

文書管理に係るAI等先端技術応用研究 2018-2020年度研究成果 最終報告

2021.9

鎌倉市

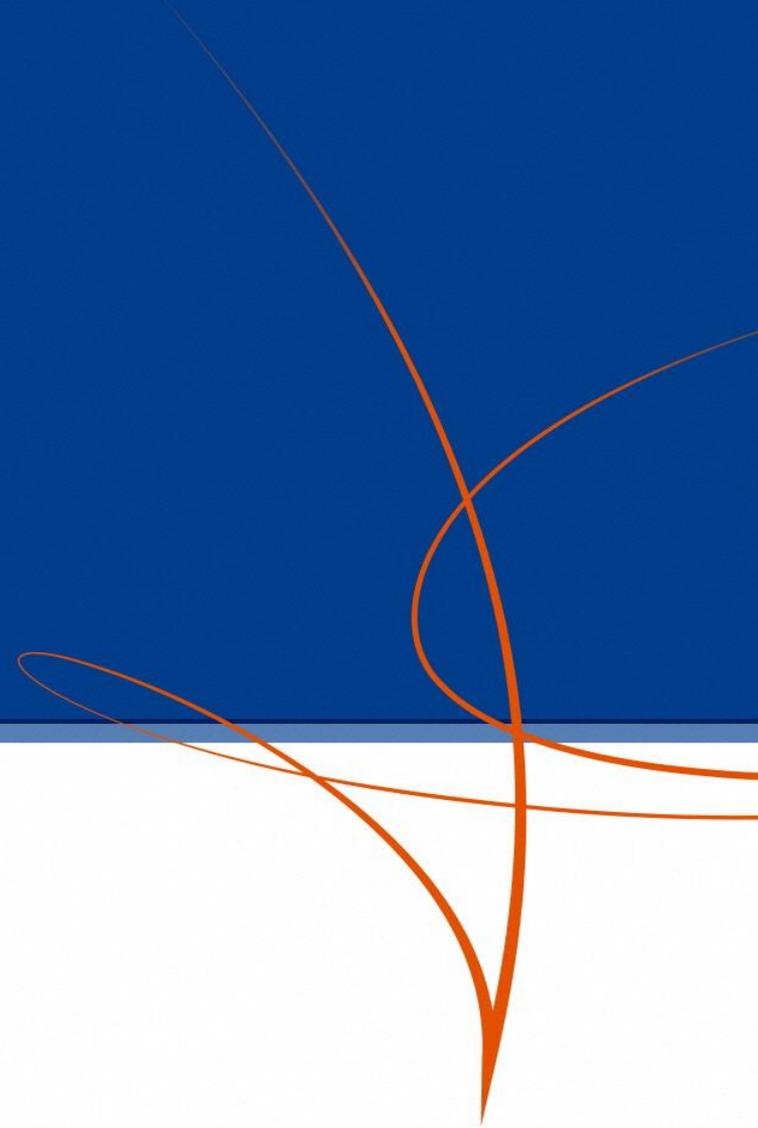
日本電気株式会社

NECソリューションイノベータ株式会社

目次

1. 実証実験の目的
2. 保存期間判定
3. 歴史的公文書判定
4. まとめ
5. 【巻末資料】用語解説

1. 実証実験の目的



1. 実証実験の目的

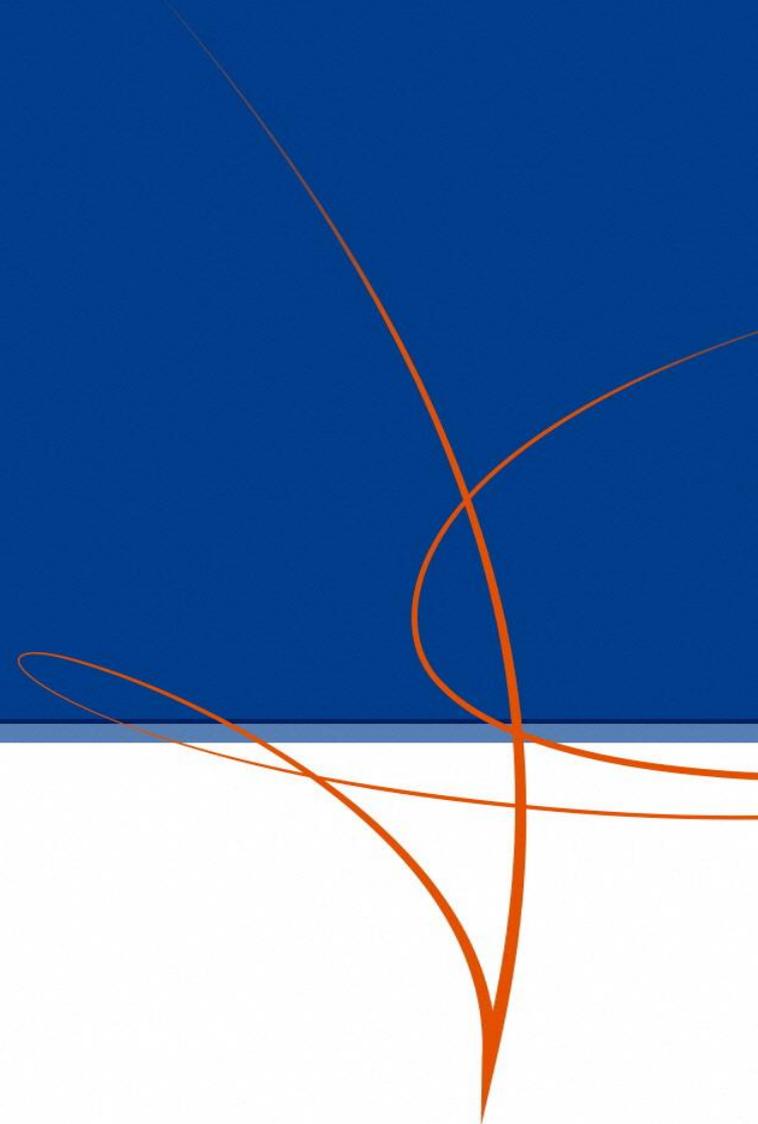
■ 今回、**行政文書の管理業務における統制強化、及び職員の負荷軽減・効率化**を目的として、AI技術を用いた実証実験を実施した。

■ 研究テーマは関係各課へのヒアリングを踏まえ、以下の2つとした。

1. 文書保存期間判定の研究

2. 歴史的公文書判定の研究

2. 保存期間判定



2. 保存期間判定（研究概要）

背景

市にて年間数十万件にも上り発生する行政文書は、徹底した保存・保管の運用がなされている一方で、その管理過程において一定の業務負荷が発生している。文書の管理基準となる「保存期間」をAIにて判定し判断材料とすることで、文書保存後の管理過程負荷を軽減する狙いがある。

目的

- 鎌倉市行政文書における保存期間設定誤りの検知。
- 文書管理業務（特に、保存・保管に係る業務）の効率化を目的とした保存期間レコメンド機能の実用性等の検証。

内容

- AIによる文書の保存期間判定を実施する。
- AI学習には、過去約10年分（2009年～2019年）の文書データを使用する。
- 廃棄予定日が2020年3月31日の文書データを判定対象とする。
- 判定結果をリスト化し市職員にてAI判定の正誤判断をすることで実証を行う。

目標

- モデル精度 90%以上

利用技術

- Google BERT

