

中小企業庁 御中

令和3年度AIを活用した総括発注業務の効率化 に係る実現可能性調査等事業 調査報告書

株式会社PKSHA Technology
2022年3月31日



事業サマリー

3つの事業内容を以下の通り実施

事業内容	実施概要	調査結果
(1) 中小企業庁の総括発注業務のAI活用のデータ作成に係る調査	16課室のうち、読み取り可能なデータが提供された12課室について、 a. メール：重複の除去、総務課から原課への発注メールの抽出を実施、最終的に振り分けられた班を特定 b. 添付ファイル：wordファイルに限定し、上記メールと添付ファイルの紐付けを行い、班を特定	a. 4,050件の振分メールを抽出 ➢ 人事異動の影響を受けない2021/8/1以降の期間に限定 ➢ from/to/ccや本文中の発注メールに特徴的な用語を手掛かりに自動抽出 b. 174件の添付ファイルを紐付け ➢ 添付ファイル情報がメールから欠落していたため、メール件名と添付ファイルの格納されたフォルダ名との紐づけにより特定
(2) 中小企業庁の総括発注業務のAIエンジンのプロトタイプ構築にかかる調査	a. 判別の振分根拠となるナレッジを構築 (1) で作成されたデータ、及び「担当班・担当者情報 & 担当内容情報」、「所管事項説明」、「閣議決定文書」、弊社保有辞書からの特徴語抽出と重み付け（意思決定への寄与度付け） b. 自動振り分けエンジンの構築 本格的な深層学習ベースエンジンの事前検証を行う重み付けキーワード抽出エンジンの構築 c. 振分結果の評価 特定の課を対象に貴庁に評価データを選定いただき、RecallとPrecisionによる精度評価と今度の精度向上見立ての整理	a. 計4,281個の特徴語を班別に構築 ➢ 班別に「1. 単体で振り分けを決定づける」「2. 振り分けに一定寄与」の2段階の重み付け b. aに基づき、メール・添付ファイルに対して、班別の振分スコアリングの上、振分候補と判断根拠を出力 c. 最終的な精度はRecall 78%、Precision 66%。 今後深層学習等エンジンの高度化により、Recall 90%、Precision 75%程度へ改善見立て
(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査	機能要件と非機能要件の定義 a. 総括係員の業務効率化実現に向けたUI/UX（機能要件） b. 貴庁セキュリティポリシーやパフォーマンスの充足（非機能要件）	a. Webアプリとして、メールと添付資料を取り込みで振分候補と判断根拠を提示 ➢ 添付ファイルはword, ppt, PDF, excelのテキスト部分を対象 b. 弊社クラウド環境にフロントエンド、バックエンドを構築。 2022年度は貴庁NWとインターネット接続のうえ、IP制限等セキュリティ対策を実施 ➢ Teams/outlook連携はセキュリティ要件から、困難な見立て

実施スケジュール

2ヶ月で以下の通り3つの事業を実施

		2022年2月				2022年3月			
スケジュール		2/7週	2/14週	2/21週	2/28週	3/7週	3/14週	3/21週	3/28週
プロジェクトマネジメント		★kickoff		★実務定例		★実務定例		★実務定例	★最終報告
(1) 中小企業庁の総括 発注業務のA I 活用のデ ータ作成に係る調査	データ受渡								
	作業								
	内容確認・議論								
(2) 中小企業庁の総括 発注業務のA I エンジンの プロトタイプ構築にかかる調 査	モデル構築								
	内容確認・議論								
(3) 中小企業庁の総括 発注業務の効率化に係る システム開発の要件定義に かかる調査	要件ヒアリング								
	要件定義								

・通常、モデル構築→確認→改善の期間を確保
・今回は期間内での再構築期間困難なため、
モデル改善は要件定義の計画へ反映

目次

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査

(2) 中小企業庁の総括発注業務のA I エンジンのプロトタイプ構築にかかる調査

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査

目次

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査

(2) 中小企業庁の総括発注業務のA I エンジンのプロトタイプ構築にかかる調査

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査 事業理解とアプローチ(1/2)

