

# *sansan* DSOC

Data Strategy & Operation Center

## データ統括部門DSOCに関する説明資料

Sansan株式会社

2021年4月14日



## 常楽 諭 Satoru Joraku

Sansan株式会社  
取締役／CISO<sup>(1)</sup>／DPO<sup>(2)</sup>／DSOC<sup>(3)</sup> 担当

2007年にSansan株式会社を共同創業し、法人向けクラウド名刺管理サービス「Sansan」のプロダクト開発を統括。現在は名刺のデータ化やデータ活用の研究開発部門であるDSOCの担当役員を務めながら、CISOとして社内のセキュリティ施策を推進する。

- (1) CISO: Chief Information Security Officer
- (2) DPO: Data Protection Officer
- (3) DSOC: Data Strategy & Operation Center

# 目次

---

**1** DSOCの概要

**2** DSOCの役割

**3** DSOCの将来像

## 1 DSOCの概要

- ミッション
- 沿革
- 組織体制
- 優秀な人材

## 2 DSOCの役割

## 3 DSOCの将来像

データ統括部門として市場における競争優位性を生み出す

## ミッション

---

# Activating Business Data

名刺、書類、業績や株価など企業の公開情報

あらゆるビジネスデータから「出会いのデータベース」を構築し、  
ビジネスや社会、未来へとつながる新しい可能性を生み出す。

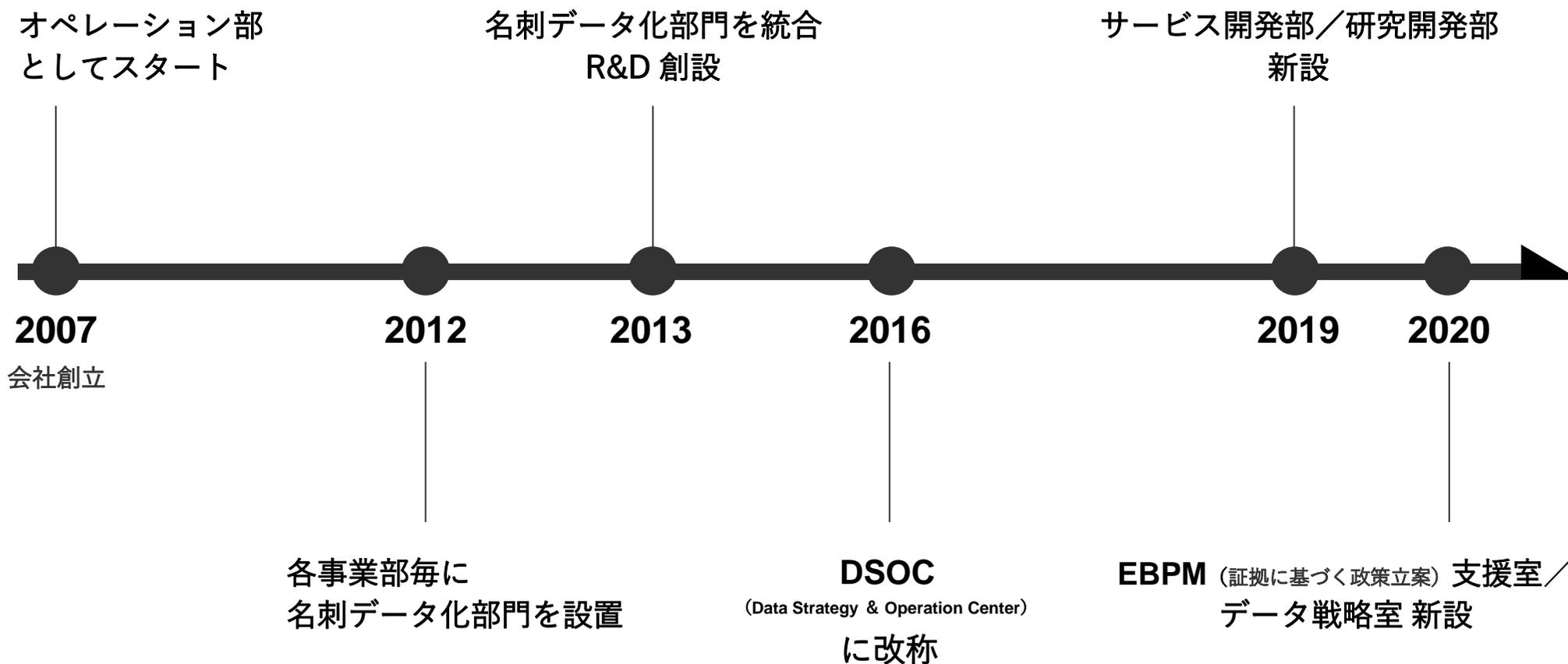
## DSOCの3つの役割

---

データの「生成」 「整理」 「活用」

# 沿革

名刺データ化のオペレーション部門から、あらゆるビジネスデータを扱うデータ統括部門へ



## 組織体制・人員構成

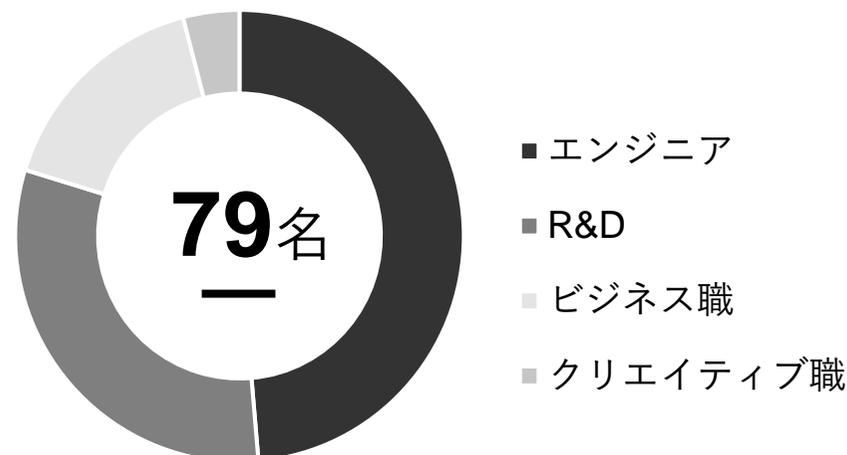
サービス運営・開発とデータ活用に関わるR&Dを中心とした組織体制  
エンジニアと研究員が人員構成の大半を占める

