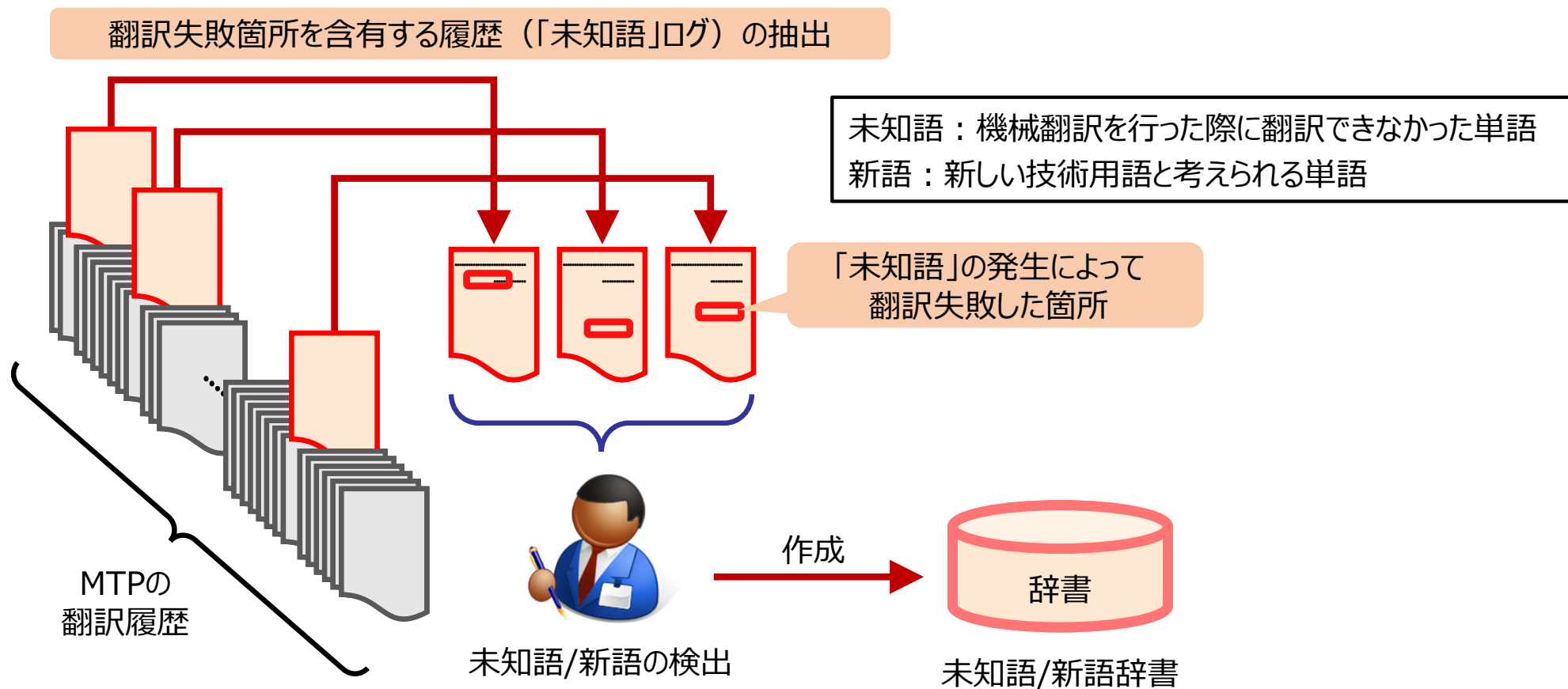
カテゴリ	無学習	第5回	第10回	第15回	第20回	1万増減
特許誤訳	90	21	13	16	16	-74
技術／一般誤訳	37	20	16	11	11	-26
未知語	2	1	1	1	0	-2
係り受けエラー	49	15	20	4	6	-43
日本語表現エラー	42	18	21	17	14	-28
湧き出しエラー	30	9	8	0	0	-30
欠落エラー	45	13	13	1	2	-43
数値エラー	22	12	10	5	5	-17
記号エラー	8	1	1	0	0	-8
その他エラー	4	2	1	0	0	-4

2. 日英 (審査書類)

カテゴリ	無学習	第5回	第10回	第15回	第18回	1万増減
特許誤訳	287	54	34	16	11	-276
技術／一般誤訳	108	55	56	20	25	-83
未知語	4	3	0	0	0	-4
係り受けエラー	84	70	73	42	39	-45
英語表現エラー	110	96	88	95	91	-19
湧き出しエラー	75	79	94	1	1	-74
欠落エラー	186	72	93	16	23	-163
数値エラー	32	33	23	5	5	-27
記号エラー	25	16	15	12	11	-14
その他エラー	13	9	16	1	1	-12

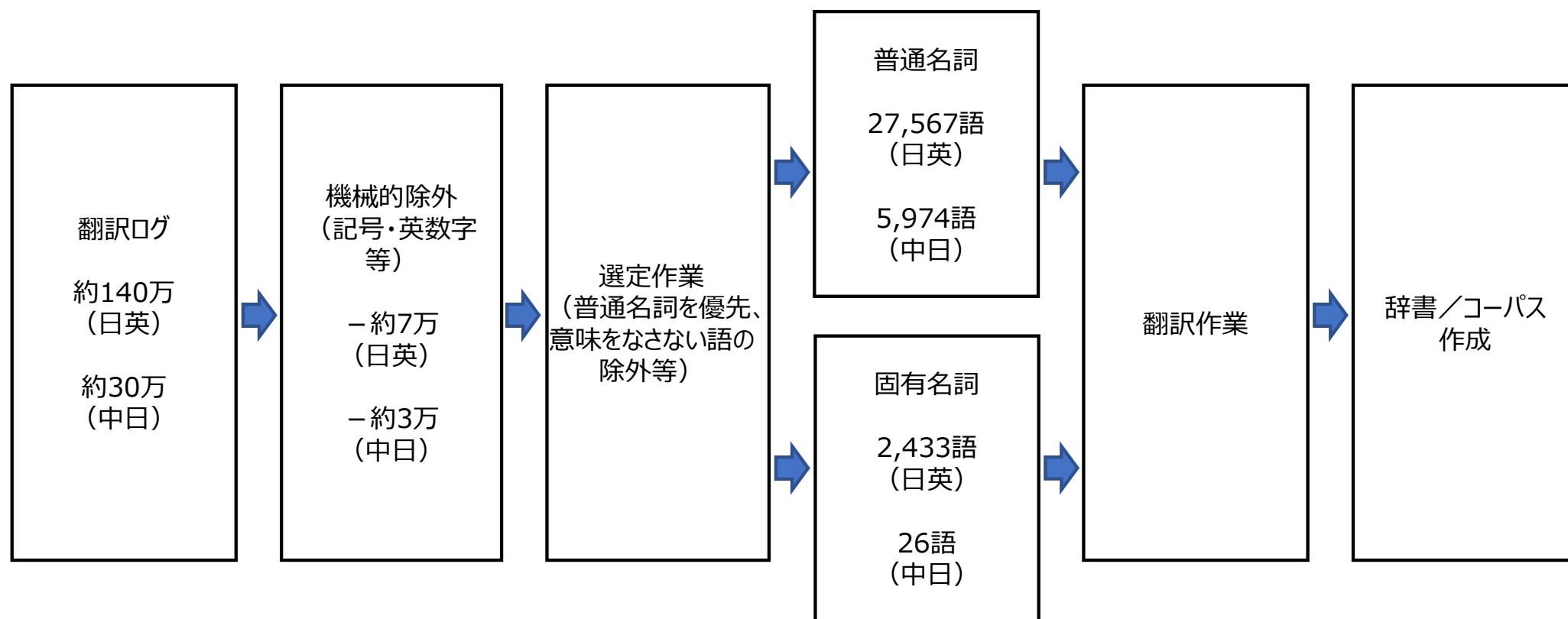
新語対応による機械翻訳精度向上に関する調査（令和1,3年度事業）

- MTPの日英翻訳／中日翻訳によって翻訳できなかった未知語／新語の辞書を作成すると共に、機械翻訳の品質維持向上のため、未知語／新語の技術分野別の出現傾向や、辞書登録すべき未知語の効率的な選定方法について調査。



新語対応による機械翻訳精度向上に関する調査（令和1,3年度事業）

- 翻訳対象語を以下の手順で選定し、辞書／コーパスを作成。
- 未知語（機械翻訳を行った際に翻訳できなかった単語）の中で新語（新しい技術用語と考えられる単語）の割合は低く、例えば文章分割の誤りで未知語が検出されてしまうケースなどが多い。



特許情報提供サービスに関する調査報告書（令和3年度事業）（1）

- 企業、特許事務所、及び特許情報を提供する事業者を対象としたアンケート調査。
- 海外の特許情報を何らかの手段で検索・照会する方々から、**利用頻度の高い公報情報の国・地域**を伺った回答結果は以下のとおりで、アメリカ、中国、欧州（EPO）の順。

回答内容	件数	割合
アメリカ	231	30%
欧州（EPO）	168	22%
WIPO	124	16%
中国	184	24%
韓国	28	4%
イギリス	0	0%
ドイツ	18	2%
フランス	0	0%
ロシア	0	0%
ASEAN	14	2%
インド	3	0%
ブラジル	0	0%
その他	4	1%
合計	774	100%

特許情報提供サービスに関する調査報告書（令和3年度事業）（2）

- 企業、特許事務所、及び特許情報を提供する事業者を対象としたアンケート調査。
- 海外の特許情報を何らかの手段で検索・照会する方々から、**利用頻度の高い公報情報の入手先**を伺った回答結果は以下のとおりで、国・公共機関のサービスとしては、J-PlatPatに次いでEspacenetが使われている。

回答内容	件数	割合
J-PlatPat	146	23%
Espacenet	124	20%
PATENTSCOPE	45	7%
その他海外特許庁の機械翻訳サービス	13	2%
WEB上の無料翻訳サービス	153	24%
民間翻訳業者	77	12%
その他	71	11%
合計	629	100%

特許情報提供サービスに関する調査報告書（令和3年度事業）（3）

- 企業、特許事務所、及び特許情報を提供する事業者を対象としたアンケート調査。
- 海外の特許情報を何らかの手段で検索・照会する方々から、**利用頻度の高い翻訳文入手機能**を伺った回答結果は以下のとおりで、全文機械翻訳文（中国・韓国）、和文抄録（アメリカ・EPO）の順。

回答内容	件数	割合
全文機械翻訳文（中国(CN)・韓国(KR)）の利用	119	34%
和文抄録（アメリカ(US)・EPO(EP)）の利用	92	27%
和文抄録（中国）の利用	75	22%
英語以外の公報について英文要約／抄録の利用	20	6%
パテントファミリーの参照	37	11%
その他	3	1%
合計	346	100%

1

特許情報提供の意義・目的

2

特許情報提供における機械翻訳の活用

3

機械翻訳に関する調査事業

4

特許庁が提供する機械翻訳文の利用方法

バルクダウンロードサービスでの中韓文献翻訳文の提供

- 特許情報のバルクダウンロードサービスは、利用者登録後に、中韓文献翻訳文を含むデータをダウンロードできる。

特許情報の一括ダウンロードサービスについて

令和3年3月1日

特許庁総務部総務課情報技術統括室

特許情報の一括ダウンロードサービスでは、特許情報標準データや諸外国の公報データなどをダウンロードいただけます。利用者登録が必要となりますので、利用を希望される方は、「2. 申込方法について」をご参照のうえ、「お問い合わせフォーム」にてお知らせください。

1. 提供データ一覧

[提供データ一覧 \(PDF : 84KB\)](#)

2. 申込方法について

以下の「お問い合わせフォーム」から、「特許情報の一括ダウンロードサービス利用希望」の旨を明記して、特許情報企画調査班へお知らせください。特許情報企画調査班から、利用登録に必要な書類の送付を含め、必要な手続きをご案内します。

[更新日 2021年3月1日]

お問い合わせ

総務課 情報技術統括室 特許情報企画調査班
TEL : 03-3581-1101 内線2361

[👉 お問い合わせフォーム](#)

特許情報プラットフォーム（J-PlatPat）での中韓公報の提供

- 中韓の特許公報・実用新案公報に対して日本語によるキーワード検索が可能。

特許情報プラットフォーム（J-PlatPat）

🔍 特許・実用新案検索 [ヘルプ](#)

書誌的事項・要約・請求の範囲のキーワード、分類(FI・Fターム、IPC)等から、特許・実用新案公報、外国文献、非特許文献を検索できます。
対象の文献種別や検索キーワードを入力してください。(検索のキーワード内は、スペース区切りでOR検索します。)
分類情報については、[特許・実用新案分類照会\(PMGS\)](#)を参照ください。

選択入力 論理式入力

テキスト検索対象
☒ 和文 ☐ 英文

文献種別 閉じる

☐ 国内文献 ☒ 外国文献 2 ☐ 非特許文献 ☐ J-GLOBAL

国内文献

- ☐ 特許(特開・特表(A)、再公表(A1)、特公・特許(B))
- ☐ 特許発明明細書(C)
- ☐ 実用新案(実開・実表・登実(U)、実全(U1)、再公表(A1)、実公・実登(Y))
- ☐ 登録実用新案明細書(Z)

外国文献

外国文献の指定は、全て、または 5カ国以内の指定が可能です。

☐ アメリカ(US)	☐ EPO(EP)	☐ WIPO(WO)	☒ 中国(CN)	☒ 韓国(KR)
☐ イギリス(GB)	☐ ドイツ(DE)	☐ フランス(FR)	☐ スイス(CH)	☐ カナダ(CA)

注釈：国・地域ごとに指定可能な検索項目については [コチラ](#) をご覧ください。

非特許文献

☐ 公開技報(N1)	☐ マニュアル	☐ 単行本	☐ 国内技術雑誌	☐ 非技術雑誌
☐ 外国学会論文	☐ 国内学会論文	☐ 企業技報	☐ 団体機関誌	☐ 予稿集

J-GLOBAL

☐ 文献	☐ 科学技術用語	☐ 化学物質	☐ 資料
-----------------------------	---------------------------------	-------------------------------	-----------------------------

中国(CN)／韓国(KR)
を選択し、日本語でテキスト検索

外国特許情報サービス（FOPISER）での中国審決の提供

- 中国の特許・実用新案に関する審決の和訳を提供（書誌情報での検索で照会）。

外国特許情報サービス（FOPISER）

特許庁 JAPAN PATENT OFFICE 審決番号索引照会(中国審決) [検索可能範囲](#) [ヘルプ](#) [メニューに戻る](#)

審決番号索引照会(中国審決)

出願・審決番号 ?
番号クリア

審決決定日 201101 ~
出願日 201101 ~

種別 ?
☒ 不服審判 ☒ 無効審判

決定の要旨
日本語で入力

発明の名称
日本語で入力

請求人
中国語で入力

被請求人
中国語で入力

根拠条文 (A=中国専利法、B=実施細則)
A-022-03

検索 ※部分一致検索 条件クリア

照会

入力番号フォーマット
■中国審決番号 制度説明

- 出願番号
 - nnnnnnnn.x
 - yyyynnnnnnnn.x
- 審決番号
 - nnnnn
 - nnnnnn

日本語テキスト等で
中国審決を検索

ありがとうございました



An Advanced Multilingual MT System: Facilitating Patent Examination and Search

Cao Jingcheng

China Patent Information Center, CNIPA

February 2023

CNPAT & MT History

For 34 years, CNPAT has been continuously committed to the development of patent information automation systems and related services including patent information processing, dissemination and consultation.

About CNPAT



PARTNERS



Development of patent information automation systems

- Information sharing platform for patents on epidemic prevention against COVID-19
- CPRS patent search platform
-



Patent information processing, dissemination, and consultation

- Multilingual machine translation system
- Patent data processing service
- Patent information retrieval and analysis service
-



Patent data management

26M+
Chinese Patent Full-text Data

32M+
Chinese Patent Full-text Image Data

5.5M+
Design Patent Image Data

8.5M+
Design Patent Graphic Data

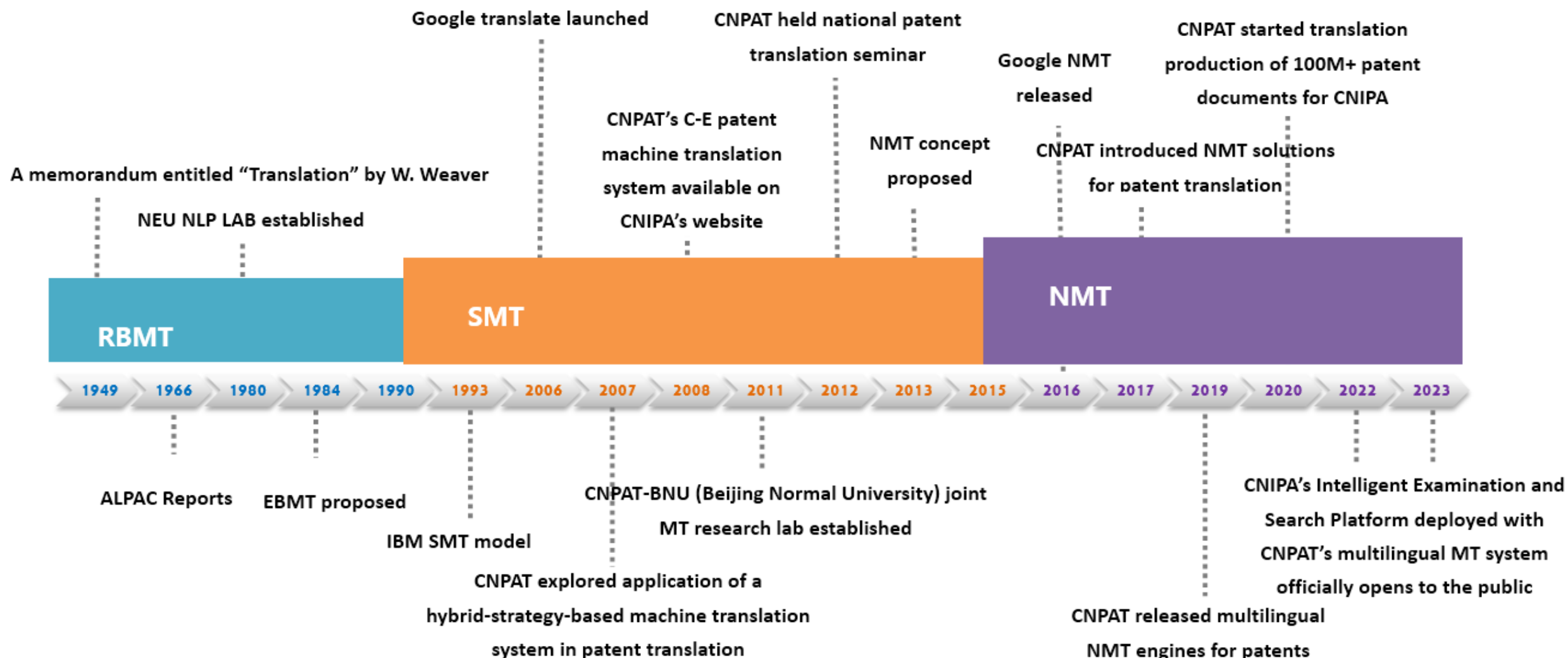
18M+
Patent Full-text coded Data

58M+
Patent Legal Status Data

120M+
Worldwide Patent Abstract Data

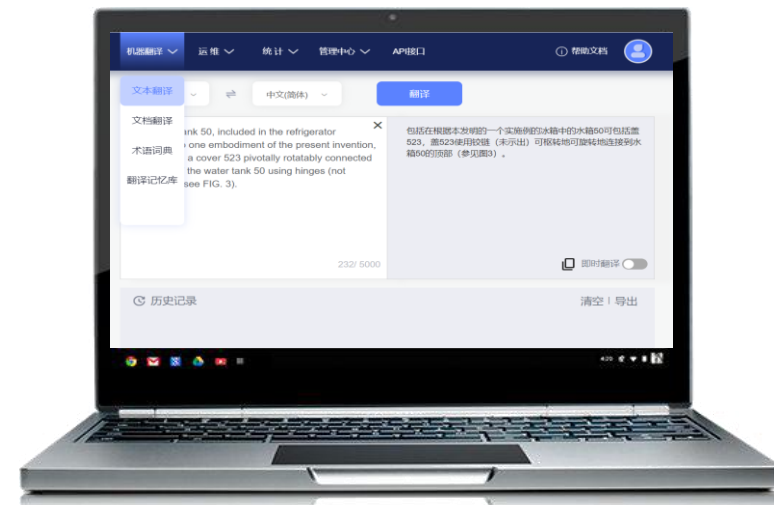
73M+
Worldwide Patent Image Data

CNPAT Machine Translation History



Multilingual MT system in application

MT application in
the new generation intelligent patent
examination and search system
of CNIPA



Multilingual MT system in application



7 languages

Chinese, English, Germany, French, Japanese, Korean, Russian



Multiple data formats

XML/Word/PPT/Excel/PDF



Patent and non-patent literature

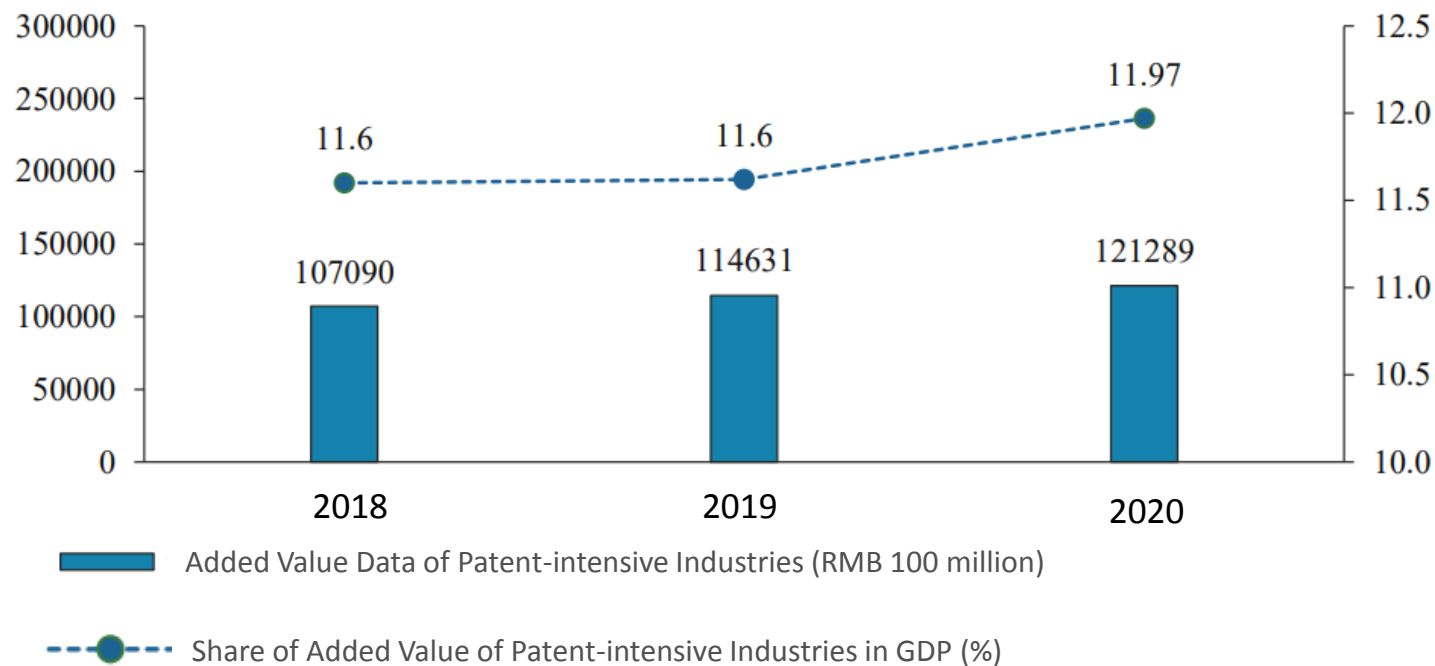
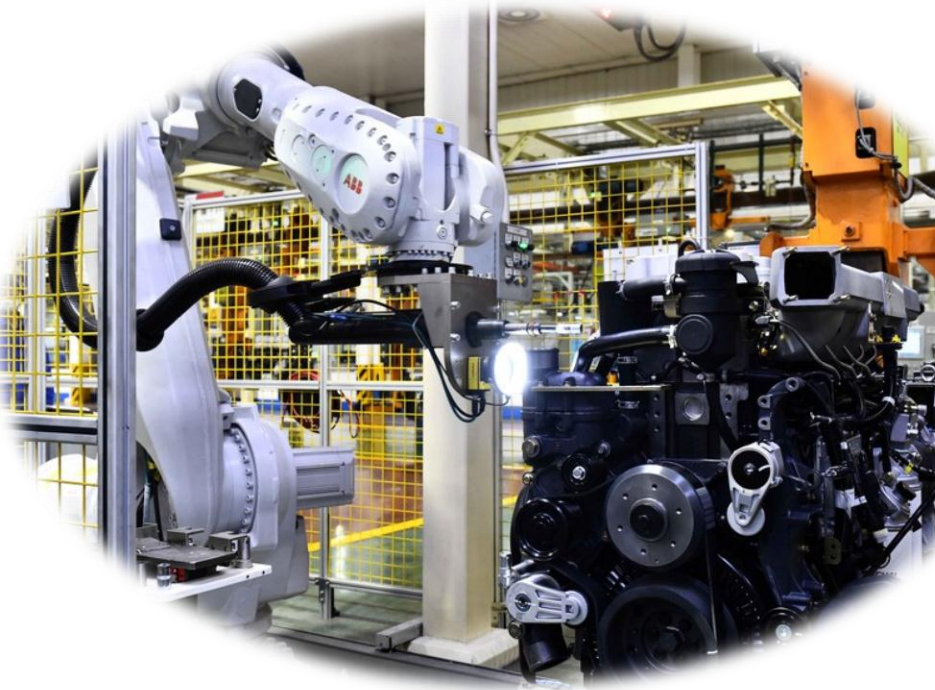
100 million+ from over 100 countries, regions and organizations

Clarifying Needs

Facing Pain Points

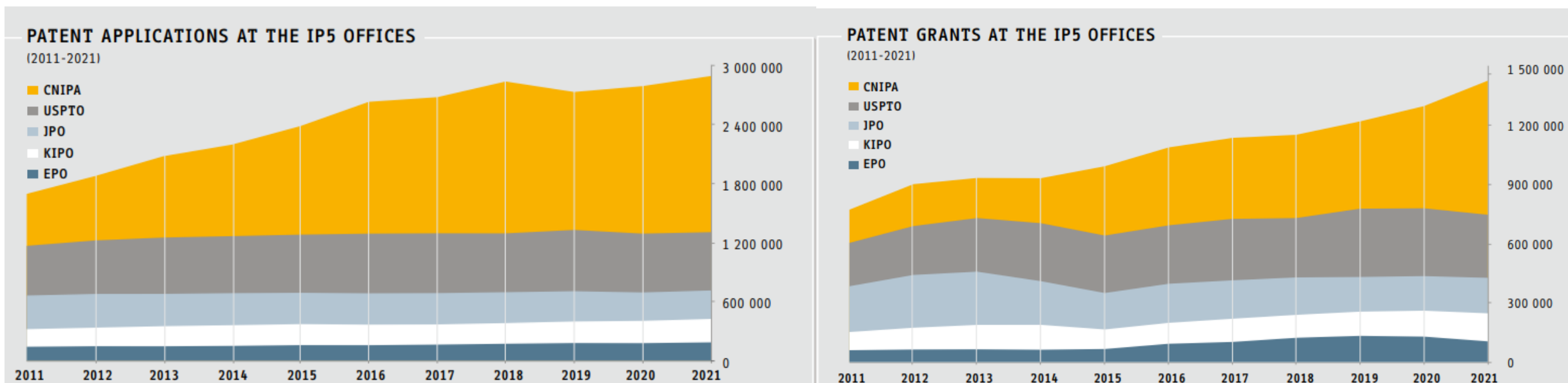
IP for innovation

2022年
中国知识产权发
展状况评价报告



(Source: 2022 China's IP Development Status Evaluation Report, released by CNIPA on December, 2022)

Urgent need for high-quality patent machine translation

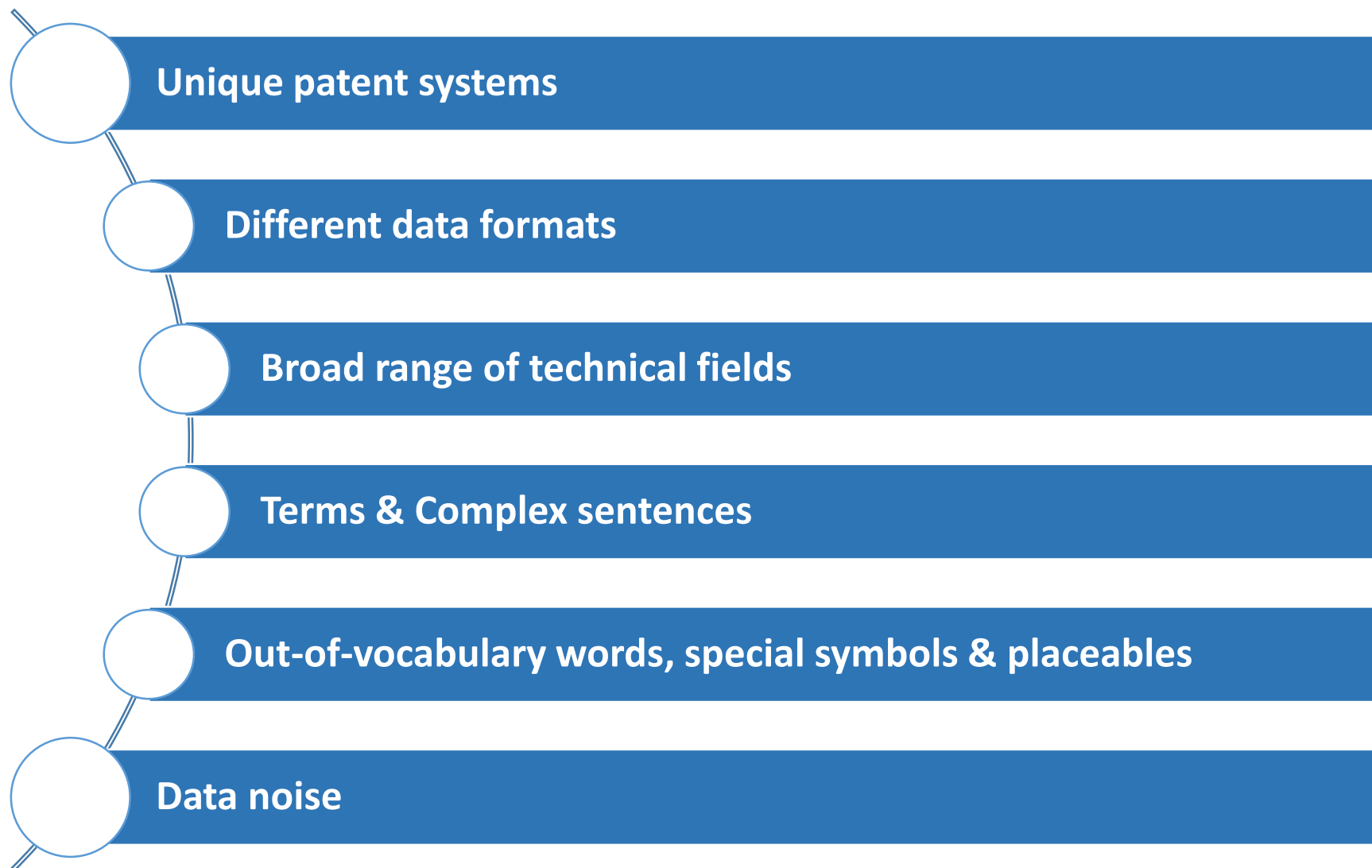


(Source: KEY IP5 STATISTICAL INDICATORS 2021)

Patents	2020	2021	Growth rate (%)	Share of world total (%)
Applications worldwide	3,281,900	3,401,100	3.6	100.0
China	1,497,159	1,585,663	5.9	46.6
US	597,172	591,473	-1.0	17.4
Japan	288,472	289,200	0.3	8.5

(Source: World Intellectual Property Indicators 2022)

Challenges in patent machine translation



Customizing MT System

Focusing on Translation Quality

01

AI technology

- NMT technology
- Advanced word segmentation technology
- Machine learning error correction model

02

Corpus resources

- Global patents
- Advanced corpus processing techniques

03

**30+ years
patent translation
experience**

- Linguistic analysis team
- Experienced patent translation and quality assessment team

04

**System
customization
service**

- Quick response
- Good extensibility

Our work in multilingual MT system construction

Corpus construction and optimization

- ✓ Obtaining high quality training data as a foundation of model training



Pre-processing/post-processing tuning

- ✓ Flexibly generating integrated solutions for specific batches of patent data



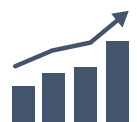
Model training and parameter adjustment

- ✓ Locating optimal model and parameters through iterative training



Cloud-based engine construction

- ✓ Achieving distributed processing capabilities and good extensibility



External translation memory

- ✓ Enhancing translation quality with multilingual MT glossaries and translation memories

About the multilingual MT system



Benefits and features

- Translation quality competitiveness for all language pairs
- Language scalability
- Real-time translation and batch translation
- Automatic language recognition
- Translation memory & glossaries
- Supporting application in multiple scenarios
- Supporting local and cloud deployment
- Confidentiality
- Work standalone and can also be fully integrated

Embracing Innovation

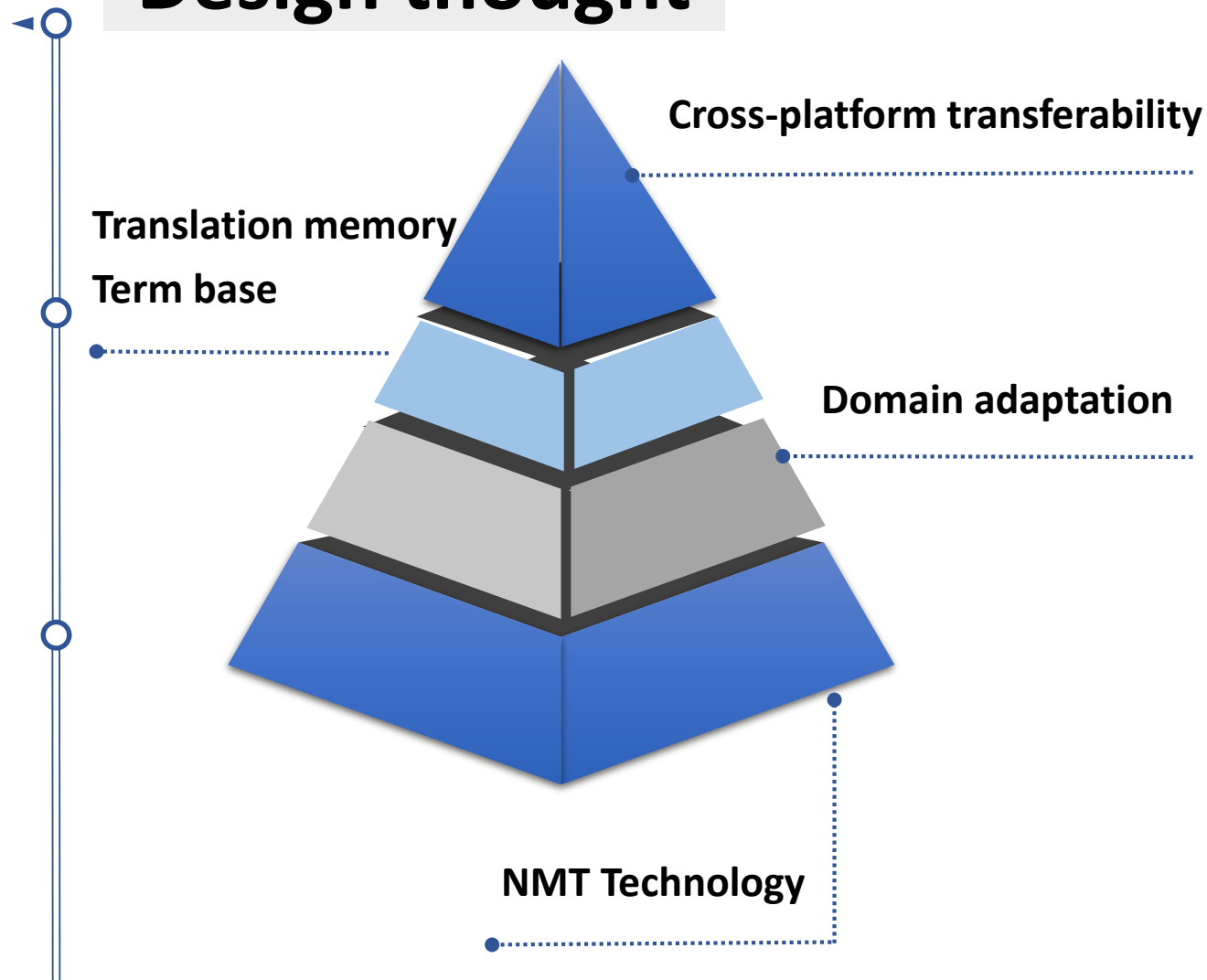
Always on The Road

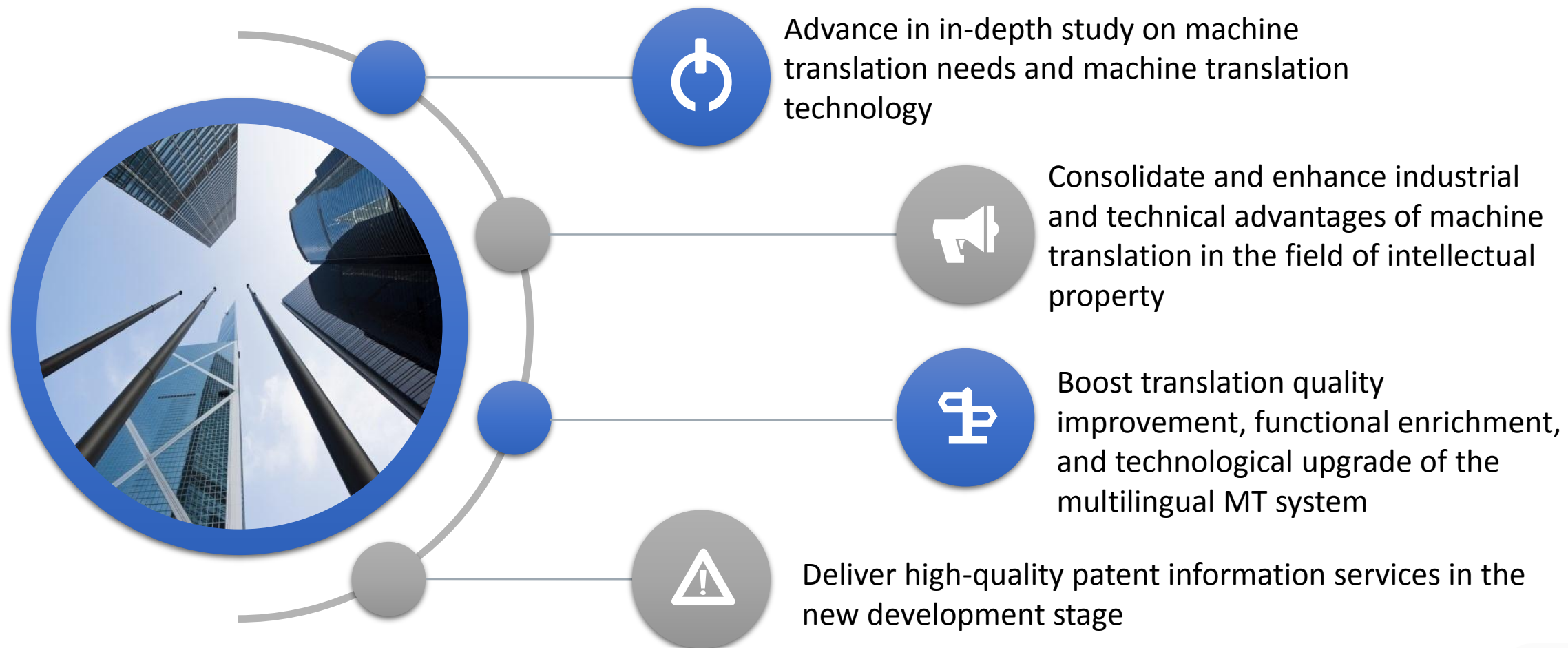
Goals

- To provide fast and accurate translation of multilingual patent documents on the premise of ensuring data security
- To expand the scope of patent retrieval and improving patent browsing efficiency for patent professionals and the public
- To offer translators high-quality machine translation output for quick post-editing



Design thought





THANK YOU!

Contact info: caojingcheng@cnpat.com.cn



守正 创新 专业 乐业

New Generation of AI Patent Machine Translation



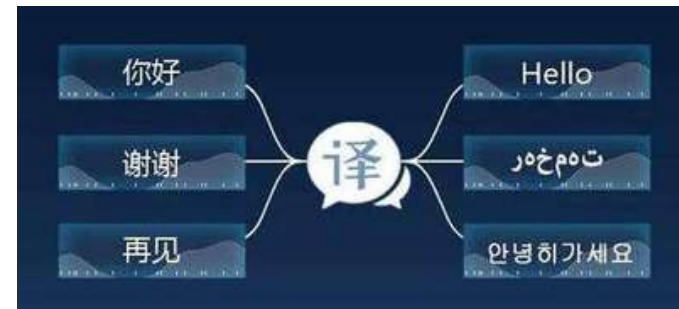
知识产权出版社 有限责任公司
Intellectual Property Publishing House Co., Ltd.



Machine Translation

Machine translation (MT) is an **information processing technology** that uses computer algorithms to automatically translate one natural language (such as English) into another (such as Chinese).

Machine translation is a typical **computer artificial intelligence technology** integrating computational science, mathematics, cognitive science, linguistics and others.