現状の振分業務を機械で再現するために必要なデータ作成作業を実施

機械での再現に向けた対応

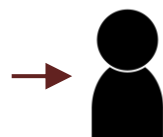
現状のヒトによる業務

頂いているデータ

学習に必要な前処理・形式

<添付ファイル付きメール>

送受信先
件名
本文
添付ファイル 

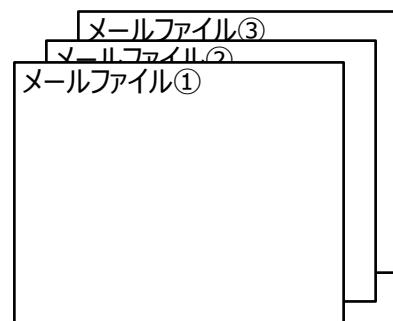


・総務課
・各課室総括係員

XX課AA班

YY課BB班

ZZ課CC班



紐づき無し

<添付ファイル/格納フォルダ>



1 振り分けられる課・班の特定
(= 正解ラベリング)

2 メールと添付ファイルの
紐づけ

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査 事業理解とアプローチ (2/2)

正解ラベリングと添付ファイルの紐付けを以下のアプローチにて実施

1 振り分けられる課・班の特定 (= 正解ラベリング)

2 メールと添付ファイルの紐付け

目的

班迄の特定がされたメールの抽出

学習として活用できる添付ファイルがメールに対してどの程度あるかの調査

実施内容

各課室の受領メールを対象に、内容が重複するメールスレッドを除外、人事異動後の期間に限定という処理を行った後、以下の①-②の条件で班までの振分メールをプログラムで自動抽出

- ① 総務課から原課に振り分けられている、または発注メールに特徴的な文言が含まれる
- ② 原課総括班から班へ振り分けられている

- ① 添付ファイルを受領した各課室のメール及び添付ファイルにおいて、以下3手法で紐づけ可能性の検証を実施
 - a. メール件名と、添付ファイル格納フォルダ名の突合
 - b. メール本文と添付ファイル名の突合
 - c. メール件名と添付ファイル名の突合
- ② 特に、総括係員の実際の運用に近く、貴庁にもご推奨頂いたaのアプローチを採用

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査 作成結果(1/2)

課室別の作成結果は以下の通り

	No.	目的	処理	A課		B課		C課		D課		E課		F課	
メール 抽出	①	人事異動後の期間に、総務課から原課に振り分けられている、または発注メールに特徴的な文言が含まれる	初回スレッドの送信日が2021/8/1以降で、①-1と①-2の和集合	527	100%	445	100%	73	100%	1218	100%	45	100%	91	100%
	①-1	内、総務課から原課に振り分けられている	「送信者に総務課所属員のアドレスが含まれている」and「To or CCに各課総括メーリス or 総括グループ中企庁メーリス がある」	96	18%	210	47%	32	44%	494	41%	11	24%	19	21%
	①-2	内、発注メールに特徴的な文言が含まれる	「発注」or「作業」or「依頼」or「締切（締め切り）」or「協議」or「×」or「各位」といった単語が件名もしくはメール本文（特に件名）に含まれている	525	100%	444	100%	73	100%	1212	100%	45	100%	90	99%
	②	内、原課総括班から班へ振り分けられている	「本文中に 振分対象の班の直後に 着く← or < がある」	414	79%	442	99%	64	88%	1141	94%	36	80%	78	86%
添付 ファイル 紐付け	③	件名とフォルダ名が一致する	②の中で、件名と添付フォルダ名がマッチングする（重複を除いた数）	8 (8)	2%	79 (65)	15%	0 (0)	0%	55 (44)	4%	2 (2)	6%	6 (6)	8%

注意事項：

- 各課室ごとに、各処理を行った際に抽出されたメール件数、及び①時点のメール件数を100%とした時の各処理後に抽出されたメール件数の割合を記載

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査 作成結果(2/2)

