

Sansanの研究開発部は、あらゆる種類のビジネスデータに向き合い事業を支え続けるとともに、研究開発を通して新たな価値を創出することで事業成長をけん引する組織として活動しています。

Team:

多様な分野から高い技術力を持ったメンバーが集う組織

## Automation Group

これまでにない速度と限りなく100%に近い精度、セキュリティーの確保を目指し、画像認識や機械学習などの技術を用いて、名刺やビジネス文書といった非定型でアナログな情報のデータ化を効率化・自動化する技術の研究開発に取り組む。

# 画像領域 # OCR # データ化 # 自動化

## SocSci Group

社会科学の各領域を専攻するメンバーで構成。「リアルとバーチャルのはざまで、社会理論を武器に名刺交換の価値を拡張する」をミッションとして掲げ、研究を通して得られた知見や成果を事業やサービス、社会へと還元することに取り組む。また、外部有識者との共同研究も積極的に展開する。

# 社会科学領域 # 新規プロダクト開発 # 効果検証

## Data Analysis Group

機械学習、自然言語処理、画像処理といった技術を用いて、あらゆるビジネスデータを収集・解析する。多様なビジネスデータを組み合わせることで、レコメンデーションエンジンやニューロス配信エンジンなどの構築に取り組む。

# 自然言語処理領域 # レコメンド # 予測

Sansan株式会社  
技術本部  
研究開発部

## Architect Group

「研究開発部が提供するサービスに責任を持つ」をミッションとして掲げ、サービスの可用性、拡張性、保守性の担保に取り組む。また、研究開発部の開発生産性に向き合いながら、データ基盤の整備、データの品質、データの取り扱いにおける安全性の担保にも取り組む。

# DevOps # MLOps # データ分析基盤

Service:

Sansan株式会社が提供するDXサービス

営業DXサービス

営業を強くする  
データベース

キャリアプロフィール

名刺管理に、転職に

クラウド請求書受領サービス

請求書受領から、  
月次決算を加速する

クラウド契約業務サービス

契約業務のDXから、  
リスクを管理する

その他、提供サービス多数

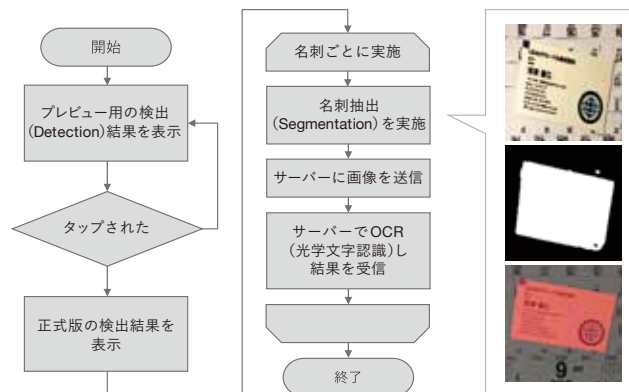
Outline of the Presentation by Sansan R&D

# 自然言語、画像、社会科学領域における研究開発とビジネス応用

画像領域

## 名刺領域のSmart Capture

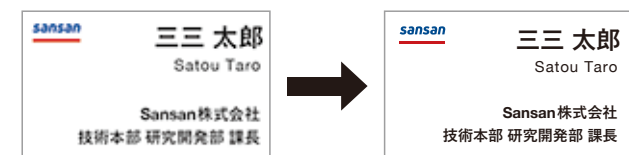
モバイルによるリアルタイム画像認識技術。従来は困難だった白い背景と白い名刺との同化、複雑な背景、影の重なりといった問題に対応。



画像領域

## 名刺画像の超解像

CNNベースの超解像モデルを用いて、低解像度の名刺画像から高解像度の名刺画像を復元。OCRの認識精度の向上につながった。



社会科学領域

## ネットワーク解析アルゴリズムのOSS公開

ジョーンズホプキンス大学のAngelo Mele准教授と共同開発したネットワークの生成メカニズムを解析するためのアルゴリズムをオープンソースとして公開。計算コストのボトルネックを解決することでSansanが保有する大規模な名刺交換ネットワークの分析に活用。



<https://github.com/sansan-inc/lighthergm>

画像、自然言語処理領域

## 請求書 / 契約書からの情報抽出

請求書や契約書のPDFをOCRと文書解析を組み合わせてデータ化するアルゴリズムを開発。加えてオペレータによる入力補助により、検出漏れを防ぐワークフローになっている。



※守秘義務を負ったオペレータがデータ化を行っています

自然言語処理領域

## メール署名から自動名刺作成

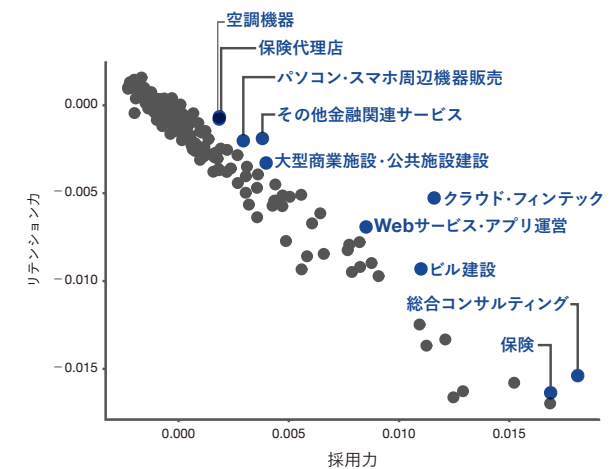
BERTによる固有表現抽出をメール向けに最適化。メールの文章中から署名を抽出し、署名から氏名・組織名などの名刺相当な情報を抽出するアルゴリズムを開発。



社会科学領域

## 転職ネットワークの分析

企業の特徴を表現した node embedding をレーティング分析に導入し、各業種が転職市場でもつアドバンテージを推定。



[https://speakerdeck.com/sansan\\_randd/dsr13](https://speakerdeck.com/sansan_randd/dsr13)

インターン情報

## Sansan R&D Summer Internship 2022

2週間の実務体験型インターンシップ。実際のサービスに基づくデータ分析を通して、研究開発の実務を体験いただけます。名刺をはじめとした出会いのデータを生かし、ユーザーへ新しい価値を提供するような探索と既存課題のさらなる深化につながる研究に取り組んでいただけます。

例年の様子など  
詳細は下記URLへ



最新情報



Sansan R&D Twitter  
@SansanRandD

R&Dの技術・研究の情報やイベント情報、  
普段の様子までいち早くお知らせしています。



Sansan Builders Blog  
<https://buildersbox.corp-sansan.com/>

ものづくりに関わるメンバーによるテックブログ。  
さまざまな技術的な知見を共有しています。