知识产权出版社
Intellectual Property Publishing House



R&D Purposes of Patent Machine Translation

1

**Facilitate patent translation,
search, and examination**

Different professions in the patent industry

2

**Address communication problems
caused by language difference**

Patents written in different languages

3

**Satisfy professional and
language needs**

Expertise in technical knowledge and language

4

Realize standardization

Unified and standardized format of patents

5

**Reduce costs and improve
efficiency**

A large number of patents

6

Promote industry development

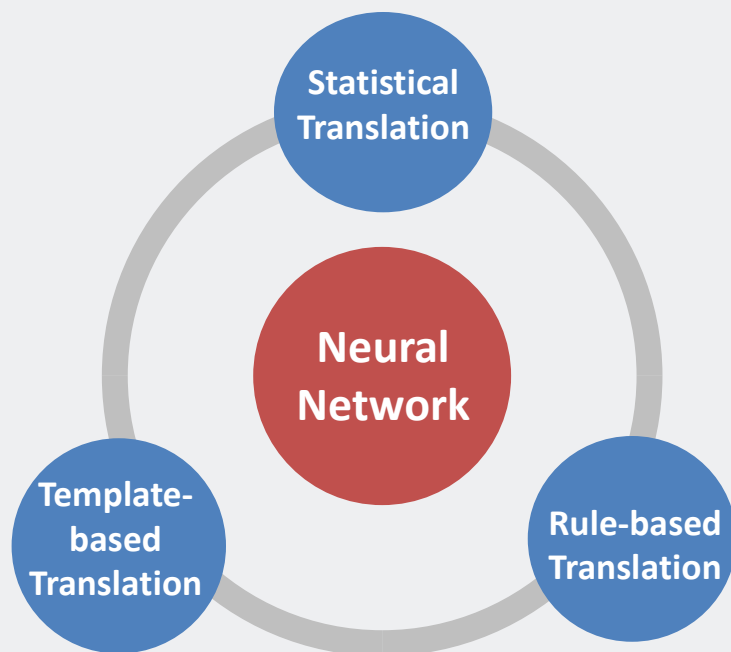
Rapid and healthy development of the patent industry



知识产权出版社
Intellectual Property Publishing House



Zhongzhi Huiyi (中知慧译) - Core Technology of Multi-strategy Neural Network Machine Translation



Neural Network + Statistical Translation

Deep Learning + Disambiguation

Template-based Translation

Analysis of Long and Difficult Sentences +
Sentence Pattern Consistency

Rule-based Translation

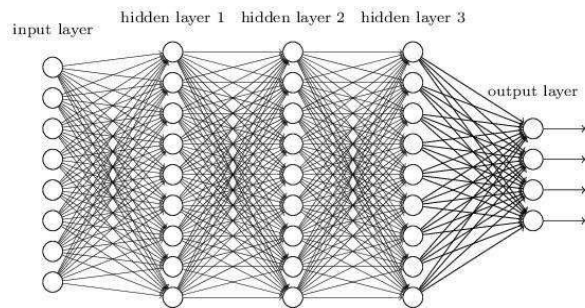
Syntax + Grammar Learning



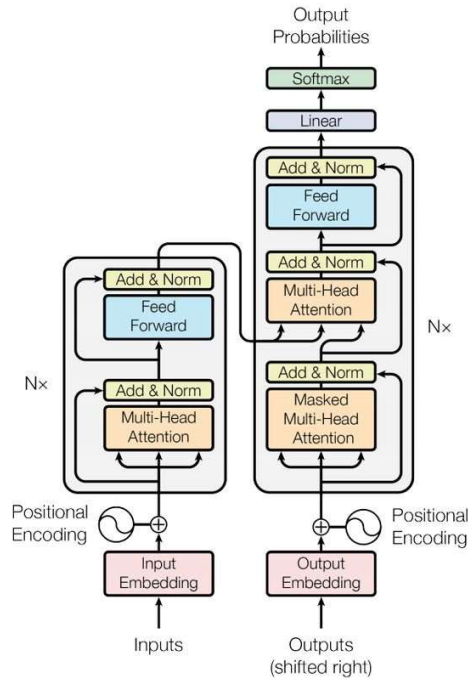
知识产权出版社
Intellectual Property Publishing House



Transformer(ANN) VS Brain Neural Network



Artificial Neural Network



Transformer



Brain Neural Network





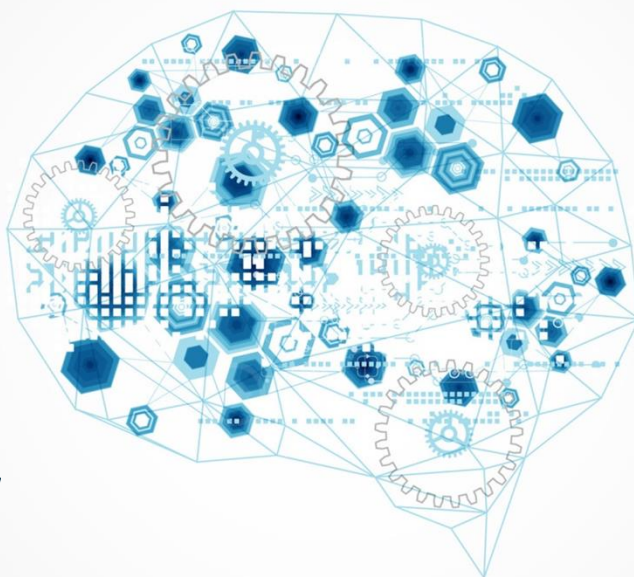
Advantages in Patent Data

01

Hundreds of thousands of patent translation templates

03

Tens of millions of high-quality bilingual patent sentences



02

Millions of patent terms

04

Pre-trained patent models based on **billions of** words



知识产权出版社
Intellectual Property Publishing House



Features of the Technology

1

Core technology

Multi-strategy neural network machine translation

3

Consistency

Terminological consistency of translation

5

Term searching

Searching terms in global patents



2

Quality and speed

High quality and speed of translation

4

Standardization

Normalization and consistency of sentence patterns

6

Corpus service

Mass data



知识产权出版社
Intellectual Property Publishing House



Challenges of Patent Machine Translation

1

Processing in **omission** to ensure information integrity

2

Processing in **long sentences, complex sentences** and specific patent sentence patterns

3

Processing in **terminological consistency**

4

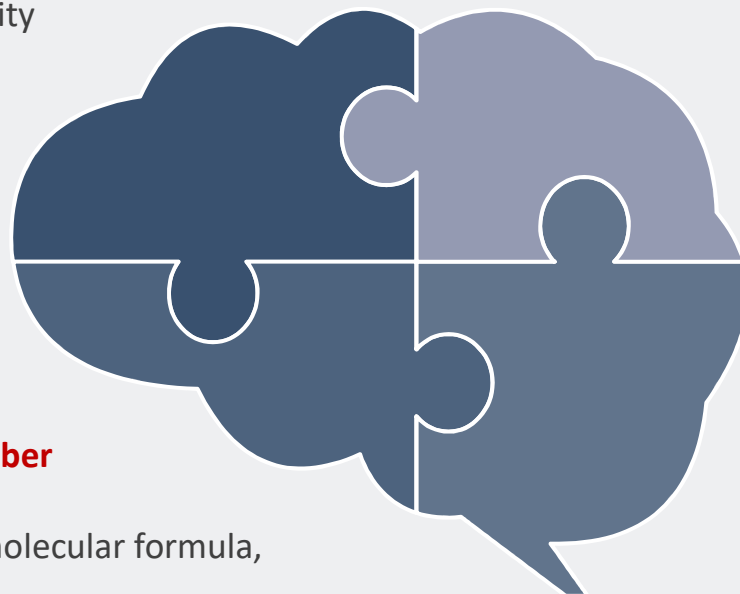
Processing in **claims** and other **special formats**

5

Processing in **number, time, name, and patent number**

6

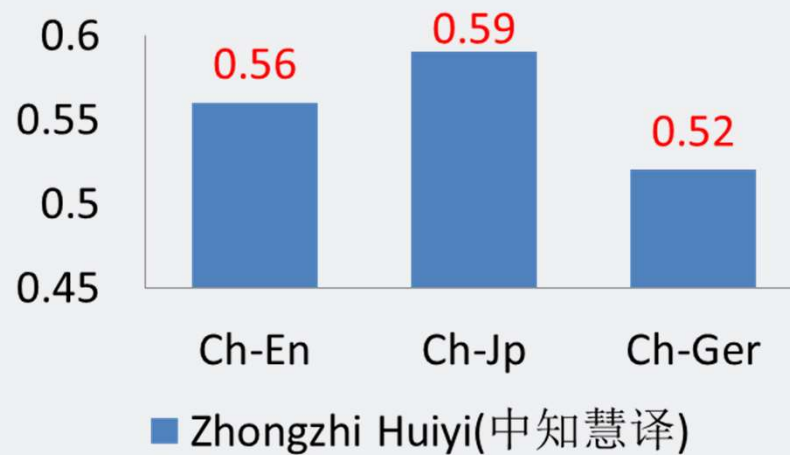
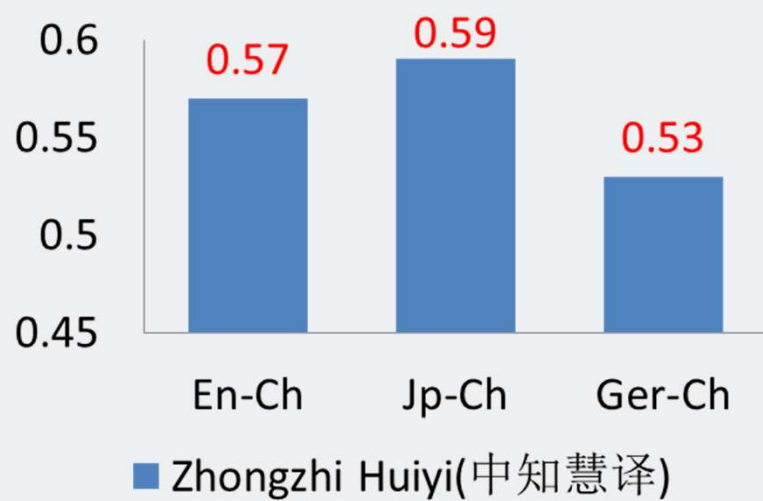
Processing in **special characters**, such as chemical molecular formula, superscript and subscript



知识产权出版社
Intellectual Property Publishing House

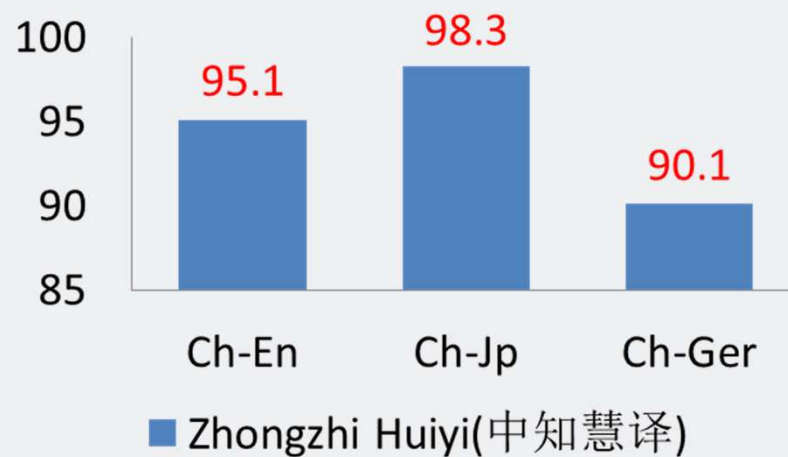
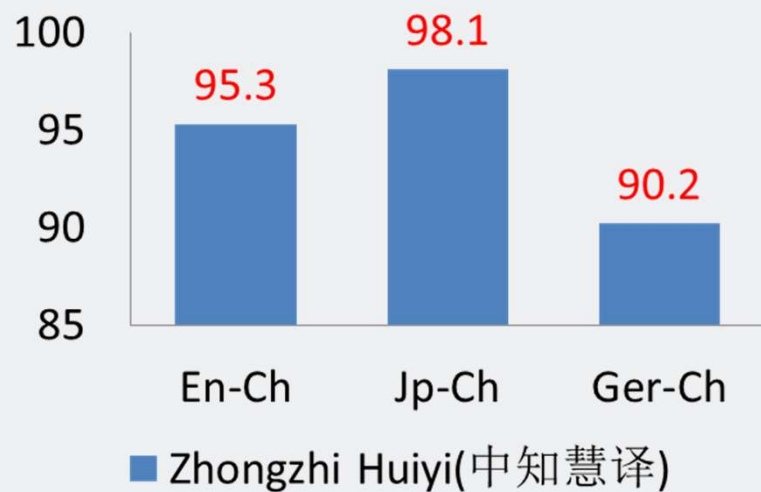


BLEU





Human Evaluation on Machine Translation





Example of NMT – English

In addition, according to a further aspect of the present invention , it is an object of the present invention to provide an improved method for supplying a network element with a charging policy to be enforced at said network element for charging of data reaching said network element of a communication network during a data session.

Machine Translation

此外，根据本发明的另一方面，本发明的目的是提供一种改进的方法，用于向网络单元提供要在所述网络单元处实施的用于对在数据会话期间到达通信网络的所述网络单元的数据进行计费的计费策略。

1. The logical structure of the long and complex sentence is processed correctly and conforms to the expression and language features of patents;
2. The translation is complete, accurate and professional.





Example of NMT – Japanese

ローヤルゼリーの水懸濁液を調製し、これに基質に対する作用部位の異なる二種類以上のプロテアーゼを同時又は逐次添加して室温以上の温度に保持し、酵素反応により不溶性物質を可溶化させることを特徴とする透明なローヤルゼリー溶液の製造法。

Machine Translation

一种透明蜂王浆溶液的制造方法，其特征在于，调制蜂王浆的水悬浮液，同时或逐次向其中添加对基质作用部位不同的二种以上的蛋白酶，保持在室温以上的温度，通过酶反应使不溶性物质可溶化。

1. The structure of the long sentence is accurately adjusted to ensure the language fluency of the translation;
2. The translation is complete, accurate and professional.





Example of NMT - German

Vorzugsweise ist die Mitte der Antriebsscheibe mit einem zweiten Rundloch versehen, wobei eine Seite der Antriebsscheibe mit einer ersten zu ihrer Achse parallelen Drehwelle versehen ist, und die erste Drehwelle nahe der äußeren Umfangswand der Antriebsscheibe angeordnet ist, wobei die erste Drehwelle dem ersten Rundloch entspricht.

Machine Translation

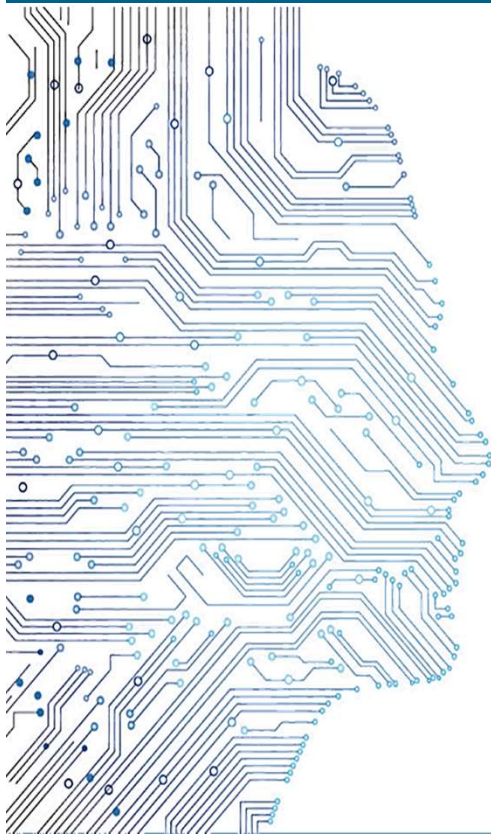
优选地，驱动盘的中心设置有第二圆孔，驱动盘的一侧设置有与其轴线平行的第一旋转轴线，并且第一旋转轴线设置在驱动盘的外周边缘附近，第一旋转轴线对应于第一圆孔。

1. The terms are accurately translated;
2. The translation is fluent, and clear in logical relations and sentence meaning.





Service for Patent Examiners



Real-time machine
translation service for
patent examiners to
facilitate their work

1

Number of Users

10,000+

2

Number of visits per day

50,000+

3

Number of words translated per day

10,000,000+



知识产权出版社
Intellectual Property Publishing House



Translation of Mass Data

Translation of Global Patent Data



01

Number of patents translated per day: **100,000**

02

Number of words translated per day: **1 billion**

03

Number of languages: **10**



知识产权出版社
Intellectual Property Publishing House



Our Partners

01

Online translation
system has more than **220,000**
registered users

02

Signed cooperation
agreements with more than **30**
companies and institutions



中咨律师事务所
ZHONGZI LAW OFFICE



中科專利商標
代理有限公司
CHINA SCIENCE PATENT & TRADEMARK AGENT LTD



商汤
sensetime

Questel

IBM

PetroTECH
中油太科



知识产权出版社
Intellectual Property Publishing House



Official Platforms

Official Website: www.iptranslation.net



Wechat Official Account



Thank You!



知识产权出版社 有限责任公司
Intellectual Property Publishing House Co., Ltd.

技術翻訳のための 機械翻訳のカスタマイズ手法

2023年2月20日
(株) クロスランゲージ
開発部
高橋 博之

カスタマイズの必要性

近年、NMT（ニューラル機械翻訳）により機械翻訳の訳質が飛躍的に向上している。

しかし、技術翻訳に使用するにはまだ問題がある。

- ✕ 専門用語を正しく訳せない。
- ✕ 文体や言い回しが不自然。



翻訳したい文書に合わせた機械翻訳エンジンのカスタマイズが必要

カスタマイズの手法

- 辞書登録

- ⚠ 名詞のみ

- ⚠ 常にその訳が使用される（文脈による訳し分けができない）
（人名、商品名などの固有名詞には有効な方法）

- 対訳文学習

（NMTは対訳文からの学習で「**翻訳モデル**」を作り、それを用いて翻訳）
翻訳したい分野の対訳文を用意して追加学習させることで翻訳モデルを強化

- ✓ 単語の訳だけでなく、文体や言い回しも学習

- ✓ 文脈に応じた訳し分けが可能

- ⚠ 学習用の対訳文が必要

NMTは数千万～数億文レベルの対訳文学習で作られるが、
カスタマイズには**数万文程度で十分**



「アダプテーション」
「ファインチューニング」
とも呼ばれる技術

対訳文学習によるカスタマイズの実証実験

以下のサービスについて調査

- Google AutoML Translation
- Microsoft Custom Translator

両社とも一般向け機械翻訳サービスを提供している。

一般向け機械翻訳をベースに対訳文学習によるカスタマイズが可能。

クラウドコンピューティングサービス上で実行。

(Google Cloud Platform/Microsoft Azure)

カスタマイズ手順

両社ともクラウド上で作業可能でWeb ベースのUIがある。



Step1 学習用対訳文用意

- 1文ずつの対訳にする
- テキスト形式への変換



Step4 翻訳モデル学習

- 学習は完全に自動化されている



Step2 クラウド環境設定

- アカウント作成等



Step5 自動評価(BLEU)

- 学習前、学習後の翻訳結果のBLEU値が算出される。



Step3 対訳文をサーバにアップロード



Step6 翻訳

- 学習後のカスタム翻訳モデルをサーバにデプロイして、APIを使って翻訳

画面イメージ(MS)：データセット

Cognitive Services | Language Studio | Custom Translator Preview

<<

Workspaces

cross-dev

med-ej-30k

Manage documents

Train model

Model details

Test model

Publish model

Workspace settings

Manage documents

Documents are stored inside your workspace and are used when training models. All documents in the relevant languages are automatically available for use within a project. [Learn more about documents](#)

+ Add document set

Download

Delete

Name	Created on	Type	English sentences	Japanese sentences	Created by
long_train_30k_je	06/09/2022	Training	30,000	30,000	dev cross
long_test_je	06/09/2022	Testing	2,000	2,000	dev cross
long_dev_je	06/09/2022	Tuning	2,000	2,000	dev cross
long_train_30k_ej	06/10/2022	Training	30,000	30,000	dev cross
long_test_ej	06/10/2022	Testing	2,000	2,000	dev cross
long_dev_ej	06/10/2022	Tuning	2,000	2,000	dev cross

画面イメージ(MS)：学習（トレーニング）実行

Cognitive Services | Language Studio | Custom Translator Preview

<<

Workspaces

cross-dev

med-ej-30k

Manage documents

Train model

Model details

Test model

Publish model

Workspace settings

Model details > med_30k_ej

View test Unpublish Update regions Duplicate model

Training succeeded

PUBLISHED

BLEU score

▲ 44.29
(14.84)

Baseline BLEU: 29.45

Training succeeded on: 6/10/2022 13:07PM

Total training time: 4 hour(s) 56 minute(s)

Training: 29,971 Tuning: 1,998 Testing: 2,000 Dictionary: 0

System generated test set: Yes ⓘ

Category ID:

[Learn how to evaluate a model's BLEU score](#)

Document sets used for training

Name	Created on	Type	English sentences	Japanese sentences	Aligned sentences	Used sentences
long_train_30k_ej	06/10/2022	Training	30,000	30,000	30,000	29,971
long_test_ej	06/10/2022	Testing	2,000	2,000	2,000	2,000
long_dev_ej	06/10/2022	Tuning	2,000	2,000	2,000	1,998

画面イメージ(MS)：翻訳結果

Cognitive Services | Language Studio | Custom Translator Preview

Q

<<

Workspaces

cross-dev

med-ej-30k

Manage documents

Train model

Model details

Test model

Publish model

Workspace settings

Test model > med_30k_ej

Download results

Showing 1-20 of 2000 results

	English	Japanese
1	Magnetic resonance imaging (MRI) scans of the thorax and upper abdomen do not appear to yield advantages over CT scans.	Reference: 胸部および上腹部の磁気共鳴画像法（MRI）スキャンは、CTスキャンよりも優れていないようである。 New model: 胸郭および腹部上部の磁気共鳴画像法(MRI)スキャンは、CTスキャンよりも優れているとは思われない。 Baseline model: 胸郭と上腹部の磁気共鳴画像(MRI)スキャンは、CTスキャンよりも優位性を生み出さないようです。
2	Be aware that marrow plasma cell distribution may vary in different sites.	Reference: 骨髓形質細胞の分布は、別の部位においては異なる場合があることに注意する。 New model: 骨髓形質細胞の分布は部位によって異なる可能性があることに注意すること。 Baseline model: 骨髓の血漿細胞の分布は、異なる部位で異なる可能性があることに注意してください。
3	Of 1057 patients enrolled, 298 had a history of sP_ carinii pneumonia.	Reference: 登録した患者 1,057 人中 298 人にニューモシスチスカリニ肺炎の既往を認めた。 New model: 登録した患者1,057人中、298人がカリニ肺炎ofsP_病歴があった。 Baseline model: 登録された1057人の患者のうち、298人がカリニ肺炎ofsP_歴史を持っていた。

対訳文数、時間、費用

- 対訳文数:
 - Google: 1000文以上(5000文以上が望ましい)
 - MS: 10,000文以上
 - 今回は **3 万文** 使用
- 時間: 3万文学習で **3～5時間** 程度
- 費用: 3万文学習で **\$70～\$130** 程度
- 学習結果（カスタム翻訳モデル）を用いた場合の翻訳料金:
 - ⚠ 通常翻訳よりも割高になる
 - Google: **\$80/100 万文字**（通常: \$20/100 万文字）
 - MS: **\$40/100 万文字**（通常: \$10/100 万文字）

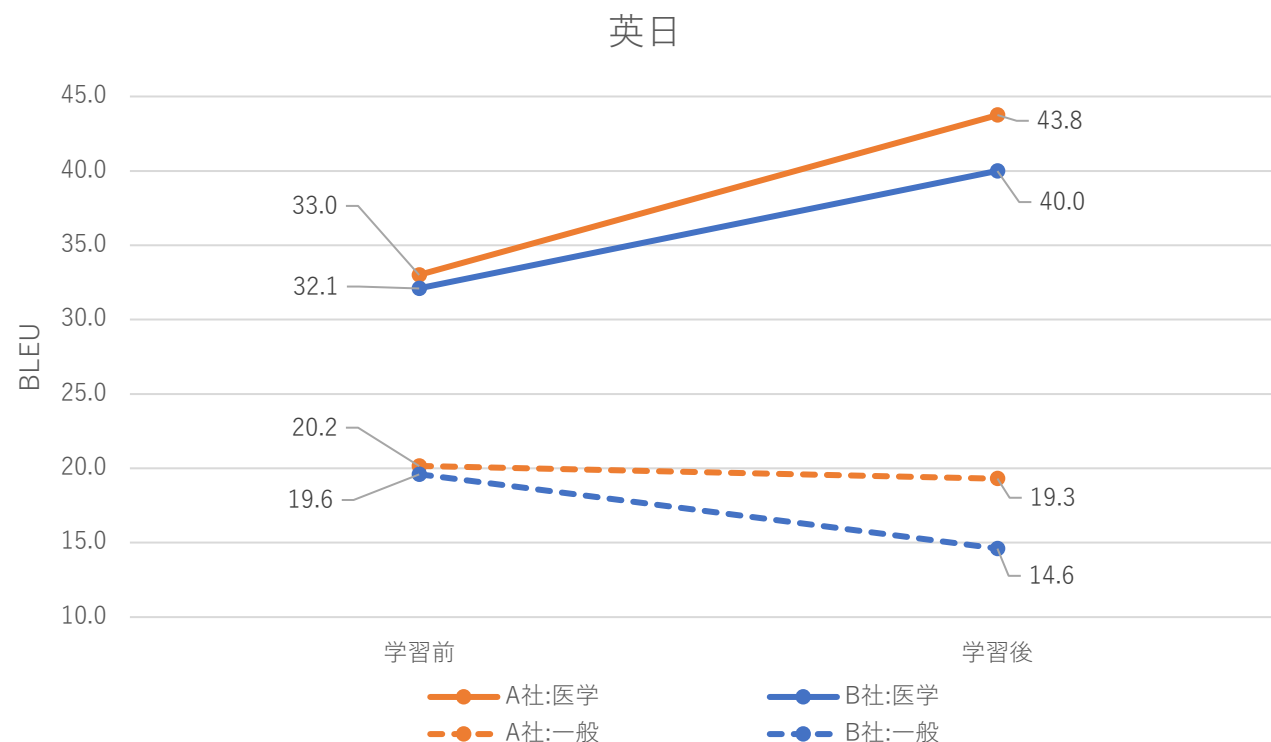
※2022年6月時点

医学論文を用いた学習（英日）

医学論文 3 万文を用いて英日・日英で学習実施
医学だけでなく一般分野の訳質の変化も見る

英日	評価文	学習前	学習後
A社	医学	33.0	43.8
	一般文	20.2	19.3
B社	医学	32.1	40.0
	一般文	19.6	14.6

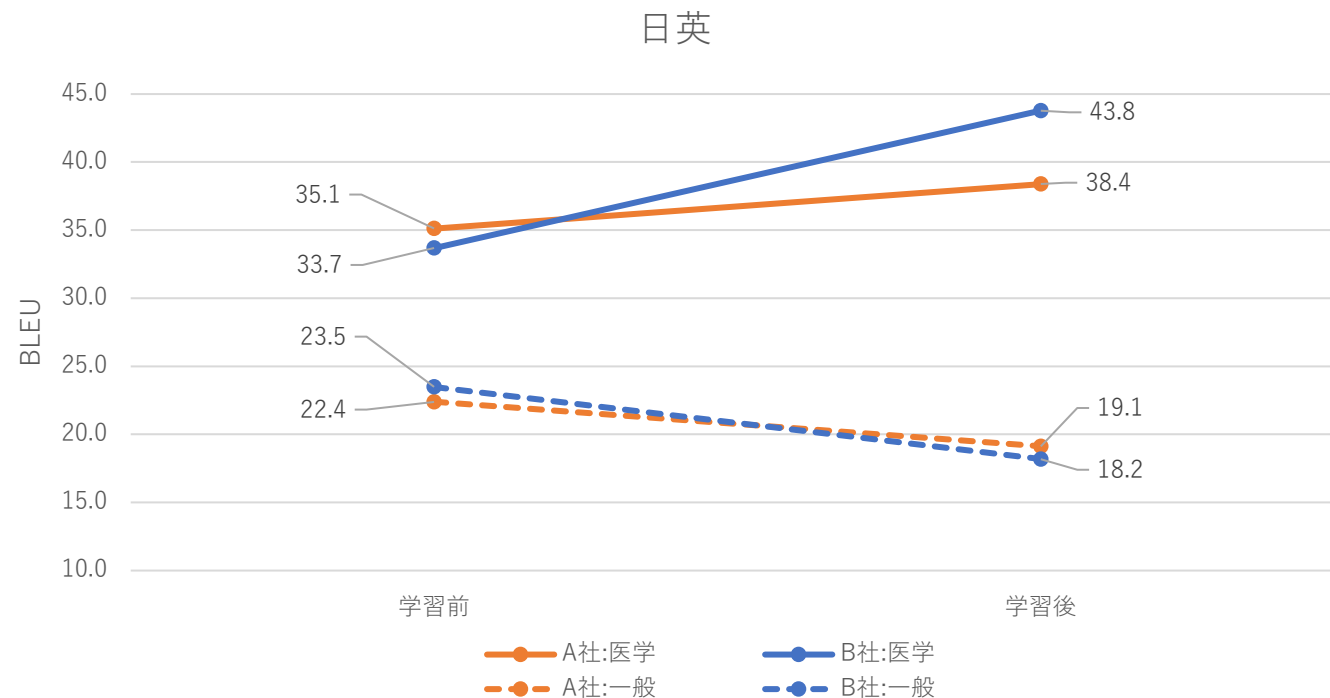
（数値は BLEU）



医学論文を用いた学習（日英）

日英	評価文	学習前	学習後
A社	医学	35.11	38.38
	一般文	22.39	19.12
B社	医学	33.68	43.77
	一般文	23.48	18.18

（数値は BLEU）



BLEU

- BLEU は機械翻訳結果の自動評価手法の一つ
- 参照訳（人間が作った正しい訳）と機械翻訳結果がどのくらい似ているかで評価する。

BLEU スコア	解釈
0～10	ほとんど役に立たない
10～20	主旨を理解するのが困難である
20～30	主旨は明白であるが、文法上の重大なエラーがある
30～40	理解できる、適度な品質の翻訳
40～50	高品質な翻訳
50～60	非常に高品質で、適切かつ流暢な翻訳
60～100	人が翻訳した場合よりも高品質であることが多い



カスタマイズにより
BLEU30台（理解できる）
から40台（高品質）に！

（<https://cloud.google.com/translate/automl/docs/evaluate?hl=ja> より）

改善例

医学用語が正しく訳せるようになる。

原文	IV期の口唇癌および口腔癌
学習前	Stage IV lip and oral cancer ⚠
学習後	Stage IV Lip and Oral Cavity Cancer ✓

医学論文では患者を「人」ではなく「例」で数える。

原文	Thirty patients (81%) had healthcare-associated infections; 18 (60%) of these 30 were infected with isolates carrying markers of CA-MRSA strains.
学習前	30人の患者 (81%) が医療関連の感染症を患っていました。これら30のうち18 (60%) は、CA-MRSA株のマーカを保有する分離株に感染していた。⚠
学習後	患者30例 (81%) が医療関連感染症であった；この30例中18例 (60%) が、CA-MRSA株のマーカを保有する分離株に感染していた。✓

悪化例

• 分野外の文の訳質は悪化することがある

用語選択がその分野寄りになる

原文	To ensure that the noodles will go perfectly with the soup, their thickness and length are carefully calibrated .
学習前	麺がスープと完璧に合うように、その太さと長さは慎重に 調整 されています。
学習後	スープの厚みと長さを注意深く キャリブレーション して、完璧なスープを作り上げます。

対象分野外の表現が訳せなくなる

原文	She's been putting in more hours at the office recently.
学習前	彼女は最近、オフィスでより 多くの時間を費やして います。
学習後	彼女は最近オフィスに もっと時間を置いて いる。

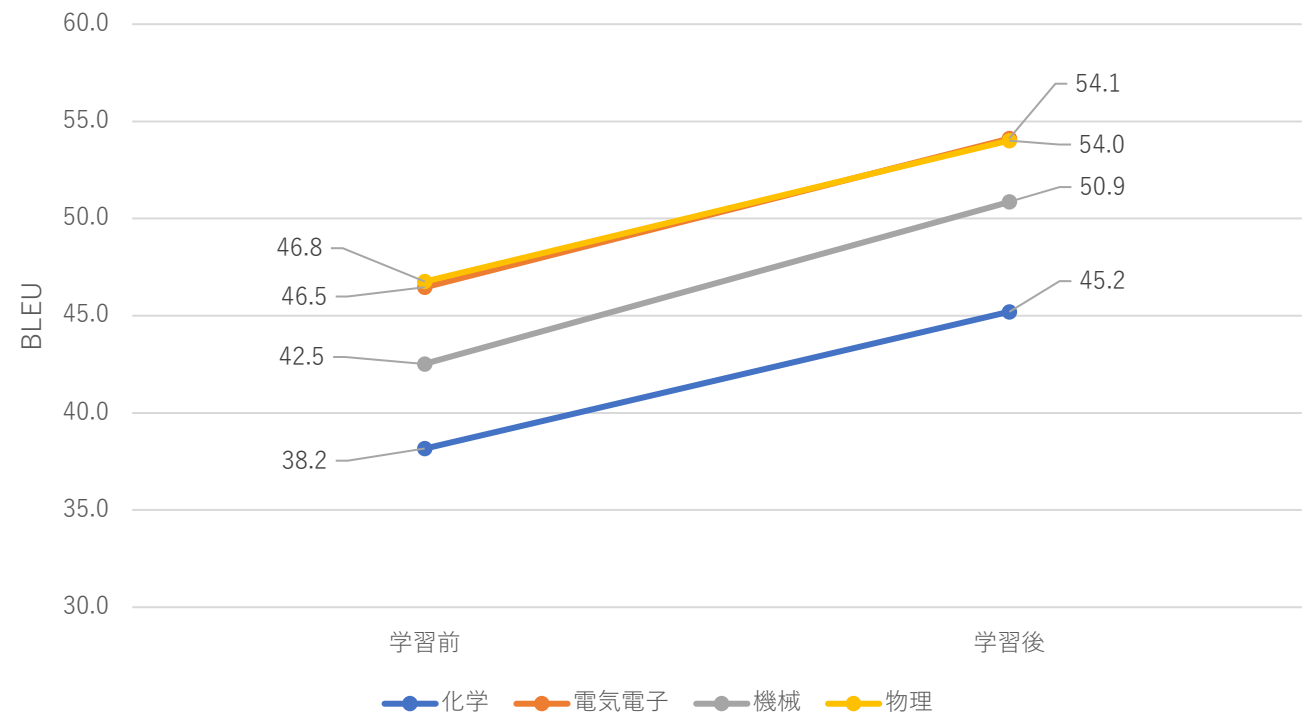
医学以外の分野での実証

- 各種の専門的な文のソースとして特許公報を利用（日→中翻訳）
- 化学、電気電子、機械、物理の4分野でそれぞれ10万文学習

分野	学習前	学習後
化学	38.2	45.2
電気電子	46.5	54.1
機械	42.5	50.9
物理	46.8	54.0

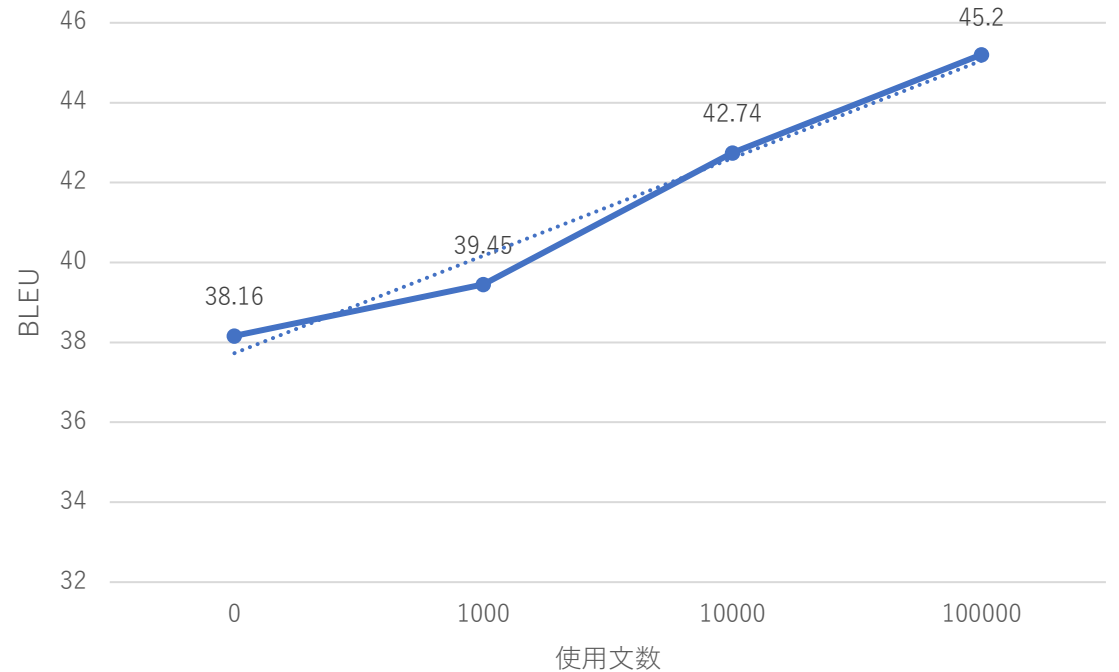
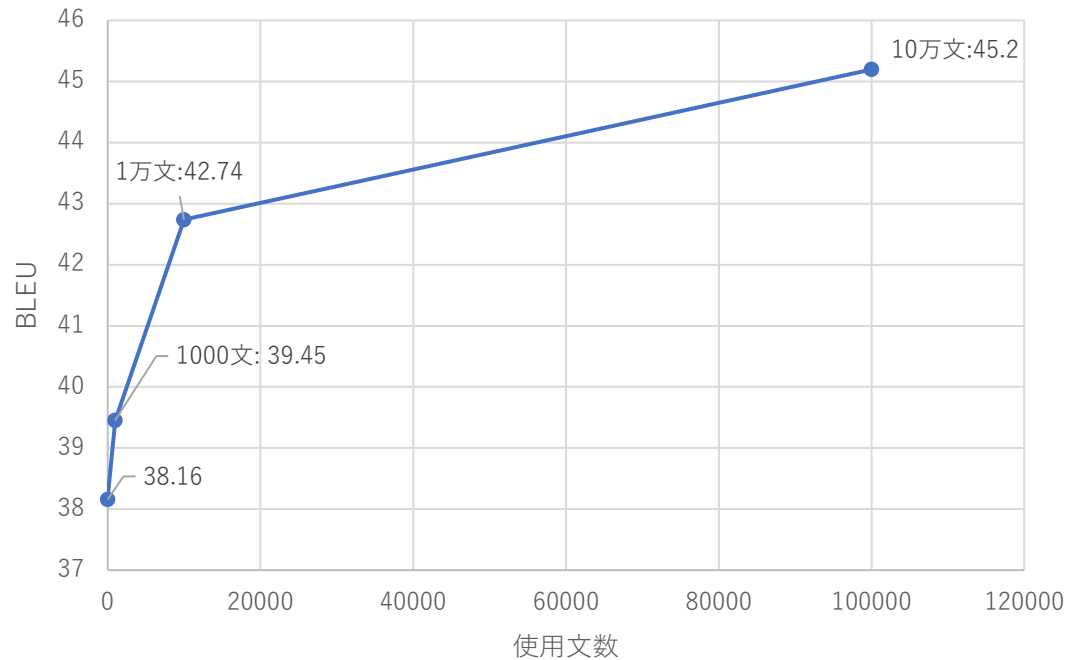
（数値は BLEU）

各分野とも +7~8 程度の向上



学習に使用する対訳文の数と訳質の関係

- 化学分野で学習に使用する文数を1000、1万、10万と変化させた



対訳文が多いほど訳質は向上するが、伸びは鈍化する。

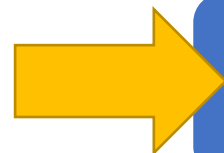
適切な分野の学習文を使うことの効果

分野分けされていない文で学習するとどうなるか？

特許文 10 万文で学習した一つのモデルで 4 分野のテスト文を翻訳

分野	学習前	学習後 (分野分けなし学習)	学習後 (分野別学習)
化学	38.2	43.8	45.2
電気電子	46.5	53.4	54.1
機械	42.5	49.8	50.9
物理	46.8	53.4	54.0

分野別に学習すると+7~8 程度向上するのに比べ、
分野分け無しだと+6~7程度の向上に留まる



学習用データは翻訳対象と同じ分野の文を使うとより効果的

まとめ

- NMT は対訳文を用いた学習によりカスタマイズできる。
 - 翻訳したい文と同分野の学習文を用意する必要がある。
 - 1 万文程度の学習文で一定の効果が得られる。
 - 学習後はその分野に特化した翻訳モデルとなり、対象以外の分野の訳質は悪化することがある。

特許情報の進歩的な活用方法

講師：正林 真之

（正林国際特許商標事務所
所長 弁理士）

2023年2月20日（月）
14：50～15：15

正林 真之

Shobayashi Masayuki

正林国際特許商標事務所 所長・弁理士

- 日本弁理士会 2020年度副会長
- 東京大学先端科学技術研究センター 知的財産分野客員研究員
- 知的財産管理技能検定 指定試験期間技能検定委員
- 元東京理科大学 専門職大学院 客員教授