課室別の作成結果は以下の通り

	No.	目的	処理	G課		H課		I課		J課		K課		L課	
メール 抽出	①	人事異動後の期間に、総務課から原課に振り分けられている、または発注メールに特徴的な文言が含まれる	初回スレッドの送信日が2021/8/1以降で、①-1と①-2の和集合	586	100%	8	100%	624	100%	3	100%	109	100%	893	100%
	①-1	内、総務課から原課に振り分けられている	「送信者に総務課所属員のアドレスが含まれている」and「To or CCに各課総括メーリス or 総括グループ中企庁メーリス がある」	112	19%	2	25%	223	36%	0	0%	44	40%	180	20%
	①-2	内、発注メールに特徴的な文言が含まれる	「発注」or「作業」or「依頼」or「締切（締め切り）」or「協議」or「×」or「各位」といった単語が件名もしくはメール本文（特に件名）に含まれている	577	98%	8	100%	617	99%	3	100%	109	100%	892	100%
	②	内、原課総括班から班へ振り分けられている	「本文中に 振分対象の班の直後に 着く← or < がある」	445	76%	7	88%	589	94%	3	100%	107	98%	724	81%
添付 ファイル 紐付け	③	件名とフォルダ名が一致する	②の中で、件名と添付フォルダ名がマッチングする（重複を除いた数）	31 (30)	7%	N/A	N/A	50 (46)	8%	0 (0)	0%	1 (1)	1%	10 (6)	1%

注意事項：

- 各課室ごとに、各処理を行った際に抽出されたメール件数、及び①時点のメール件数を100%とした時の各処理後に抽出されたメール件数の割合を記載

目次

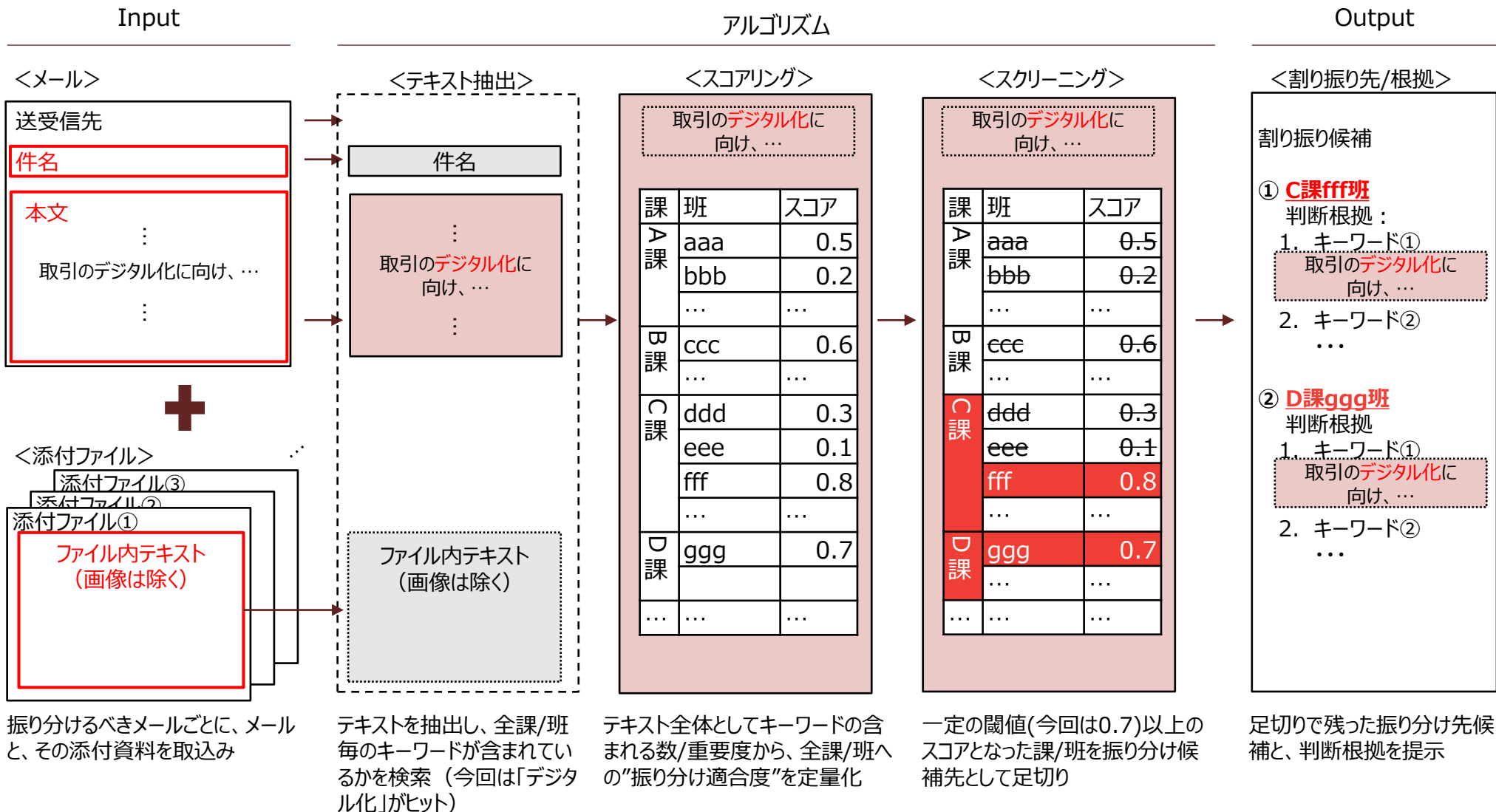
(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査