DSOCの組織体制<sup>(1)</sup>

### sansan DSOC



DSOCの人員構成<sup>(2)</sup>



(1) 2021年3月1日現在  
(2) 2021年2月28日現在

## 優秀なエンジニアリング人材と研究員

データ化技術に関わるエンジニアのほか、社会科学や経済学専門の研究員といった多様なバックグラウンドを持つスペシャリストが在籍

画像処理

自然言語処理

深層学習

機械学習

複雑ネットワーク

経済学

経営学

社会学

# kaggle



Grandmaster 在籍

### 博士学位

博士（理学）2名、博士（経済） 1名  
博士（工学）2名

# 目次

---

## 1 DSOCの概要

## 2 DSOCの役割

- データの生成
- データの整理
- データの活用

## 3 DSOCの将来像

## データの生成：名刺データ化における特有の課題

データ活用のためには、名刺情報が正確にデータ化されることが必須  
特有の課題のため、名刺をOCRのみで正しくデータ化するのは現実的に困難

### 名刺データ化に求められる精度



氏名や会社・組織・役職名、連絡先等の人物情報

1文字でもデータに間違いがあると、  
情報としての価値が毀損する

名刺のデータ化は、  
限りなく100%の精度が求められる

### 名刺データ化における課題



- 非定型  
様々なデザイン、形状、文字サイズ



- 画像のクオリティ  
撮影環境に左右される



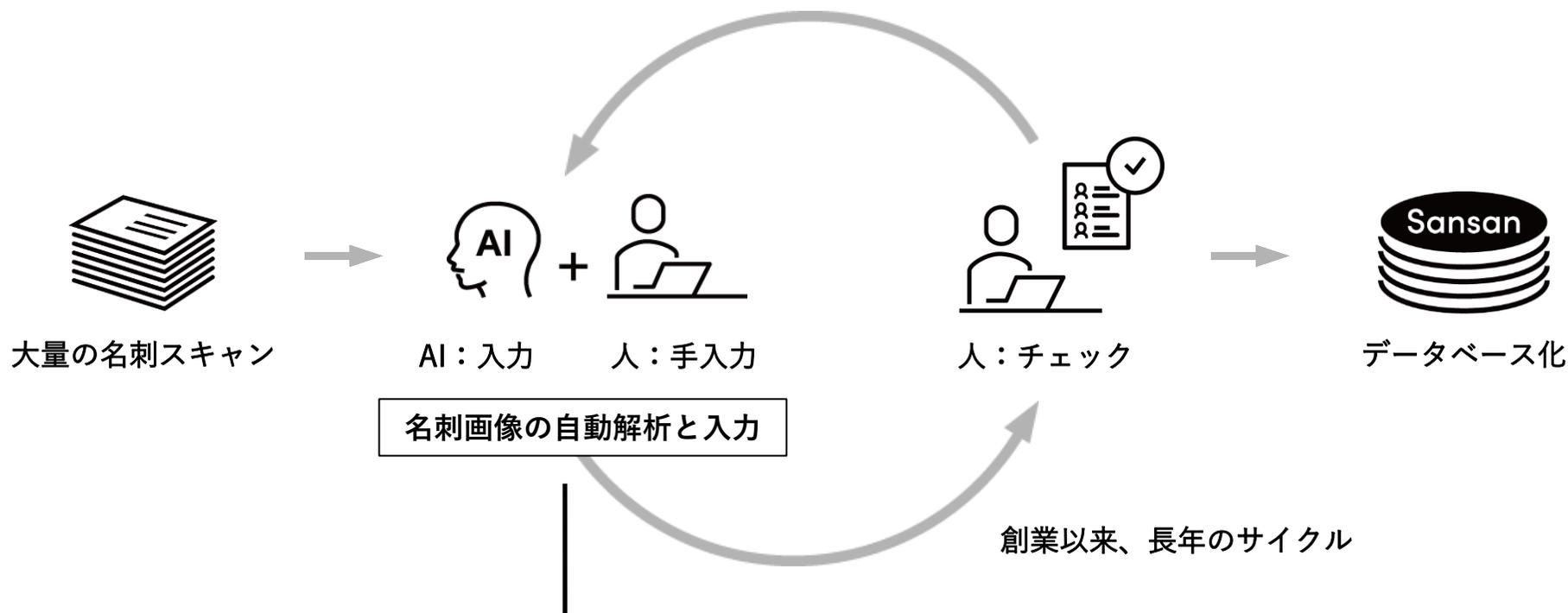
- 項目特定  