〒100-0005

東京都千代田区丸の内1-7-12 サピアタワー

TEL 03-6895-4500

e-mail: shobayashi@sho-pat.com

URL: www.sho-pat.com



IP情報を見える化して活用する

IPランドスケープとは？

- IPは使わないともったいない
- 情報はみんなで作らないともったいない
- IP情報はみんなで作えないともったいない

① **IP(特許)情報は企業の技術的な本音が反映されたビッグデータ**

特許情報はうそをつかない

∴ 出願するのは結構お金がかかる

② **分類やテーマ等がよく整理されているので分析しやすい**

∴ 審査や調査のし易さのために緻密に分類分けがなされている

③ **企業・技術・商品開発の動向（時系列情報）が分かる**

∴ 企業の研究開発や商品開発の成果・将来の動向が表れている

(参考) 特許出願は、日本だけでも、毎年30～40万件
外国も合わせると毎年数百万件

**「特許情報」分析の結果は、
企業の経営、事業、研究開発等の戦略策定に活用できる**

従来型の「特許調査」

自社の弱み回避の「消極型」

目的＝「特許出願」戦略立案
Report to 知財部門・R&D

特許情報

特許性調査、侵害抵触調査

(R&Dテーマ探索、
知財ポートフォリオ構築)

IPランドスケープ

自社の強みを活かす「積極型」

目的＝「事業」戦略立案
Report to 経営企画・事業企画

特許情報×マーケット情報

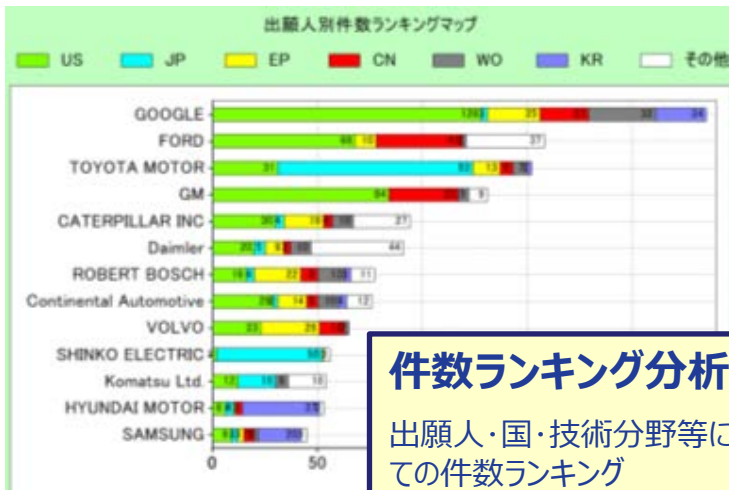
新規用途開発調査、アライアンス先候補・顧客候補の探索調査、
M&A・投資対象候補調査

※金沢工業大学大学院 杉光教授の講座より

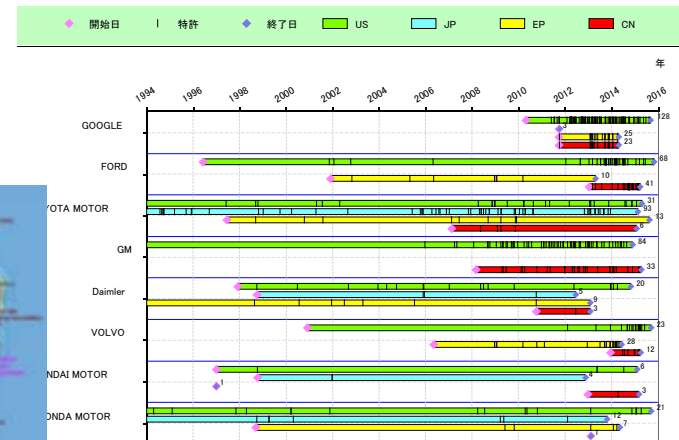
当所の考えるIPランドスケープとは

『IP情報を軸として事業、製品、サービス、財務、その他の非IP情報を総合的に加味しながら、対象企業の現在置かれている**位置づけ**とその**未来を提言し、経営に役立てるもの**』

経営層に対して、IP情報 & 非IP情報を
わかりやすく図示し、
真実を伝える、
意思決定を促す道具



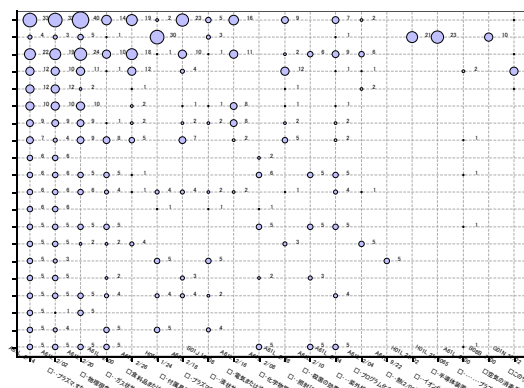
技術俯瞰マップ



参入分析

新規参入時期や継続期間を把握し、主要出願人の参画実態を把握

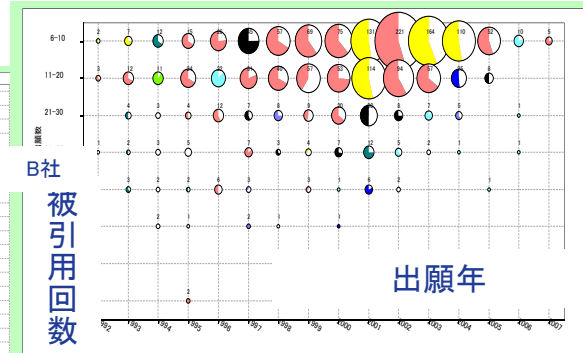
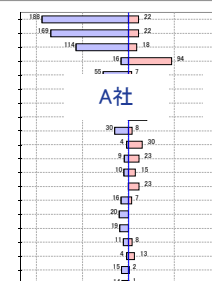
出願人



技術分類

技術分類

HOIL 51/1
H05B 33/14
H05B 33/22
HOIL 51/00
H05B 33/10
HOIL 27/32
HOIL 27/00
H05B 33/26
G01C 21/1
G01C 21/1
HOIL 51/30
HOIL 51/40
G01F 15/00



マトリクス分析

課題・解決手段、材料・用途等のマトリクスから開発アプローチを把握

競合比較

サイテーション分析

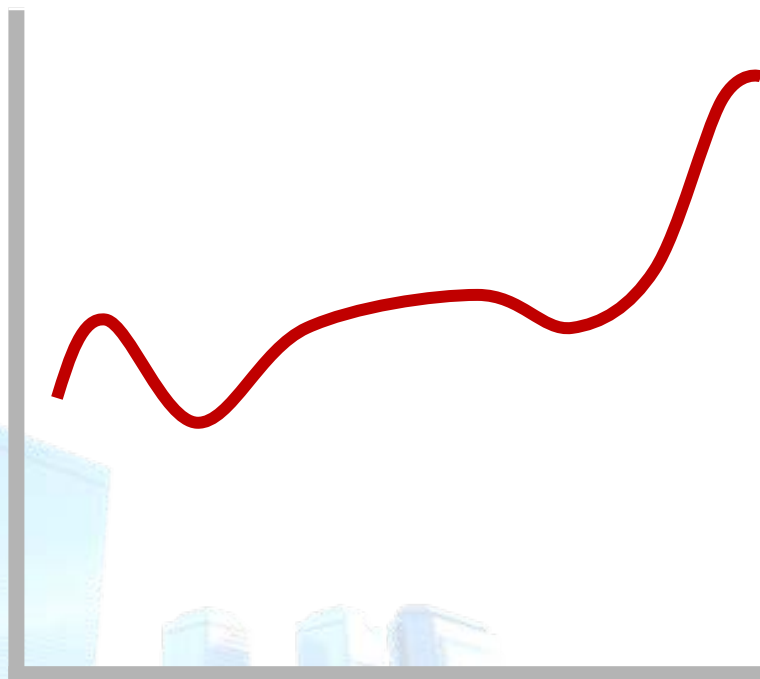
重要特許の把握、棚卸し、類似技術抽出

IPランドスケープでは 仮説思考が重要

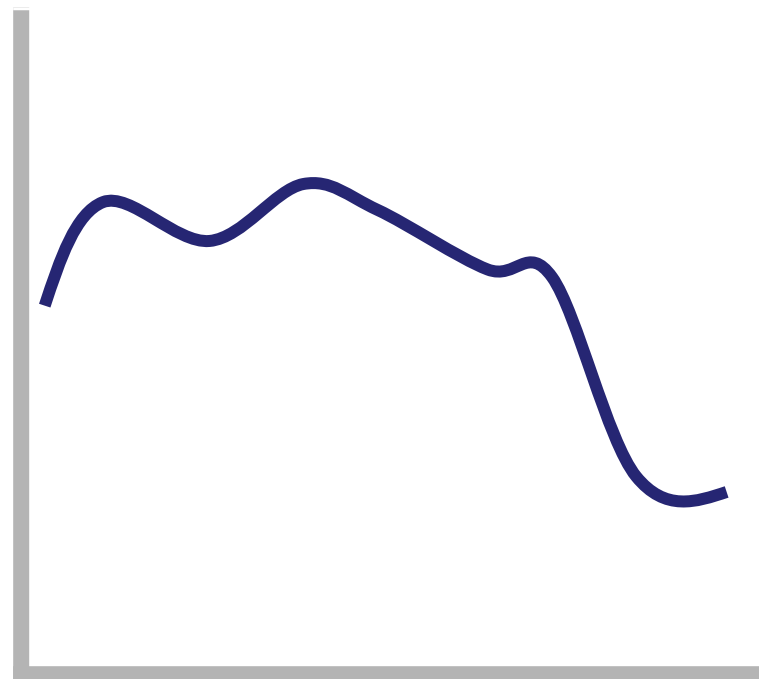
出願数の推移を見てみると

ある領域・企業で

出願数が増加した



出願数が減少した



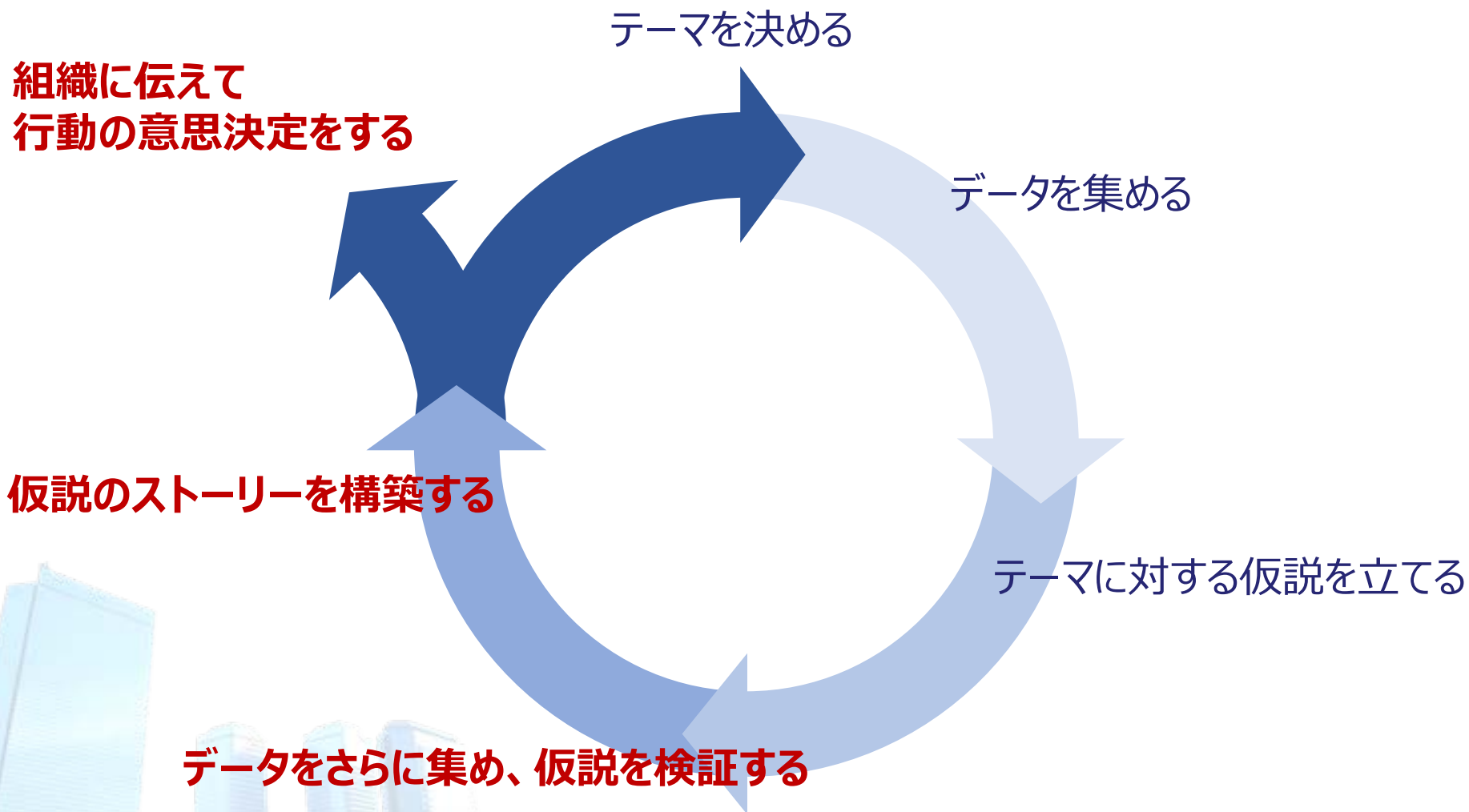
様々な仮説が成り立つ

ある領域・企業で
出願数が増加した

研究開発を始めた
新製品・新サービスを開発
知財注力への方針転換
出願ノルマが増えた
事業統合、発明者の異動
共同研究・開発の増加

出願数が減少した

研究開発の終了
製品のコモディティ化
内製化の限界、外部調達化
発明者の離脱・退職
別領域へのリソース集中
事業撤退・縮小



将来に備えて、特許戦略はどうするのか？

将来の脅威・機会に備えた今からの行動
事業が先か？ 特許が先か？

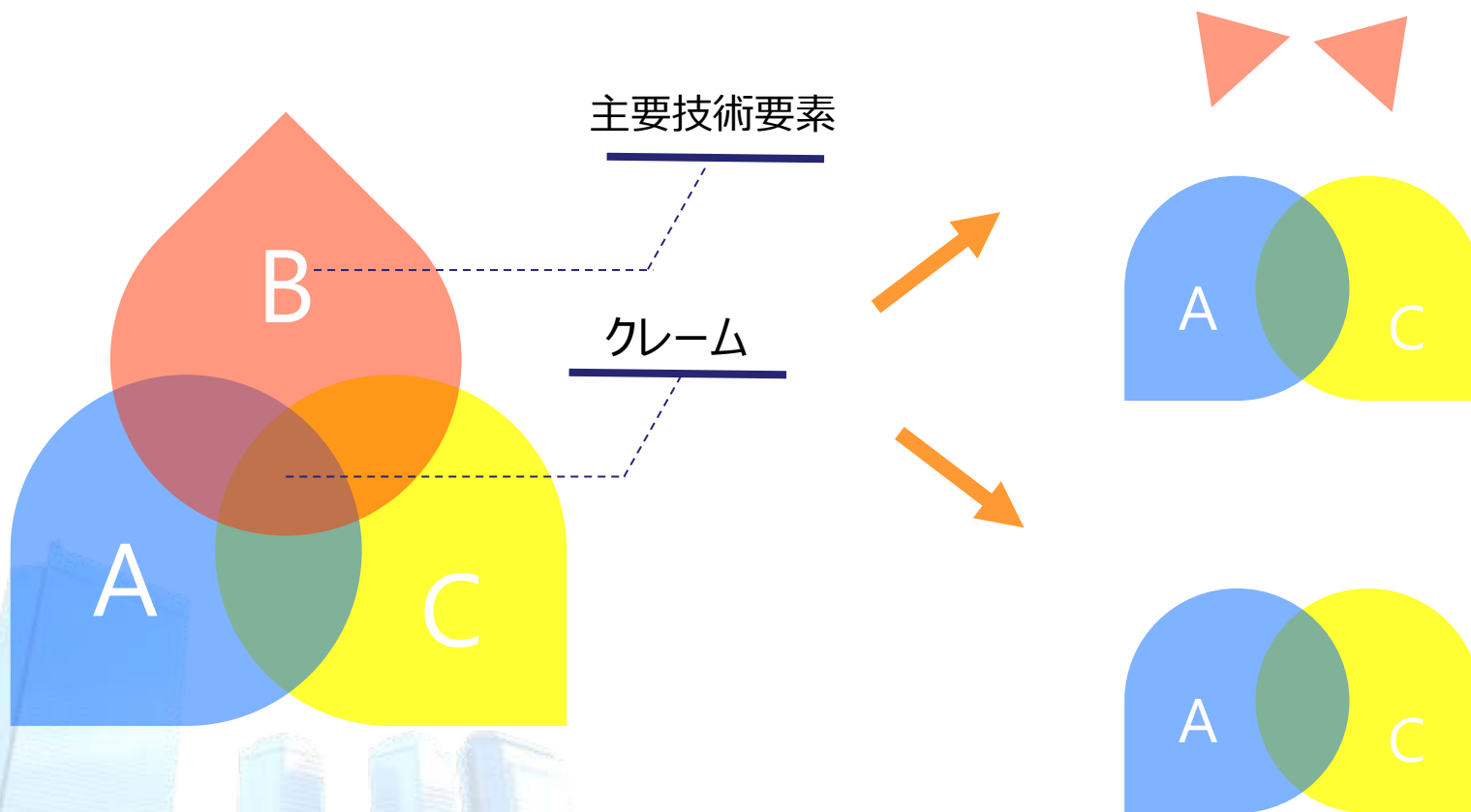
トンネル掘削では
岩盤と砂地で行動が異なる

モノづくりとコトづくりでは
特許の最適な取得時期が異なる

大深度・大水圧下で施工された
世界最大級トンネル

シールドマシンカッター（海ほたる）

特許戦略の例) 模倣防止を狙う特許クレームの考え方



主要技術要素をあえて変え、模倣を防ぐ

知的財産経営と知的財産戦略

～IPランドスケープを用いた新しい知財戦略へ～

コーポレートガバナンス・コード

～会社の持続的な成長と中長期的な企業価値の向上のために～



2021年6月11日

株式会社東京証券取引所

コーポレートガバナンス・コード
『知的財産投資・情報開示関係』

補充原則:3-1③

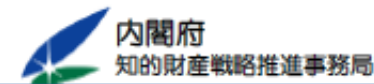
上場会社は、(中略)人的資本や**知的財産への投資等**についても、
自社の経営戦略・経営課題との整合性を意識しつつ
分かりやすく**具体的に情報を開示・提供**すべき

補充原則:4-2②

取締役会は、(中略)人的資本・**知的財産への投資等**の重要性に鑑み、
これらをはじめとする経営資源の配分や、
事業ポートフォリオに関する戦略の実行が、
**企業の持続的な成長に資するよう、
実効的に監督**を行うべき

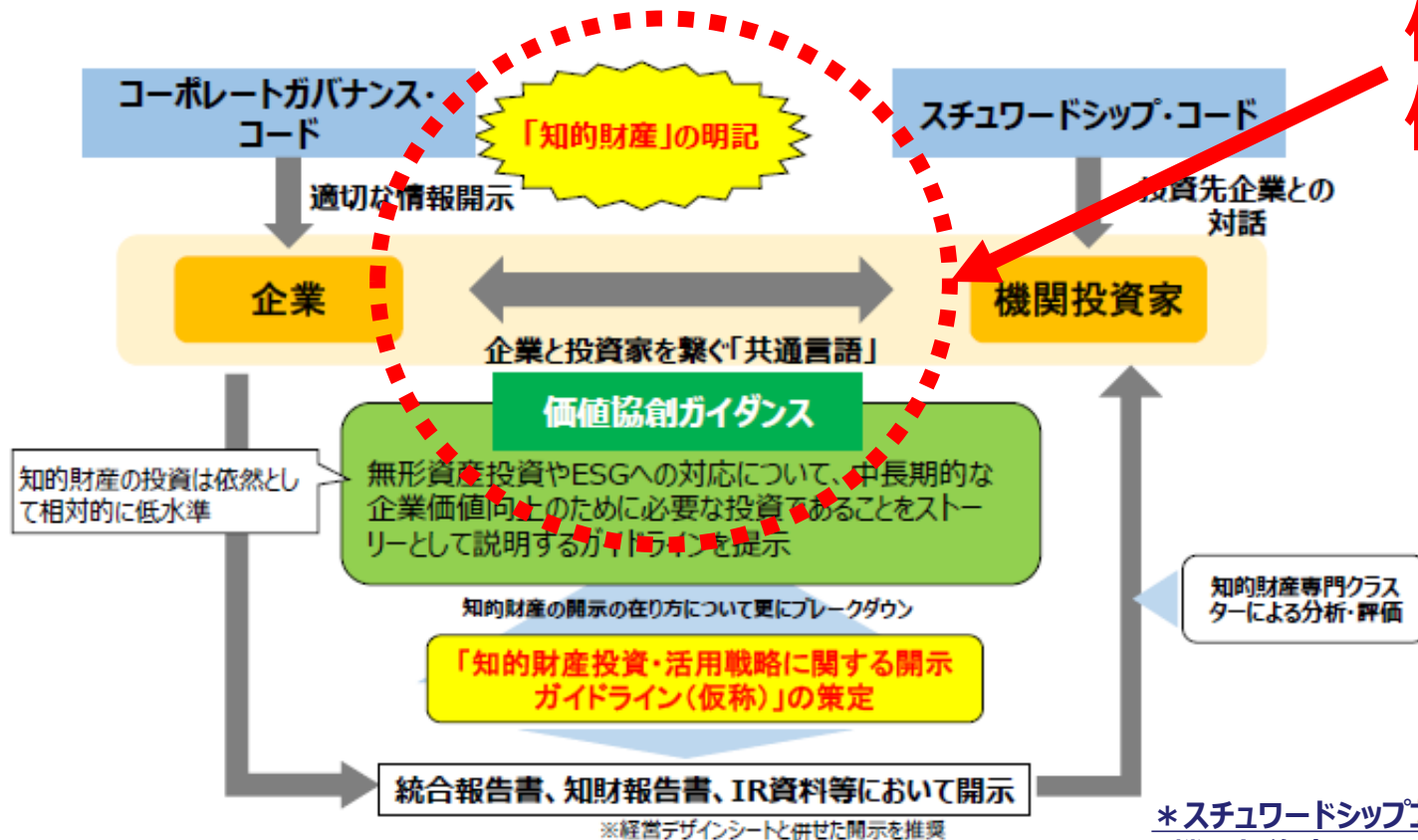
出典:<https://www.jpx.co.jp/equities/listing/cg/index.html>

知的財産の投資・活用戦略の開示に向けた方策案



- 企業における知的財産の投資・活用戦略を開示するにあたり、**知的財産投資・活用戦略に関する開示の在り方を示すガイドラインの作成**が必要。
- 知財等無形資産の開示の在り方について、価値協創ガイダンスの中で明確化。

**価値の協創
価値向上**



***スチュワードシップコード
機関投資家のあるべき姿を規定したガイダンス**

知財ガバナンス時代における知的財産部門や知財サポート機関の役割

		知財の重要性		
		小	中	大
知的財産部門		管理型	コンサル・アドバイス型	経営関与型
知財サポート機関 (特許事務所)		アウトソーシング	アウトソーシング コンサル・アドバイス	アウトソーシング コンサル・アドバイス 第三者評価

ビジネスの本質は、 「モノ(コト)づくり」と「カネづくり」の両側面

理系的



モノ(コト)づくり
ビジネス

- ・モノを活用する労働力ビジネス
- ・モノ（技術と労働）の生産を活かし、質の創造と量の拡大により、高価値に転換
- ・技術開発あり＝ヒトと知恵(考えること) が介在
- ・農・林・漁・工・商業など、モノの生産を管理
- ・富裕者会になると、カネづくりビジネスに移行しがち
⇒ 本質依存＝技術＝知力

時代の傾き

文系的



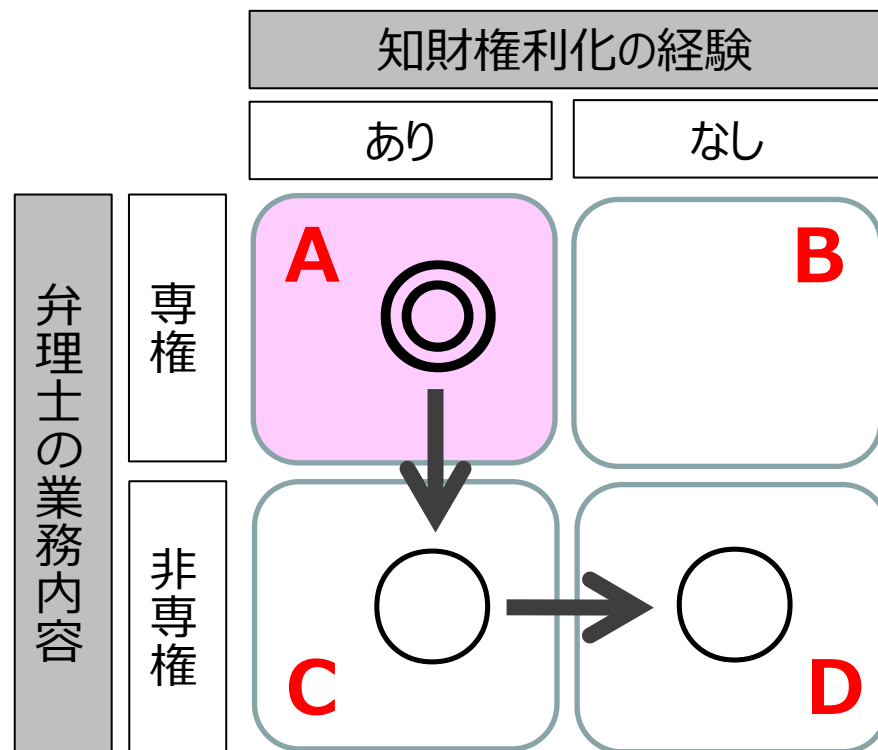
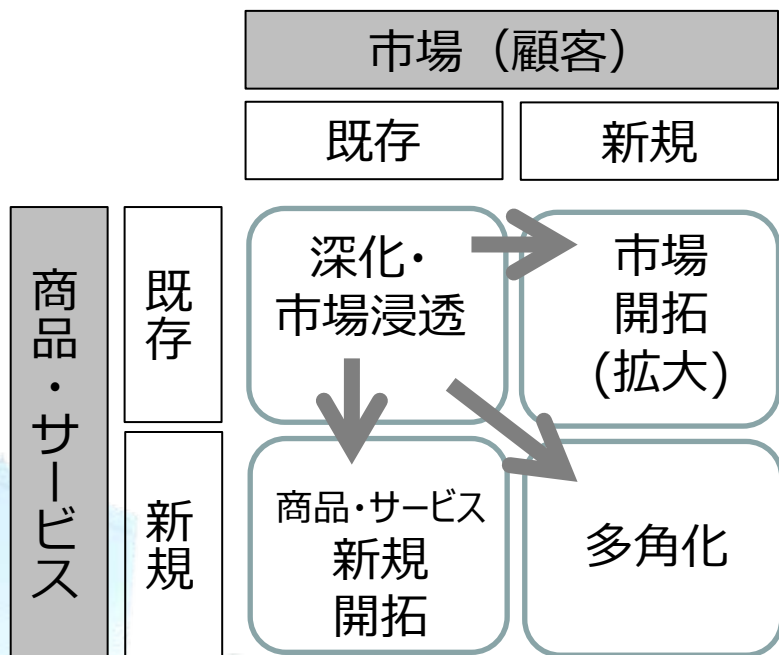
カネづくり
ビジネス

- ・契約を利用する手数料ビジネス
- ・モノを必要としない＝機械化したヒトが介在
- ・ルール・仕組みに依存し、カネにカネを生ませる
- ・成熟社会での儲ける仕組みづくりの中心
- ・銀行・証券など、手数料ビジネスを管理
- ・どのような社会でも廃れずに存在
⇒ 現象依存＝ルール化＝権力

どちらもビジネスとして重要

（出典）『ソニー破壊者の系譜』 原田節雄（さくら舎）

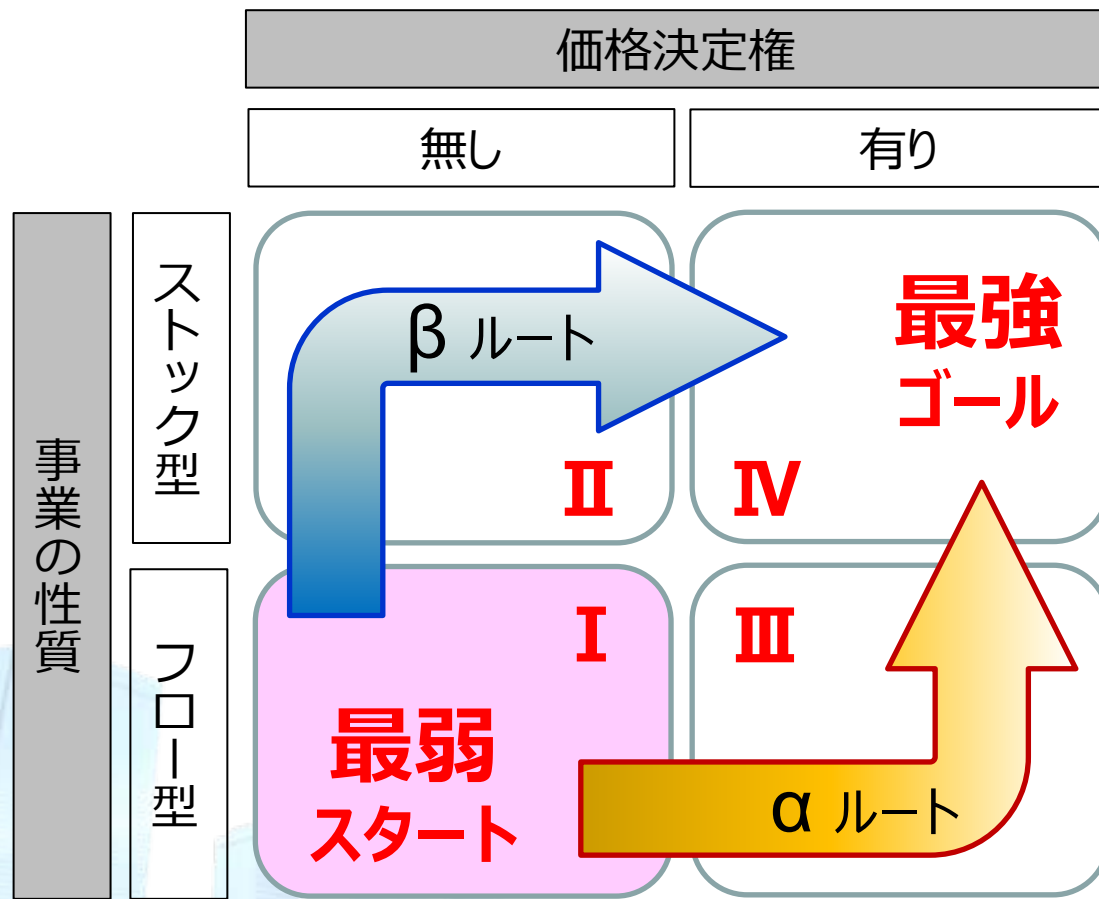
●製品・市場の戦略マトリクス アンゾフの成長マトリクス



- A** : これまでの業務の中心
B : 新たな市場の開拓
C : 既存市場での浸透
D : 多角化された市場

戦略
↓

● 儲けるビジネスのためのステップ



- ストック型ビジネス：仕組みを作り、その仕組みによって成り立つ形のビジネス（携帯契約者の数で儲ける）
- フロー型ビジネス：単発でその都度仕事を請け負いを繰り返す形のビジネス（飲食店など、稼動することで儲ける）

β ルート

知財権による独占を通じて
ストック化してゆくルート

α ルート

知財権で交渉力を
強化してゆくルート

● 知財ポートフォリオマトリクス

- 特許権の有無と実施の有無でマトリクスを構成
- マトリクスのどの象限に注力するかで、ポートフォリオの戦略は異なる

α : 従来からの知財 = 強い特許により参入障壁を構築

β : 利益率を判断し、特許取得のコストをかけずに事業実施

γ : 知財純資産の活用

特許権			
		あり	なし
権利の実施 (実施予定)	あり	α	β
	なし	γ	δ ノウハウなど

● 注力領域を選択し、戦略を策定する

(ア)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

(イ)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

(ウ)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

(エ)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

(オ)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

(カ)

		特許権	
		あり	なし
実施 (実施予定)	あり	α	β
	なし	γ	δ

●オープン/クローズ戦略の4視点

- オープン化
- 権利化
- 見える化（有形的秘密化）
- 秘匿化、ブラックボックス化
 - 完全ブラックボックス化
（文書化、管理、管理指針、契約）



● オープンとクローズを使いこなす

$\alpha \Leftrightarrow \beta$ のバランス :

オープン／クローズ戦略

		特許権	
		あり	なし
権利の実施	あり	従来の知財の 発想・戦略 (大半はこの意識) α	事業の利益率を 判断し、あえて 特許を取得しない β

情報

オープン

クローズ

実施

オープン

クローズ

広告
宣伝

特許出願
権利取得

標準化
(サブマリン)

出願せず
ノウハウ
として維持

(出典)
知的財産政策ビジョン



秘匿化	知財の 独占 実施	知財の 独占的 ライセンス	パテントプール クロスライセンス	無償 実施	オープン ソース ソフト	オープン コラボ レーション
-----	-----------------	---------------------	---------------------	----------	--------------------	----------------------

●これまでの知財戦略と新しい知財戦略

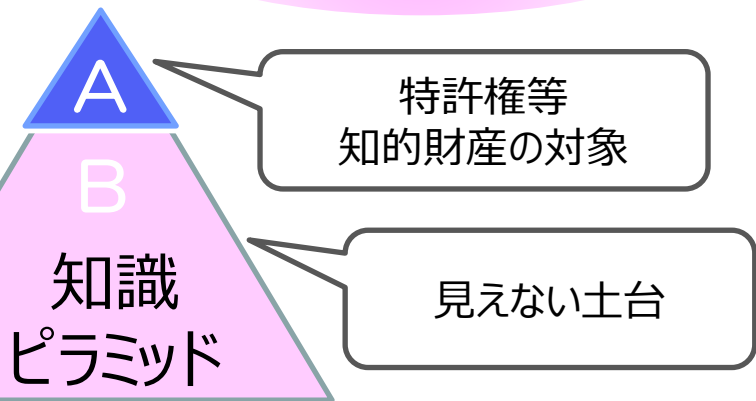
		特許権	
		あり	なし
権利の実施 (実施予定)	あり	従来の知財の 発想・戦略 (大半はこの意識)  α	事業の利益率を 判断し、あえて 特許を取得しない β 
	なし	権利を使った事業 を伴わないが、戦略 により純資産に  γ	この部分の 無形資産を いかに発掘 (マイニング) するか δ