(2) 中小企業庁の総括発注業務のA I エンジンのプロトタイプ構築にかかる調査

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査

(2) 中小企業庁の総括発注業務のAIエンジンのプロトタイプ構築にかかる調査 AIエンジンの全体像

以下のインプット、アルゴリズム、アウトプットを要件定義



(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査 本調査の対象スコープ

本事業においては、特定の課班を対象に、添付ファイルタイプはwordに絞り、精度を検証

		本事業の実施内容	2022年度計画
対象課室		特定の課（X課）及びその班	全課室
インプット		•メールの本文・タイトル •添付資料(word)	左記に加え、 添付資料（excel, ppt, PDF）
ナレッジ構築		•担当班・担当者情報&担当内容情報 •所管事項説明、閣議決定文書 •メール本文 •添付資料（班まで紐付けられた量が限定的なため、ファイル名が班の名前とセットで書かれているケースを重点取り込み）	左記をベースに 貴庁内追加データソースや、外部の辞書等も検討
アルゴリズム		重みづけキーワード抽出	重みづけキーワード抽出 + 深層学習モデル（予定）
アウトプット		X課内の班の候補と判断根拠	左記に加え、判断根拠の含まれる パッセージ箇所
評価	指標	Recall（取りこぼしの無いよう、多めに候補を出し、人が取捨選択する）を、Precision（確実なもののみ厳選表示し、人が不足を補う）より重視しつつ、F値（recallとprecisionのバランス）も重視（＝recall偏重となり、毎回全課室班を表示しては意味がない）	
	データ	貴庁より、評価データ（X課の添付ファイル付き振分メール）を26件頂戴	（追って、適切な評価データ選定ご相談）

(2) 中小企業庁の総括発注業務のAIエンジンのプロトタイプ構築にかかる調査 ナレッジ構築

X課について、弊社特徴語抽出結果をベースにキーワードを取捨選択、優先度付頂いた上で弊社にてさらに選定

班名	「05_担当班・担当者 情報 & 担当内容情報_最新」のキーワード (件)	「03_所管事項説明資料」、「04_過去の閣議決定文書の担当箇所抽出」からの特徴語抽出			メール本文からの特徴語抽出			弊社にて精度比較の上、さらに取捨選択
		弊社で実施 (X課全体)	貴庁により取捨選択、重みづけ頂いた結果*		弊社で実施 (班ごと)	貴庁により取捨選択、重みづけ頂いた結果*		
			1	2		1	2	
a班	6	3729	1	0	上位500件ずつ共有	0	0	11
b班	4		0	0		0	2	5
c班	2		5	20		5	2	77
d班	2		2	20		4	3	93
e班	1		10	3		2	0	21
f班	9		4	12		2	2	66
g班	4		1	36		10	4	4,008

- 重みづけのラベルは以下の通り
 - そのキーワードのみで振り分け可能な決定的なもの（全省庁、経産省で見てX課（主担当でなくても少なくともX課に情報共有必要）を含む）
 - 振り分けに一定寄与する事が多いもの（庁内ではX課（主担当でなくても少なくともX課に情報共有必要）を含む）