2. 保存期間判定（分析方法）

① 鎌倉市の文書データから
AI学習に用いる入力データを作成

- 文書件名
- 要旨
- 添付文書情報

保存期間

② モデル学習



③ AIモデルによる判定処理・結果のリスト化

鎌倉市統一データ										AI予測結果	作業	
所属名	大分類	中分類	小分類	フォルダ名	文書年	和暦文書年	決裁種別	文書番号	文書件名	保存期	AI予測保存期	職員確認
政策創造課	(010)政策創造部	(010)庶務	(020)通知	001.通知 (庁内)	2018	平成30年度	電子4紙	28簿政第177号	『市民の声』の送付について	1	3	
政策創造課	(010)政策創造部	(070)人事	(110)研修	001.研修	2016	平成28年度	電子4紙	28簿政第108号	第2回プラチナイブニングセミナー『産学官の協働によるイノベーションを創出するリンクーズの取り組み』について（後発）	3	5	
政策創造課	(010)政策創造部	(140)所管事業	(010)所管事業	001.研修	2016	平成28年度	電子4紙	28簿政第148号	第4回鎌倉市まじづくりプラシコンテスト決選大会の開催について	3	10	
政策創造課	(010)政策創造部	(070)人事	(110)研修	001.研修	2016	平成28年度	電子4紙	28簿政第119号	平成28年度「戦い療法（データ・リサーチリテラシー）」について（後発）	3	5	
政策創造課	(010)政策創造部	(070)人事	(110)研修	001.研修	2016	平成28年度	電子4紙	28簿政第169号	自治体職員向けオープンデータ研修への出席について	3	5	
政策創造課	(010)政策創造部	(070)人事	(130)使命	001.使命書	2016	平成28年度	電子4紙	28簿政第234号	シェア経済サミットについて（後発）	3	30	