α : 従来からの知財 = 強い特許により参入障壁を構築

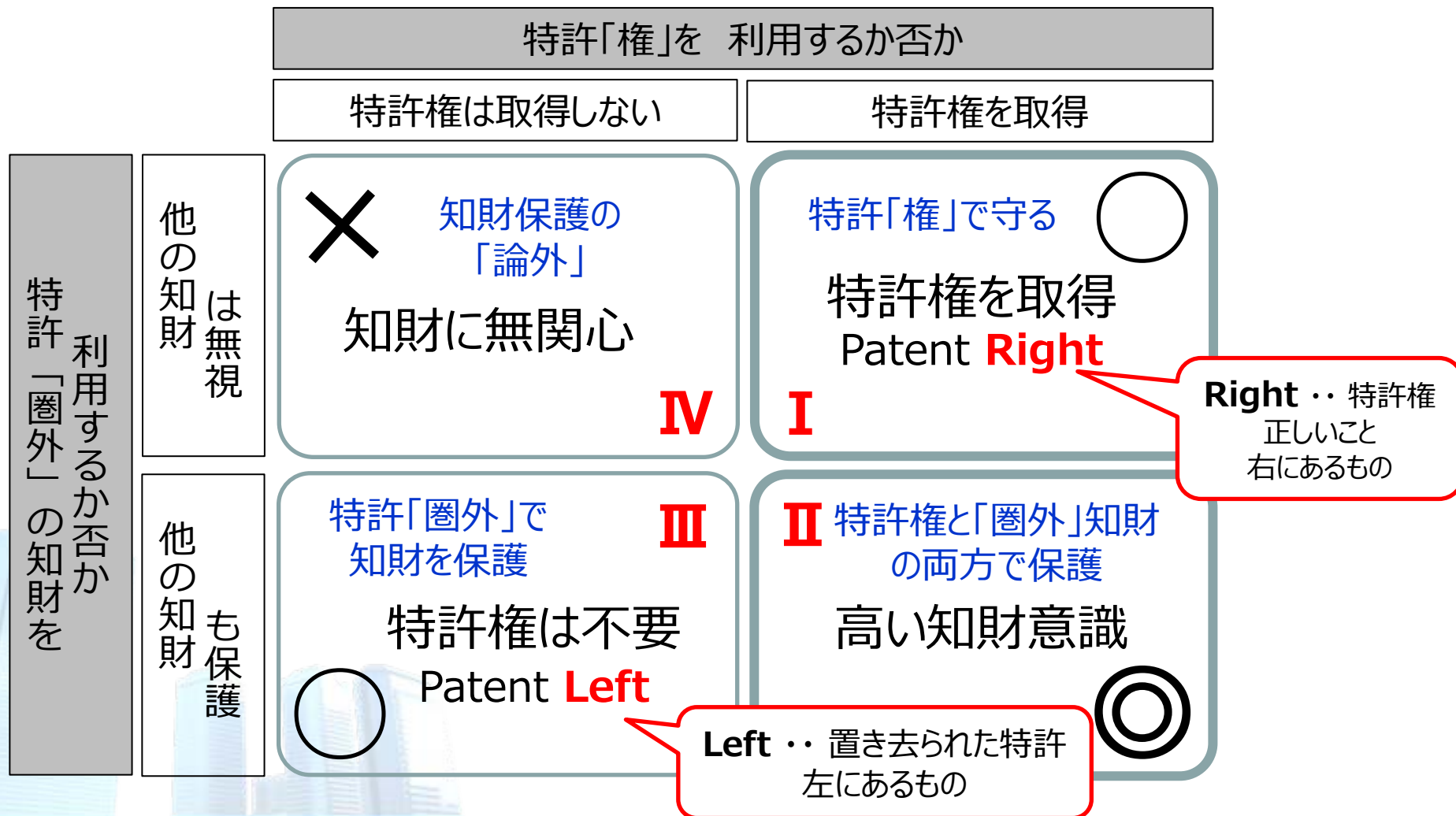
β : 利益率を判断し、特許取得のコストをかけずに事業実施

γ : 知財純資産の活用

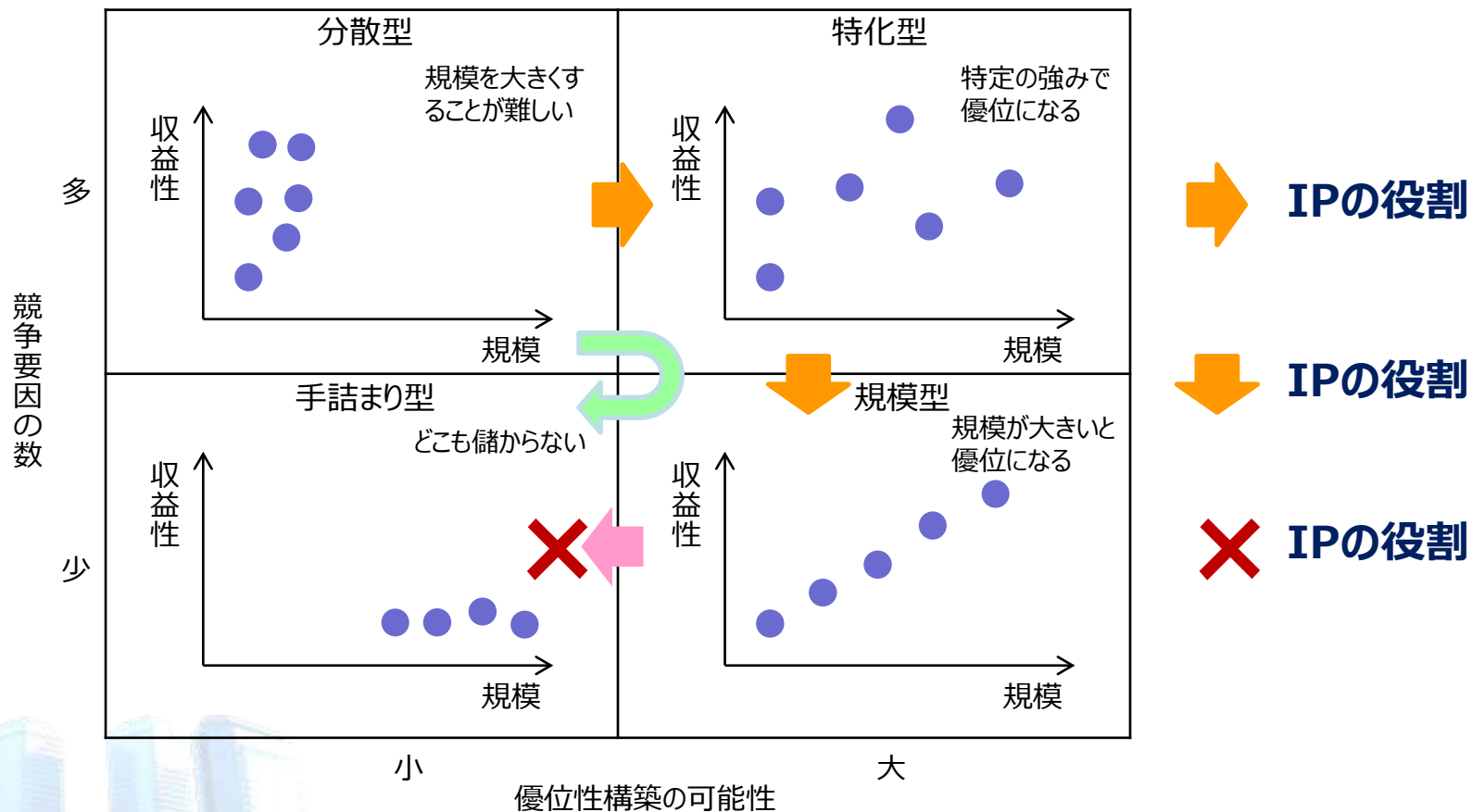
δ : 過去の失敗例・ノウハウ・他社情報などの無形資産が、この部分に存在する



● 特許「権」と特許「圏外」：～ 特許権以外にもある、知財の保護 ～



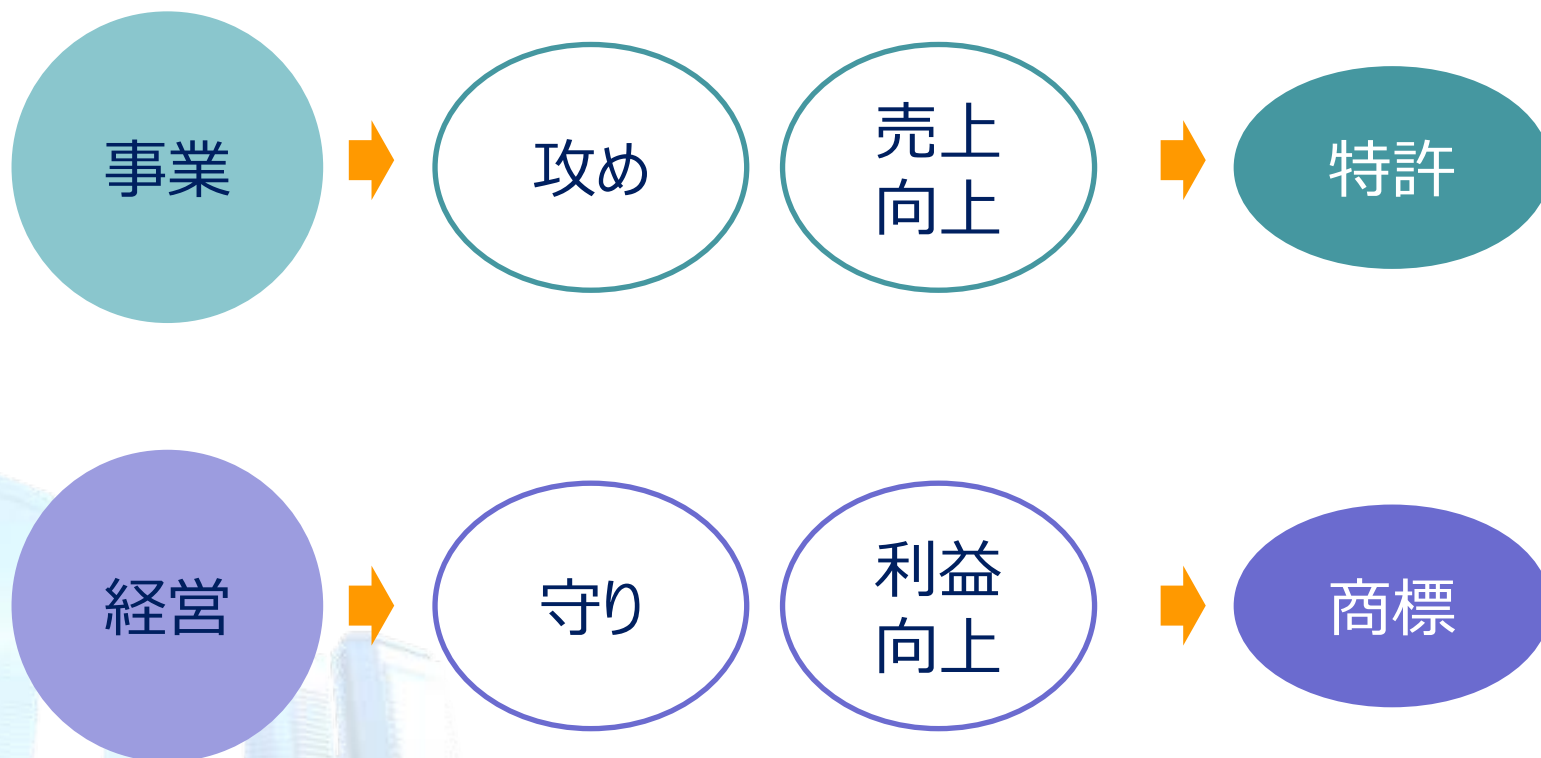
規模型戦略への移行とIPの活用



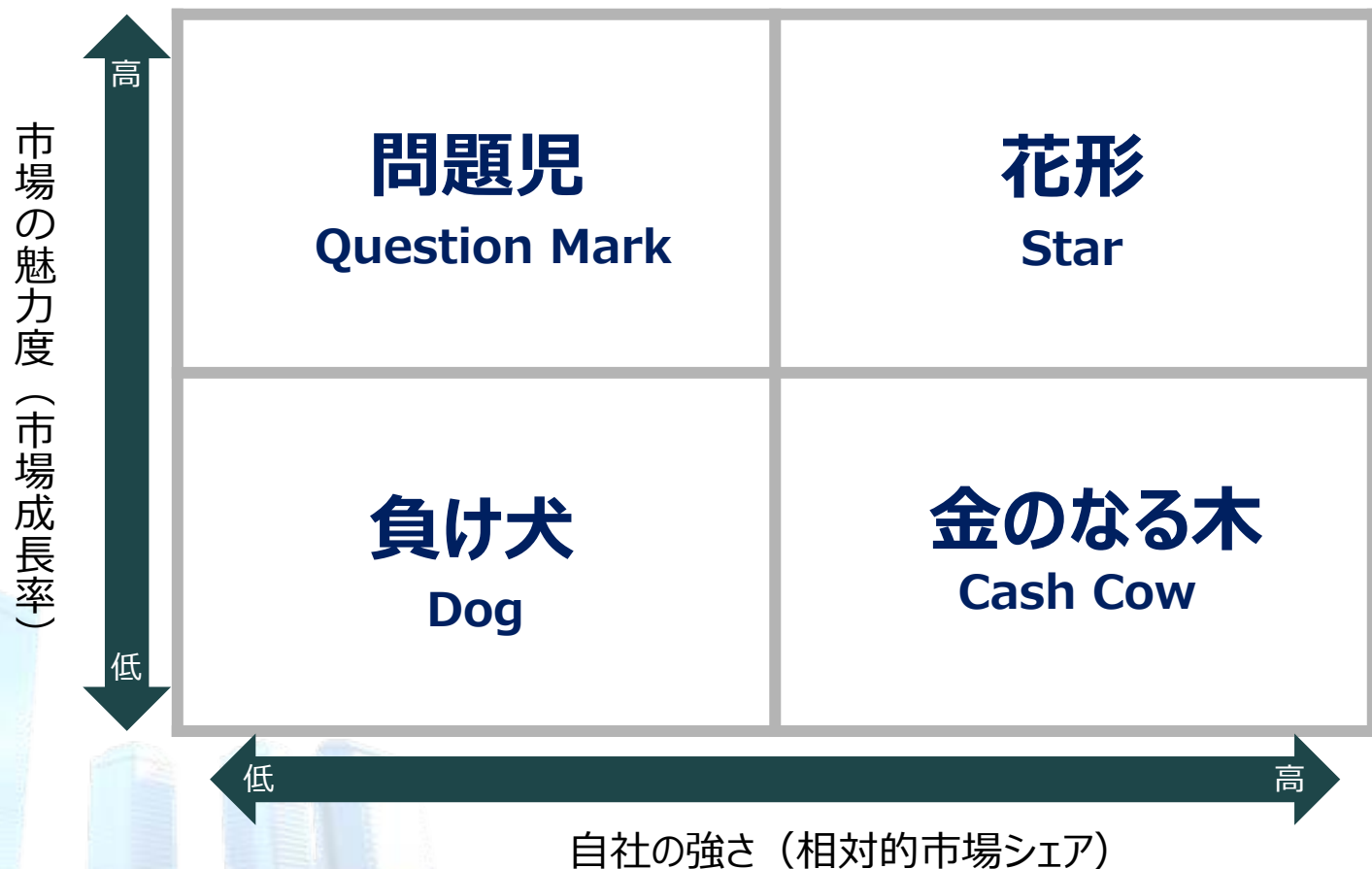
※ボストンコンサルティンググループが考案したアドバンテージマトリクス

「事業」と「経営」

- 「事業」と「経営」の領域は異なる
- 「特許」が「経営」のmatterとなるには？

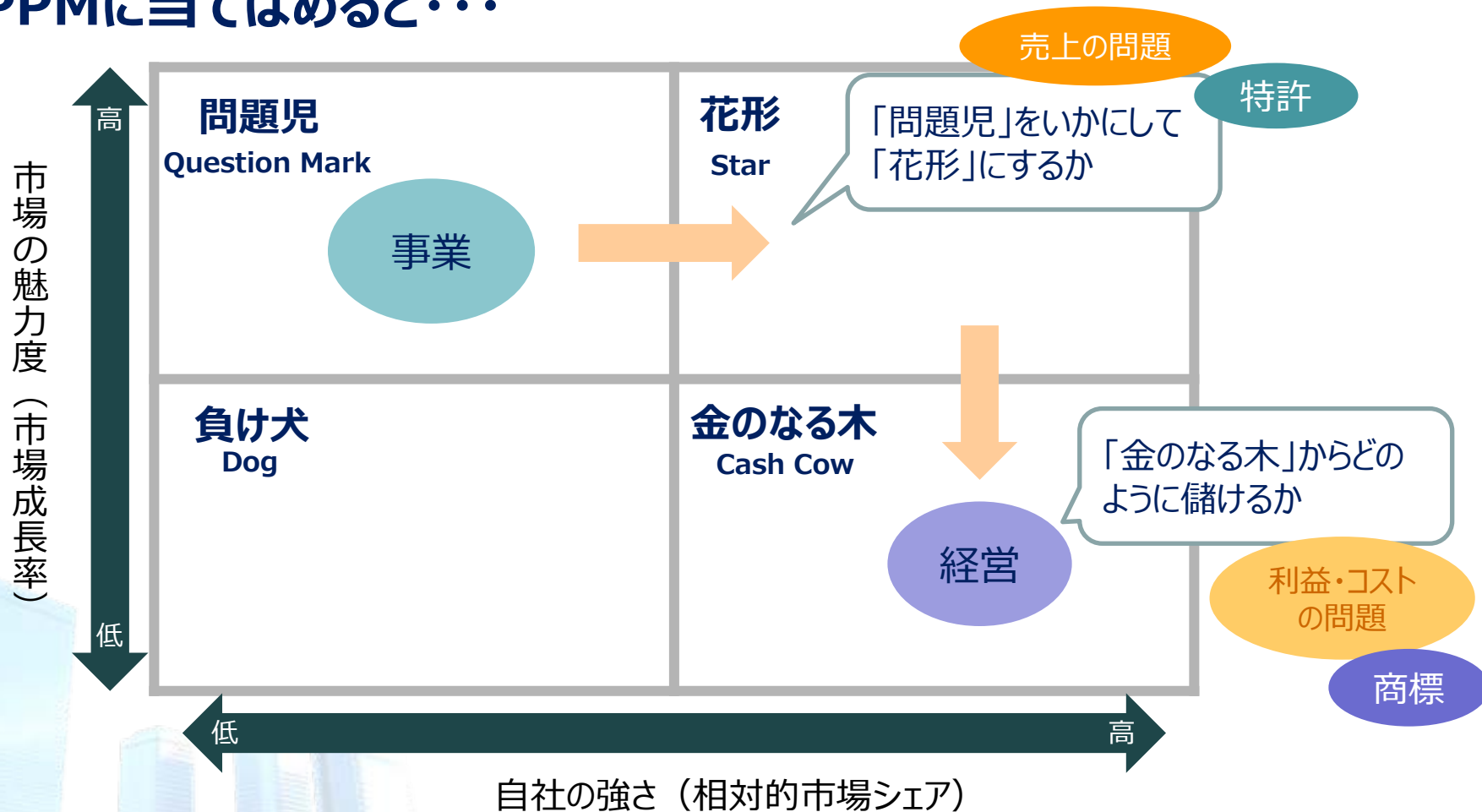


PPMマトリクス



「事業」と「経営」

PPMに当てはめると・・・



勝てる戦略と、勝てる市場

		勝てる戦略	
		Yes	No
勝てる市場	Y	誰でも勝てる	戦略がなくても勝てる場合が多い
	N	天才でないと勝てない	必ず負ける

S:セグメンテーション

T:ターゲティング

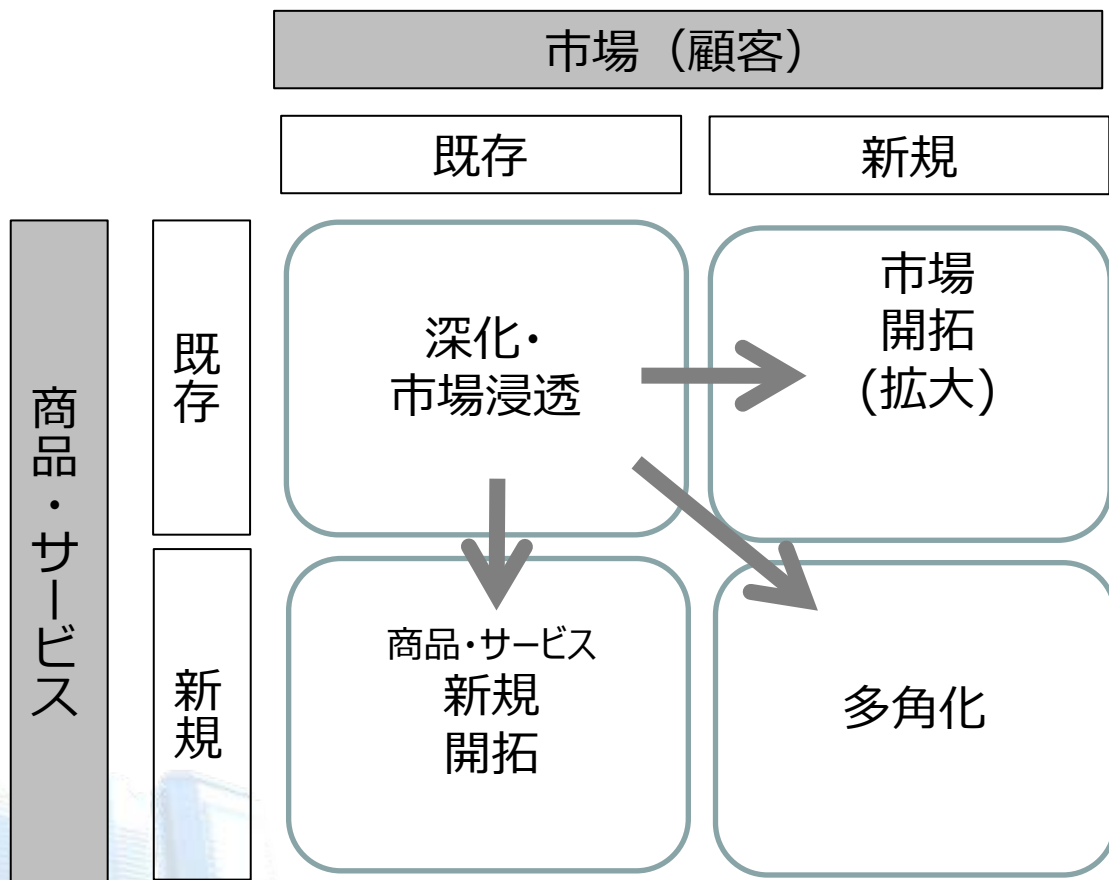
P:ポジショニング

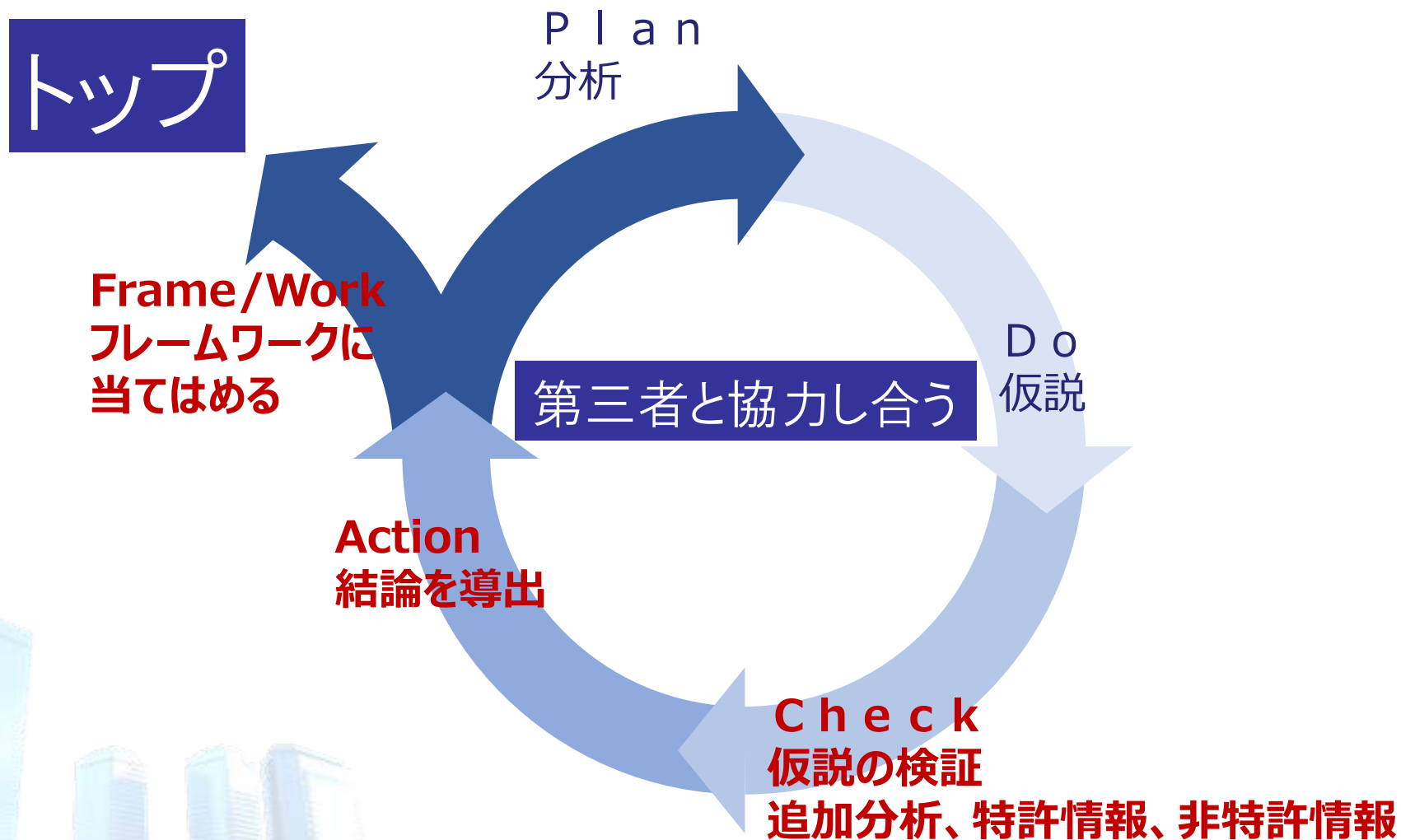
Value 経済価値はあるか	Rarity 希少性は高いか	Inimitability (模倣困難性は高いか)	Organization (組織に優位性はあるか)	
YES	YES	YES	YES	持続的競争優位
YES	YES	YES	NO	潜在的競争優位
YES	YES	NO		一時的競争優位
YES	NO			競争均衡
NO				競争劣位

SWOT分析

	プラス	マイナス
内部	強み (S)	弱み (W)
外部	機会 (O)	脅威 (T)

アンゾフの成長マトリクス





**豊臣秀吉が偉いのは、
一夜城の築城を提案をしたからではなく、
一夜城を実際に作ったからである。**



戦略だけに陥らない。適切な領域を選び、行動をする。

正林国際特許商標事務所が考えるIPランドスケープとは

『IP情報を軸として事業、製品、サービス、財務、その他の非IP情報を総合的に加味しながら、対象企業の現在置かれている**位置づけ**とその**未来を提言し、経営に役立てるもの**』

経営層に対して、IP情報&非IP情報を
わかりやすく図示し、
真実を伝える、
意思決定を促す道具

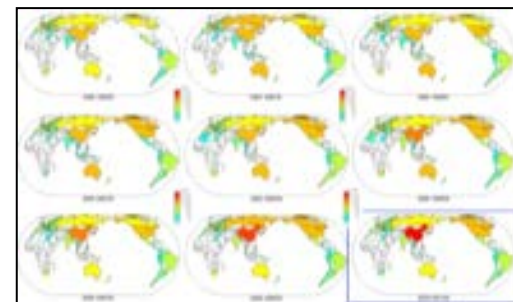
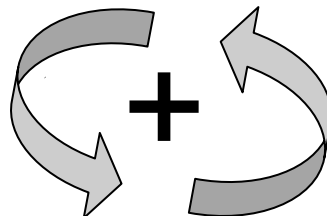
● 特許情報の解析の手法（当事務所の特長的な手法）：

的確なマクロ分析



鳥の目

知財アナリストの的確な
分析観点を設定した分析

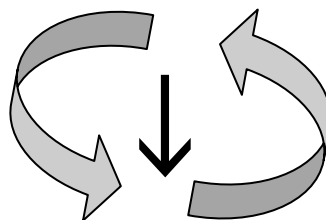


的確なミクロ分析



虫の目

特許庁OB、知財アナリストによる
技術の的確な把握

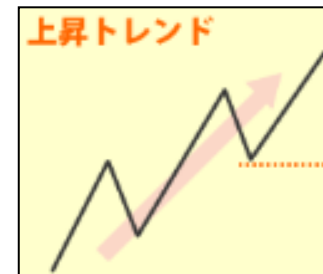


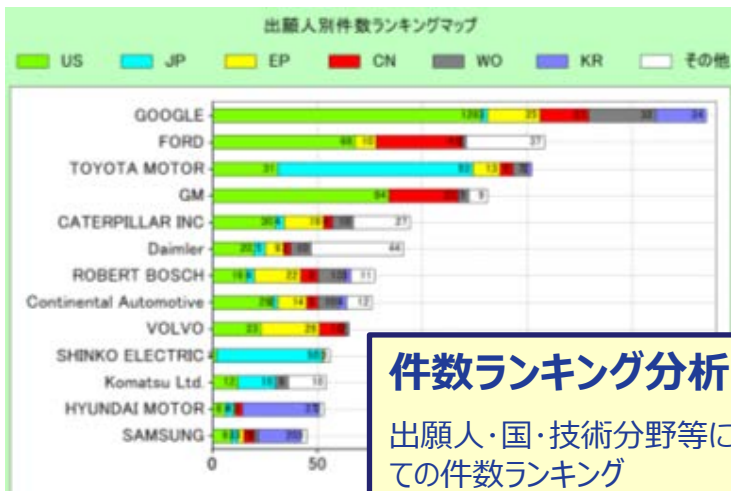
予測（流れを読む）



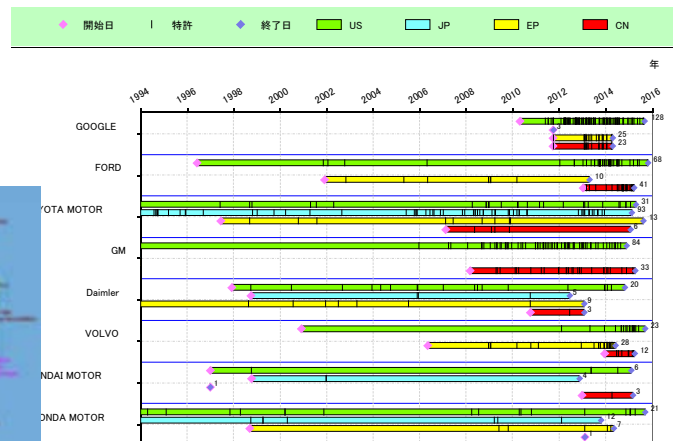
魚の目

知財アナリストの他、
必要に応じて会計士、
MBA保有者も含めて検討





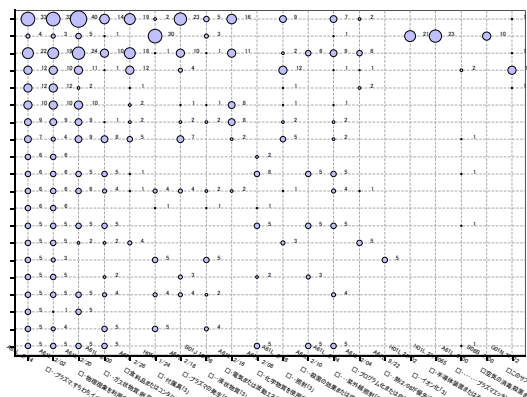
技術俯瞰マップ



参入分析

新規参入時期や継続期間を把握し、主要出願人の参画実態を把握

出願人

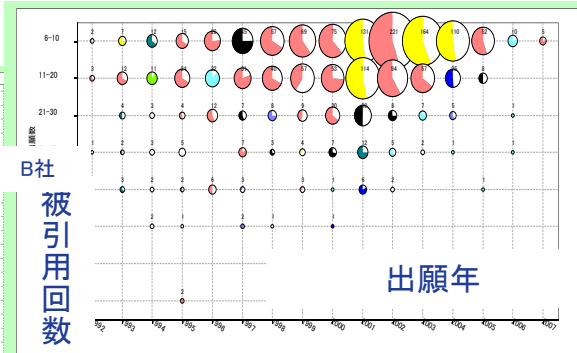
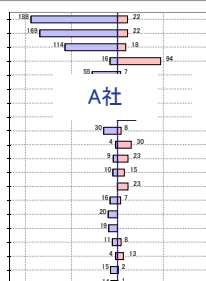


技術分類

技術分類

HOIL 51/1
H05B 33/14
H05B 33/22
HOIL 51/00
H05B 33/10
HOIL 27/32
HOIL 27/00
H05B 33/26
G01C 21/1
G01C 21/1
HOIL 51/40
G01F 15/00

---有機発光物質を含むもの(7)
用されるもの 有機発光ダイオード
あるいは化学的または物理的組
成要素の配置あるいは化学的
材料を用いたまたは移動部分として
---有機的に形成された膜層を含むもの
2.32 ---半導体素子(73)
に物に適用する装置または工程
HOIL 51/20 ---装置(74)
用される構成部品を有するもの 例
HOIL 51/30 ---材料の選択(76)
6層上に形成された複数の半導
体層物質の配向あるいは組成
100%純度の有機材料を含むもの(7)
有機層が膜厚以上の膜で形成
製造または処理中に適合した
H05B 33/28 ---半透明電極
G01F 15/00 ---測定対象の光学的性質を有する化合物



マトリクス分析

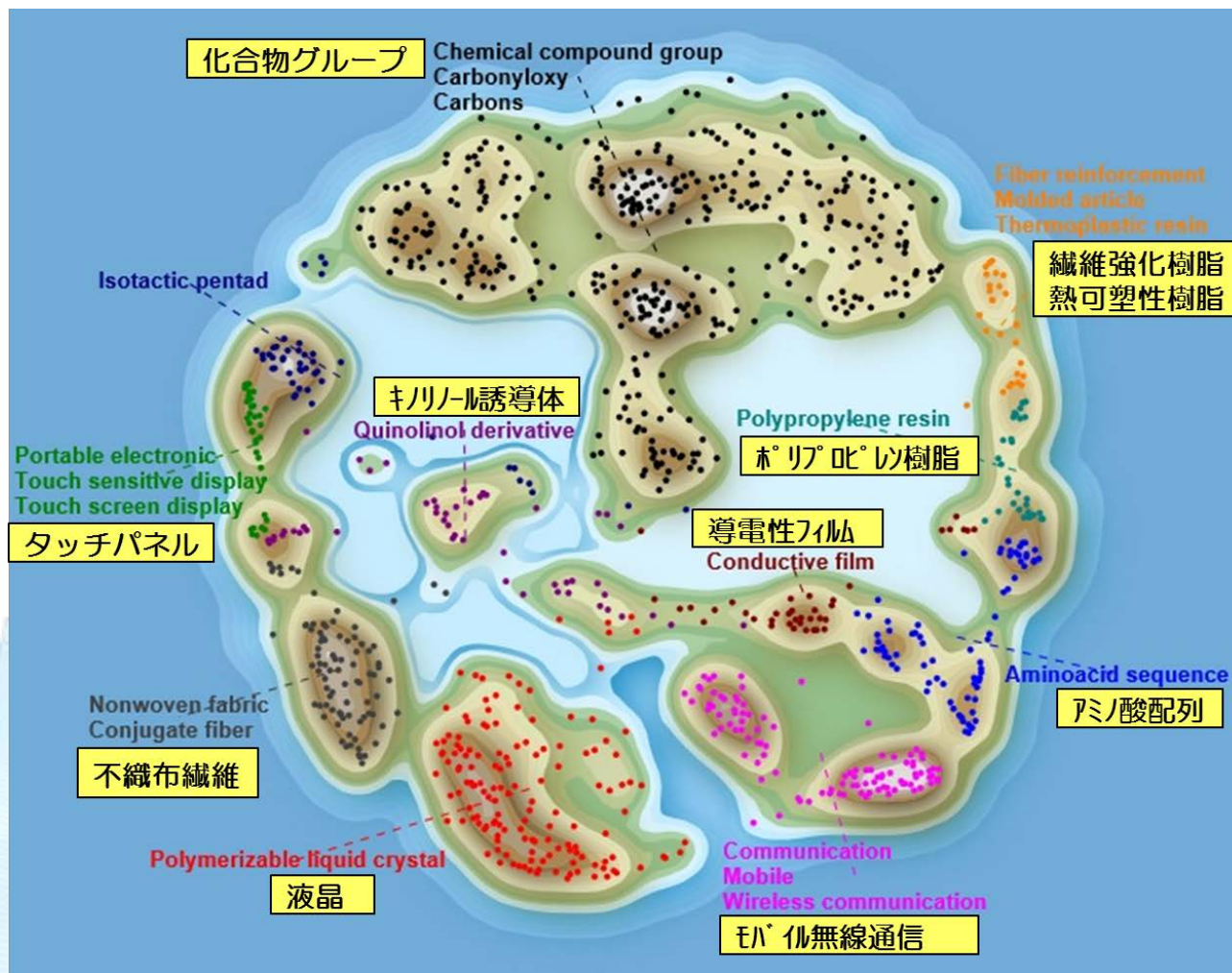
課題・解決手段、材料・用途等のマトリクスから開発アプローチを把握

競合比較

サイテーション分析

重要特許の把握、棚卸し、類似技術抽出

● 特許情報の積極的な活用 ～ 技術の2次元マップ化：

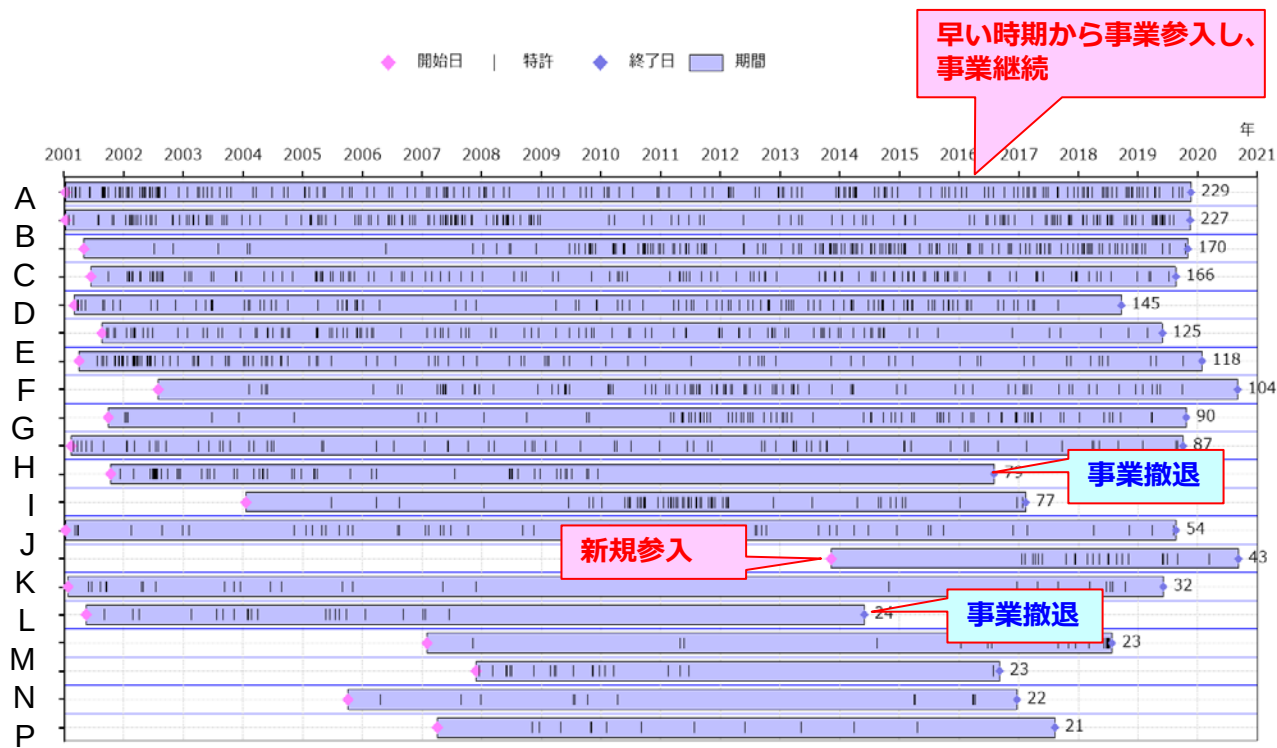


各点は出願を示し、技術的に近いものほど近い距離にプロットされている。クラスター分析により、分類されたクラスター毎に点の色を変えている。

JNC保有特許
(出願審査中含む)
検索式 ① * ②
= 1 6 9 2 ファミリ
① 出願人：
「JNC corporation」
② 出願日：
1995年1月1日以降
発行国：全世界

●参入分析 各社の事業参入時期や継続状況を把握

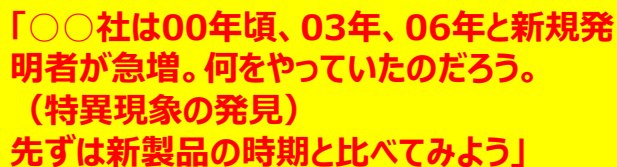
- 各出願人の出願開始日及び最新出願日を視覚的に把握
- 各出願人の事業参入時期（出願開始日）及び事業継続状況（最新出願日）を視覚的に把握可能。



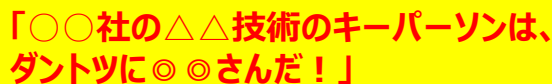
新規発明者分析

発明者と技術分類の分析 (特定分野のキーパーソンの探索)

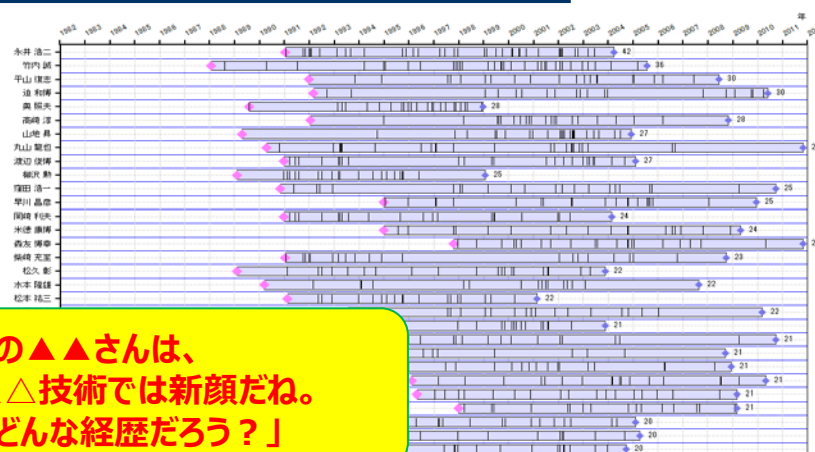
 東芝(株)
  富士フイルム(株)
  GE グループ
  日立グループ
  キヤノン(株)



発明者を切り口にした
競合会社の特徴（戦略、体制？）を
分析できる



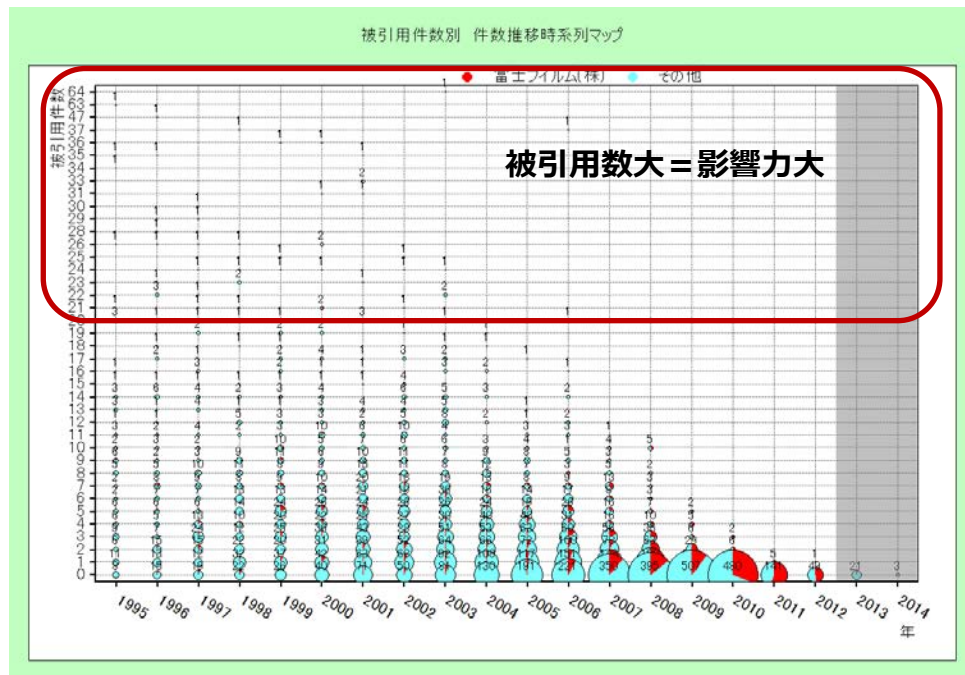
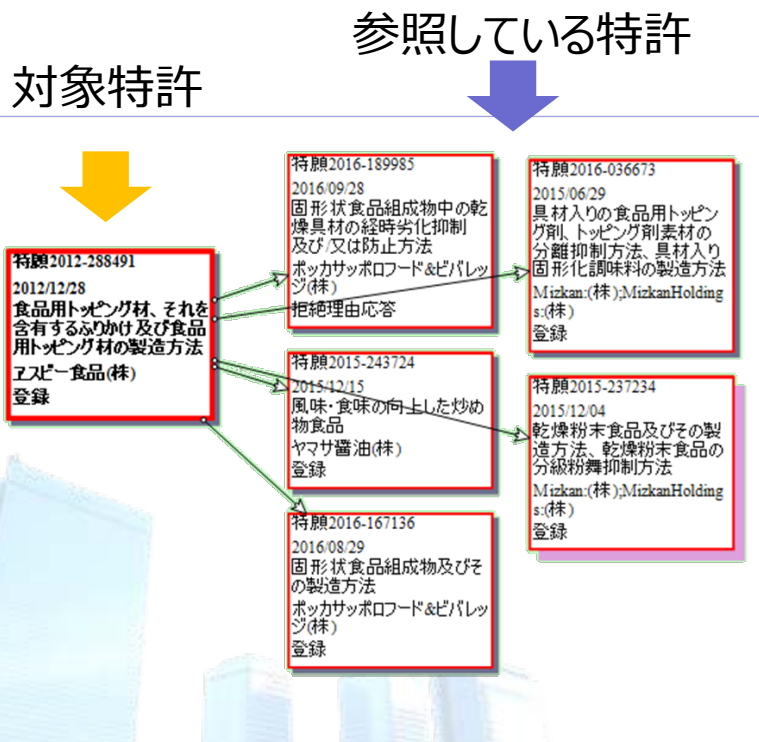
各発明者の時系列分析