(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査 アルゴリズム

精度と実装の実現性を鑑みて、本事業では重み付けキーワードマッチングを選択

判断のポイント	キーワードマッチング		深層学習	
	単語情報に基づくマッチング	重み付けされた単語情報に基づくマッチング	Sentence BERT	DensePassageRetrieval, etc
精度	△	○	△	△～◎
本事業内（～3月末） 実装可能性	○	○	○	×

今回のアプローチ

- ・ 課室・班への振り分けの基準となる情報として、単語・フレーズに分解する方法が有力として見立てており、それに基づくマッチングアプローチが有効である見立て
- ・ 深層学習ベースの手法の候補として挙げていたSentence BERTは、文の意味的類似度に基づくマッチング手法であるが、受領したメールデータの傾向およびマッチングの判断基準を鑑みると、当該モデルの単純適用では十分な精度が期待できない見通しとなった（当該手法は、課室・班の振り分け基準と、メール・添付ファイル中のテキスト情報が文として類似している場合に有効）
- ・ 文脈の中の一部に着目した判断の実現には、異なる種類の深層学習モデルの適用が必要であり、それには当初想定よりも工数が必要

(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査 アルゴリズム

重み付けされた単語情報に基づくマッチングの詳細は以下の通り

	アプローチ
概要	各課・班の担当・活動等の特徴をキーワードの集合として扱い、メールおよび添付ファイルにおけるキーワードの出現量に基づくスコアリングを通じて課・班への振り分け先スコアリングを行う
アプローチ	事前準備 (分類モデルの構築) 1. 各課室・班に関連する情報源（関係者が有する知識、文書集合等）から、キーワードの人手による列挙、自動獲得したワードの選別等を通じて、それぞれを特徴付ける重み付きキーワード集合（知識）を構築 2. 得られた情報に基づいて、各課室・班に関する特徴ベクトルを構成 メールの振り分け 3. 各メールスレッドに対して、「解析対象のメール/添付ファイル全体」を一つの特徴ベクトルに変換する 4. 3で得られたメールスレッドの特徴ベクトルと、各課室・班の特徴ベクトルを用いて関連度スコアを計算 5. 上記の結果に基づいて上位N件を根拠情報付きで出力 根拠情報：各課室・班を特徴づける言語表現の出現箇所
評価方法	評価データの振り分けの結果、スコアがスコアが閾値を超えた班を出力とし、定量的な精度を算出（後述）

(2) 中小企業庁の総括発注業務のAIエンジンのプロトタイプ構築にかかる調査評価

確実に候補を厳選するよりも、取りこぼしを減らすことに比重をおいたモデルを想定

具体例

		実際の割り振りの正解	
		正	負
モデルによる予測	正	A a班 b班 c班	B d班 e班
	負	C f班	D g班

精度の考え方

モデルの定量的な精度指標

- Recall (再現率)
 - 「実際に正であったもののうち、どれだけ正と予測できたか」という予測の網羅性を示す指標
 - 左図では $\text{Recall} = A / (A+C) = 75\%$
 - 今回は、実際の割り振り先のうちどの程度モデルが正しく予測できたか
- Precision (適合率)
 - 「正と予測したものが、どれだけ正しかったか」という誤予測の少なさを示す指標
 - 左図では $\text{Precision} = A / (A+B) = 60\%$
 - 今回は、モデルが予測した割り振り先のうち、実際どの程度が正しかったか
- F値 (適合率と再現率の調和平均)
 - $2 / \{(1/\text{Recall}) + (1/\text{Precision})\}$ で計算される、0~1の間の値
 - 左図では、0.67

留意点

- RecallとPrecisionはトレードオフの関係にあり、業務要件に合わせたバランスが大切



今回のモデルの方向性

- Recall (取りこぼしの無いよう、多めに候補を出し、人が取捨選択する) を、Precision (確実なもののみ厳選表示し、人が不足を補う) より重視しつつ、そのバランス、F値も重視 (毎回全課室班を表示しては意味がない)