2. 保存期間判定（モデル精度の評価）

下記より 一定水準には到達したものの、モデル精度目標（90%）未達

対象データ：廃棄予定日が2020年3月31日の文書データ

モデル精度：約 **79%** 100%に近いほど性能が高いといえる

ラベル	適合率	再現率	F値	判定件数
1年未満(即時廃棄)	11%	5%	7%	525件
1年	58%	69%	63%	29702件
3年	70%	70%	70%	54850件
5年	89%	85%	87%	118861件
7年	81%	67%	73%	400件
10年	82%	79%	80%	31376件
13年	0%	0%	0%	38件
15年	92%	68%	78%	529件
30年+99年	74%	82%	78%	15029件
平均	79%	78%	79%	251310件

※表中の適合率、再現率、F値等の用語について、詳細は巻末の用語解説資料を参照。

図1：保存期間毎の結果

2. 保存期間判定（モデル精度観点の評価）

- ・テスト時のモデル精度では87%をマークしたものの、本番時にはモデル精度が79%まで低下
- ・テスト時と本番時の判定対象データ件数差異が影響したと考えられる

	学習データ件数	判定対象データ件数	モデル精度
テスト時	約98万件	1万件	87%
本番時	約105万件	約 25 万件	79%

テスト時と比較して本番時にはモデル精度が低下する結果となった。
テスト時と本番時で、学習データ件数と判定対象データ件数の比率に大きく差異があったことが原因のひとつと考えられる。

<学習データ件数：判定対象データ件数>

- ・テスト時 98 : 1
- ・本番時 4.2 : 1

2. 保存期間判定（得られた知見・気づき）

- ・データ件数が少ない文書は、データ傾向の影響を大きく受ける
- ・混同行列をもとに現在の設定・運用経緯を検証することも可能か

図2から、以下の傾向を見て取れる。

- ・1年未満、13年保存文書は精度が極端に悪い。
- ・7年保存、15年保存文書は精度が比較的良い。

ラベル	適合率	再現率	F値	判定件数
1年未満(即時廃棄)	11%	5%	7%	525件
7年	81%	67%	73%	400件
13年	0%	0%	0%	38件
15年	92%	68%	78%	529件

図2：保存期間毎の結果（件数が少ないもの）

図3より、保存期間13年文書は全てが「3年保存」と判定されていることがわかる。

単純なAIによる判定誤りともいえるが、同一案件の文書でも、運用途中で文書規程・運用の見直し等により、当該案件の規定の保存期間が変更された可能性もあるため、文書管理業務におけるAIモデルの精度については、実際の運用経緯を踏まえ判断することも必要と思われる。

		判定結果								
		1年未満	1年	3年	5年	7年	10年	13年	15年	30年
実際の値	13年	0件	0件	38件	0件	0件	0件	0件	0件	0件

図3：混同行列（保存期間13年）

赤色となっている箇所の数字が、保存期間13年と判定された件数。
今回の実証では13年と判定された（的中した）件数は0件で、38件をすべて「保存期間3年」として判定している。

2. 保存期間判定（実証結果の評価・定量）

- ・ 正答率（市職員判定にて「AI判定が正当」となった率）に課題
- ・ ただし、実際に**81件**の文書の保存期間見直しに寄与できたことは成果

92課中、回答のあった26課分を対象に集計を行った。(図4)

市職員チェックの結果、AIが保存期間誤りの可能性があるかと判定した文書**1,220件**のうち、実際に訂正の必要があるとされた文書は**81件**で、正答率は**6.64%**という結果であった。