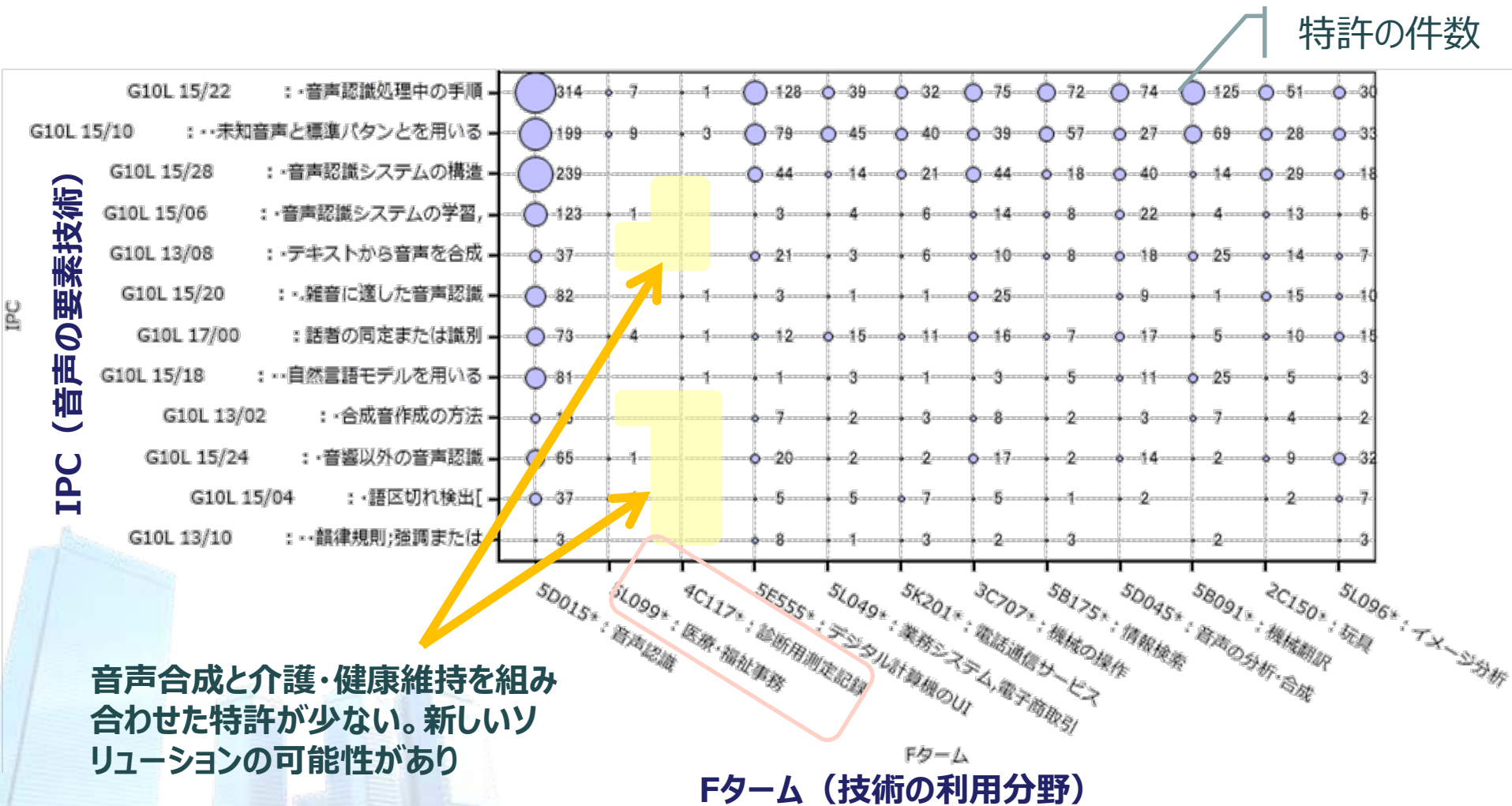
「〇〇社の▲▲さんは、
意外に△△技術では新顔だね。
その前はどんな経歴だろう？」

●サイテーションマップ 自社の技術の影響度を測る

- 対象特許を参照している特許が多いほど影響力が高い
- 参照している特許の出願人は潜在的な競合の可能性もある



●マトリクスマップのホワイトスペースからチャンスを発見する



解析のサンプル

対象会社と競合会社との 特許力の強みと弱みの比較

指定出願人 FI別 件数コンパスマップ

対象会社

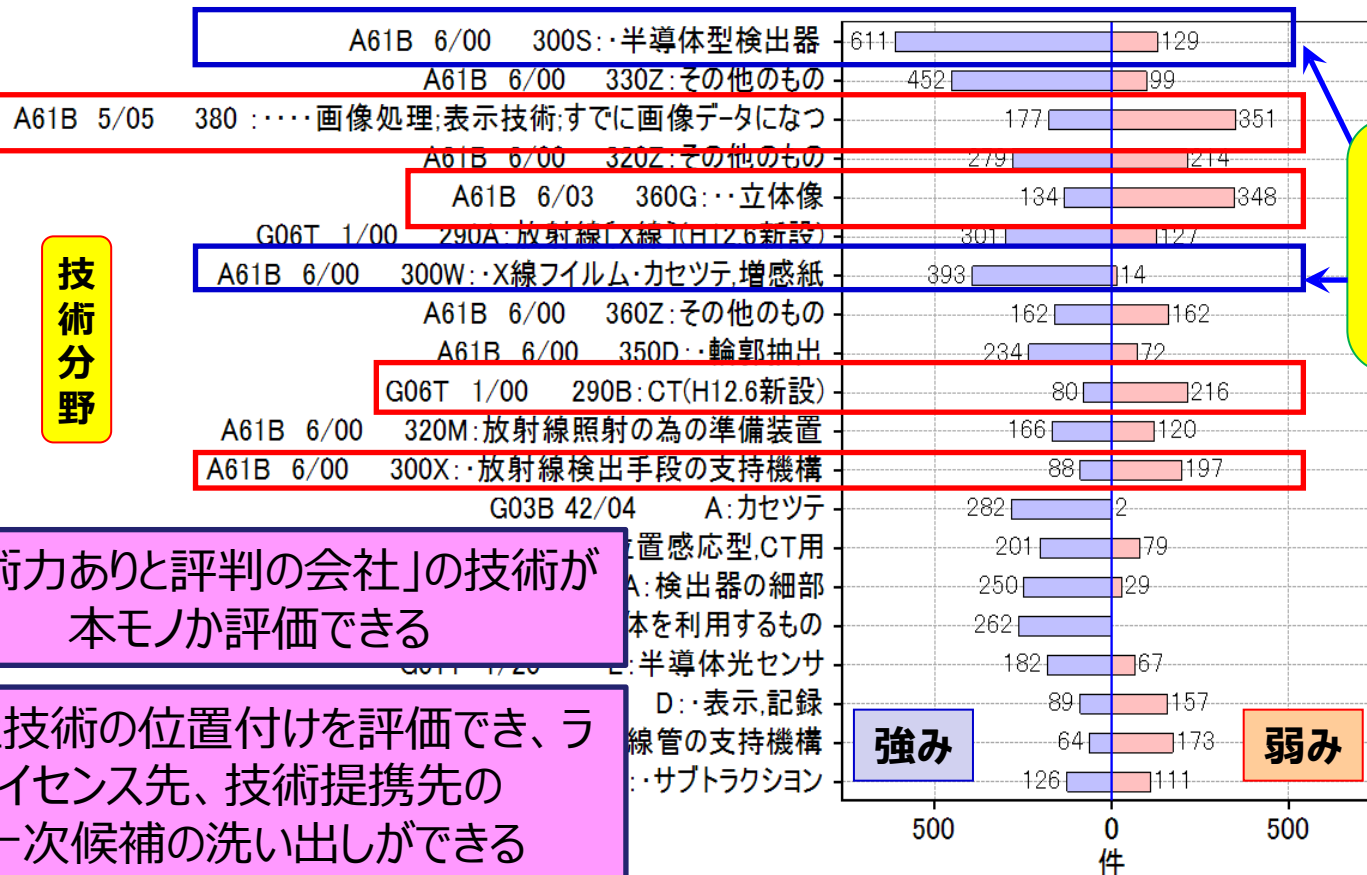
検討他社

技術分野

「技術力ありと評判の会社」の技術が
本モノか評価できる

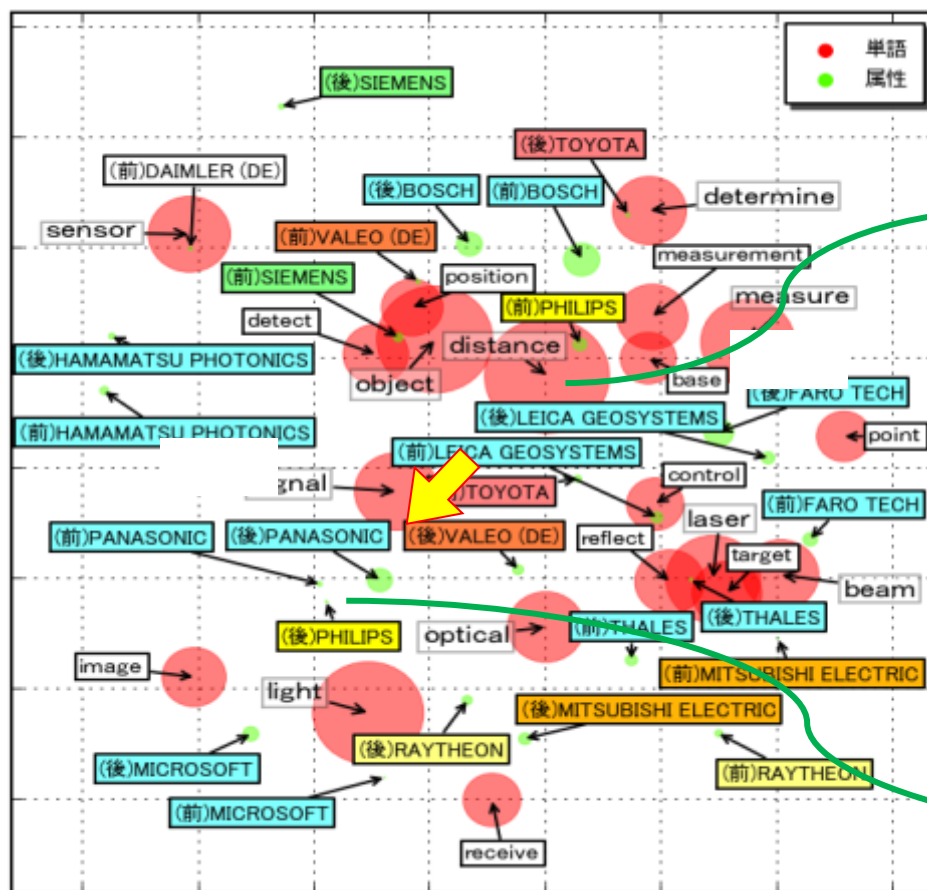
自社技術の位置付けを評価でき、ラ
イセンス先、技術提携先の
一次候補の洗い出しができる

「自分が強く、競合
会社1が弱い分野
もあるので、これを
何とか活用した
い！」



強み

弱み



(10) United States
(12) Patent Application Publication
(11) Inventor: Bester et al.

(13) Pub. No.: US 2010/0315827 A1
(14) Pub. Date: Dec. 16, 2010

(54) VEHICLE LIGHT

(71) Applicant: Nix Bester, Düsseldorf (DE); Bester, Nix, Nix, Nix (DE); Luis Bester, San Francisco (CA); Augustina Gregorina Bester, Nix, Nix, Nix (DE); Walter Peter Bester, Nix, Nix, Nix (DE)

(72) Assignee: KUNENKUN PHILIPS ELECTRONICS N.V., DORDRECHT (NL)

(21) Appl. No.: 12000000
(22) PCT Filed: May 15, 2008
(23) PCT No.: PCT/EP2008/000000

(37) (US) (Date): May 4, 2010

(56) Foreign Application Priority Data
May 25, 2007 (EP) 0700707.4

Publication Classification
(51) Int. Cl. F21F 7/00 (2006.01)
(52) U.S. Cl. 362/004

ABSTRACT
The invention relates to a motor vehicle lighting device which comprises at least a light source comprising at least one LED, a light-transmitting part that is arranged such that it causes light from the light source and a reflector that is arranged such that it reflects that portion of the light from the light source that passes through the light-transmitting part and complex it into the beam of the lighting device, a reflector portion and a reflector portion being provided within the light-transmitting part, wherein said reflector portion

(10) United States
(12) Patent Application Publication
(11) Inventor: Jorman et al.

(13) Pub. No.: US 2013/0114019 A1
(14) Pub. Date: May 9, 2013

(54) MULTIVIEW DISPLAY

(71) Applicant: KUNENKUN PHILIPS ELECTRONICS N.V., DORDRECHT (NL)

(72) Inventors: Willem J. Jorman, Dordrecht (NL); Hugo J. Cornelissen, Nix, Nix, Nix (NL); Tim Bester, Dordrecht (NL); Karmen P. Karmen, Nix, Nix, Nix (NL)

(73) Assignee: KUNENKUN PHILIPS ELECTRONICS N.V., DORDRECHT (NL)

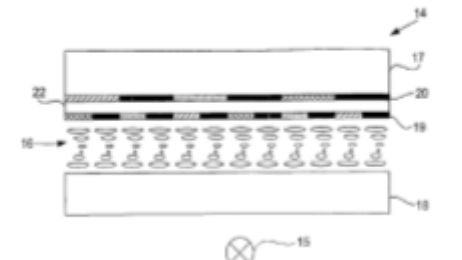
(21) Appl. No.: 12000000
(22) Filed: Sep. 16, 2012

Related U.S. Application Data
(62) Division of application No. 12/254,000, filed on Sep. 16, 2011, now Pat. No. 8,200,000, which is a division of application No. 10/771,000, filed on May 25, 2006, now Pat. No. 8,040,000, filed on application No. PCT/EP2006/000000 on Sep. 23, 2006.

(56) Foreign Application Priority Data
Sep. 27, 2005 (EP) 0520000.0

Publication Classification
(51) Int. Cl. G02F 2/00 (2006.01)
(52) U.S. Cl. 362/004
(53) CPC 362/004 (2013.01)
(54) USPC 362/004

ABSTRACT
A multi-view display is arranged to display views directed to respective viewing areas, comprising pluralities of imaging units and color filters, each of said color filters is associated with one of said imaging units, the color filters being arranged according to a first pitch and in a first sequence of colors, and a barrier including a plurality of color portions comprising color filter material, the color portions being arranged according to a second pitch that is substantially equal to twice the first pitch and in a second sequence of colors that corresponds to the first sequence of colors when viewed in white, the barrier is positioned so that light exits the display panel after passing through one or more portions of one color filter and the color portions of the barrier are configured to cooperate with the color filters to selectively direct light to the first and second viewing areas.



出願人名の前に追加されている、(前)は2006-2010年出願、(後)は2011年以降出願であり、前後の期間で、PHILIPSが最も移動が大きい
→ ミクロ的な分析

IPランドスケープの事例

【水処理膜分野の例】

日東電工のアライアンス先として
三菱レイヨンを選定した事例

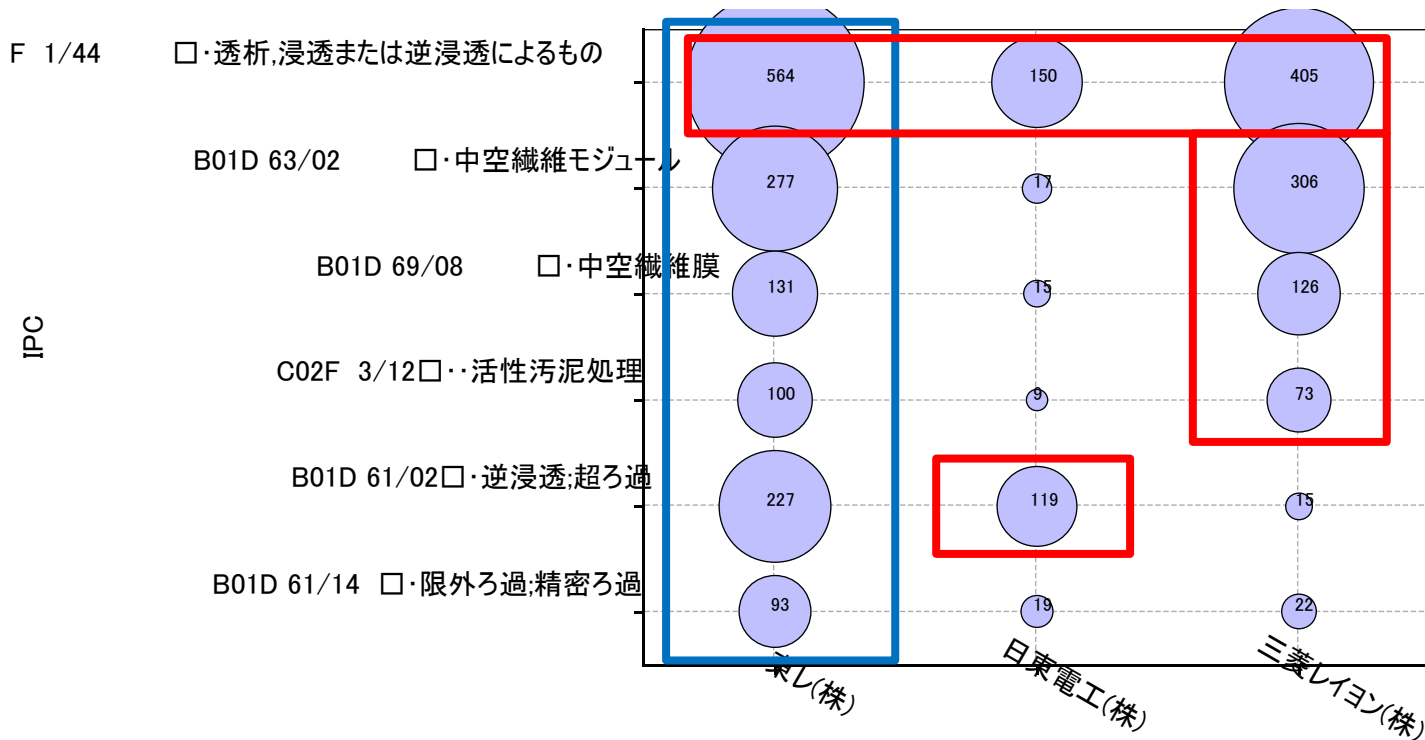
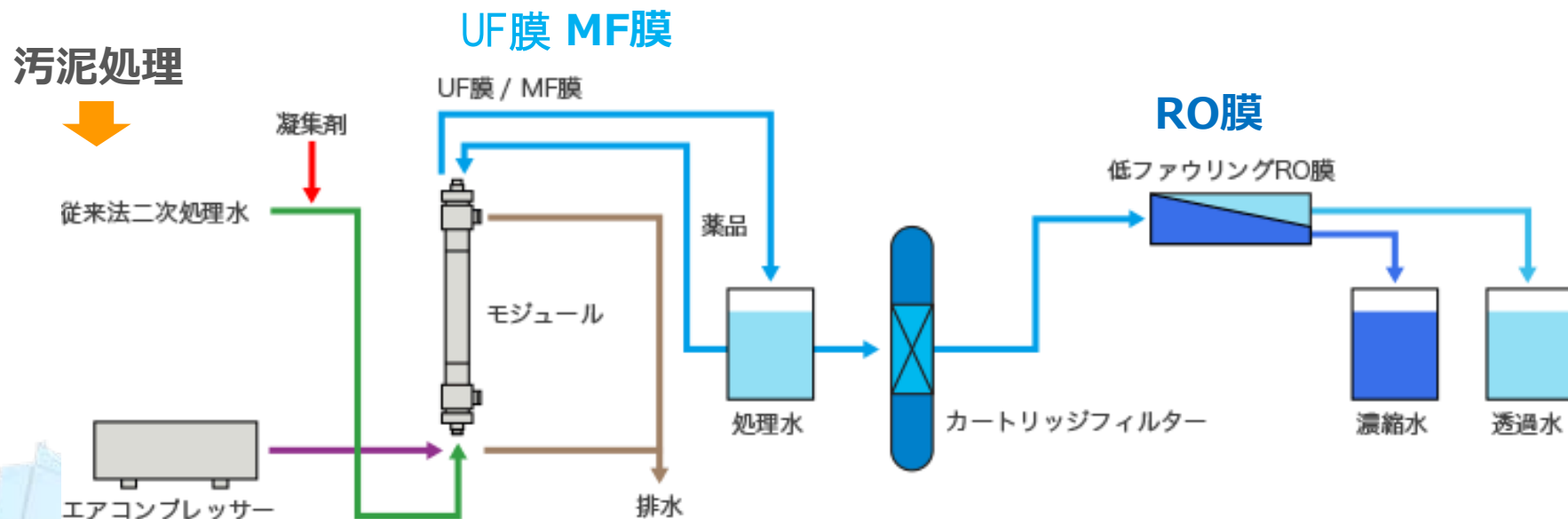


図5 IPC×出願人の件数マトリクス

(自社: **日東電工**) 日東電工の強み・・・逆浸透膜
 (提携候補) 三菱レイヨン・・・中空繊維モジュール、活性汚泥処理
 (競合) 東レ・・・全ての分野に強い

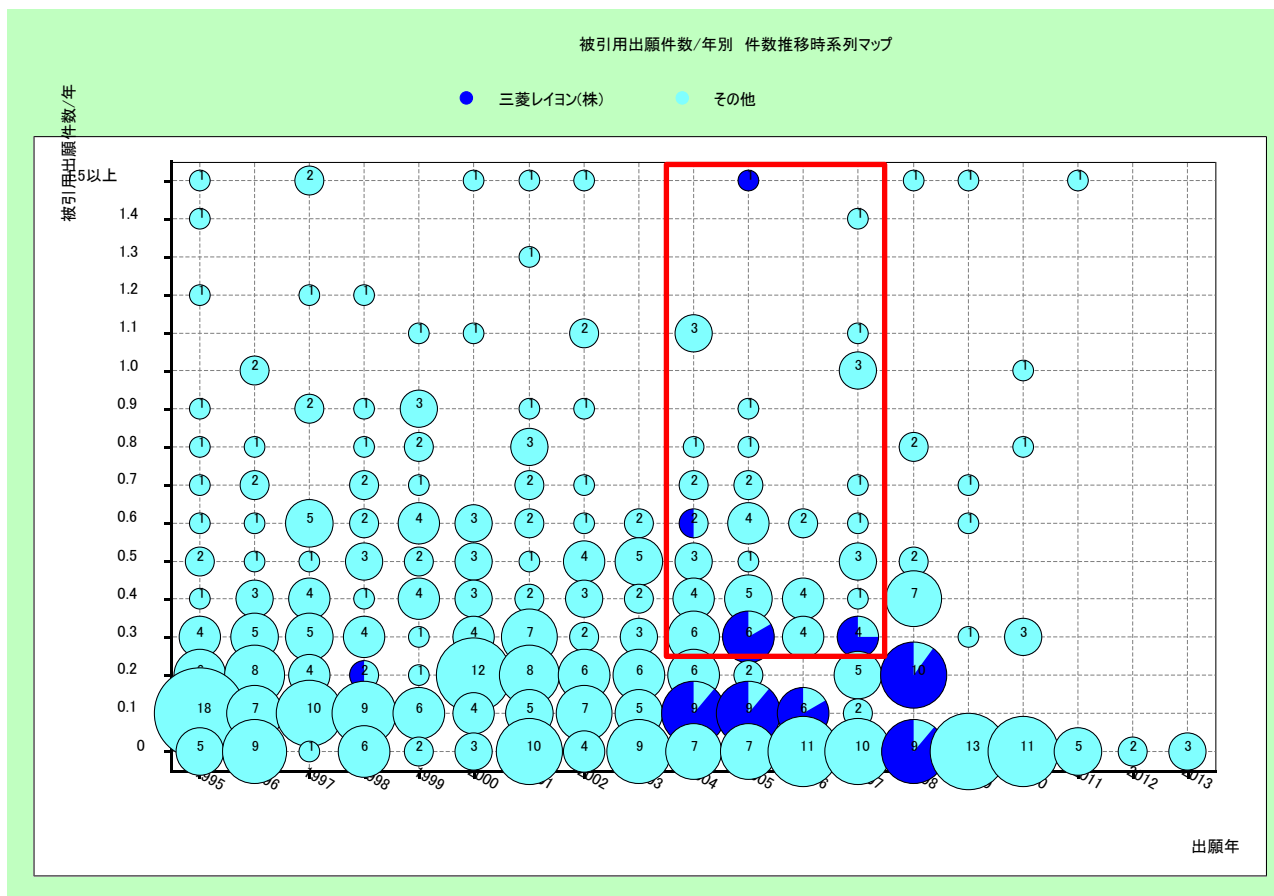
水処理膜分野において、東レは、RO膜～MF膜全ての細孔サイズの膜技術を保有しているため、これら複数の膜を組み合わせた統合システムを提供することが可能



UF膜 限界ろ過膜
MF膜 精密ろ過膜
RO膜 逆浸透ろ過膜

← 自社の特許

出願年と1年間当たりの被引用出願件数のマトリクス



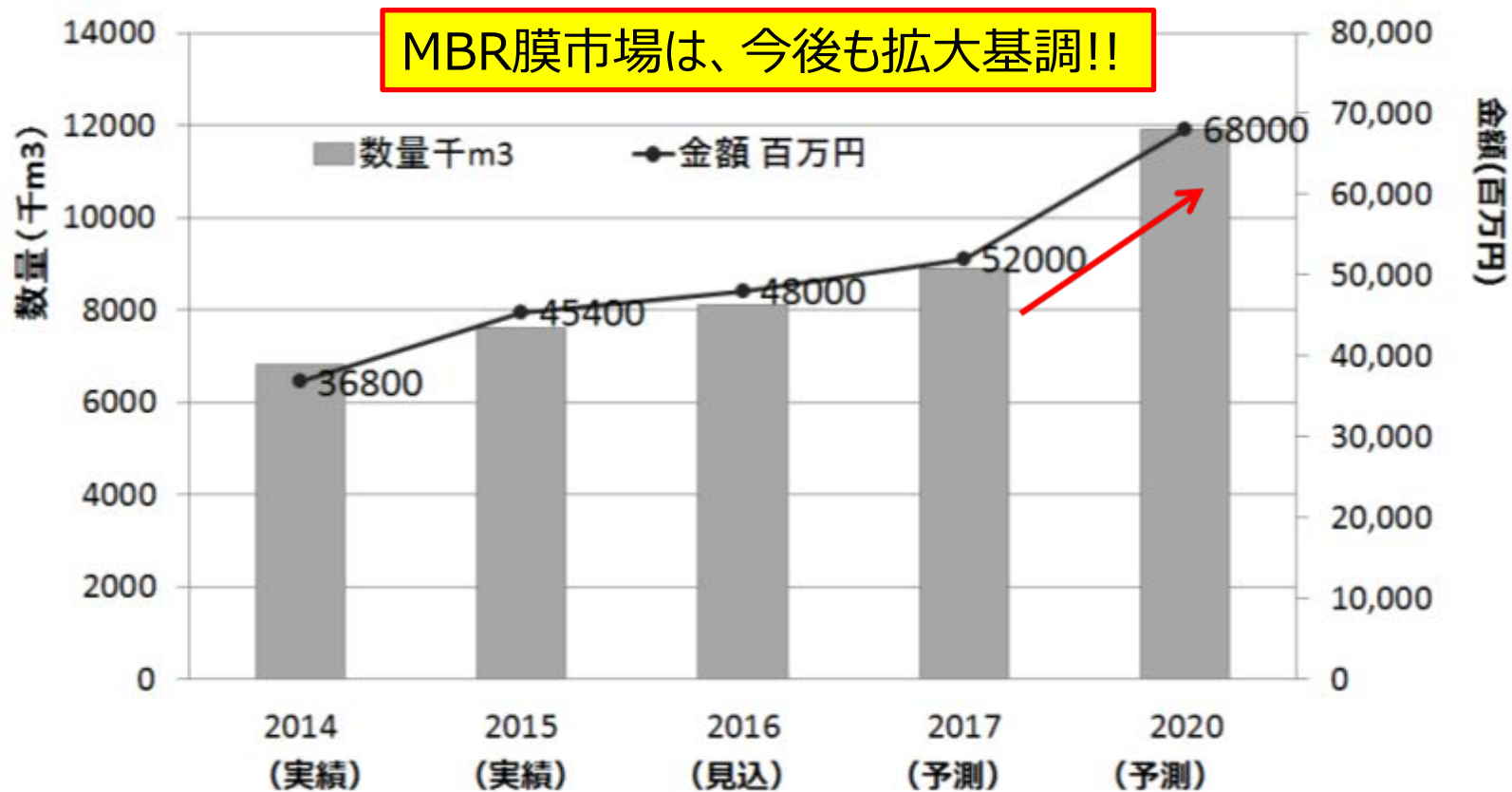
事業実施するうえで、重要特許を確保しておくことは大切。

三菱レイヨンは、現在のところ重要特許を複数保有しており、当分の間はこれらの重要特許が権利切れになることはなさそう。

→日東電工にとって、しっかりと重要知財を保有する相手(三菱レイヨン)と組むことができる

MBR用膜の世界市場

(富士経済 2016年版水資源関連市場の現状と将来展望)



ニュース アクア 2015年01月21日

日東電工株式会社との戦略的パートナーシップ合意書の締結について

三菱レイヨン株式会社

三菱レイヨン株式会社(本社:東京都千代田区、社長:越智仁、以下「当社」)は、世界市場における水環境事業の発展拡大を図るため、日東電工株式会社(本社:大阪市、社長:高崎秀雄、以下「日東電工」)と排水処理及び排水再利用市場における戦略的パートナーシップ合意書を締結しました。

当社と日東電工は長年、水処理事業において両者の保有する技術、ノウハウを共有し、新製品の共同開発や精密ろ過膜の海外販売を軸とした協力関係を築いてきました。こうした連携をさらに強化するため、このたび、再生水市場において高い成長率が見込まれる中国ならびにインド・中東地域をはじめとしたグローバル・バリューチェーンの強化を推進していた当社と日東電工の狙いが合致し、両社関係を全面的な戦略的パートナーシップへと押し上げることで合意しました。

水処理市場は世界的な人口増加や経済成長による水不足、水質悪化を背景に拡大しております。その中でも再生水市場は海水淡水化に比べ低コストで導入可能であり環境負荷も少ないことから工業排水、生活排水向けに需要が伸びています。このような需要の拡大に合わせて、両社の共同開発する製品を今後5年以内に水処理量5,000m³/日以上的大型受注50件を目標に世界の再生水市場への展開を目指します。

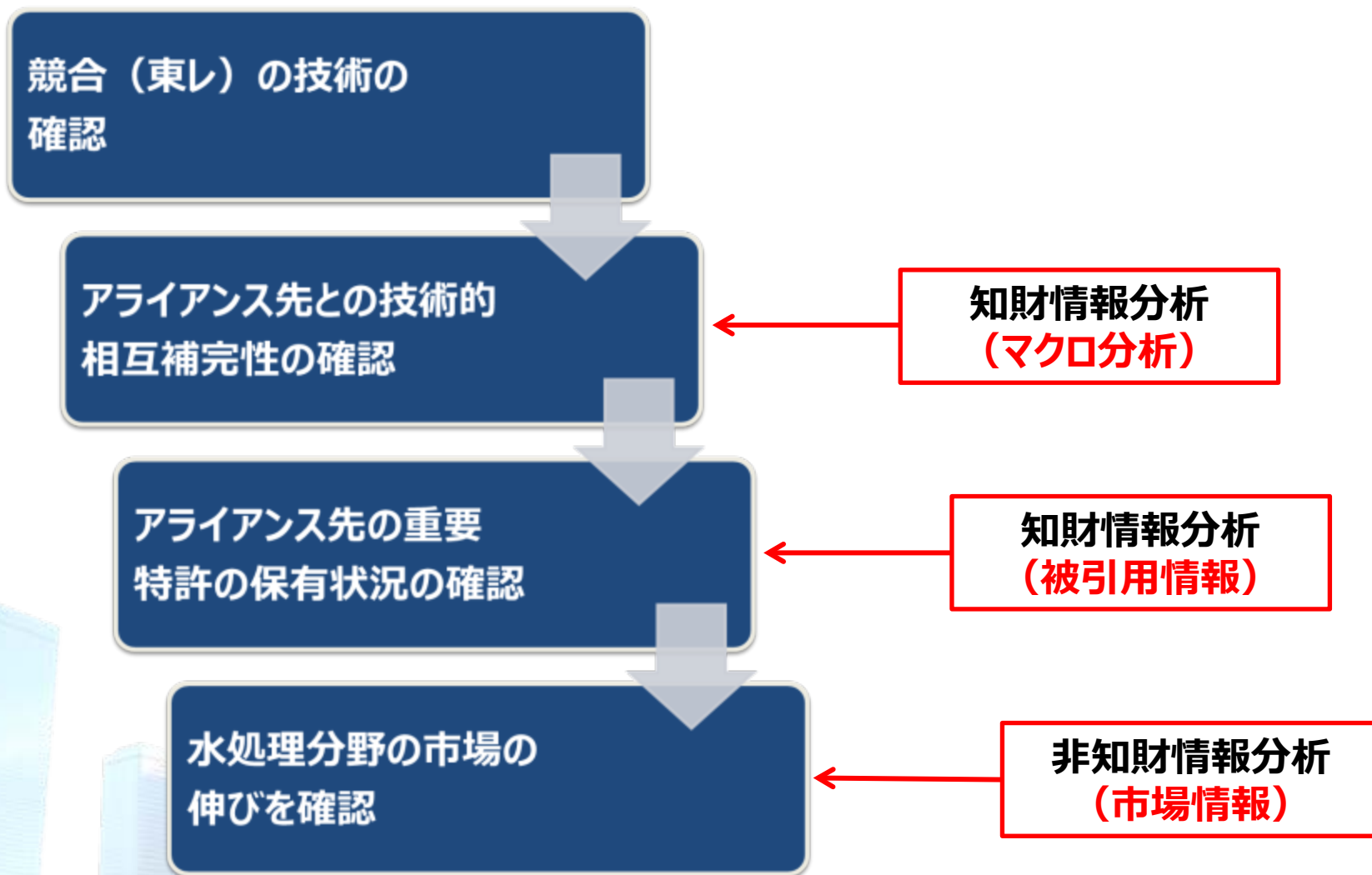
当社は、水環境事業をグローバルに展開し、MBR(膜分離活性汚泥法)による再生水事業、産業・医療分野で活用される中空糸膜フィルター《ステラボア》、浄水器《クリンスイ》を主力事業・主力製品としています。今回の合意により、このうち再生水事業において、当社MBR製品と日東電工が世界的に高い市場シェアを誇る逆浸透膜の組み合わせによる製品・サービスの共同開発や各地域への販売ルート等のリソースを最大限活用し、世界の水環境の改善に貢献して参ります。

日東電工と三菱レイヨンの提携についてニュースリリース有り、
前述までの予測通りの結果。

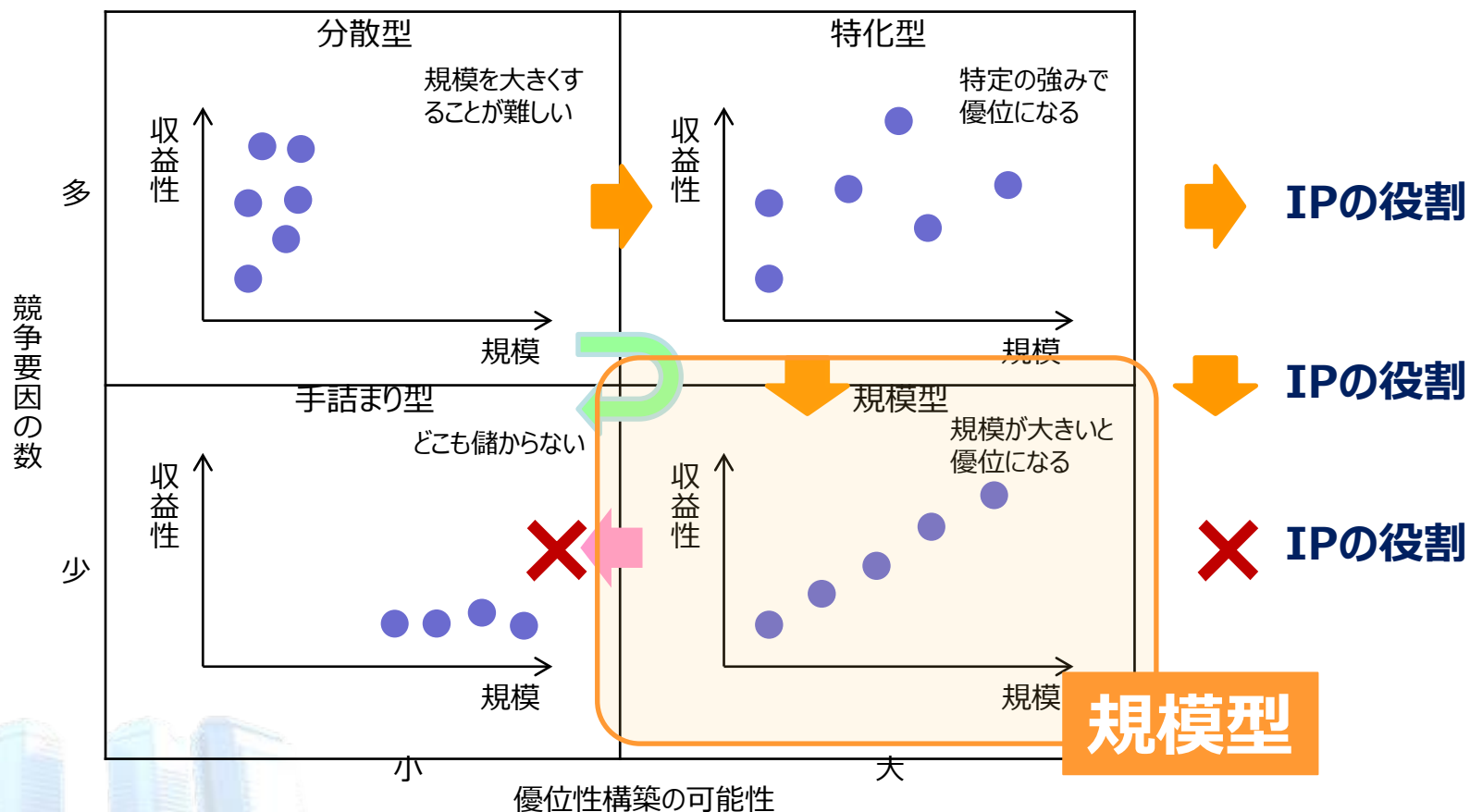
日東電工の逆浸透膜
+
三菱レイヨンのMBRと中空繊維モジュール

↓
アライアンスの結果、相乗効果を生み、複数の種類の膜から最適な組み合わせによる統合システムを提案できる

【アライアンス先探索事例のまとめ】



IPを活用することで、事業環境を変える

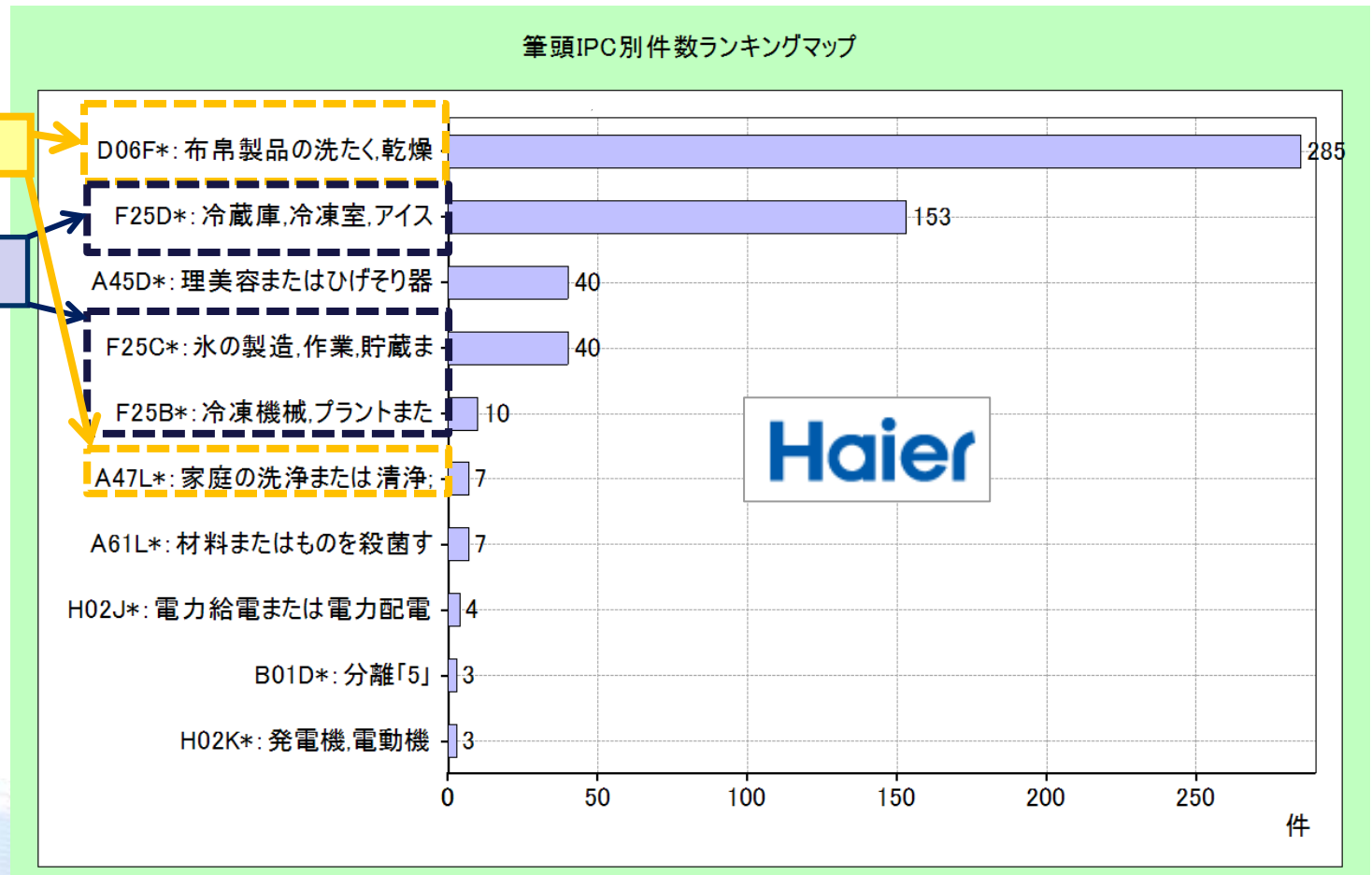


※ボストンコンサルティンググループが考案したアドバンテージマトリクス

【技術提携・買収候補探索の例】

**ハイアールグループの冷蔵庫分野における
技術提携・買収の候補を
特許情報から探索した事例**

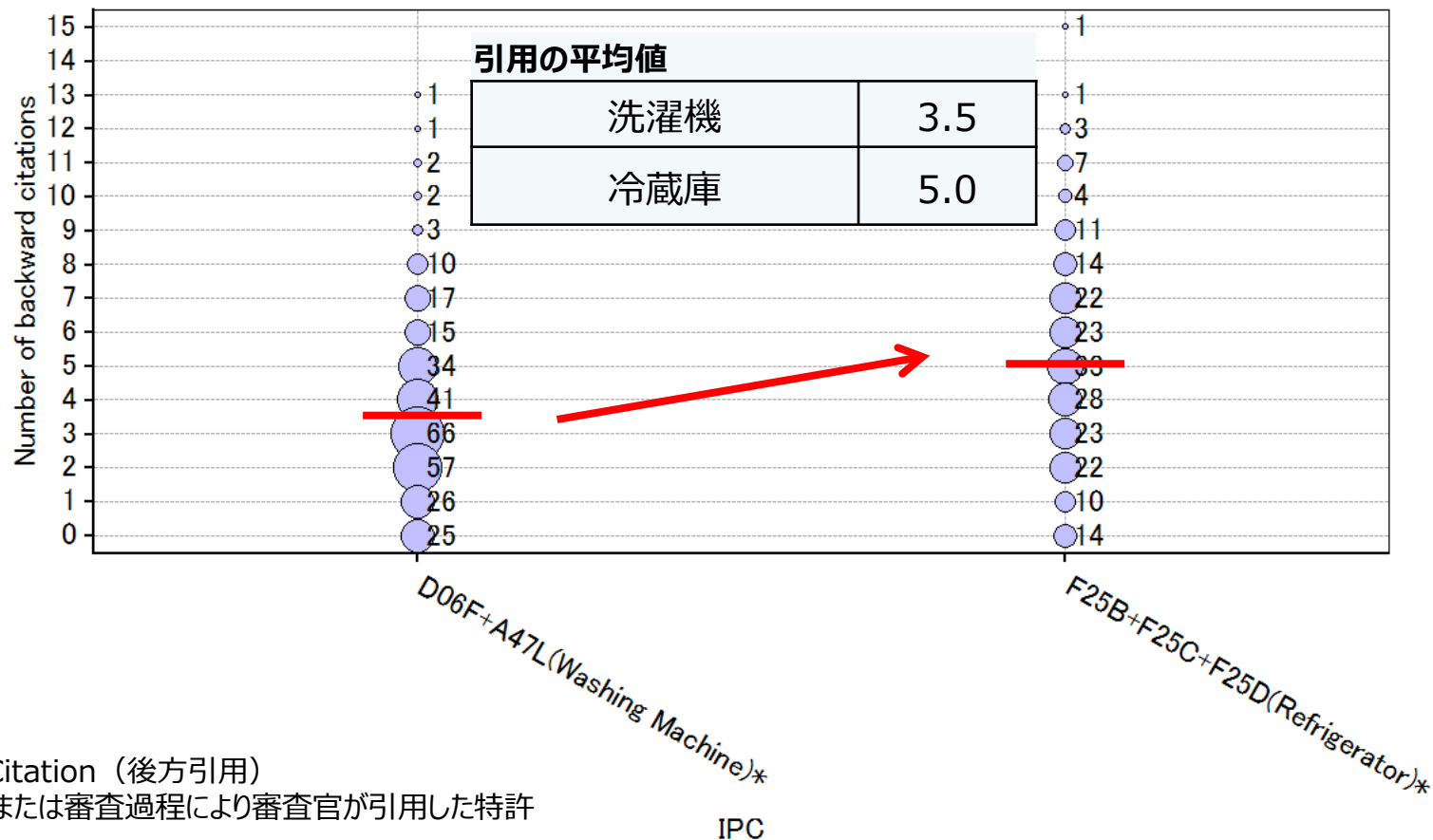
ハイアールの日本出願の分野は、ほとんど洗濯機か冷蔵庫に属している



引用が多い特許ポートフォリオは技術提携が必要と判断

Bubble Map (Number of backward citation × IPC)

