

# 書類整理の自動化

青山学院大学理工学部  
情報テクノロジー学科 Dürst 研究室

学籍番号：15820094

嘉松 一汰

# 目次

第1章	はじめに	2
1.1	研究背景	2
1.2	研究目的	3
1.3	構成	3
第2章	関連研究	5
2.1	既存手法やクラスタリングの課題	5
2.2	基礎技術	5
第3章	提案手法	7
3.1	提案する手法・アプローチ	7
3.2	システム設計	7
第4章	実験・評価	10
4.1	様々なドキュメントでの実験	10
4.2	精度の分析	10
4.3	既存手法との比較	11
第5章	考察	12
5.1	ドキュメントごとの結果の解釈	12
5.2	改善点	12
第6章	おわりに	14
6.1	今後の展望	14
6.2	研究のまとめ	14
	参考文献	15

# 第1章 はじめに

本章では、本研究を始めるにあたっての動機、および本論文の構成を示す。

## 1.1 研究背景

日本企業のRPA(Robotic Process Automation [3]) 導入率は全体で38%、中堅・中小企業では25%となっており、非常に少ないことがわかる。(図 1.1) また、大企業と中堅・中小企業の間に20%以上の差があり、技術や規模による格差も見取れる。これらの原因となっており要因として考えられることは、大きく分けて2つある。1つ目は、RPAには専門領域と非専門領域が存在するということである。専門領域はPC上の操作や、デジタルの領域における処理である。いっぽうで、紙媒体の処理等の、アナログの世界で行われる処理は非専門領域としている。特に、手書きの文字や画像の認識を高い精度を保ちながら自動で行うことは、現代では非常に困難なことである。縦書き文字横書き文字が混在していたり、旧字体や特殊文字等の組み合わせも考えられるため。例外的な処理までを自動で行う必要があるからである [1]。2つ目の原因は、紙媒体の業務を行っているコミュニティのIT知識の乏しさにある。詳しくは次のセクションで説明する。このような現状を踏まえて、次章からは、OCR技術を使用したアプローチを提案する。



図 1.1: 企業の RPA 導入率

## 1.2 研究目的

本研究の目的は、IT 知識が乏しく、紙媒体の業務を行っているコミュニティを中心に、RPA を使用して紙媒体の処理を自動で行うシステムを作成することである。具体的な内容に入る前に、紙媒体の業務と IT 知識の関連性について深掘りする。日本で働く人事・総務担当者に、「紙媒体中心の業務で不便を感じたことがあるか？」とアンケートを取ったところ、61%が不便を感じたことがあると回答した。(図 1.2)。上記の理由として、システム障害への恐怖感や、IT 知識の乏しさが挙げられます。しかし、一連の流れを RPA にすることで、IT 知識の有無に関わらず、システムを運用することを目指す。

## 1.3 構成

本論文は 6 章構成になっている。第 1 章では、本研究の背景や目的を述べる。第 2 章では既存手法との比較や課題を述べる。本研究での実装手法を第 3 章で提案し、第 4 章で実験と評価、第 5 章で結果の考察を述べ、第 6 章で本研究の総括を行う。

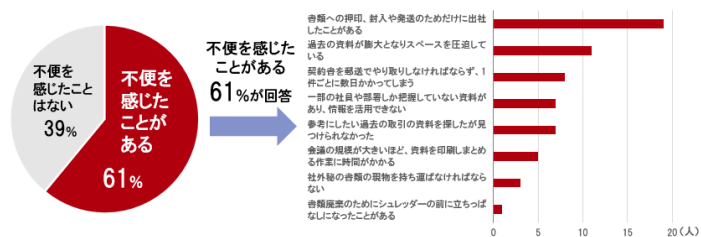


図 1.2: 紙媒体中心業務に関するアンケート

## 第2章 関連研究

本章では、本研究の課題や、使用する基礎技術を示す。

## 2.1 既存手法やクラスタリングの課題

[illegible]

メタな記号\$.

数式

$$\sum_{k=0}^{100} (|x^k| + 1)$$

モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。  
モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。

## 2.2 基礎技術

模気という構造を用いる。模気という構造を用いる。模気という構造を用い  
る [?]。模気という構造を用いる。模気という構造を用いる。模気という構造を  
用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を  
用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を  
用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を  
用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を

用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。

模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。  
 模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる [2]。  
 模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。

?? 節 で述べた○○は黒い。特に ??項 のはとても黒い。

## 第3章 提案手法

アノテーションの定義と、アノテーションの内部データ構造を説明する [?].