(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査評価

現状、8割の精度は確認でき、今後深層学習ベースのエンジン高度化等によるさらなる上昇を想定

	途中経過	最終結果	今後のさらなる改善余地
アルゴリズム	重みづけ キーワード抽出 ver.1	重みづけキーワード抽出 ver.2 (キーワードとのマッチを取る際に、複合語だけでなく 単語ベースでもマッチを取るよう改善)	マッチング手法の高度化 (文脈や重要箇所の特 定によるマッチング高度化 等)
評価データ	X課の班に振り分けられて いるメールの中からランダム 抽出された、90件 (内、添付ファイル6件)	X課について、追加でいただいたデータを含めて、振 分精度の評価として適切な26件* (全て添付ファイル付き)	
精度	Recall : 30% Precision : 40% F値 : 30	Recall : 78% Precision : 66% F値 : 71	Recall : 90%~ Precision : 75%~ F値 : 82~

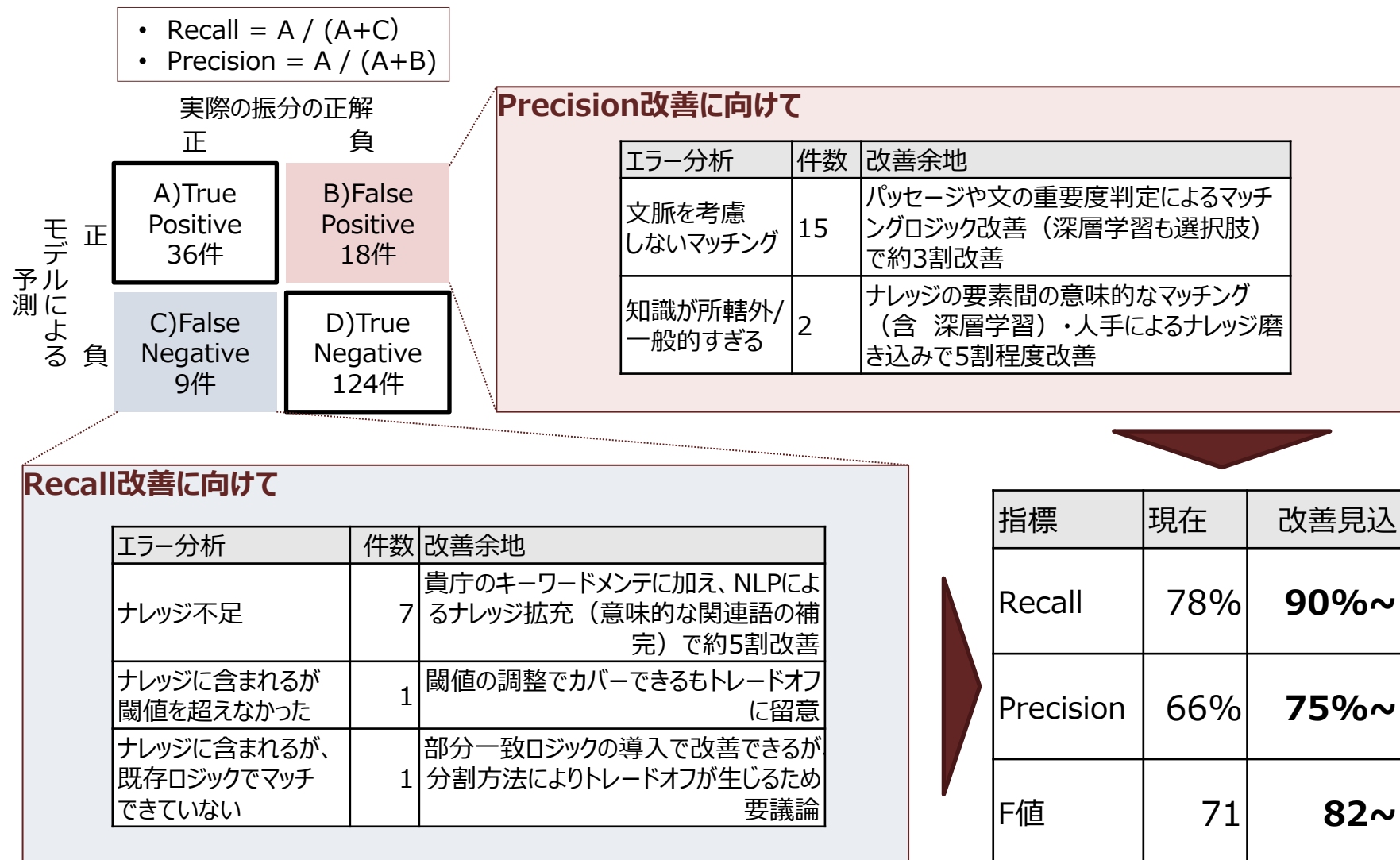
		実際の振分の正解	
		正	負
モデル 予測による	正	A)True Positive	B)False Positive
	負	C)False Negative	D)True Negative