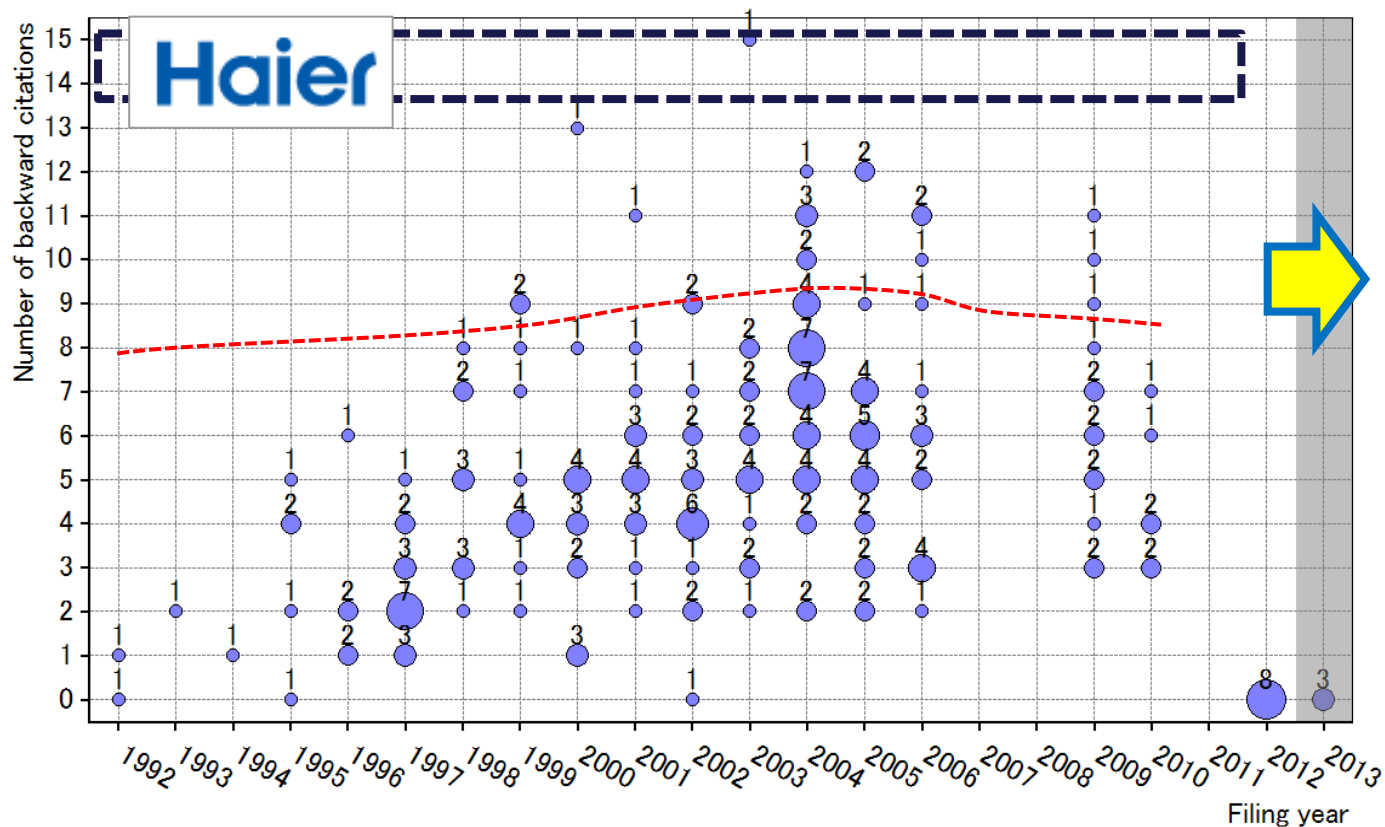
Haier



Backward Citation (後方引用)
出願人自ら、または審査過程により審査官が引用した特許

ハイアールグループの引用の時系列変化

Bubble Map (Number of backward citations × Filing year Matrix)



ハイアールグループの特許に出願人・審査官により引用された特許の出願人の中には、技術提携の候補が存在する可能性がある。

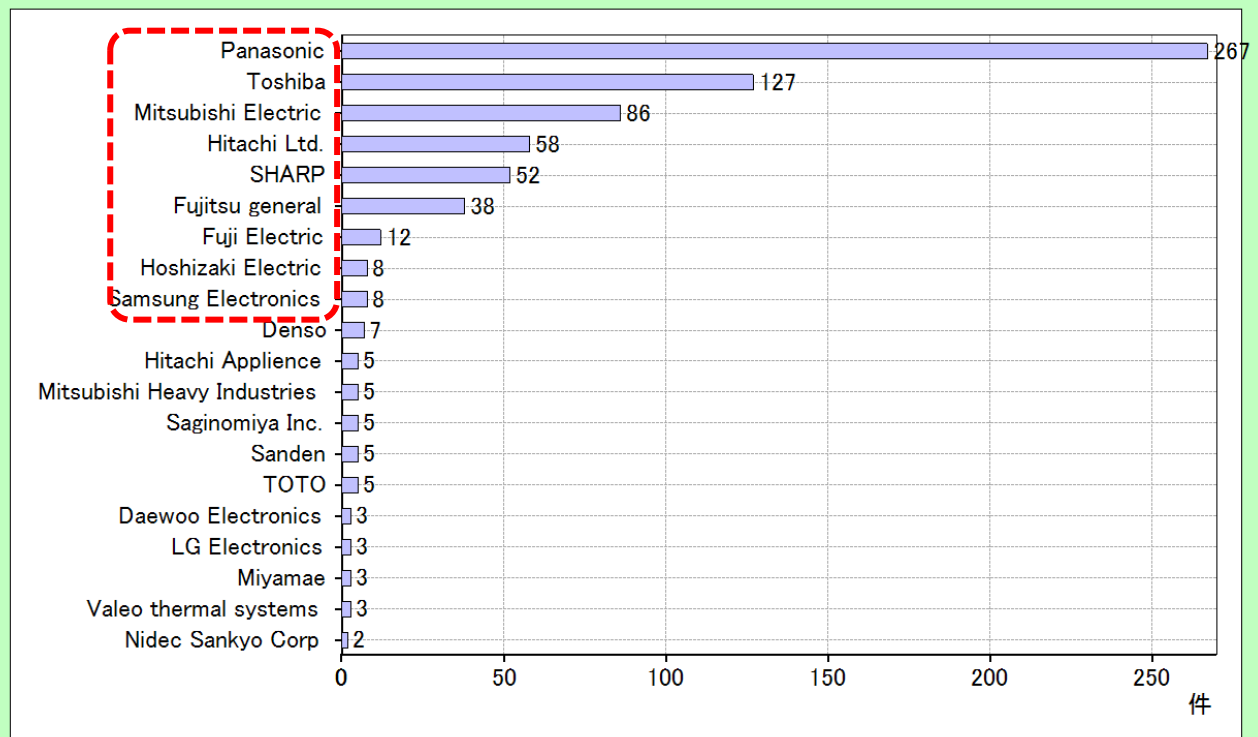
ハイアールの日本特許に関連している特許を母集団として出願人をランキング

ハイアールグループの技術提携の候補をランキング情報から可視化



それぞれの候補が当該技術領域で出願している特許を再度抽出し、分析を行い技術提携の可能性について検討する

出願人別 件数ランキングマップ

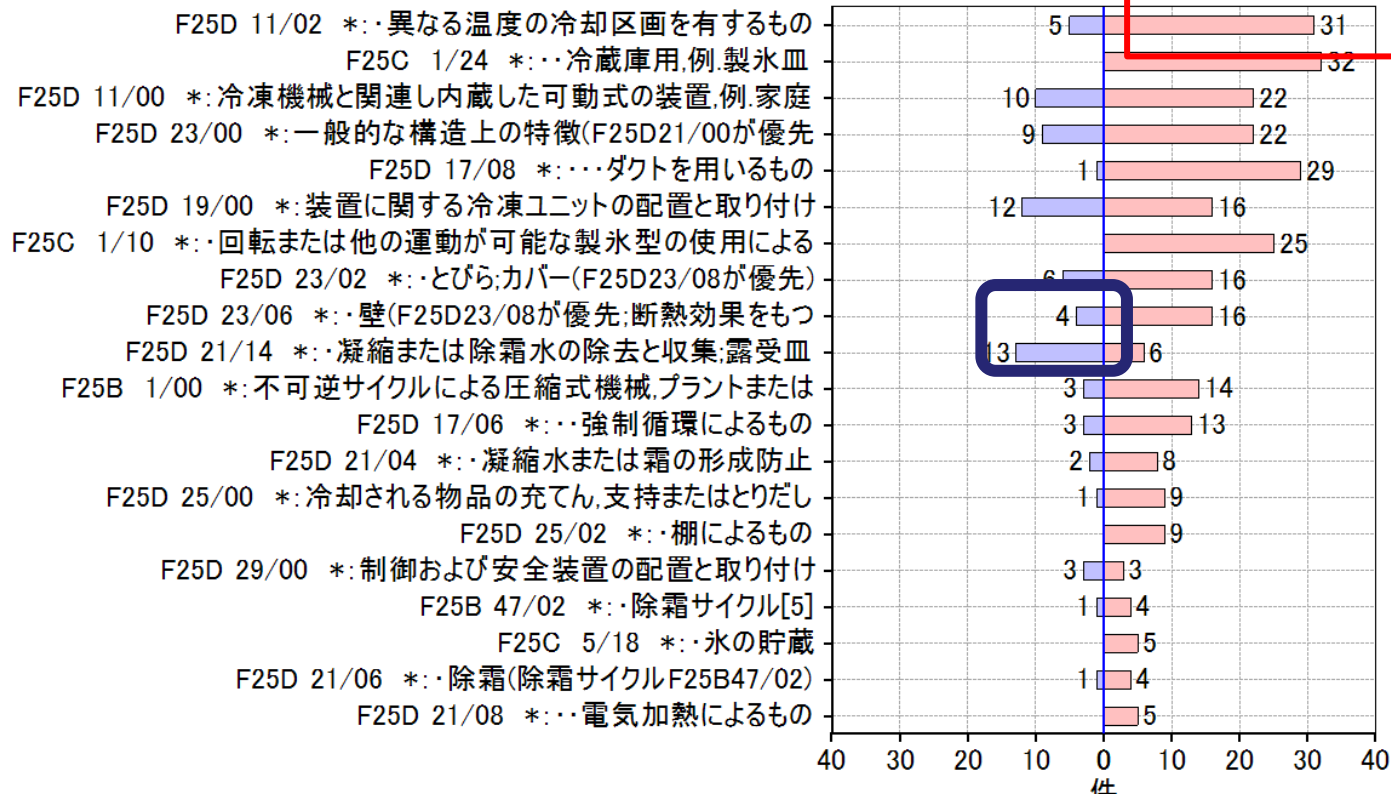


例としてホシザキとハイアールグループの冷蔵庫領域における特許を比較

指定出願人別 件数コンパマップ

■ A: ホシザキ

■ B: ハイアールグループ



ハイアールグループがホシザキ(旧ホシザキ電機)と提携もしくは買収をするメリットは、除霜水の除去や収集(露受皿)の技術にある。逆にホシザキのメリットは製氷皿などの製氷技術である。両社の技術を合わせることで、より冷蔵庫の領域で技術的に質の高い製品を創れる可能性が生まれる。

Haierの冷蔵庫に関し、特許侵害訴訟が頻発

2014年4月 Heifei Meiling **V S** Haier

2016年3月 Hangzhou Naide Refrigeration Electrical Apparatus Factory **V S** Haier

2016年8月 Hangzhou Naide Refrigeration Electrical Apparatus Factory **V S** Haier

☐

Qingdao **Haier** Group Electric Appliance Industry

vs

Hangzhou Naide Refrigeration Electrical Apparatus Factory

CN201030153865 / Front filter (DJS-1-3T)
All cases on this patent (50) / family (50)

China

KATZAROV'S MANUAL

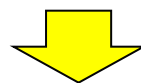


Front filters (前置过滤器)

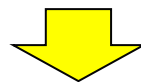


技術提携・買収の候補を探索する方法として、対象領域の特許を調べる。技術的なメリット・デメリットを比較し、候補を抽出できる。

引用情報から技術的に関連が深い複数の候補企業を抽出する



候補企業と特許ポートフォリオを比較し、技術提携のメリット・デメリットをそれぞれ分析する



技術的なメリットやデメリットを特許の詳細、またその他技術情報から分析し、特許情報から導かれる候補企業を絞り込む

IPランドスケープによるセグメンテーションから、ターゲット領域（勝てる市場）がわかり、勝てる戦略を考えることができる。

	勝てる戦略	勝てない戦略
勝てる市場	 誰でも勝てる	 戦略がなくても勝てる場合が多い
勝てない市場	 天才でないと勝てない	 必ず負ける

S:セグメンテーション

T:ターゲティング

P:ポジショニング

【事業探索の例】

**対象企業の現行製品である
電磁波シールドフィルムに関し、
現在のディスプレイ用途以外の
応用展開先を探索した事例**

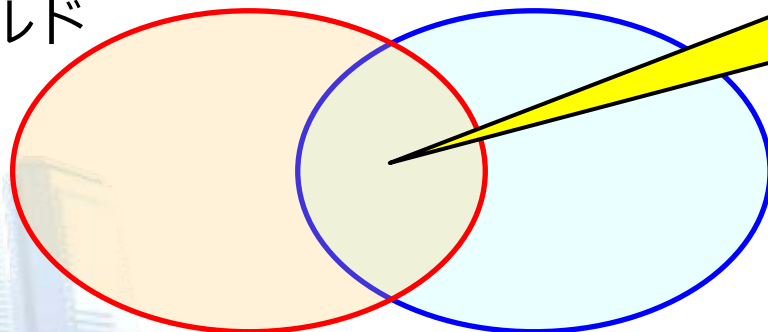
- ◆ 使用データベース
SRPERTNER（日立システムズ）

新用途探索のために「電磁」と
「シールド」を持つ特許を広く探索

- ◆ 母集団の設定（検索実施日：2019/3/20）
要約+請求+発明名：電磁 near3（シールド+遮蔽） [17,128件]
要約+請求+発明名：フィルム+シート [1,215,615件]
出願日：20120101:20171231 [2,405,780件] ---過去6年

⇒614件／代表のみ（764件／すべて）

電磁波シールド



対象母集団

614件

フィルム、シート

特許分類からの調査

- ◆ 国際特許分類（IPC）ランキング
- ◆ IPC別成長率比較（特許件数の増加率）

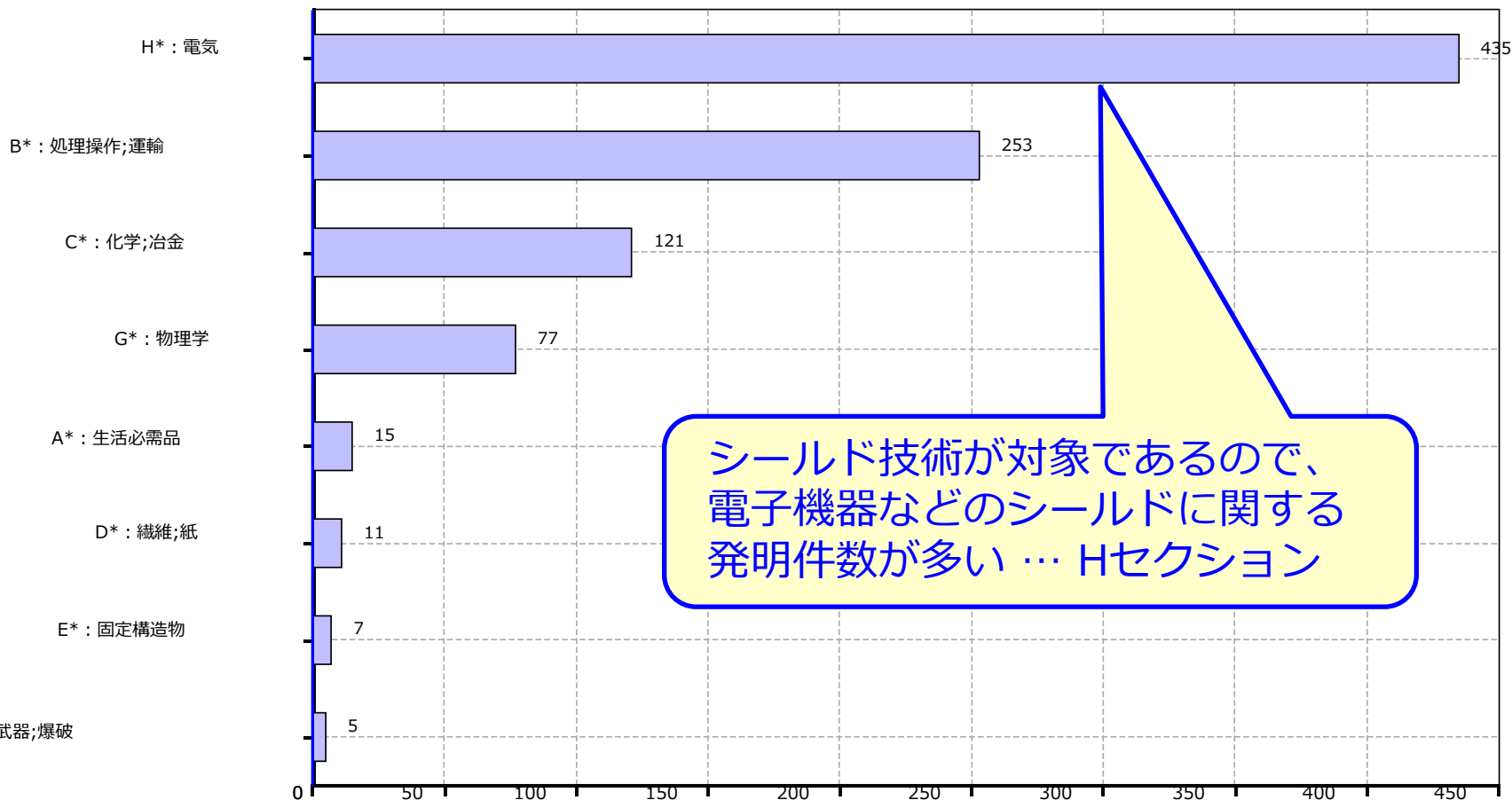


特許出願数が伸びている
IPCに着目し、内容調査

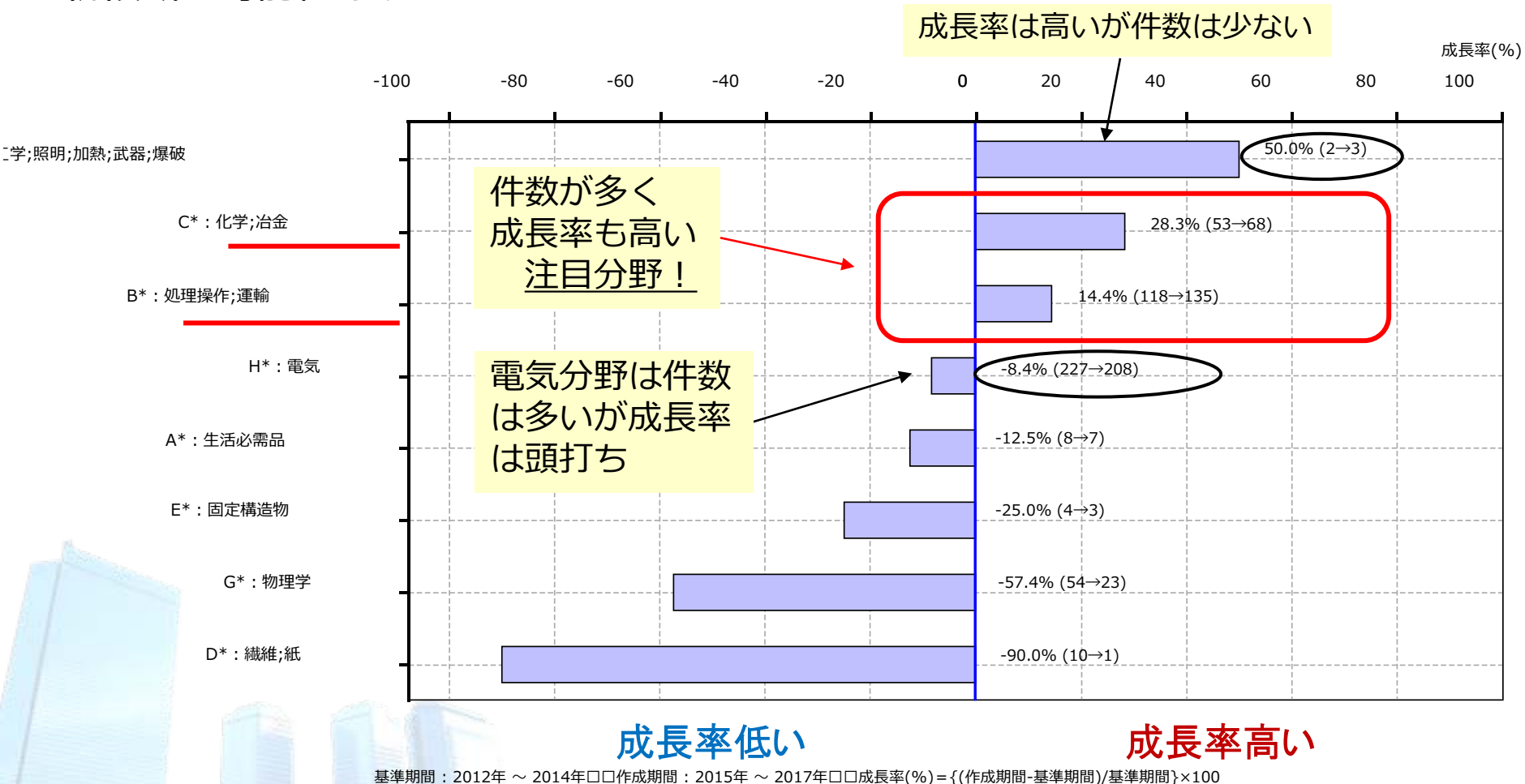


新規用途の抽出

国際特許分類（IPC）で母集団を分類すると、電気分野の技術が多いように見えるが・・

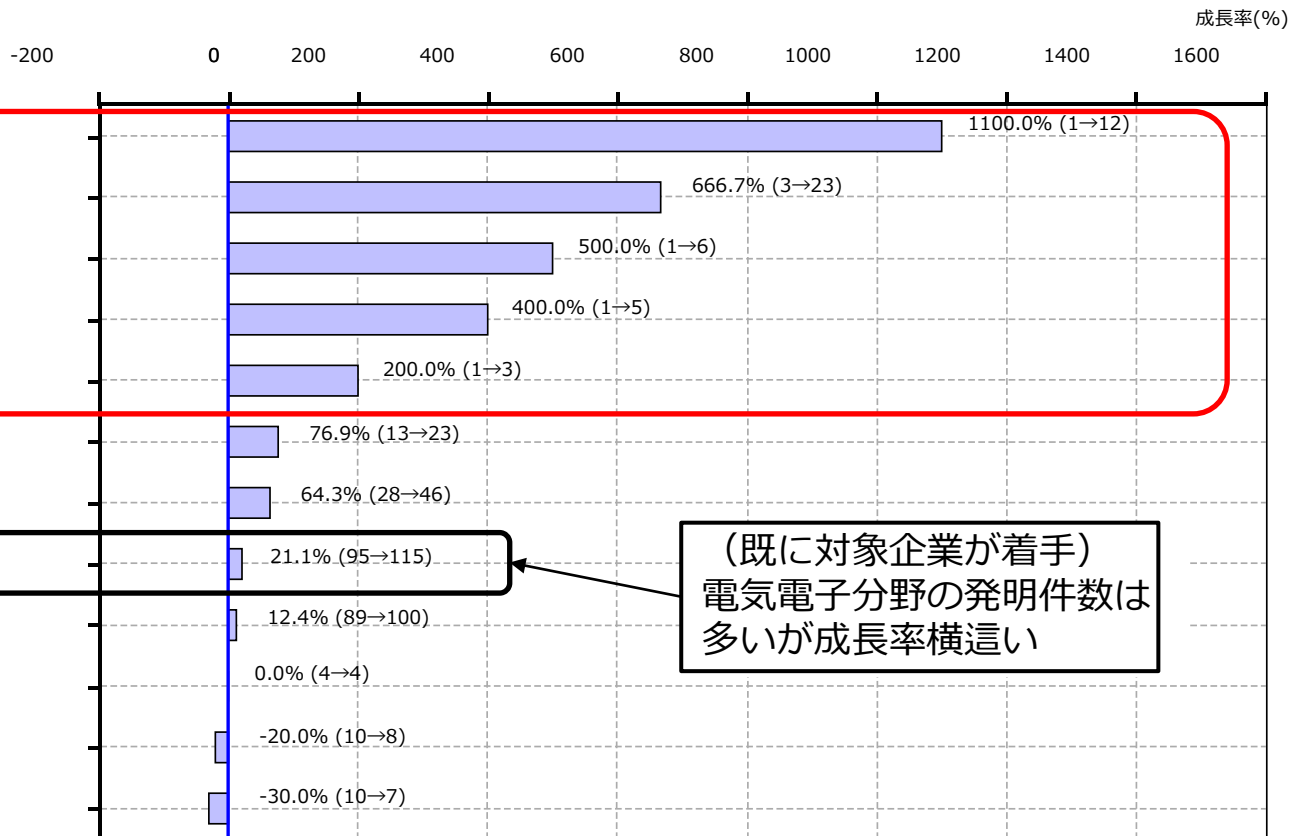


特許件数の増加率である成長率で見ると、電気とは異なる分野が高い。
新領域の可能性ありか？



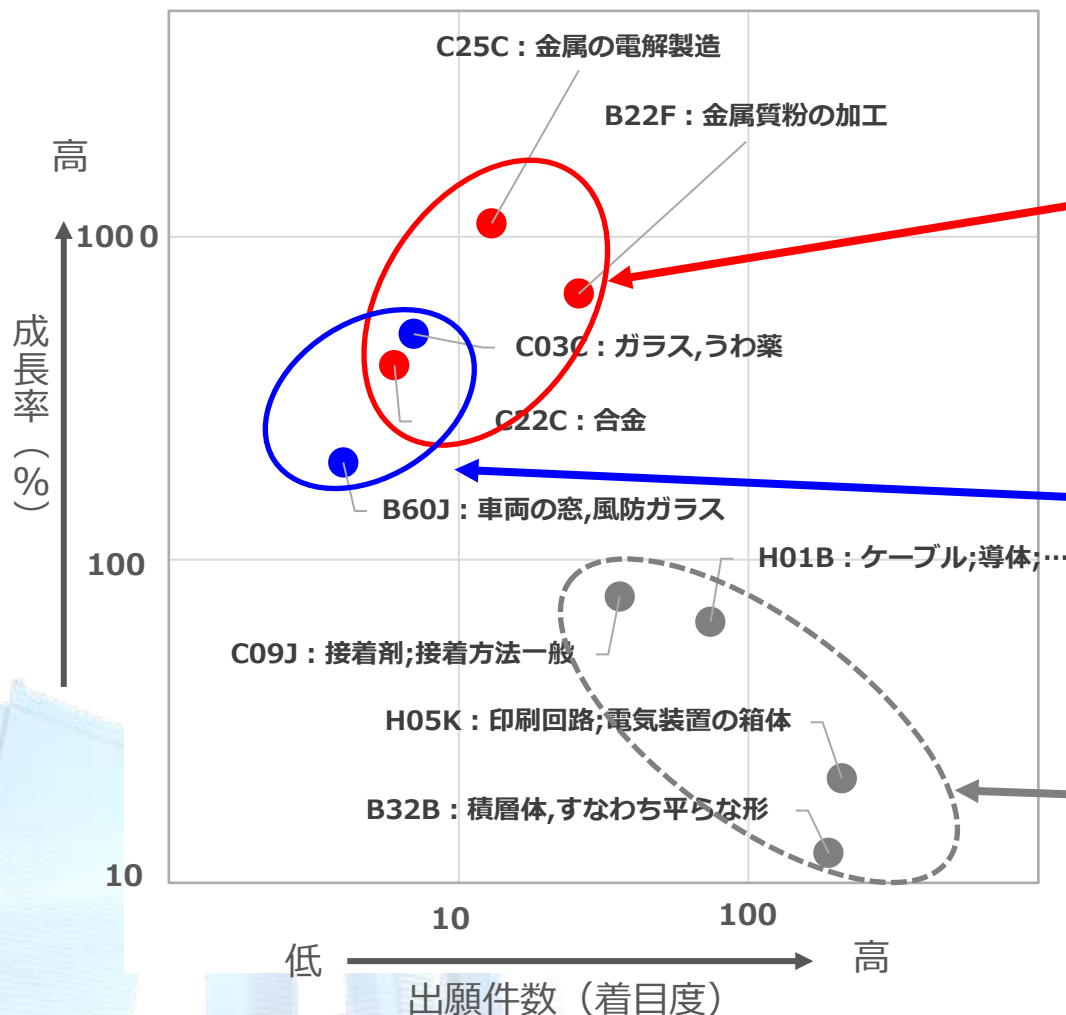
成長率に注目し、特許分類をより深掘して、新しい領域のヒントを探す

ランキング上位に着目



基準期間：2012年～2014年 □□作成期間：2015年～2017年 □□成長率(%) = {(作成期間-基準期間)/基準期間} × 100

縦軸に特許件数（着目度）、横軸に増加率（成長率）をプロットして比較をする
着目度は低くても、成長率が高い領域は有用な候補



出願件数中くらい、成長率大

電磁波シールドに用いる銅粉
(新たな付加価値を有する銅粉)
--- PC/携帯の筐体、配線板用途

出願件数少、成長率大

自動車のフロントガラスなどに
使用される張り合わせガラス

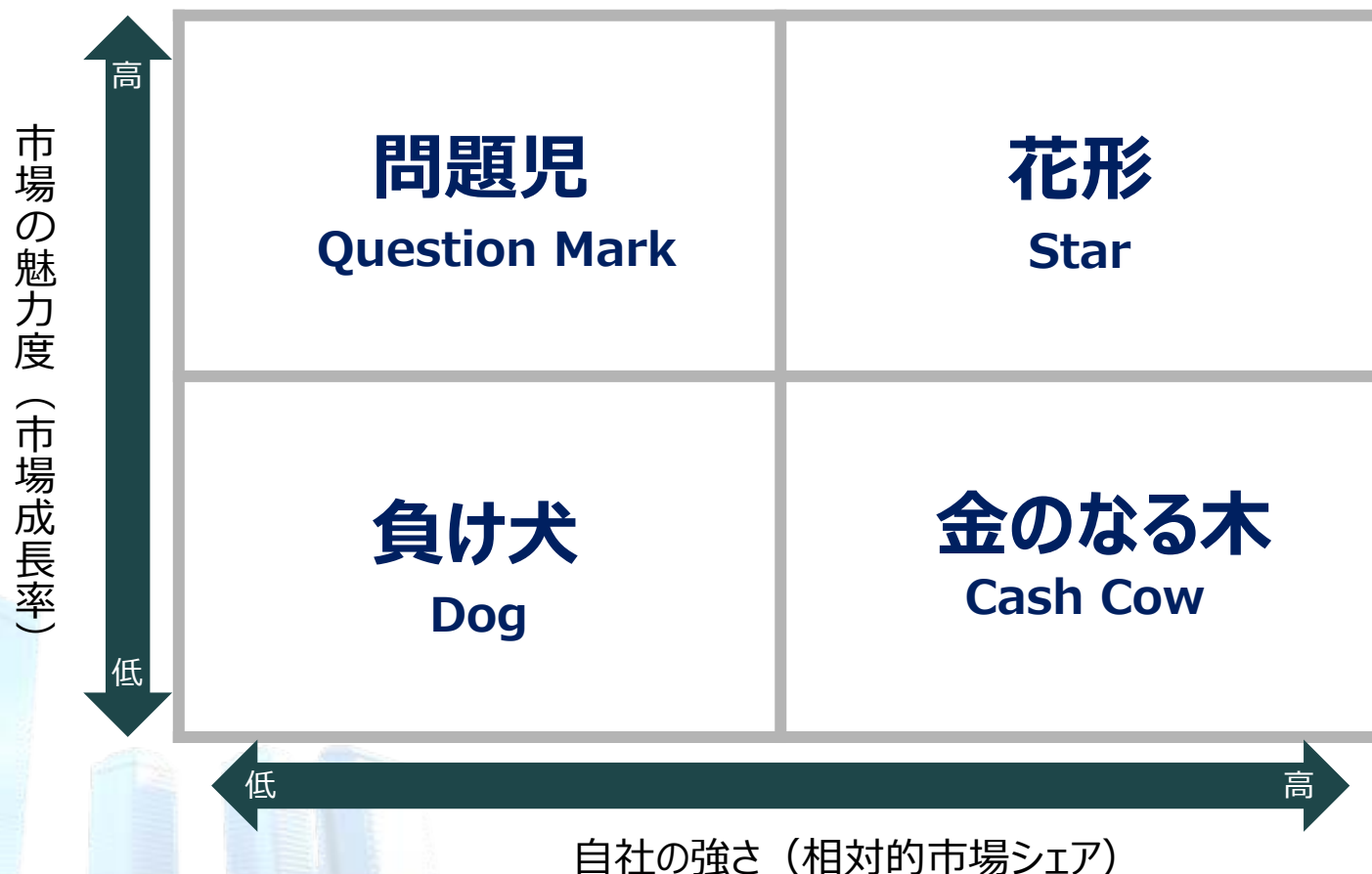
- 電磁遮蔽機能あり
- 電波は遮蔽しない

出願件数多、成長率中

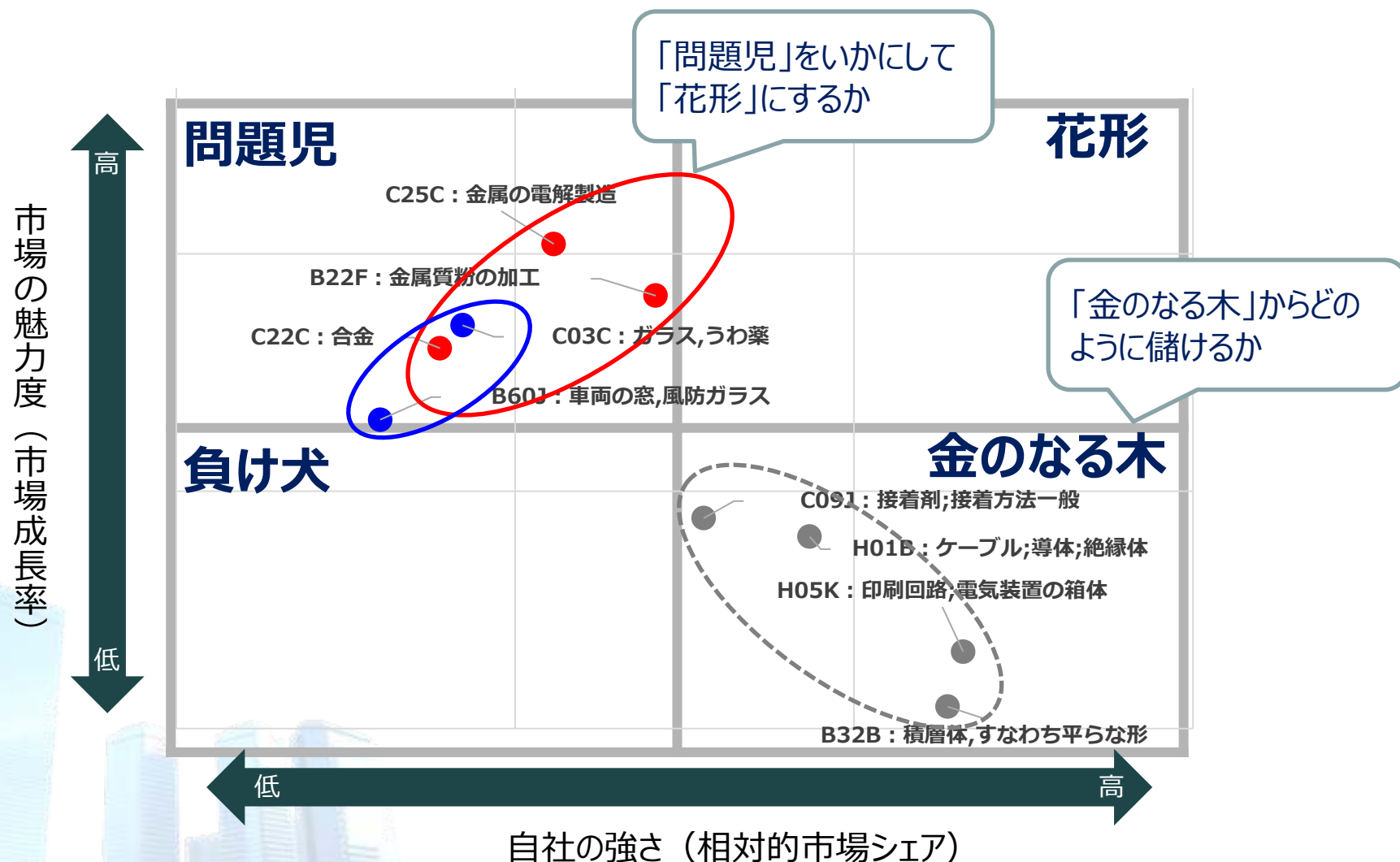
主に電磁波シールドに関わる
従来型の配線板及びケーブル
--- 既に出願多く、技術的に飽和