所属毎の正答率算出の結果、正答率が60%を超えたのは「納税課」「地域のつながり課」「固定資産評価審査委員会（納税課内部組織）」の3所属となった。

この3所属の正答率が高かった要因として、「業務特性上、取り扱う案件が一定しており、文書件名の命名規則にもある程度の法則性がある」ことが影響しているのではないかと推察される。

地域のつながり課においては、30年保存とAIが判定した文書**7件中6件が正しい結果であった。**

所属名（略称）	総件数	○	×	正答率
企画計画	57件	1件	56件	1.75%
政策創造課	6件	2件	4件	33.33%
秘書課	15件	3件	12件	20%
歴史まち	62件	15件	47件	24.19%
行政経営	7件	0件	7件	0%
IT	10件	1件	9件	10%
コンプラ	22件	0件	22件	0%
公的	51件	1件	50件	1.96%
総務	9件	0件	9件	0%
納税	5件	4件	1件	80%
地域	25件	19件	6件	76%
商工	93件	1件	92件	1.08%
深沢	34件	0件	34件	0%
都市計画	7件	0件	7件	0%
都市調整	36件	4件	32件	11.11%
都市景観	74件	0件	74件	0%
みどり	55件	4件	51件	7.27%
下水	234件	16件	218件	6.84%
公園	32件	0件	32件	0%
作業センター	2件	1件	1件	50%
議会総務	6件	0件	6件	0%
議事調査	10件	0件	10件	0%
学務	318件	6件	312件	1.89%
選挙	33件	0件	33件	0%
固定資産	5件	3件	2件	60%
指令情報	12件	0件	12件	0%
合計	1220件	81件	1139件	6.64%

図4：○×集計

2. 保存期間判定（実証結果の評価・定性）

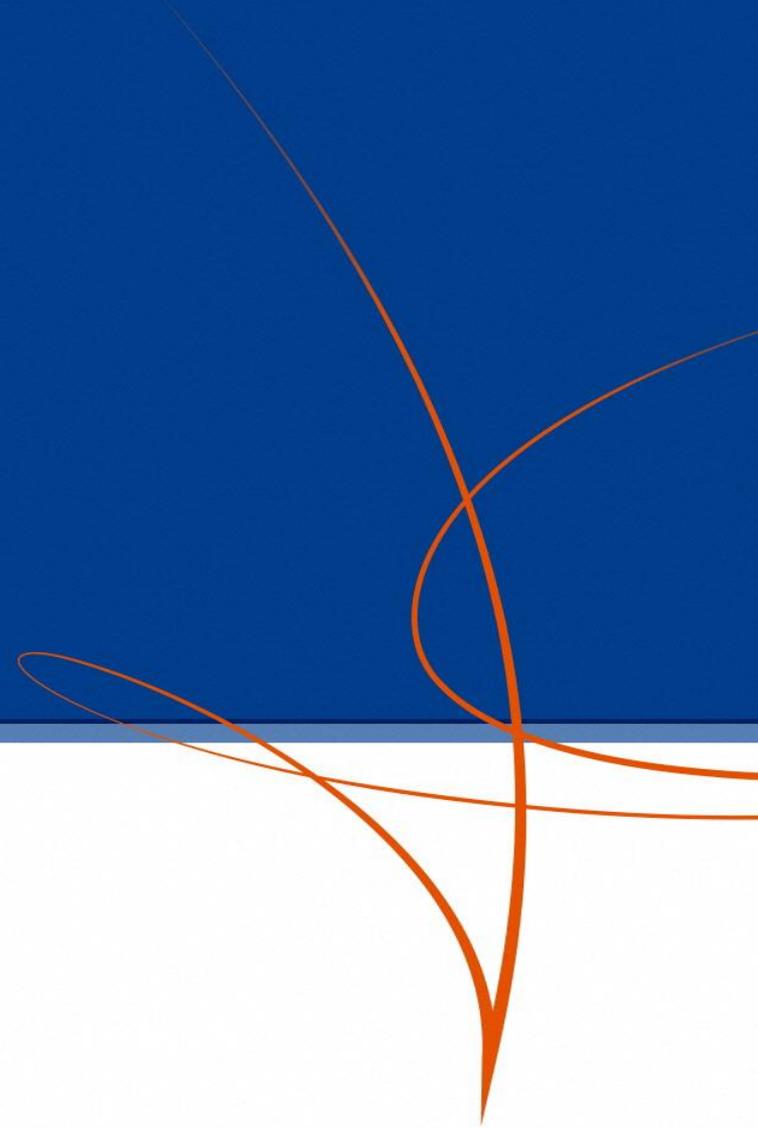
- ・ 保存期間の見直しには一定の効果
- ・ 保存期間をリコメンドする際は、判定根拠（過去文書事例）を同時に示す必要も