### 3.1 提案する手法 ・ アプローチ

モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる [?]. モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる [?]. モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる [?].

モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。  
 モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
 ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モゲ  
 という概念を用いる。モゲという概念を用いる。モゲという概念を用いる。図 3.1  
 の長方形に角は四つである。本文中から必ず参照する。オプションについても調  
 べた。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用  
 いる。モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用  
 いる。

PowerPoint で図を作るときは、スライドサイズを先に指定すると余白が少なくなる。

モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。  
モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。

### 3.2 システム設計

模気という構造を用いる。模気という構造を用いる。模気という構造を用いる [?]. 模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。表 3.1 は豹ではなく表である。本文中から必ず参照する。オプションについては解説しない。



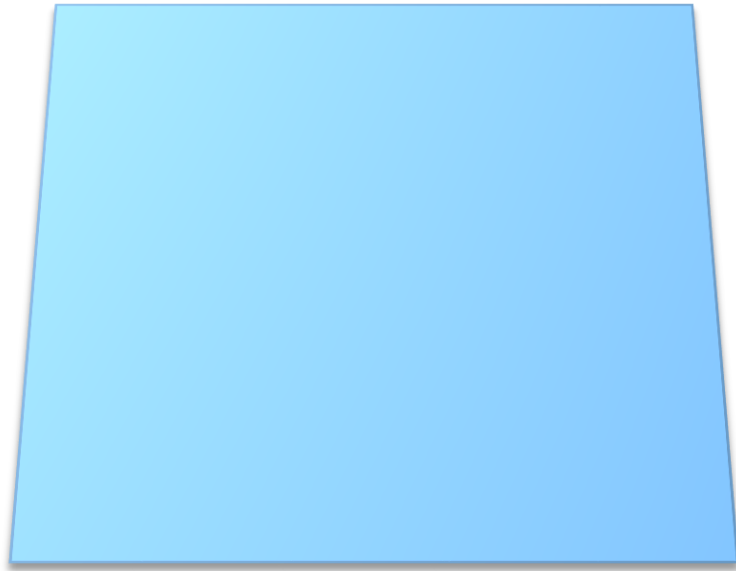


図 3.1: PowerPoint 上での矩形

表 3.1: 情報テクノロジー学科線路側研究室〇〇情報

研究室	場所	人数
佐久田研	O-YYY	5
原田研	O-XXX	7
Dürst 研	O-527	3
大原研	O-ZZZ	2

模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。  
模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる [2]。  
模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。

## 第4章 実験・評価

アノテーションの定義と、アノテーションの内部データ構造を説明する [?].

## 4.1 様々なドキュメントでの実験

[illegible]

メタな記号\$.

## 数式

$$\sum_{k=0}^{100} (|x^k| + 1)$$

モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。  
モゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。モゲという概念を用いる。モ  
ゲという概念を用いる。モゲという概念を用いる。

## 4.2 精度の分析

[illegible]

用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。

模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。  
 模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる [2]。  
 模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
 気という構造を用いる。模気という構造を用いる。模気という構造を用いる。

### 4.3 既存手法との比較

模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。  
模気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる。模  
気という構造を用いる。模気という構造を用いる。模気という構造を用いる [?]。

?? 節 で述べた〇〇は黒い。特に ?? 項 のはとても黒い。?? 項 のはやや黒い。

## 第5章 考察

アノテーションの定義と、アノテーションの内部データ構造を説明する [?].

### 5.1 ドキュメントごとの結果の解釈

モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる [?]. モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる [?]. モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．

- 1991 年の×△△△△の AAA は秀逸であると感じた。
- 2001 年の△△△△△の BBB がよかった。
- 1997 年の△○×の CCC がややよかった。

箇条書きは便利ですが、○○なので控える。論理的なごまかしや妥協につながる可能性があるかもしれないかもしれないと思うのかもしれない。表や地の文で書く。

モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．モゲという概念を用いる．

### 5.2 改善点

模気という構造を用いる．模気という構造を用いる．模気という構造を用いる [?]. 模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．模気という構造を用いる．



## 第6章 おわりに

[illegible][illegible]

## 6.1 今後の展望

## 6.2 研究のまとめ

## 参考文献

- [1] D-Analyzer. RPA にできないこと、不得意な業務はなにか？ , 2019.
- [2] Koichi Sasada. Yav: yet another rubyvm: innovating the ruby interpreter. In *Companion to the 20th annual ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*, pages 158–159. ACM, 2005.
- [3] yasuihiro sasaki. RPA(Robotics Process Automation) の可能性. *2017年春季全国研究発表大会*, 2017.