・ 2014年□□作成期間：2015年～2017年□□成長率(%) = {(作成期間-基

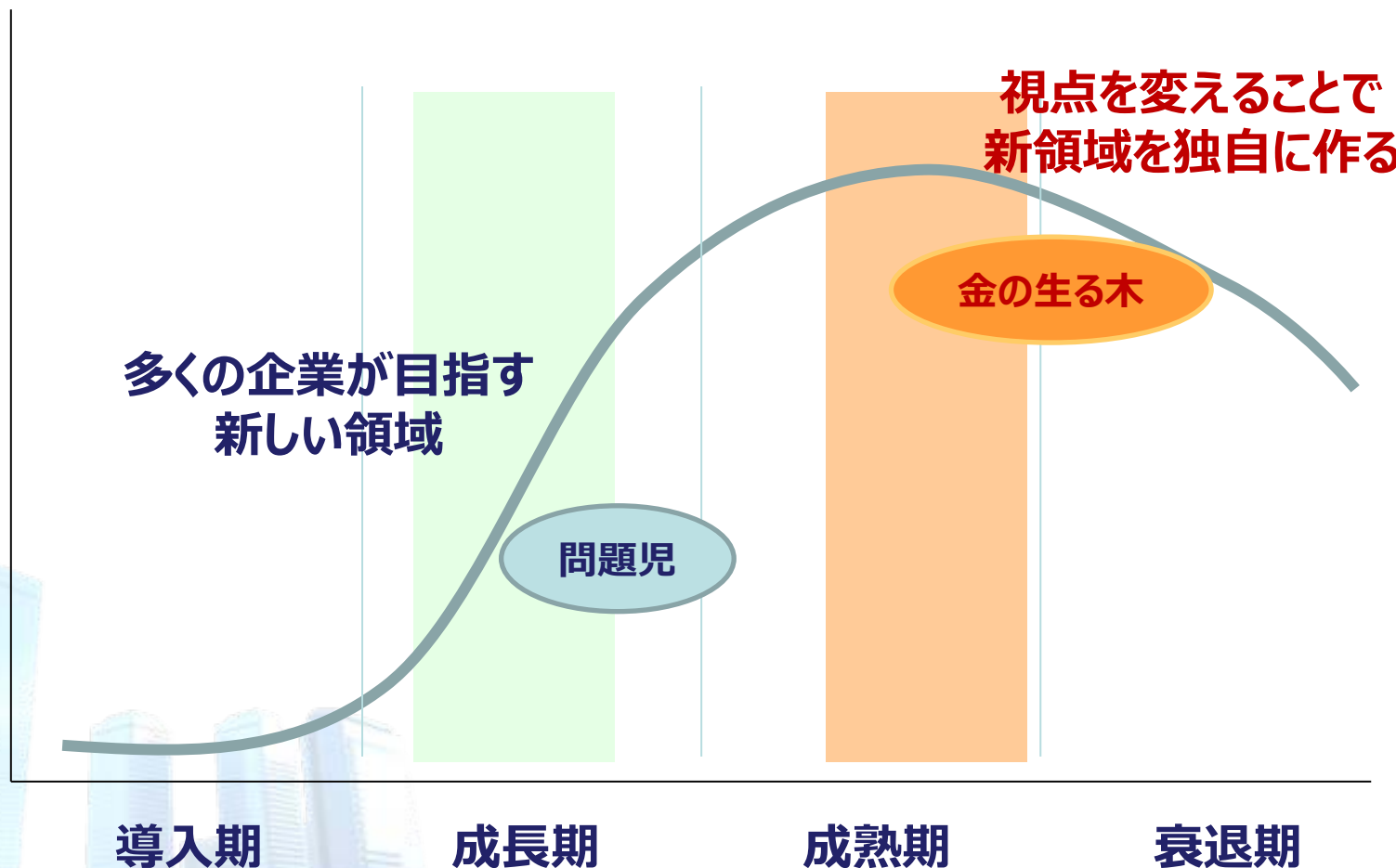
特許情報から見えることは、経営視点での技術ポートフォリオマネジメント



問題児を見つけることで、花形へ成長させる。金のなる木から儲ける。



成長率が高い分野は新領域の可能性は高い
しかし、成熟期にも、視点を変えることで、独自の領域を作る可能性がある



本事例では以下の分野を抽出した；

1. 多層板を含むリジッド配線板及びパソコンや携帯の筐体
 - ◆ 効率的な形状を有する銅粉がコア技術の出願が多い
 - ◆ 多層化、微細配線化の進むリジッド配線板の分野ではクロストークやセンシティブなデバイス実装の観点から電磁波シールドフィルムの重要性は増す
2. 自動車関連などの合わせガラスの部材
 - ◆ 導電性フィルムを挟んだ合わせガラス：電磁波遮蔽タイプ
 - ◆ 熱線などは遮蔽するがラジオ・テレビなどの電波は透過させる合わせガラス：電波透過タイプ
 - ◆ 波長選択性（遮断／透過）は新たな付加価値となり得る

抽出した視点をきっかけに、より深い検討につなげる

- ・ 成長率がこれからの分野は他にもあるのか
- ・ 成長率が低い領域にも、まだ残っている市場はないのか
- ・ 今までのビジネスモデルを変えるヒントは存在するか

【代表例-1】電磁波シールドに用いる銅粉

文献

発明の名称	Niコート銅粉、及びそれを用いた導電性ペースト、導電性塗料、導電性シート、並びにNiコート銅粉の製造方法
出願番号	PCT/JP20**/XXXXXX
公報番号	WO20**/XXXXXX
出願日	20**/**/**
ステータス	出願中

【背景技術】

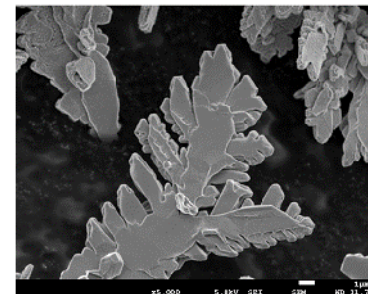
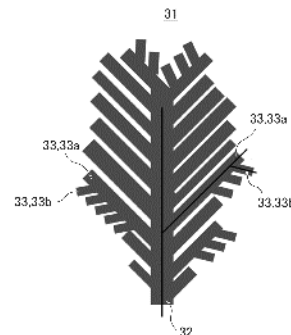
【0005】

電磁波シールドは、**電子機器からの電磁気的なノイズの発生を防止するために**使用されるもので、特に近年では、**パソコンや携帯の筐体**が樹脂製になったことから、**筐体に導電性を確保するために**、蒸着法やスパッタ法で薄い金属皮膜を形成する方法や、導電性の塗料を塗布する方法、導電性のシートを必要な箇所に貼り付けて電磁波をシールドする方法等が提案されている。その中でも、樹脂中に金属フィラーを分散させて塗布する方法や**樹脂中に金属フィラーを分散させてシート状に加工してそれを筐体に貼り付ける方法では、加工工程において特殊な設備を必要とせず自由度に優れており多用されている。**

要約

【要約】

Niを被覆した銅粉同士が接触する際における接点を多くして優れた導電性を確保しつつ、銅粉同士の凝集を防止して、導電性ペーストや電磁波シールド等の用途として好適に利用することができるNiコート銅粉を提供する。本発明に係るNiコート銅粉は、表面にNi又はNi合金が被覆された**特定の形状の銅粒子が集合して、樹枝状形状あるいはその銅粒子の凝集体の形態を有してなるNiコート銅粉である。**このようなNiコート銅粉によれば、その銅粉同士の接点が多くなり、優れた導電性を示す。



Cu粉に特徴のある導電性組成物

【代表例-2】自動車用遮熱合わせガラス

文献

発明の名称	自動車用遮熱合わせガラス
出願番号	特願20**-XXXXX
公報番号	特開20**-XXXXX
出願日	20**/**/**
ステータス	出願中

要約

【要約】

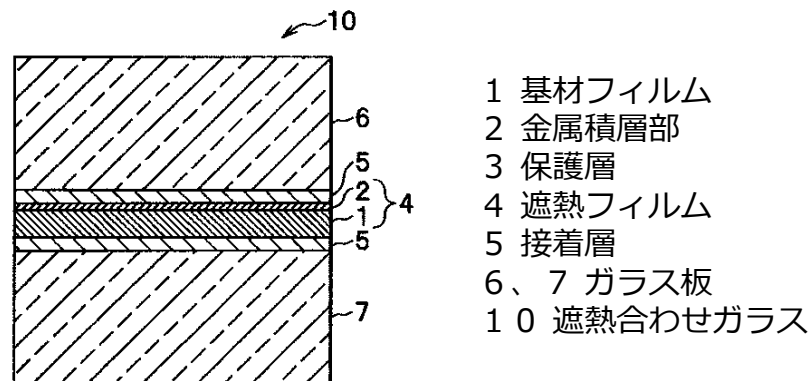
【課題】熱線の遮蔽性能、可視光線の透過性能、ヘイズ、電磁波の透過性能、選択的な分光透過性能に優れ、生産性にも優れた自動車用遮熱合わせガラスを提供する。

【解決手段】2枚のガラス板によって遮熱フィルムが挟まれた構成を有する自動車用遮熱合わせガラスである。遮熱フィルムは、**基材フィルムの表面に金属層と金属酸化物層とからなる金属積層部を有するものであり**、金属層が銀、または銀と他の金属からなる化合物を含有する層であり、金属酸化物層が亜鉛ドーパド酸化インジウムまたは錫ドーパド酸化インジウムを含有する層である。金属層と金属酸化物層は交互に形成されており、金属層と金属酸化物層の層数の合計が5以上である。金属積層部は、格子状のクラックによって複数の島状部に分断されており、島状部の一辺の長さの平均値が80～500μmである。可視光線透過率70%以上、波長850nmの分光透過率35%以上、ヘイズ2%以下、**電磁波遮蔽率10dB以下**である。

【背景技術】

【0002】従来から、自動車等の交通機関の省エネルギー対策の一つとして、**熱線遮蔽性能を有した透明材料**の開発が進められている。例えば、車窓から降り注ぐ太陽光線のうちの**可視光線は透過するが、熱線は遮蔽し**、かつ車内の熱を外部へ逃がさないための断熱機能を有した窓板用透明材料が開発されている。

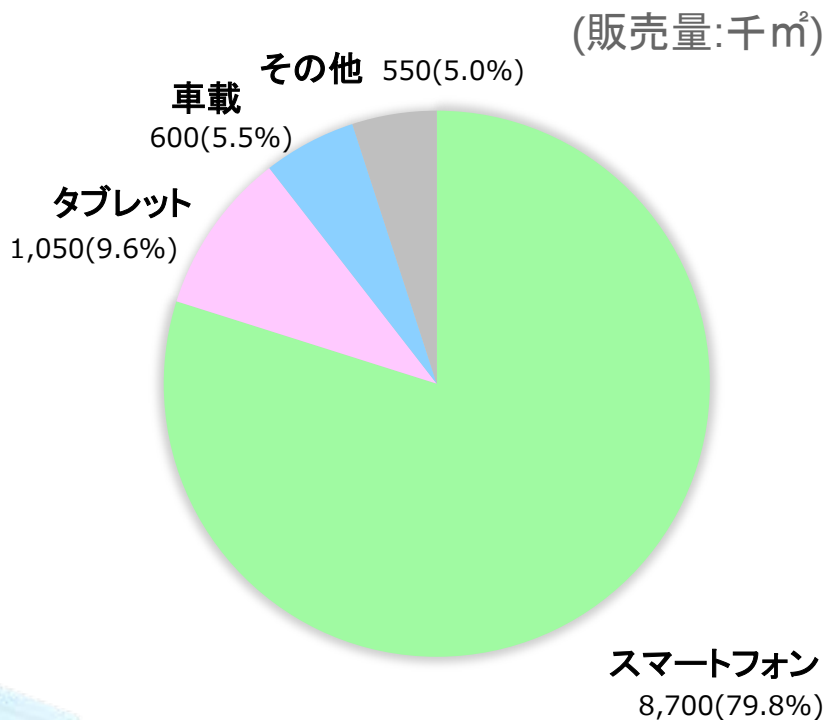
【0004】ところが、このような均一な金属層は、一般に電磁波を反射するため、車内において携帯電話や携帯テレビ等を使用することが困難になるといった問題が生じることがある。そこで、**熱線は遮蔽するが、電磁波は透過させるといった機能を有したガラス板やフィルム**の開発が進められてきている。



- 1 基材フィルム
- 2 金属積層部
- 3 保護層
- 4 遮熱フィルム
- 5 接着層
- 6、7 ガラス板
- 10 遮熱合わせガラス

電波は透過する合わせガラス

FPC用電磁シールドフィルム：用途別ウェイト



2017年見込

需要の大半はスマートフォン向けである。ディスプレイ、タッチパネル、カメラモジュール、指紋認証などFPCの使用箇所に応じて使用される。

スマートフォン以外の用途では、カーナビなど車載向けが拡大傾向にあり、今後車載電装機器のFPC化の進展による需要拡大が期待されている。

出典：2018年 エレクトロニクス先端材料の現状と将来展望
株式会社富士キメラ総研（2018年1月17日発行）

電磁波シールドフィルムはスマートフォンなどのディスプレイ用途が大半を占めており、今後車載・その他への展開が期待されている

IPランドスケープを経営のコンパスに



ニュースを対象とした日英機械 翻訳システムの研究開発

N H K 放送技術研究所
後藤功雄

2023/2/20 第7回特許情報シンポジウム

NHK

内容

- NHKでの日英機械翻訳の利用状況
- 英語ニュース制作の課題
- これまでの研究開発

内容

- NHKでの日英機械翻訳の利用状況
- 英語ニュース制作の課題
- これまでの研究開発

NHKの英語ニュース配信

■ 国際放送 **TV&Radio**

■ インターネットサービス

- **NHK WORLD-JAPAN News**

- スマホアプリによる通知

- **英語字幕**付き総合テレビ特設ニュースのライブ配信

制作支援に機械翻訳を利用

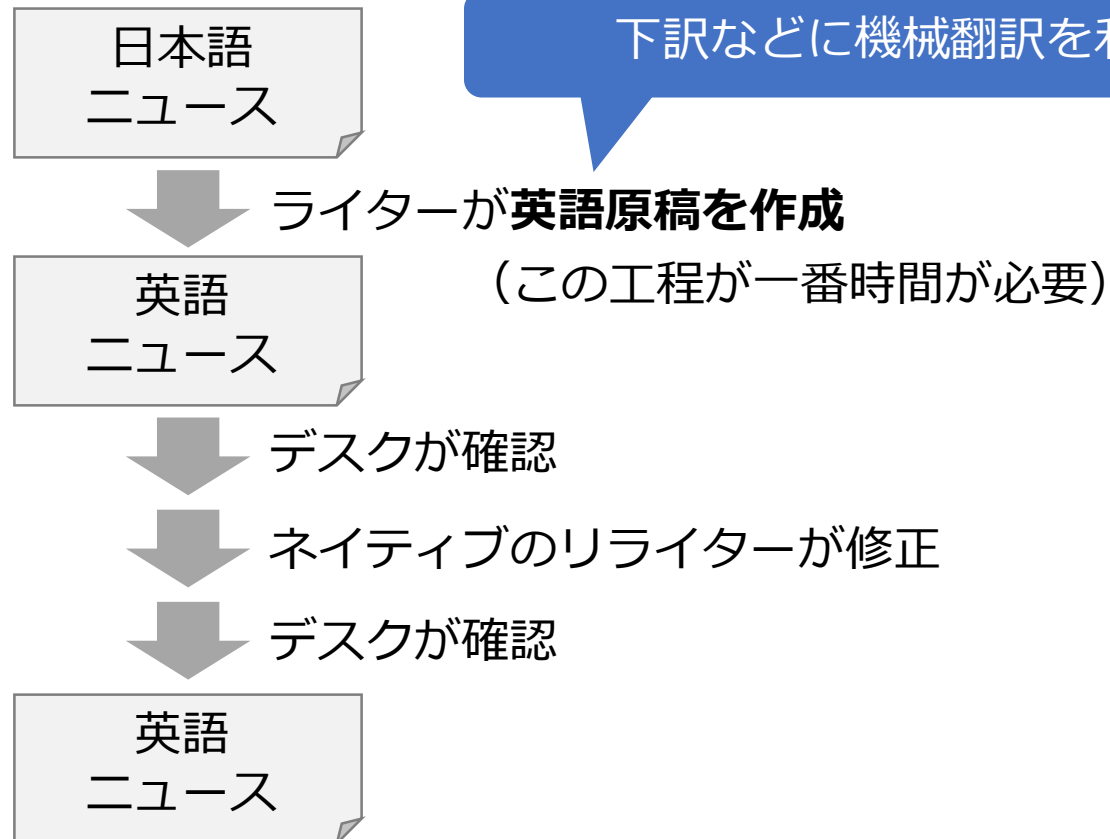
英語ニュース原稿を利用

日本語字幕を機械翻訳で英訳

■ テレビニュース 2ヶ国語放送 英語副音声

英語ニュース制作支援での利用

制作工程



国際放送局担当者からのコメント

2019年台風19号の際、機械翻訳の利用で通常より1.5～2倍の出稿数

短時間で出稿できることで、刻々と変わる状況をきめ細かく出せる

泊まり、土日のライターが手薄で多くの出稿が必要になった時に助かる

特設ニュースの英語字幕ライブ配信での利用

(2022年6月開始)

- 総合テレビ特設ニュースにA I 翻訳の英語字幕を付けてインターネットでライブ配信
- 総合テレビの日本語字幕データをA I 翻訳で英訳し、英語字幕として利用
- 震度5弱以上の地震発生時、津波注意報・津波警報・大津波警報・大雨特別警報などの発表時に原則実施
- ライブ配信ページには、AI翻訳のため、正確な表現ではない場合もある「おことわり」を掲載



内容

- NHKでの日英機械翻訳の利用状況
- 英語ニュース制作の課題
- これまでの研究開発

英語ニュース原稿の一般的な特徴（一部）

[Papper, 2020; David and Peter, 2019; Block, 2011; Strunk and White, 1999]

① 逆ピラミッド型

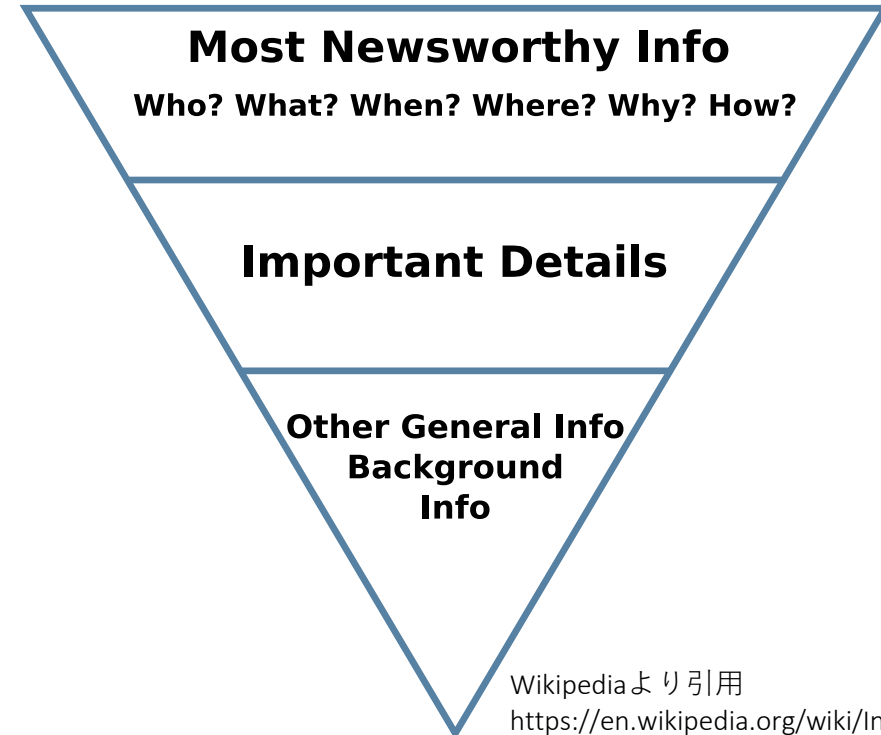
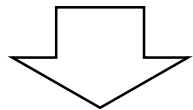
最初に、ニュースで最も重要なことを伝え、
つぎに、重要なことの詳細を伝え、
その後、その他の背景情報などを伝える

② 短く、シンプル

同じ情報の重複がない。

③ 同じ表現の繰り返しを避ける

etc...



Wikipediaより引用
[https://en.wikipedia.org/wiki/Inverted_pyramid_\(journalism\)](https://en.wikipedia.org/wiki/Inverted_pyramid_(journalism))

英語ニュースは日本語ニュースの単なる翻訳ではない

NHK日英ニュースの違いの例 [後藤, 2022]

気象庁「東京でサクラ開花」発表 平年より12日早く

Start of cherry blossom season declared in Tokyo

リード

- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
- [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.

本文

- [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
- [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
- [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
- [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
- [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

2021年3月15日のニュース

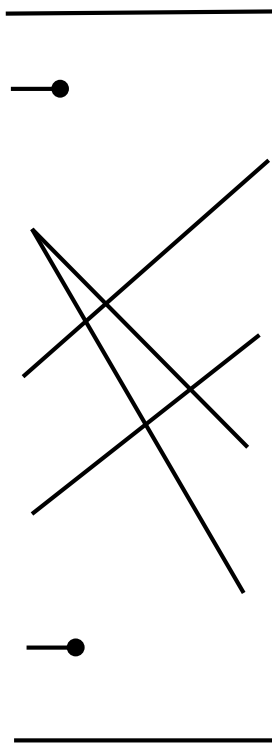
NHK日英ニュースの違いの例

[後藤, 2022]

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
 - [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

青：対応あり 紫：対応あり（意識・変換） 赤：対応なし

リードの違い

東京で桜が開花した
去年同様最も早い開花

発表 平年より12日早く

[1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。

[2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。

[3] 気象庁によりますが、東日本や西日本など、春の陽気となりまし

[4] 東京 千代田区の担当者が訪れ、開花の日々となっているサクラに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。

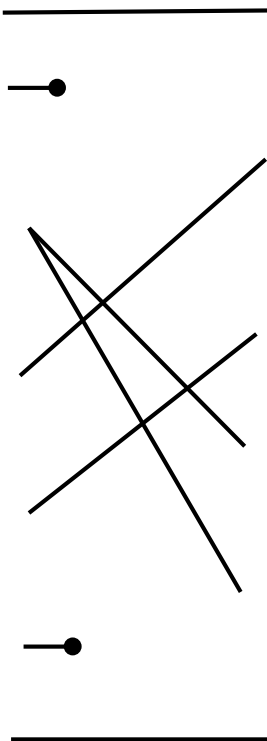
[5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。

[6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。

[7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

リード：全体の概要

対応関係



東京で桜が開花した

Start of cherry blossom season declared in Tokyo

[1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.

[2] Agency says that at least five blooming cherry trees in Somei-ya, a district in central Tokyo.

[3] The declaration comes as the agency says it was the first time since statistics were first kept in 1953.

[4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.

[5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.

[6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リード：ニュースで最も重要なことだけ

青：対応あり 紫：対応あり（意識・変換） 赤：対応なし

リードの次の文の違い

気象庁「東京でサクラ開花」発表 平年

各地で暖かい気温

対応関係

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表した。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区千代田の担当者が、上野公園の桜の花が咲いているのを確認したと発表しました。
 - [5] 東京のサクラの開花は平年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

リードの次の文：
背景情報

Start of cherry blossom

靖国神社で5輪以上咲いた

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration was based on statistics. It was the first time since 1945 that cherry blossoms were first seen in Tokyo.
- [4] Temperatures were in the 20s Celsius in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リードの次の文：
リードで伝えた内容を詳しく

青：対応あり 紫：対応あり（意識・変換） 赤：対応なし

リードの次の文の違い

気象庁「東京でサクラ開花」発表 平年

各地で暖かい気温

対応関係

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表した。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区千代田の担当者が、上野公園の桜の花が咲いているのを確認したと発表しました。
 - [5] 東京のサクラの開花は平年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

リードの次の文：
背景情報

順番の変更

Start of cherry blossom

靖国神社で5輪以上咲いた

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration was based on statistics. It was the first time since 1945 that cherry blossoms were first seen at the shrine.
- [4] Temperatures were in the 20s Celsius in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リードの次の文：
リードで伝えた内容を詳しく

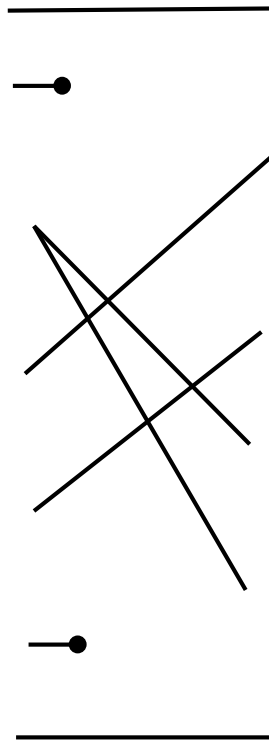
青：対応あり 紫：対応あり（意識・変換） 赤：対応なし

リードと本文との重複の違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
- [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
- 本文
- [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
- [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
- [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
- [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
- [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リードと本文との重複の違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- [1] 気象庁は**14日午後、「東京でサクラが開花した」と発表しました。**
- [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
- 本文
- [3] 気象庁によりますと、14日は東日本や西日本を中心に晴れ、**14日午後**2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、**「東京でサクラが開花した」と発表しました。**
- [4] 東京 千代田区の靖国神社には、**14日午後**2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、**「東京でサクラが開花した」と発表しました。**
- [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
- [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
- [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

重複

対応関係

Start of cherry blossom season declared in Tokyo

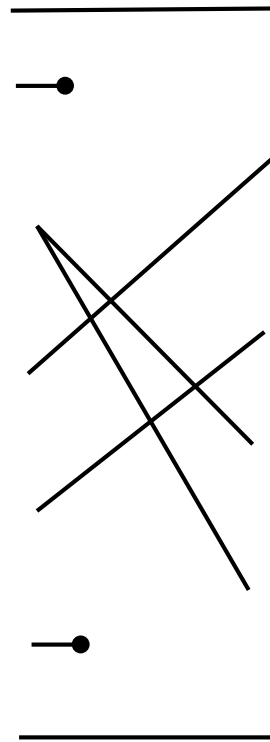
- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リードと本文との重複の違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- [1] 気象庁は**14日午後、「東京でサクラが開花した」と発表しました。**
- [2] **平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。**
- 本文
- [3] 気象庁によりますと、14日は東日本や西日本を中心に晴れ、日中の気温を上回る春の陽気となり、**重複**
- [4] 東京の青国神社には、**14日午後2時**に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、**「東京でサクラが開花した」と発表しました。**
- [5] 東京のサクラの開花の発表は、**平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。**
- [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
- [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

リードと本文との重複の違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- [1] 気象庁は**14日午後、「東京でサクラが開花した」と発表しました。**
- [2] **平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。**
- 本文
- [3] 気象庁によりますと、14日は東部地方に雲が広がりましたが、東日本や西日本を中心に晴れ、日中の気温を上回る春の陽気となりました。
- [4] 東京の靖国神社には、**14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。**
- [5] 東京のサクラの開花の発表は、**平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。**
- [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
- [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

重複

重複

重複無し

削除

削除

Start of cherry blossom season declared in Tokyo

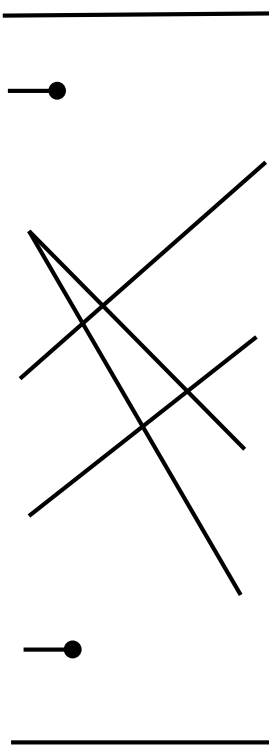
- [1] The Japan Meteorological Agency **has declared the start of cherry blossom season in Tokyo.**
- Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] **The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.**
- [4] Temperatures of some parts of the country were above the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

同じ表現の繰り返しの違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
 - [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

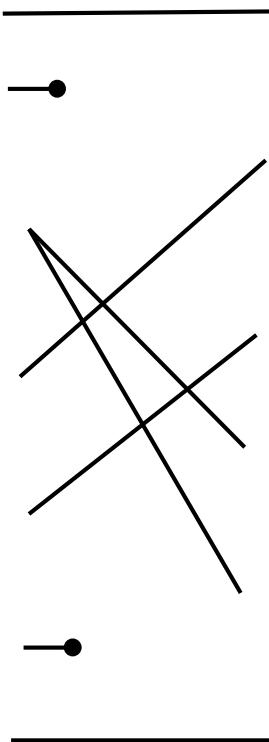
- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

同じ表現の繰り返しの違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
 - [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

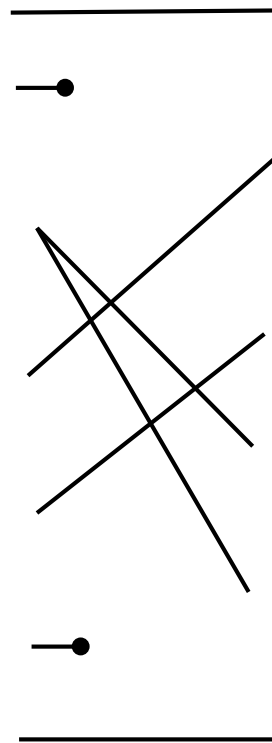
繰り返しが多い

同じ表現の繰り返しの違い

気象庁「東京でサクラ開花」発表 平年より12日早く

- リード
- 本文
- [1] 気象庁は14日午後、「東京でサクラが開花した」と発表しました。
 - [2] 平年より12日早く、去年と並んで統計を取り始めてから最も早い開花となりました。
 - [3] 気象庁によりますと、14日は東北などで雨雲が広がりましたが、東日本や西日本を中心に晴れ、各地で平年の気温を上回る春の陽気となりました。
 - [4] 東京 千代田区の靖国神社には、14日午後2時に気象庁の担当者が訪れ、開花の目安となっているソメイヨシノに5輪以上の花が咲いているのを確認し、「東京でサクラが開花した」と発表しました。
 - [5] 東京のサクラの開花の発表は、平年より12日早く、去年と並んで、昭和28年に統計を取り始めてから最も早い開花となりました。
 - [6] 気象庁の担当者は「去年と同じように、このところ平年より暖かい日が続いたことから早い開花となった」と話しています。
 - [7] 民間の気象会社によりますと、今後も西日本や東日本の各地で、平年よりも早くサクラが開花する見込みです。

対応関係



Start of cherry blossom season declared in Tokyo

- [1] The Japan Meteorological Agency has declared the start of cherry blossom season in Tokyo.
- [2] Agency officials confirmed on Sunday afternoon that at least five blossoms had opened on the benchmark tree of the Somei-yoshino variety at Yasukuni Shrine in central Tokyo.
- [3] The declaration came 12 days earlier than average. It was the same as last year, and the earliest since statistics were first kept in 1953.
- [4] Temperatures on Sunday exceeded the seasonal average in some parts of the country.
- [5] Although rain clouds spread over the northeastern Tohoku region and other areas, eastern and western Japan were mainly sunny.
- [6] A commercial weather-information firm says cherry trees will likely start blooming earlier than usual in many western and eastern parts of the country.

繰り返しが多い

繰り返しを避ける

内容

- NHKでの日英機械翻訳の利用状況
- 英語ニュース制作の課題
- これまでの研究開発

英語ニュース制作支援の目標

- 現段階：内容を正確に翻訳する
- 今後：日本語ニュースから英語ニュースを生成する

日英機械翻訳の研究開発

- 日英ニュース記事は文レベルでは正確な対訳になっていないものが多い
- 新聞・通信社の日英報道記事もある程度同様の傾向
(訳抜けを多く含む対訳データ → これで学習すると訳抜けが頻出)



ニュース分野の日英対訳で高品質な対訳文対が存在しない

高性能な機械翻訳に最も重要なものは高品質な対訳データ
→ **人手でNHKニュースの高品質な対訳データを構築**

ニュースの日英対訳文の構築

- NHK日本語ニュース文 → 人手英訳
- NHK日本語ニュース文 → 日英機械翻訳 → 人手ポストエディット
- NHK英語ニュース文 → 英日機械翻訳 → 人手ポストエディット

合計100万文対を構築



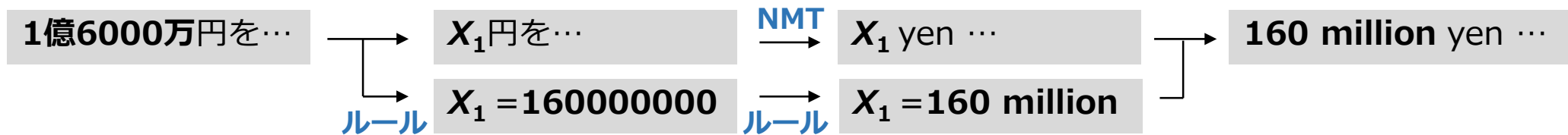
NHKニュースの翻訳品質が向上

ニュース翻訳の課題は対訳文とNMTだけで解決できるわけではない

数字の翻訳

- ニュースで数字は重要
- 数の単位の変換が必要：＜例＞ 1億6000万 → 160 million
 - NMTで単語列として翻訳すると間違える時がある

- ルールベースとの融合（大きな数字・小数点を含む数字）



- ニュース特有の表現への対応
 - 風速：メートル／秒 → キロメートル／時に変換
- ＜例＞ 風速30メートル → Winds of 108 kilometers per hour

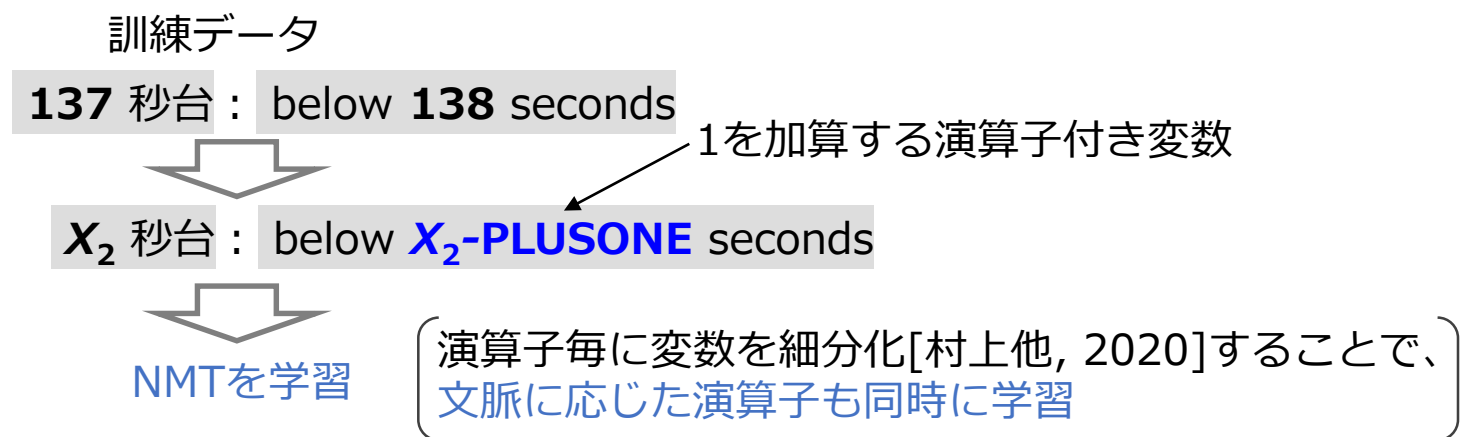
数字の翻訳（文脈依存）

[後藤, 2022]

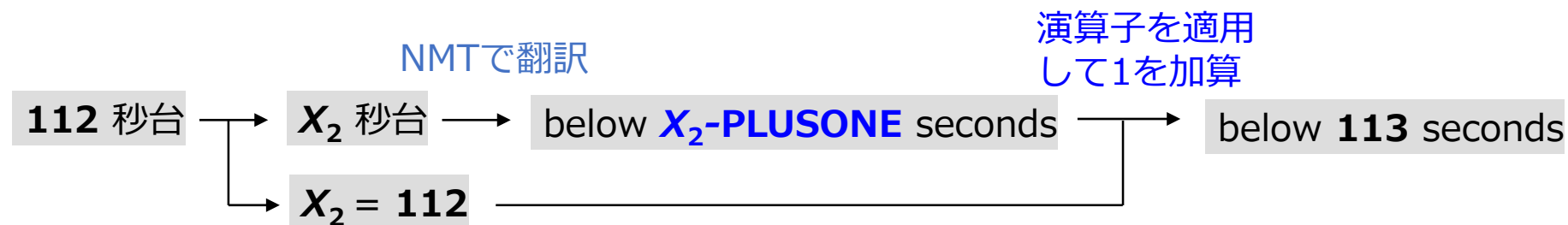
■ 目的言語文脈に依存する値の変化への対応

<例> **112** 秒台 → below **113** seconds

学習時

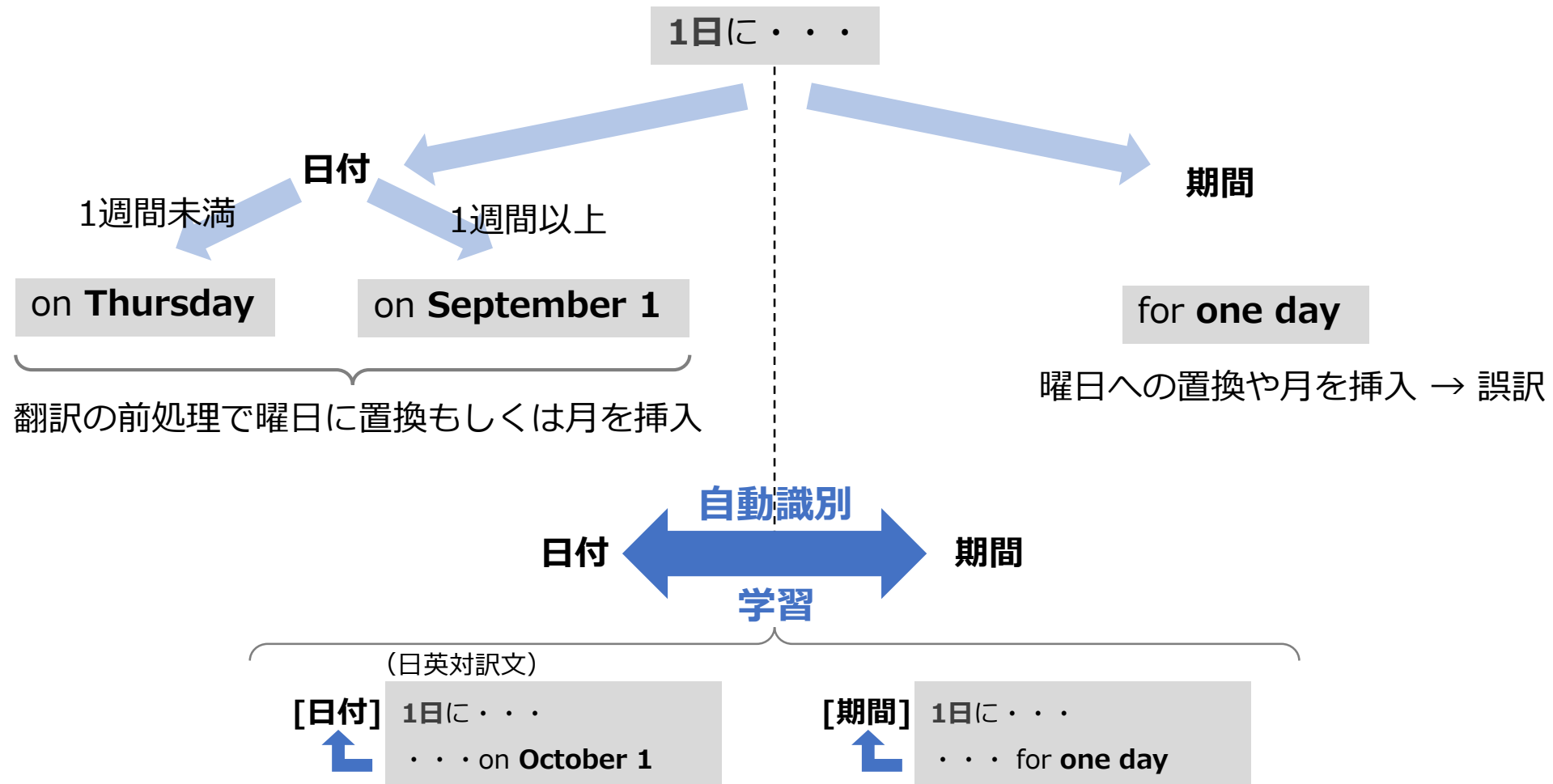


翻訳時



日付表現の翻訳

[Kinugawa et al., 2022]



翻訳の前処理で曜日に置換もしくは月を挿入

曜日への置換や月を挿入 → 誤訳

期間と日付で表層が異なる英語表現から対訳の日本語日付表現の日付・期間の分類を特定

ユーザー辞書機能の導入

■ 辞書機能として、相補関係にある2つの手法を組み合わせる利用

- 入力に訳を挿入することで出力を制御[Dinu et al., 2019]

<入力例> 自由民主党は・・・

→ <t> 自由民主党 <d/> Liberal Democratic Party of Japan </t> は・・・

語彙の情報を翻訳で利用できる

頻度ゼロの語と未学習の語には対応できない。低頻度でも出力されない場合あり



学習データで高頻度語を対象

- 該当表現を変数に置換して翻訳し、出力中の変数を訳語に置換[Long et al., 2016]

<入力例> ザポリージャ原発で・・・ 変数に置換 → <TERM₁>に・・・

<出力例> ... at <TERM₁> → 値に置換 ... at Zaporizhzhia Nuclear Power Plant

低頻度語、登録直後の表現を対象

文脈の利用（主語補完）（1/2） [Goto et al., 2020]

■ 既存の文脈利用手法（2-to-1） [Tiedemann and Scherrer, 2017]

- 学習時と翻訳時に、直前の文を区切り記号(<delim>)と共に翻訳元文に追加

機械翻訳への入力

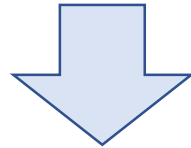
文脈	{	安倍晋三首相は24日午後、衆院本会議での代表質問で、沖縄県で米軍機の事故が相次いでいることについて「地域住民の安全確保は大前提だ。最優先の課題として、日米で協力して取り組む」と述べ、再発防止に全力を挙げる考えを強調した。
翻訳元文	{	<delim>日米地位協定の改定に関しては「事案に応じた最も適切な取り組みを切り開く」と慎重な姿勢を示した。

- 機械翻訳の学習の枠組み（正解訳からの誤差逆伝播）だけでは、多くの文脈の情報から必要な情報を推測して、そこを活用するように学習することは困難
- 2文以上前の文脈を利用できない

文脈の利用（主語補完）（2/2） [Goto et al., 2020]