図5には、各課からのコメント、及びコメントに対する分析を抜粋して示す。コメントからも、保存期間見直しにおいて、一定の効果があったものとする。「数値だけの提示では判断材料とならない」という指摘は、今後の改善案としたい。

評価	所属	コメント	（コメントの分析）
◎	政策創造課	6件中2件の保存期間の誤りが発見できたのは、研究の成果だと感じる。AIが誤っていると予測した理由が分かること納得感に繋がる。また、起案時点でAIの判定ができるようになれば、ミスの撲滅や確認作業の削減につながるため、今後本格導入を期待したい。	正答率にとらわれず「2件抽出できた」ことは研究成果として評価できる。AI予測の根拠については、複数課から同様のコメントがあり、今後の課題である。
○/△	公的不動産活用課	保存期間を延長するかどうかの判断の参考となった。ただし、延長後の適正な保存期間については、AIからの提案はほとんどが誤っていた。	判断の一助として評価された点は良かった。ただし、コメントにあるとおり、延長後の保存期間のAIによる予測精度は改善の余地があることが判明した。
×	行政経営課	AIがどのようなデータに基づいて判定したのか等、判定基準が不明なことから、予測結果が合っているかどうかの判断がつかなかった。AIの判定基準を知りたい。	AI予測の根拠については、複数課から同様のコメントがあり、今後の課題である。
×	深沢地域整備課	全般的な印象として、軽易な文書と重要な文書の判別が、AIでは困難であるように感じた。	同課においては、AIの判定が合っていた件数は0件であった。

図5：定性回答

3. 歷史的公文書判定



3. 歴史的公文書判定（研究概要）

背景

歴史的公文書に資する重要文書は複数回の選別を経て決定される。現在鎌倉市では各課による1次選別運用の定着化・高度化を図っており、AIによる補助が可能か検証する狙いがある。

目的

- 現在試行されている「各課による歴史的公文書の選定」効率化・精度向上への寄与。
- 「保存期間」以外の文書属性データに対する判定機能の検証。

内容

- AIによる歴史的公文書判定を実施する。
- AI学習には、中央図書館にて保有する歴史的公文書データ(約5400件)を使用する。
- 廃棄予定日が2020年3月31日の文書データ*1を判定対象とする。
- 実際の業務フロー*2を鑑み、歴史的公文書判定結果の市職員によるチェックは実施しない。

*1 AI判定時点において、該当データに対する歴史的公文書選別はまだ実施されていない状態であるため、「30年保存」として設定されているデータを歴史的公文書、それ以外を非歴史的公文書として見做し判定処理を実行。

*2 実際に歴史的公文書選別結果が判明するのは実証期間外である2020年7月以降のため、実証期間中の照合（**実証**）作業は不可と判断。

目標

- モデル精度 80%以上

利用技術

- Google BERT

3. 歴史的公文書判定（分析方法）

①中央図書館保有の歴史的公文書選別結果データ、及び鎌倉市文書データからAI学習に用いるインプットデータを作成。歴史的公文書に該当するかどうかをデータに付与した。

・ 文書件名

歴史的公文書に該当するものを1、それ以外を0とした歴史的公文書フラグをインプットデータに付与。

②モデル学習



③AIモデルによる判定処理

3. 歴史的公文書判定（精度観点の評価）

- ・ 下記より、**テスト時のモデル精度目標（80%）は未達。**
- ・ **本番時には実際の歴史的公文書の判定結果ではなく、保存期間が30年のデータを歴史的公文書と見なして判定を行った。このため、正確な検証には至らず、参考記録に留まった。**