- Recall = $A / (A+C) = 48\%$
- Precision = $A / (A+B) = 52\%$
- F値 = $2 / \{(1/\text{Recall}) + (1/\text{Precision})\}$

*27件受領したうち、1件は添付ファイルが読み込めない形式だったため除外

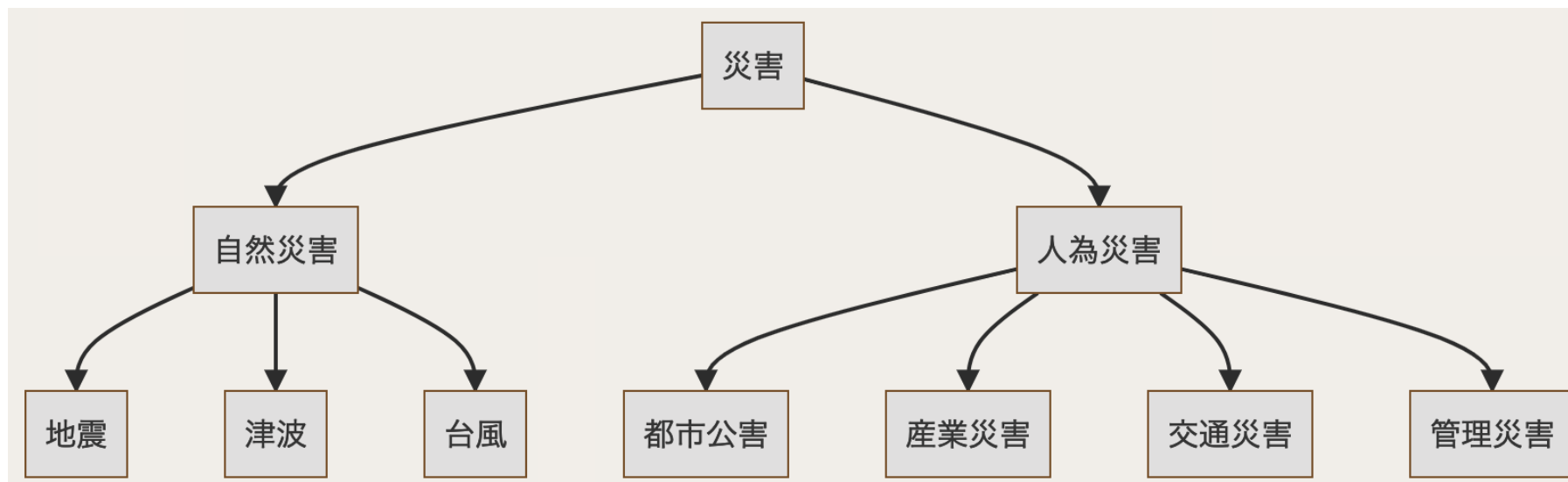
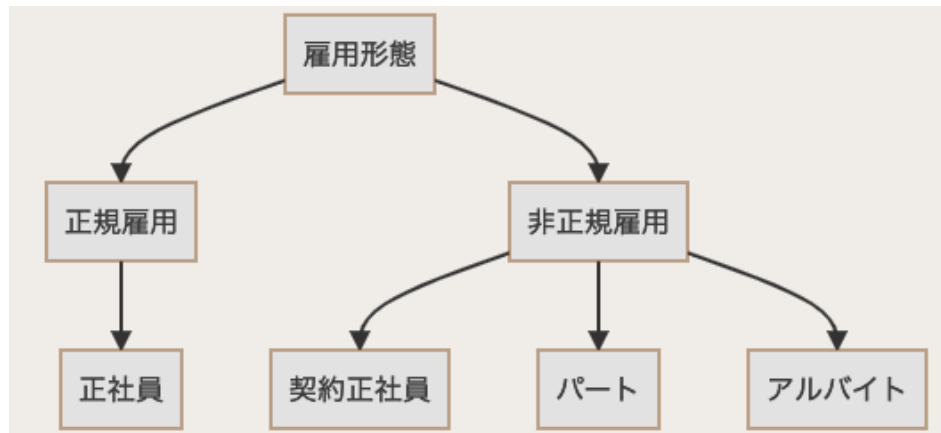
(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査 評価データ詳細