文脈

安倍晋三首相は24日午後、衆院本会議での代表質問で、沖縄県で米軍機の事故が相次いでいることについて「地域住民の安全確保は大前提だ。最優先の課題として、日米で協力して取り組む」と述べ、再発防止に全力を挙げる考えを強調した。



文脈を**述語・項構造解析**し、**主語・主題**を抽出

安倍晋三首相



翻訳元文に追加（直前の文に主語・主題がない場合はそれ以前の文から抽出して追加）

機械翻訳
への入力

安倍晋三首相<delim>日米地位協定の改定に関しては「事案に応じた最も適切な取り組みを切り開く」と慎重な姿勢を示した。

2-to-1では学習が難しい部分を構造解析に基づいて実施

文脈の利用（目的言語文脈） [Mino et al., 2020a]

■ 目的言語側文脈

- 英語側の一貫性や同一表現の繰り返し抑制の制御に必要

■ 利用方法

- 入力文に直前の英語文を追加

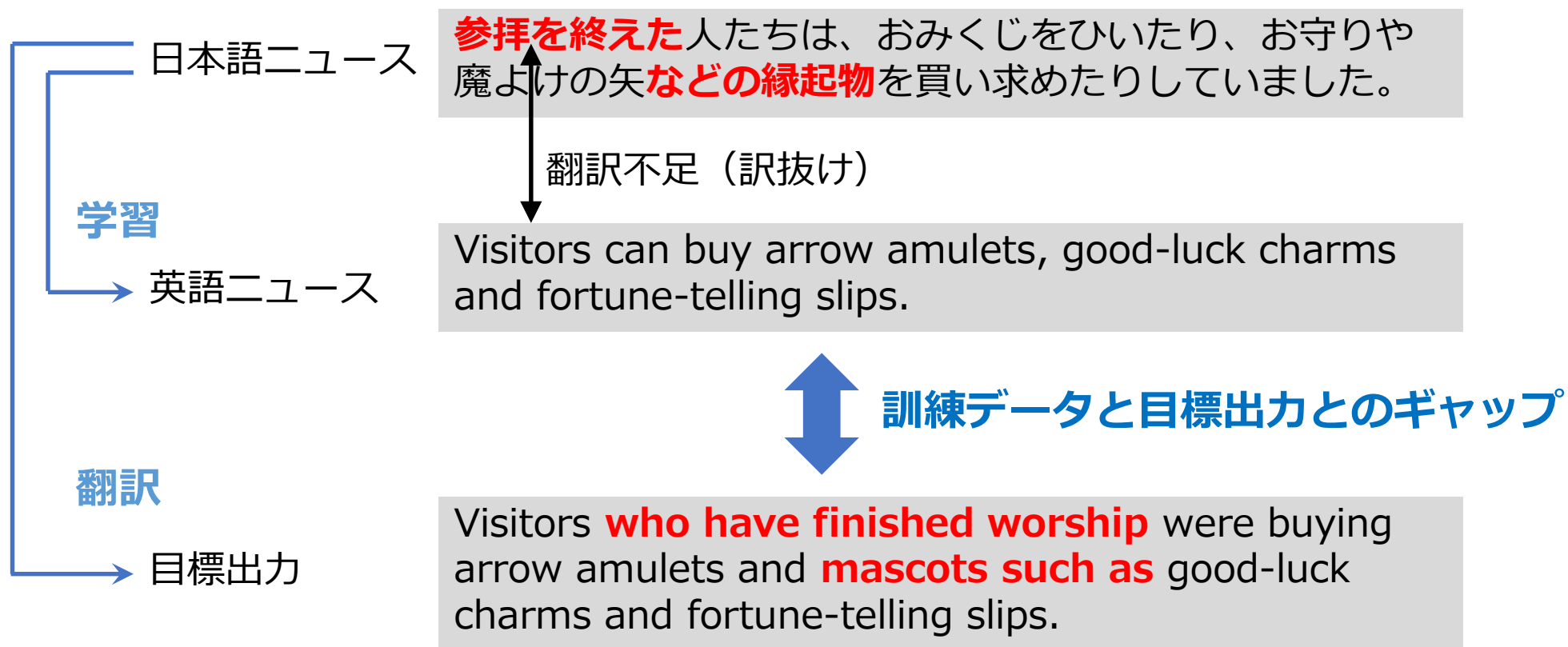
■ 従来手法[Kim et al., 2019]の課題

- 文脈として利用する直前の目的言語文の違いによる悪影響
 - ・ [学習時] : 訓練データ
 - ・ [翻訳時] : 機械翻訳出力（翻訳誤りやtranslationeseあり）

■ 対策

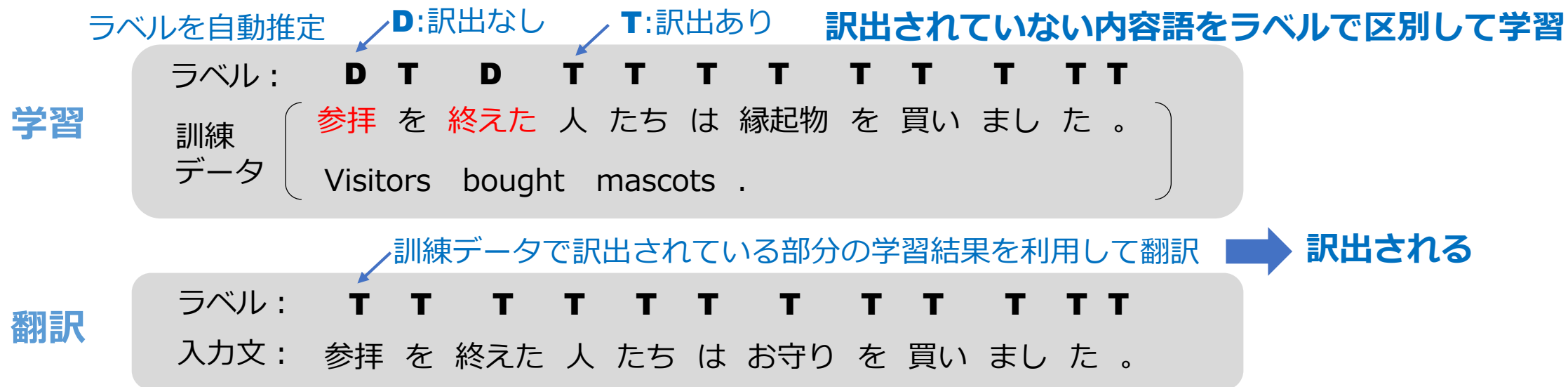
- 文脈として利用する直前の目的言語文
 - ・ [学習時] : 訓練データ（学習初期） → 訓練データ（割合小さく） & 機械翻訳出力（割合大きく）
- 学習時と翻訳時の文脈の違いを低減

日英ニュースの翻訳知識の活用(1/2) [後藤他, 2020]

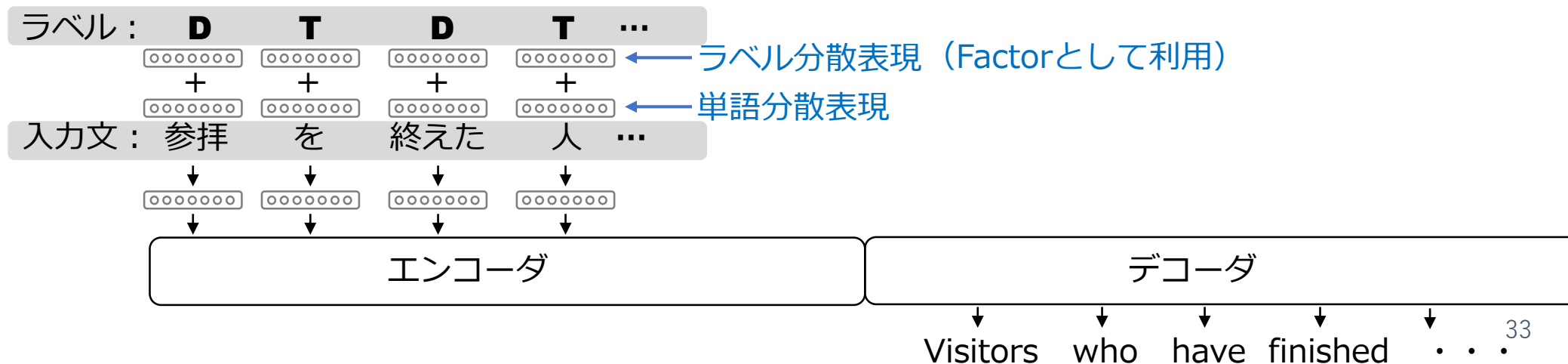


学習データでの訳抜けは、翻訳時に訳抜けの原因となる

日英ニュースの翻訳知識の活用(2/2) [後藤他, 2020]



ラベルの利用方法



複数種類の訓練データの活用 [Mino et al., 2020b]

訓練データの種類	データの特徴		
	日本語	対訳の質	英語
日英人手翻訳（英語ニュース調）	ニュース	高	目標出力
日英人手翻訳（通常の英訳）	ニュース	高	翻訳
日英ポストエディット	ニュース	高	翻訳
英日ポストエディット	翻訳	高	ニュース
MT逆翻訳	翻訳	低	ニュース
日英ニュース自動文対応	ニュース	低	ニュース

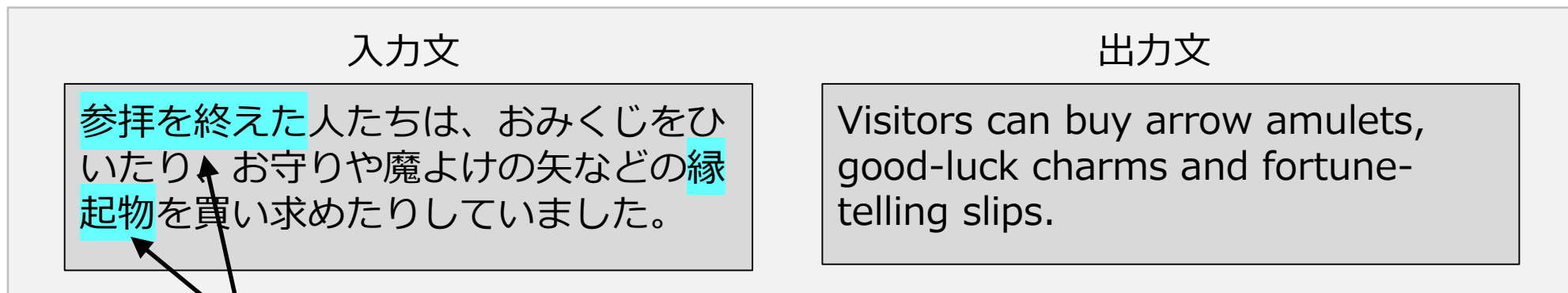
- データの**各特徴を表す複数のタグ**を訓練データの日本語文に付与
（ただし、英語の特徴タグには、目標出力とニュースは同じタグを付与）



各種訓練データのそれぞれの特徴を明示的に学習

- 翻訳時に目標とする特徴を表すタグを付与して出力を制御
翻訳時のタグ： <日本語:ニュース> <対訳の質:高> <英語:目標出力・ニュース>

制作支援向け訳抜け検出 [後藤他, 2019]



訳抜け部分が一目で分かれると便利



訳抜け箇所の検出

参拝を終えた人たちは、おみくじをひいたり、お守りや魔よけの矢などの縁起物を買って求めたりしていました。

強制的に逆翻訳



Visitors can buy arrow amulets, good-luck charms and fortune-telling slips.

英語文にない日本語部分の生成確率は低くなる → 訳抜け箇所と推定

まとめと今後の予定

■ まとめ

- NHKでの日英機械翻訳の利用状況
- 英語ニュース制作の課題
- これまでの研究開発

■ 今後の予定

- 翻訳機能・制作支援
 - ・ 新しい翻訳知識の自動獲得
 - ・ 翻訳誤りの推定
- 日本語ニュースから英語ニュースの制作自動化
(翻訳の範囲を超えたタスク)

ご清聴ありがとうございました。

本研究成果の一部は、国立研究開発法人情報通信研究機構の委託研究（課題197および課題225）により得られたものです。

参考文献

- [Papper, 2020] Robert A. Papper, Broadcast News and Writing Stylebook, seventh edition, Routledge, 2020.
- [David and Peter, 2019] David Ingram and Peter Henshall, The News Manual, <https://www.thenewsmanual.net>
- [Block, 2011] Mervin Block, Broadcast Newswriting: The RTDNA Reference Guide, A Manual for Professionals, CQ Press, 2011.
- [Strunk and White, 1999] William Strunk and E. White, The Elements of Style, Fourth Edition, Longman, 1999.
- [後藤, 2022] 後藤 功雄, ニュースを対象とした日英機械翻訳システムの研究開発, AAMT Journal, No. 77, p. 4-10, 2022.
- [村上他, 2020] 村上 聡一郎, 渡邊 亮彦, 宮澤 彬, 五島 圭一, 柳瀬 利彦, 高村 大也, 宮尾 祐介, 時系列株価データからの市況コメントの自動生成, 自然言語処理, 2020, 27 巻, 2 号, p. 299-328.
- [Kinugawa et al., 2022] Kazutaka Kinugawa, Hideya Mino, Isao Goto, Ichiro Yamada, Leveraging a Bilingual Corpus to Resolve Date–Duration Ambiguity in Japanese Numeric Day Expressions, Vol. 29, No. 2, p. 638-668, 2022.
- [Dinu et al., 2019] Georgiana Dinu, Prashant Mathur, Marcello Federico, and Yaser Al-Onaizan. Training Neural Machine Translation to Apply Terminology Constraints. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 3063–3068, 2019.
- [Long et al., 2016] Zi Long, Takehito Utsuro, Tomoharu Mitsuhashi, and Mikio Yamamoto. Translation of Patent Sentences with a Large Vocabulary of Technical Terms Using Neural Machine Translation. In Proceedings of the 3rd Workshop on Asian Translation (WAT2016), pages 47–57, 2016.
- [Tiedemann and Scherrer, 2017] Jörg Tiedemann and Yves Scherrer. Neural Machine Translation with Extended Context. In Proceedings of the Third Workshop on Discourse in Machine Translation, pages 82–92, 2017.
- [Goto et al., 2020] Isao Goto, Hideya Mino, Hitoshi Ito, Kazutaka Kinugawa, Ichiro Yamada, and Hideki Tanaka. Neural Machine Translation Using Extracted Context Based on Deep Analysis for the Japanese-English Newswire Task at WAT 2020. In Proceedings of the 7th Workshop on Asian Translation, pages 72–79, 2020.
- [Kim et al., 2019] Yunsu Kim, Duc Thanh Tran, and Hermann Ney. When and Why is Document-level Context Useful in Neural Machine Translation?. In Proceedings of the Fourth Workshop on Discourse in Machine Translation, pages 24–34, 2019.
- [Mino et al., 2020a] Hideya Mino, Hitoshi Ito, Isao Goto, Ichiro Yamada, and Takenobu Tokunaga. Effective Use of Target-side Context for Neural Machine Translation. In Proceedings of the 28th International Conference on Computational Linguistics, pages 4483–4494, 2020.
- [後藤他, 2020] 後藤 功雄, 美野 秀弥, 山田 一郎, 訳抜けを含む訓練データと訳抜けのない出力とのギャップを埋めるニューラル機械翻訳, 言語処理学会第26回年次大会, 2020.
- [Mino et al., 2020b] Hideya Mino, Hideki Tanaka, Hitoshi Ito, Isao Goto, Ichiro Yamada, Takenobu Tokunaga. Content-Equivalent Translated Parallel News Corpus and Extension of Domain Adaptation for NMT. In Proceedings of the Twelfth Language Resources and Evaluation Conference, pages 3616–3622, 2020.
- [後藤他, 2018] 後藤 功雄, 田中 英輝, ニューラル機械翻訳での訳抜けした内容の検出, 自然言語処理, 2018. 25 巻, 5 号, p. 577-597, 2018.

自然言語処理の未解決課題 ～対話、法律、医療、政治分野での 研究事例

第7回特許情報シンポジウム 招待講演

狩 野 芳 伸
Yoshinobu Kano
静岡大学 情報学部
行 動 情 報 学 科



国立大学法人
静岡大学

Kano Laboratory

National University Corporation
Shizuoka University

自己紹介: 狩野 芳伸

- 静岡大学 情報学部 行動情報学科 准教授(2014年～)
 - 浜松キャンパス
- 修士課程以来「自然言語処理」が専門
 - 学部: 東京大学理学部物理学科
 - 修士～博士～ポスドク: 情報理工学系研究科コンピュータ科学専攻(辻井潤一先生)
 - JSTさきがけ研究者「情報環境と人」領域
- 人间的に言葉を「理解・生成・対話」するシステムの構築
 - 音声・テキストを統合したリアルタイム解析器
 - 機械学習一発ではできない、過程を説明可能なホワイトボックス
 - 仕組みも人間に近づける: 人間(と脳)自体の仕組みを考察
 - 言語だけでなく背景となるすべての知的処理を包含する
 - 箱庭ではなく実応用で評価: 対話システム、質問応答、文生成、法律(試験解答)、医療(診断支援)、政治など
 - ※特許を扱ったこともあるのですが、今日はあえてそれ以外を
- 詳細は <http://kanolab.net/kano/index.ja.html>

進行中の主な研究テーマ・プロジェクト

- 心理学的に妥当な言語処理モデルと応用 (JSTさきがけ、萌芽)
 - 脳科学・認知科学による人間に近いモデルに基づく日本語話し言葉解析器の構築と検証 (挑戦的開拓 東京大・上智大)
- 対話システム: 人狼知能プロジェクト (東京大 鳥海先生ほか) 自然言語部門オーガナイザ
- 意味の重ね合わせによる高度な言語生成 (吉田英雄記念事業財団)
- 法律文書の処理と裁判の自動化支援
 - 試験問題の自動解答・質問応答 (東ロボ: 大学入試の自動解答): 終了
 - COLIEE 司法試験自動解答の国際タスク (基盤S NII 佐藤健先生)
- SNS・新聞記事・議会議事録を用いたAIによる世論形成過程と政治家の応答性の分析 (基盤B 学習院大・慶応大・大阪大)
 - 議会議事録の自動分析 (挑戦的開拓 学習院大 野中先生)
 - SNSにおける欺瞞とその広がり自動検出・推測と政治学・社会学的分析および予防的介入 (セコム科学技術財団)
- 医療言語処理: 電子カルテの処理と自動診断
 - 精神疾患の自動診断 (CREST/AIP加速 共同研究G 慶応義塾大学病院岸本先生/NII 佐藤真一先生)
 - 効果・副作用の発見と治療方針 (浜松医科大・国立病院機構)
 - 放射線学会・病理学会の収集する全国所見データ (AMED JP-AID)

言語処理に残された課題は何か

- 超大規模深層学習モデルによる性能向上
 - 特に近年の滑らかな文生成は目を見張るものがある
 - しかし、訓練データもパラメータ数も膨大
 - 人間には一生かけても不可能な学習
 - 人に近い振る舞い・結果を得たいならば、仕組み(制約)も人間に近づけるべき
 - 自然言語は人間の制約の下で形づくられているはず
- 課題: 計算資源と実行時間のコスト
 - トータルで人間よりはるかに計算量を要する
- 課題: 計算過程の解釈が難しい
 - 応用によっては説明や根拠が必須
- 課題: 首尾一貫した生成ができていない
 - 当然、論理性も疑わしい
- 課題: 「常識」「世界の知識」が扱えていない
- **いまの技術で解きうる課題**にも取り組みつつ、**解けていない課題に挑戦**してきた
 - その概要をご紹介: 医療・対話・法律・政治
 - 分野の今後をご一緒に考えたいと思います

電子カルテを利用した医療支援

- 大規模機械学習によるがん治療の最適化
 - 抗がん剤の作用副作用、患者属性を用いた腫瘍マーカー値の時系列推測
 - 国立病院機構との共同研究、数百万件規模の電子カルテを集積した大規模医療集積基盤(NCDA)を利用
 - 実は絞るとデータ量が足りない
 - 数値データのほうが有効な可能性
 - その中の着目ポイントを教えてくれる？(医師の知見)
 - 電子カルテテキストは相当broken
 - 数値・テキストの解釈に各科の専門知識が必要
- 「カルテ」からの「退院サマリ」の自動生成(厚生労働省科研費)
 - 要約のようで要約でない
 - カルテに書いていないことも書かれる
 - 医師・病院により大きく異なる
 - 「よいサマリとは何か」を指導する勉強会が院内で開かれる…

自然言語処理による精神疾患自動分類

- 話し言葉
 - 会話の録音から精神疾患を予測する
 - 現状、1000時間程度の世界最大規模診断付き音声コーパス
 - 1名1visitあたり1時間程度の録音
 - 1時間は3パート(質疑自由会話・絵の説明・物語の説明)に分かれる
 - 全データ文字起こし済み
 - 一部データに詳細な音声・言語学的アノテーションを付与
- 書き言葉
 - SNS投稿の書き込みから精神疾患を予測する
- 研究プロジェクト
 - JST CREST「自然言語処理による心の病の理解:未病で精神疾患を防ぐ」(共同研究G)
 - JST CREST「精神医学×メディア解析技術による心の病の定量化・早期発見と社会サービスの創出」(加速フェーズ)(共同研究G)
 - AIP加速「精神医学×メディア解析技術の展開:精神疾患への介入の挑戦」(共同研究G)

自然言語処理による精神疾患自動分類

- 精神科の診断は問診が基本のためまさに言語処理～会話分析だが、自然言語処理としては難しい課題
 - うつ病: 悲観的発言が増え、思考が遅くなり、質問のゴールにたどり着きづらくなる(精神運動制止)
 - 躁病: 注意を集中させることができず、考えが飛躍する(観念奔逸)
 - 統合失調症: 猜疑的な発言が増え、文脈や単語のつながりが弱くなる(連合弛緩)
 - 認知症: 物の名前が出にくくなり(語想起障害)、指示語が増え、話が回りくどくなる(迂遠)
 - 一方で診断基準は必ずしも安定しない
 - できれば「未病」に対応したい
- 現在のところ、各疾患vs健常の分類でそれぞれ8～9割程度の性能
 - 発達障害(ASD)のスコアリングでも高い性能を達成

健常者との比較観察

- 共通: 長い文は難しい傾向、文が複雑になるから？
 - 接続詞が少ない、名詞の語彙
- うつ病: 「自己主張」が強い？
 - 主格の“ハ格”が多く見られ、語彙も多い
- 認知症: 対話能力が落ちている
 - 間違いが多く、フィラーとしての指示語も多いが、内容語は少なく、応答も非内容語が多い
 - 発話音量平均が大きい、おそらく加齢による聴覚の低下による
- 統合失調症: 物語説明パートで、「証拠」に乏しい？
 - 目的格である“二格”が少なく、補語も少ない

書き言葉（SNS投稿）の分析

- Twitter等のSNS投稿を分析
 - メンタルヘルスに不調のあるユーザを推定したい
- 超大規模データを収集
 - 半自動ラベル付与
 - 時系列変化を追う
 - ユーザ間の影響、社会的なイベントの影響を分析
- SNS投稿テキストの分析は難しい
 - 短文が中心
 - 省略がおおい
 - 文脈依存性が高い
- 一方で大規模データが取得可能
 - タスク設定の工夫次第で興味深い研究ができそう

アカウントを不調とみなすパターン

- 精神疾患に使用される薬剤名を含むツイート履歴がある
 - 79の抗精神病薬、抗うつ薬、抗不安薬（右上表）
- メンタルヘルスを示唆する単語（右下表）をプロフィールに含む
 - 薬剤名および示唆する表現を含むツイートは除外
- この2条件に合致するアカウントを抽出し、不調（正例）とみなす
 - アカウント数：15,266人
 - ツイート数：34,505,381件
- 非不調群もほぼ同数に揃えた
 - アカウント数：16,519人
 - ツイート数：28,685,425件
- 正例負例ともノイズが入りうるが、大規模なため統計的に意味のある処理ができると期待

表3：使用した抗精神病薬、抗うつ薬、抗不安薬の薬剤名（五十音順）

アナフラニール、アピリット、アモキシサン、イフェクサー、イフェクサーSR、インプロメン、インヴェガ、ウインタミン、エビリファイ、エミレース、エリスパン、クレミン、クロフェクトン、クロルジアゼポキシド、コレミナール、コンスタン、コントミン、コントロール、サインバルタ、ジアパックス、ジェイゾロフト、ジプレキサ、スピロピタン、セディーエル、セトウス、セニラン、セパゾン、セルシン、セレネース、セロクエル、ゼプリオン、ソラナックス、ゾテピン、タンドスピロン、ダイアップ、チミペロン、テトラミド、デジレル、デパス、デプロメール、トフラニール、トリプタノール、トレドミン、トロペロン、ドグマチール、ニューレプチル、ネオベリドール、ノバミン、ノリレン、ハロマンズ、バルネチール、パキシル、ヒルナミン、ピブレzzo、ビーゼットシー、フルデカシン、フルメジン、プロピタン、ホリゾン、ホーリット、ミラドール、メイラックス、メレックス、メンドン、リスパダール、リフレックス、ルジオミール、ルボックス、ルーラン、レキサルティ、レキソタン、レクサプロ、レスタス、レスリン、レボトミン、レメロン、ロドピン、ロナセン、ワイパックス

表4：気分（感情）障害圏、統合失調症圏及び神経症性障害、ストレス関連障がい及び身体表現性障害圏の訴えや症状に関連する単語

うつ、鬱、躁、パニック、不安、ネガティブ、精神病、精神の病、自傷、障害者手帳、倦怠、動悸、息苦、食欲不振、頭痛、吐き気、しびれ、痺れ、麻痺、過眠、食欲低下、散漫、めまい、目眩、息切れ

メンタルヘルス不調群のツイート例

- 顕著なものとうそでないものがあり、人間には明確にわからないものも検出できる可能性がある
 - 実験では250トークンor120トークンに連結して使用

メンタルヘルス不調群のツイート例

※文脈情報を失わない程度に手作業で匿名化(改変)した

顕著	死にたくても死ねない。もう何もわからない。いつも親と仕事の話してるんだけど、無意識に死にたいとか言って親泣かせてる自分がうざいし、死んだほうがまし。で、死にたいとか考えると逆に落ち着いてくる。けど生きててもなあ、っていうループにはまって結局死にたいってなる。
曖昧	風呂に入りながら娘の帰宅を待つが、帰ってくる気配なし。手元にあるスマホで聞いてみようか。ご飯も出来てるしはよ帰ってきてほしい 何してんだろう眠くなってきた。でも外のバイクうるさすぎるから寝落ちすることはない、まあリラックスタイムを満喫しますか
隠微	おはようございます！寒すぎて散歩に行くか迷ってるけど、犬つれてけば暖かいかな、、でもちょっと貧血気味だからやめとこうかな汗 でも犬は歩かないわけにはいかないって顔してるし、ちょっと出かけます！

SNSにおける欺瞞とその広がり of 自動検出・推測と政治学・社会学的分析および予防的介入

- **超スマート社会の「悪」を探るには必須：SNS**
 - 情報化社会において、テキストメッセージは作成・伝達のコストが低く、人間への影響力は大きい e.g. フェイクニュース
- **欺瞞とその広がりを自動検出し予防的介入を試みる**
 - 欺瞞：他者を意図的に誤った方向に導く
 - － 「嘘」は「欺瞞」の一種で間違っていることが分かっている情報を言って他者をだますこと
 - － 欺瞞を意図的でなく伝達してしまうことも「悪」に加担する
 - 「虚偽」かどうかは、「事実」に反するかどうかで決まるが、何が事実かの機械的判定は困難なので、**相対的な関係**に留める
 - － 同様の情報の「流れ」～「流れ」同士の矛盾～どちらかが欺瞞
- **マクロな情報の流れと変化をミクロな言語表現から組上げ可視化・推測する**
 - 既存研究は**単語統計**しか見していない
 - 基盤となる道具：文章ペアの**関係認識**(矛盾・含意・中立)と**影響力予測をミクロに計算、全体を時系列でとらえる**

SNSにおける欺瞞とその広がりの自動検出・推測と政治学・社会学的分析および予防的介入(セコム科学技術財団研究助成 代表)

(悪の) インフルエンサー
どういう人が
どう発信すると
誰にどのように届くか？

(欺瞞に) 流される大衆
どういう人が
どう受信すると
どのように変わるか？

欺瞞の発信

欺瞞の受容・伝達

欺瞞の防止

悪

自己矛盾した投稿

「事実」と矛盾した投稿

ユーザ
ネットワーク
フォロワー
フォロワー

一般ユーザ

リツイート

いいね

変容した内容
の情報伝達

訂正
発信
実験

欺瞞の伝達防止と 訂正効果分析

影響推測器を用いた
影響力の大きい
文面生成

意見変化につながる
発信者・受信者
属性と文面の分析

ツイート

ニュース記事

国会議事録

捉えたい欺瞞情報の流れ

入社禁止期ソノウドソーシング調査

心理学的属性 政治/社会的意識
学歴 職業 年齢 家庭環境 等

アノテーション付与・ファインチューニング

欺瞞と広がりの検出器

話し言葉省略補完・解析器

気分・心理状態推測器 含意関係認識器
嘘検出器 影響（リツイート）推測器

事前学習

チューニング

大規模
ツイート
データ

科研費「～AIによる世
論形成過程と政治家
の応答性の分析」
関連成果

政治学・社会学的分析

イデオロギー測定と政治的傾向
移民の参政権・生活保護問題の分析 等

アカウント 時系列変化

時間

受信
発信
(投稿)
アンケート
調査
...

基盤的言語処理ツールの構築

- 本研究の入力はSNS投稿・ニュース記事・国会議事録
- SNS投稿の主語・目的語補完システムを構築済み
 - SNS投稿は省略・崩れが多く話し言葉に近い
 - 大規模ツイートデータによる事前学習時に省略補完のバイアスをかけることで高精度で省略補完を実現
 - 述語項構造(文内照応・文外照応・談話外照応・係り受け)で80%を超える性能: 本研究で省略補完に利用する
 - 応用として対話補完性能、対話システム性能(一貫性)の主観評価において既存システムを大きく上回った(発表予定)
 - 構築した事前学習済みモデルは汎用のため、本研究のすべての言語処理タスクに利用し照応の性能向上を図る
- ほか、下記研究プロジェクトで構築している事前学習済みモデルも統合する

心理学的に妥当な言語処理モデルと応用(JSTさきがけ、科研萌芽 代表)

脳科学・認知科学による人間に近いモデルに基づく日本語話し言葉解析器の構築と
検証(科研挑戦的開拓 東京大・上智大 代表)



嘘つきを見破る：人狼知能

- 人狼知能プロジェクト(後述)
 - 人狼ゲーム：嘘つき役「人狼」とそれ以外「村人」に分かれ、会話を通じて隠された役を見抜く
 - 人狼を自動プレイさせるプロジェクトを数年にわたり主催中
- 人狼ゲームの掲示板大規模コーパスを用い、**嘘つき(人狼)の自動検出**を行った
 - 7～8割程度の予測性能
 - 内容や論理以外の部分(広い意味での文体や会話パターン)で推測
- **精神・心理状態推測器(前項)と合わせ「人物タイプ」を推測**
 - **後述のビッグファイブ・ダークトライアドの推測も合わせて学習**
 - ビッグファイブ：**パーソナリティ特性**(開放性、誠実性、外向性、協調性、神経症傾向)
 - **ダークトライアド**：**反社会性パーソナリティ**(自己愛症：ナルシシズム、権謀術数主義：マキャヴェリアニズム、精神病質：サイコパシー)



世論形成過程の推測

■ 世論形成過程と政治家の応答性の分析

- 過去の全国会データを用いた国会議事録の分析
- 政治家の発言の含意関係認識(含意:相手文の意味を包含するか)
 - 含意関係認識器の絶対的性能が不足～大学入試社会科自動解答、司法試験自動解答・裁判の自動化支援
- 「バズリ」リツイート数～影響力の回帰学習による予測
 - 政策的世論へのインフルエンサーを抽出、フォロー・フォロワー数で正規化してから予測を試みる
- Twitterユーザのクラウドソーシングによる大規模調査
 - パーソナリティ、心理学的属性、政治的傾向、基本個人属性など
- 政治学・社会学的分析と予防的介入実験
 - イデオロギー推測、対外国人意識・移民問題と虚偽情報など

SNS・新聞記事・議会議事録を用いたAIによる世論形成過程と政治家の応答性の分析
(基盤B 学習院大・慶応大・大阪大)(代表)

試験問題の自動解答・質問応答(東ロボ:大学入試の自動解答)(社会科担当)COLIEE
司法試験自動解答(基盤S分担 国立情報学研究所)(U Alberta)主催者

欺瞞の発信

悪

自己矛盾した投稿

「事実」と矛盾した投稿

ツイート

ニュース記事

国会議事録

欺瞞の受容・伝達

ユーザ
ネットワーク
フォロワー
フォロワー

一般ユーザ

リツイート

いいね

変容した内容
の情報伝達

訂正
発信
実験

大規模定期クラウドソーシング調査
心理学的属性 政治/社会的意識
学歴 職業 年齢 家庭環境 等

アノテーション付与・ファインチューニング

欺瞞と広がりの検出器

話し言葉省略補完・解析器

気分・心理状態推測器 含意関係認識器
嘘検出器 影響（リツイート）推測器

政治学・社会学的分析

イデオロギー測定と政治的傾向
移民の参政権・生活保護問題の分析 等

欺瞞の防止

欺瞞の伝達防止と
訂正効果分析

影響推測器を用いた
影響力の大きい
文面生成

意見変化につながる
発信者・受信者
属性と文面の分析

アカウント
時系列変化

時間

受信

発信
(投稿)

...

アン
ケート
調査

...

事前学習

チューニング

科研費「～AIによる世
論形成過程と政治家
の応答性の分析」
関連成果

難しい課題

COLIEE: 法律文書の自動処理タスク

- COLIEE (Competition of Legal Information Extraction/Entailment)
 - 具体的なタスクの共有による法律文書処理研究コミュニティ形成が目標: 毎年開催 (COLIEE 2023開催中!)
 - 大きく情報抽出と含意関係認識の2タスクに分割
 - Case Law (判例法、カナダ): Task1, Task2 (COLIEE 2018より新設)
 - カナダ連邦裁判所の判例DBを利用
 - vLex社の提供による
 - Statute Law (成文法、日本): Task3, Task4
 - 司法試験の民法短答式を利用
 - 問題文と法律文の含意関係を二値で答える (Task 4)
 - オーガナイザーは試験問題と法律条文を日英両言語で提供
 - 現在、約1000ペア
 - 毎年最新の問題をテストデータにし過去問は訓練データに追加

COLIEE: 実際の短答式問題の例

- 多択式の問題を、各選択肢ごとにYes/Noの二択に変換
 - 英訳も作成
 - 下記は、「簡単な」問題例

抵当不動産を買い受けた第三者が、抵当権者の請求に応じてその抵当権者にその代価を弁済したときは、抵当権は、その第三者のために消滅する。

(平成26年民法短答式第14問選択肢工)

民法 第三百七十八条

抵当不動産について所有権又は地上権を買い受けた第三者が、抵当権者の請求に応じてその抵当権者にその代価を弁済したときは、抵当権は、その第三者のために消滅する。

COLIEE Task 4 の結果

- BERT系の深層言語モデルとそのアンサンブルを使った手法が上位

Team	L	#	Accuracy	Team	L	#	Accuracy
JNLP.BERTLaw	E	81	0.7232	KIS_3	J	61	0.5446
TRC3mt	E	70	0.6250	sim_neg.OvGU	E	61	0.5446
TRC3t5	E	70	0.6250	UEC1	J	61	0.5446
UA_attention_final	?	70	0.6250	taxi_BERTXGB	E	60	0.5357
UA_roberta_final	?	70	0.6250	UECplus	J	60	0.5357
KIS_2	J	69	0.6161	CUGIVEN	E	58	0.5179
llntu	J	69	0.6161	CUPLUS	E	58	0.5179
cyber	E	69	0.6161	linearsvm_no_ngram.HONto	E	57	0.5089
UA_structure	?	68	0.6071	POS_simneg.OvGU	E	57	0.5089
GK_NLP	?	63	0.5625	taxi_le_bigru	E	57	0.5089
linearsvm.HONto	E	63	0.5625	TRC3A	E	56	0.5000
JNLP.BERT	E	63	0.5625	UEC2	J	55	0.4911
KIS	J	63	0.5625	baseline_attention.OvGU	E	54	0.4821
linearsvm_no_ngram.HONto	E	62	0.5536	AUT99-BERT-MatchPyramid	E	52	0.4643
JNLP.TfidfBERT	E	62	0.5536	AUT99-LSTM-CNN-Attention	E	50	0.4464

開催を通じた分析～主な課題点

- 法律文書特有の構造
 - ～COLIEE 2019: 古典的な言語処理手法
- 絶対的なデータ量の不足
 - ～COLIEE 2020: 事前学習(転移学習)による深層言語モデル
 - ～COLIEE 2022: アンサンブル(システム結果の混合)
- 一般的な知識源(常識、世界の知識)の不足
 - ???

問題解答要素技術ごと統計(上位抜粋)

task	# prob	11	23	24	25	26	5	14	16	27	6	7	10	13	8	12	15
正解数		81	70	70	70	70	69	69	69	68	63	63	63	63	62	62	61
言語		E	E	E	E	E	E	J	J			E	E	J	E	E	J
条件	74	47	46	43	43	39	46	44	45	43	45	42	39	44	40	44	42
人物役割	48	35	29	25	30	30	29	29	29	26	22	25	25	24	24	26	24
人物関係	48	35	29	26	30	29	29	30	29	25	23	25	26	23	24	28	23
形態素解析	33	30	20	22	22	23	21	27	21	24	23	21	26	22	22	19	22
照応解析	23	17	15	15	14	16	12	15	14	13	13	11	14	13	12	10	14
曖昧性解消	1	1	0	0	0	1	1	0	1	1	1	1	0	1	1	0	1
意味役割付与	14	9	9	11	9	9	9	9	10	8	6	5	8	8	5	9	7
動詞言換	41	32	25	26	23	20	25	26	27	22	20	20	24	21	19	21	20
一般辞書	16	12	7	9	10	10	11	9	9	12	9	7	11	9	8	9	9
述語項構造	73	57	45	47	45	48	47	45	46	44	38	39	41	40	39	37	38
否定形	69	47	46	40	45	41	48	40	41	42	35	40	34	42	37	37	40
法律事実関係	55	37	30	34	34	34	33	34	33	32	29	30	29	28	29	32	26
含意関係	20	16	11	11	12	11	9	11	12	11	12	12	7	12	12	11	12
係り受け	24	18	17	16	14	15	16	17	16	16	15	11	15	14	13	12	12
条文検索	12	7	6	9	7	6	6	9	9	5	8	3	9	3	3	10	4
一般言換	3	3	3	2	2	1	2	1	2	3	0	2	1	2	2	2	2
箇条書き	11	6	6	7	5	4	5	7	7	7	4	5	7	5	5	7	5
数値計算	1	0	1	0	1	1	1	0	0	1	0	0	0	1	0	0	24

今後のCOLIEEタスク設計

- 説明可能なAI
 - 参加者は短期間で実装させるため、安易な方向に走りがち
 - End-to-end で学習しておわり が多い
 - 結果が良かったとしても、実用につながるのか？ 本当にちゃんと解けているのか？
 - 根拠を示させるなど、「説明」の要素が入ったタスク設計をしたい
 - 正解の作成、自動評価基準をどうするか？ 説明箇所をあてさせる？
- COLIEEの目標を達成することは、**法律分野のみならず人間の知的能力を幅広くカバーする、人工知能分野の究極的な目標を包含する**
 - 参加者募集中です！

人狼知能：嘘・説得・信頼

■ 人狼のゲーム性

- 客観視点での情報不確定性(不完全情報性)
 - － 得られた情報が真かどうか不明
- 推理：他者の意図のモデル化
 - － 自分がこう思っていると相手が思っているだろう
 - － 文脈からの真偽の判定
- 説得：他者から見た自己のモデル化
 - － 相手に情報を与えて信頼を得る

■ 嘘と信頼

- 誰かが嘘をついていうという状況
- 村人側：自分が語る事をどうやって信用してもらうか
- 人狼側：自分が騙る事をどうやって信用してもらうか
 - － 襲撃する対象をどう話し合って決めるか
- 「説得する人工知能」



人間らしい対話システムの評価として

- GPTベースの大規模言語モデルにより、対話システムは近年急速に性能を上げている
 - 語彙・文法はとても流暢で、1文では人間と見分けがつかないレベルになってきた
 - 一問一答での受け答えも、相当かみ合っている
- まだまだ完璧ではない: 前提知識の共有や深い意味などが取れているかは疑わしい
 - 内容や設定に対する一貫性のなさ
 - 矛盾した応答をするということは、論理性も疑わしい
 - 長い文脈はおそらくあまりとれていない
 - 「常識」「共通基盤」といった外部知識に対応できていない
 - 大規模データの統計的な「重ね合わせ」でどこまでいけるか？
- 人狼なら相手の話を聞かねばゲームにならないはず
 - より人間に近づくほどそうなる(あるいは人間と対戦)
 - 評価は勝敗よりも、対話～ゲーム行動のリンク
- 毎年「人狼知能大会」を開催しているが、まだまだ「会話を通じたゲームがきちんとできている」とは言い難いレベル

「力業」の限界と今後

- 一般的な知識の不足：人間は自然と身に着ける
 - 法律文書に現れるのは最終的な表現であって、背後にさまざまな知識、人間特有の知的構造がある
 - 機械は何兆ものデータを利用しているが人間は小規模でできる
 - 規模もモデル構造もどんどん人間から離れていっている
- 「古典的な」言語処理と深層学習の融合を図るときか
 - 自然言語処理分野全体としても、次なるブレークスルーではないか
 - 「論理構造」「述語項構造」など
 - 深層学習は埋め込みベクトルを使っているが、もっと「決定的」な要素を含める など
 - シンボルと分散表現の融合：e.g. 多義語はどうなっている？
- 人間に近いモデルの導入
 - 逐次的・音声処理の融合、人間に近い制約を加えたモデル 等