テスト時 : 中央図書館から借用したデータ約2000件を対象に判定
モデル精度 : 約 **47%**

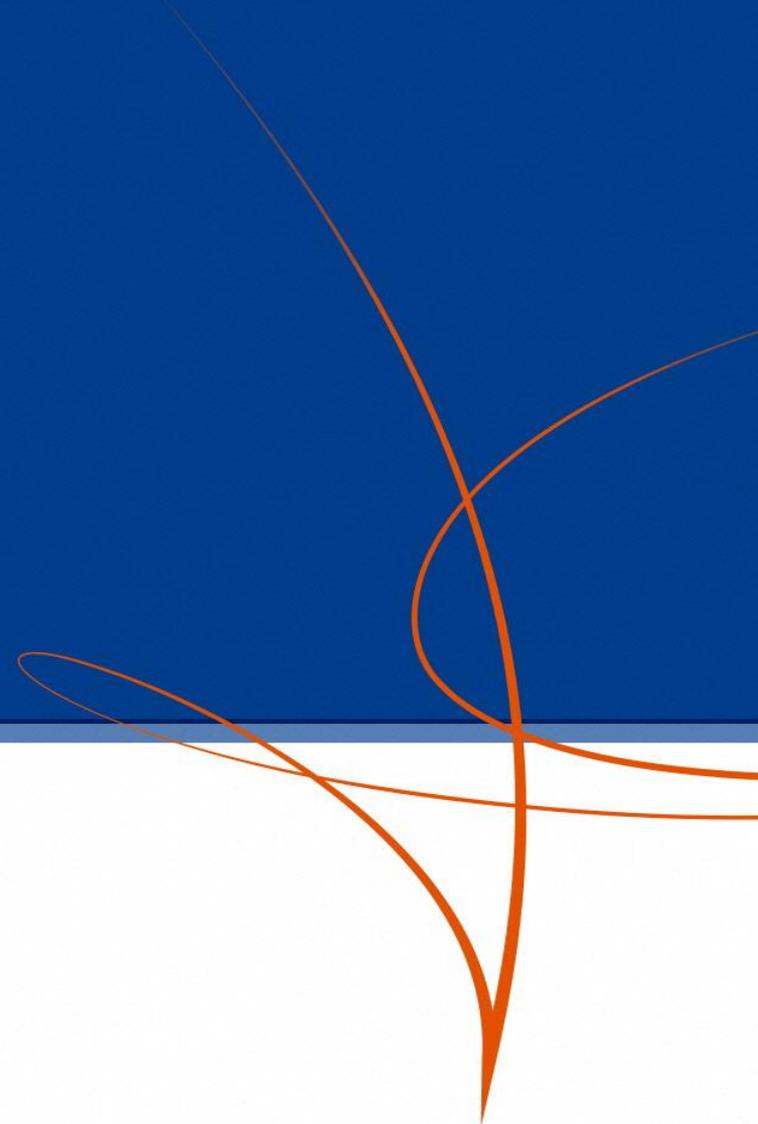
本番時 : 廃棄予定日が2020年3月31日の文書データを対象に判定
モデル精度 : 約 **83%**

テスト時のモデル精度については、モデル学習に利用できたデータ件数が少なかったこともあり、モデル精度は47%に留まった。

また、本番時の判定対象データについて、鎌倉市の運用上、実際に歴史的公文書に該当するかどうかは実証期間以降に判断されるスケジュールとなっていたため、30年保存の文書のみを用いて歴史的公文書にあたるかどうか判定を行った。