評価の厳密な妥当性を担保するため、インプットを下記のように整理した上で精度を再測定



(2) 中小企業庁の総括発注業務のA Iエンジンのプロトタイプ構築にかかる調査 ご参考) ナレッジ不足を補完するNLP技術

上位概念・下位概念の体系化によって、所管を説明するキーワードのリストアップを自動化



目次

(1) 中小企業庁の総括発注業務のA I 活用のデータ作成に係る調査

(2) 中小企業庁の総括発注業務のA I エンジンのプロトタイプ構築にかかる調査

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査機能要件

現状業務の意思決定フローを支援するため、以下のようなUI/UXを想定

振分対象メール

AI判定

Webアプリとしてブラウザから
利用

判定結果の編集

本番システム現状イメージ

機能要件

<入力>

- Webアプリとしてブラウザから利用する
- From, 件名, 本文をコピー&ペーストで入力する
- 添付ファイルをダイアログから選択、またはドラッグ&ドロップで追加
- word, ppt, excel, PDFは読取可能だが、パスワード付ファイルは対応困難)

<出力>

- 送信先の課・班の名称とメーリングリストのメールアドレス、推論スコア（複数・選択可）
- 送信先の判断根拠と添付ファイル内該当箇所を提示（ppt,PDFはページ番号まで、wordは根拠の存在する段落、excelはシート番号やシート名まで表示）

<判定結果の編集>

- 提示されなかった課・班を画面上から選択して追加できる。
- 判断根拠は参考情報であり変更不可。メール文面作成時に各担当者にて適宜記述する

<人事異動・組織変更>

- 課・班、課と班の親子関係、各組織のメーリングリストのメールアドレスを参照・更新する管理画面を設ける
- 個人単位でのマスタ管理は行わない

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査 非機能要件

貴庁との議論踏まえ、システム開発時のセキュリティ方針を整理

区分	本調査での方針
外部	接続元IPアドレスを制限して貴庁指定のIPアドレスおよび委託先のIPアドレスのみアクセス可能とする
	再学習データ蓄積のため、入力されたメールと添付ファイルをシステム内に保持する。その際、AWS ストレージサービスのサーバーサイド暗号化 (AES 256) を利用する (「電子政府推奨暗号リスト」記載の AES 128 アルゴリズムより強力な暗号化アルゴリズム)
内部	推論画面は認証*は設けず貴庁ネットワーク利用者であれば誰でも利用可能とする (機密性のある情報をユーザに提示しない想定のため、セキュリティリスクと利便性の兼ね合いから推論画面は認証不要と想定) *開発時の詳細議論が必要
	マスタ管理画面は管理者 ID, パスワードによる認証*を必要とする (多要素認証はコスト面の影響が大きく、実証期間中はパスワード認証のみで進める事が想定される。将来的に貴庁クラウド等に移行する場合、シングルサインオンの仕組みがあれば利用も検討) *開発時の詳細議論が必要
その他	ISMS 認証等、国際標準の規格に基づいた情報資産管理を実施する
	システム構築先としては ISMAP 登録クラウド事業者である Amazon Web Service 等を利用

(3) 中小企業庁の総括発注業務の効率化に係るシステム開発の要件定義にかかる調査 想定インフラ構成 (2022年度)

貴庁セキュリティポリシーに準じ、2022年度は以下のインフラ構成を想定

前提：貴庁クラウド内 (Teams, Outlook) - Internet - 委託先クラウド (アルゴリズム, 画面)

委託先開発部分