上記経緯から、モデル精度83%はあくまで参考記録となる。

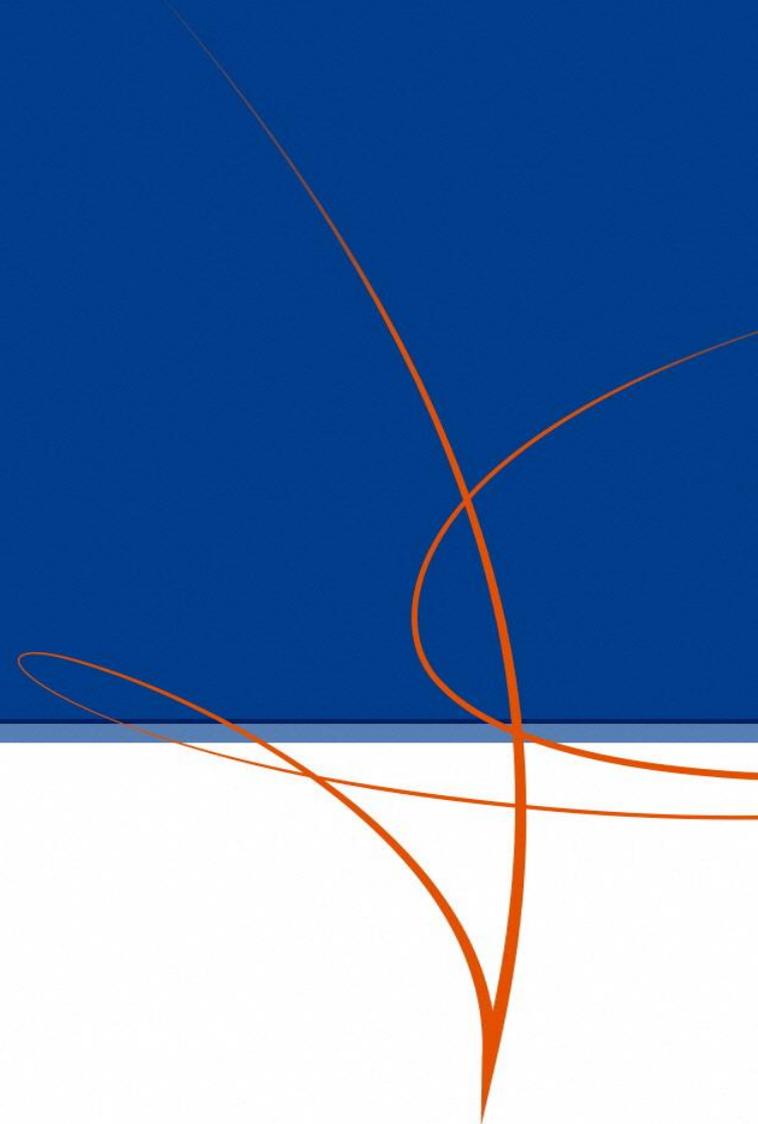
4. まとめ



4. まとめ

- 今回、実際の行政文書データの一部を活用しAI機能の実証を行えたことで、鎌倉市・NECのみならず、各自治体における文書管理業務の高度化に向けた大きな知見を得ることができた。
- 複数部局を巻き込みワークショップを開催し、業務フローや課題・利用データを洗い出す等、直接的な機能開発以外でも共創しながら取り組むことができた。
- 所管課毎の業務特性による保存期間判定機能の有用性差異や、保存期間提示の際に根拠の見える化が必要である等、AI機能を利用するにあたっての課題を発見できた。
- AI機能を利用した文書管理業務における業務負荷軽減・効率化について、その足掛かりとなる実証を実践することができた。
- 歴史的公文書判定に関しては、利用データの選定や組み合わせの考案等、AI機能開発のプロセスに課題が残った。
- 実証期間が新型コロナウイルスによる緊急事態宣言の期間と一部重複し、効果検証のための全庁への協力依頼は当初の想定よりも簡便な形式となってしまったが、関係者の連携、協力により、中止とすることなく、無事に共同研究を終了することができた点は大きな成果となった。

5. 【卷末資料】用語解説



5. 【巻末資料】用語解説

● ラベル

データに付与された正解データのこと。

一般的に、入力されたデータに対し、

この付与された正解データを導けるようにAIモデルは学習を行っていく。

今回実証では「保存期間」や「歴史的公文書にあたるかどうか」をラベルとした。

● 混同行列

あるデータを分類したときに、その正解・不正解を整理しておく表のこと。

例) データとして与えた動物の画像が「犬」の画像であるか判定するモデルがあり、判定対象データ中に犬の画像が8件、猫の画像が2件あったとする。

右に例として挙げる混同行列からは、以下のことがわかる。

- ・犬の画像5件を正しく「犬」と判定している。
- ・犬の画像3件を誤って「犬ではない」と判定している。
- ・猫の画像1件を誤って「犬」と判定している。
- ・猫の画像1件を正しく「犬ではない」と判定している。

		判定結果	
		○	×
実際のデータ	犬	5件	3件
	猫	1件	1件

図5：混同行列の例

● 適合率

AIが正解と判定したデータのうち、実際に正解であるデータの割合。

● 再現率

実際に正解であるデータのうち、AIが漏らさずに正解と判定できたデータの割合。

● F値

適合率と再現率のバランスをとった指標。今回実証の「モデル精度」はこのF値を指す。