氏名や肩書き等の項目を判定  
表記のルール統一

OCRのみで名刺の正確なデータ化は困難

# データの生成：名刺情報のデータ化オペレーション

競争優位性の源泉である名刺データ化精度99%以上を実現する仕組みとテクノロジー

## 大量の名刺情報を、迅速かつ正確にデータ化



### テクノロジーの組み合わせ

名刺項目の自動判別・  
マイクロタスク化

名刺言語の自動判定

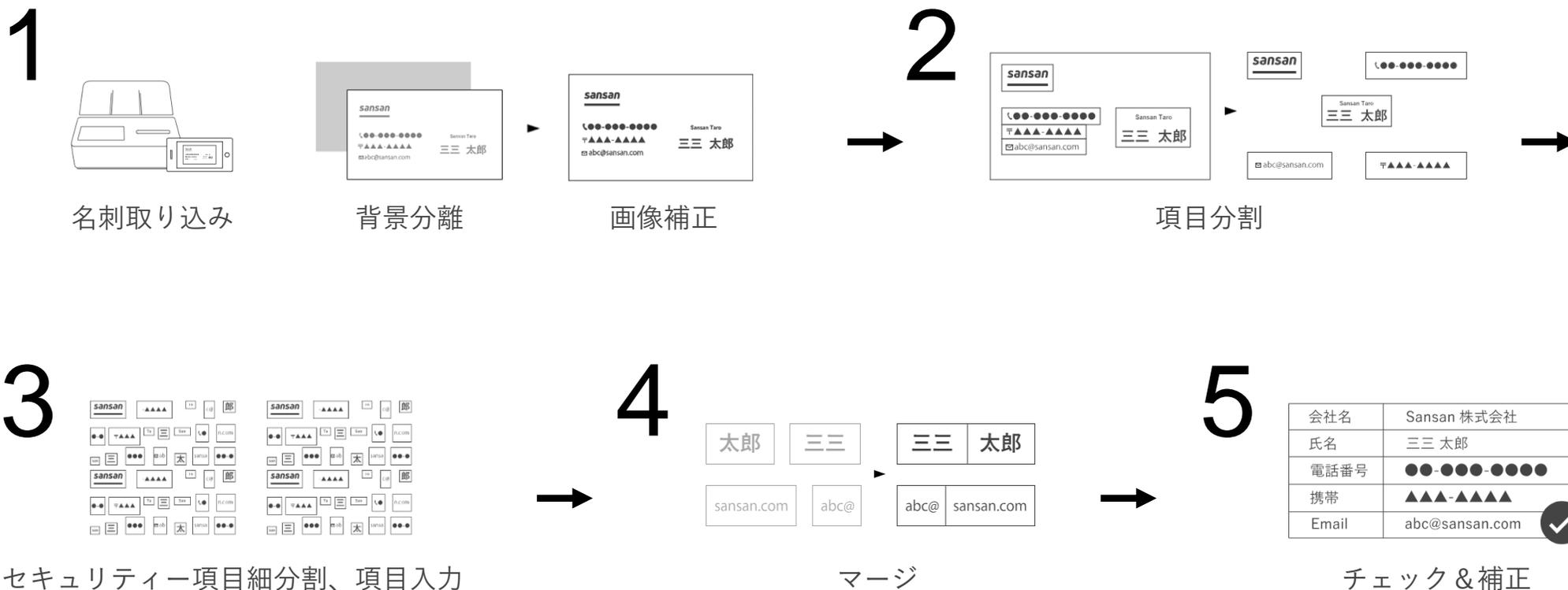
オペレーターへの  
自動振分け

分散化された入力  
オペレーターネットワーク

画像処理技術

# データの生成：OCRを補完する取り組み① - データ化システム「GEES」 -

大量の名刺を正確かつ効率的にデータ化する独自のオペレーションシステム

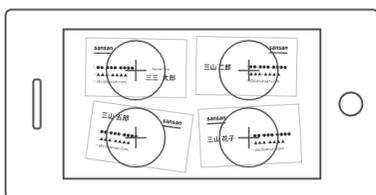


セキュアな環境を構築

会社名	Sansan 株式会社
氏名	三三 太郎
電話番号	●●-●●-●●●●
携帯	▲▲▲-▲▲▲▲
Email	abc@sansan.com



AIを活用した独自の画像認識技術によって、名刺を高速かつ高精度でデータ化



## スマートキャプチャー

撮影されてから数秒で結果をユーザーに届けることを可能にする技術



## 言語判定

文字を読み取らずに言語を判定



## 項目セグメンテーション

文字を読み取らずに、名刺のデザインから項目を見分ける



## ミスタイクディテクター

誤りの傾向を学習してミスの可能性を予測

## データの生成：OCRを補完する取り組み③ – DSOC OCR –

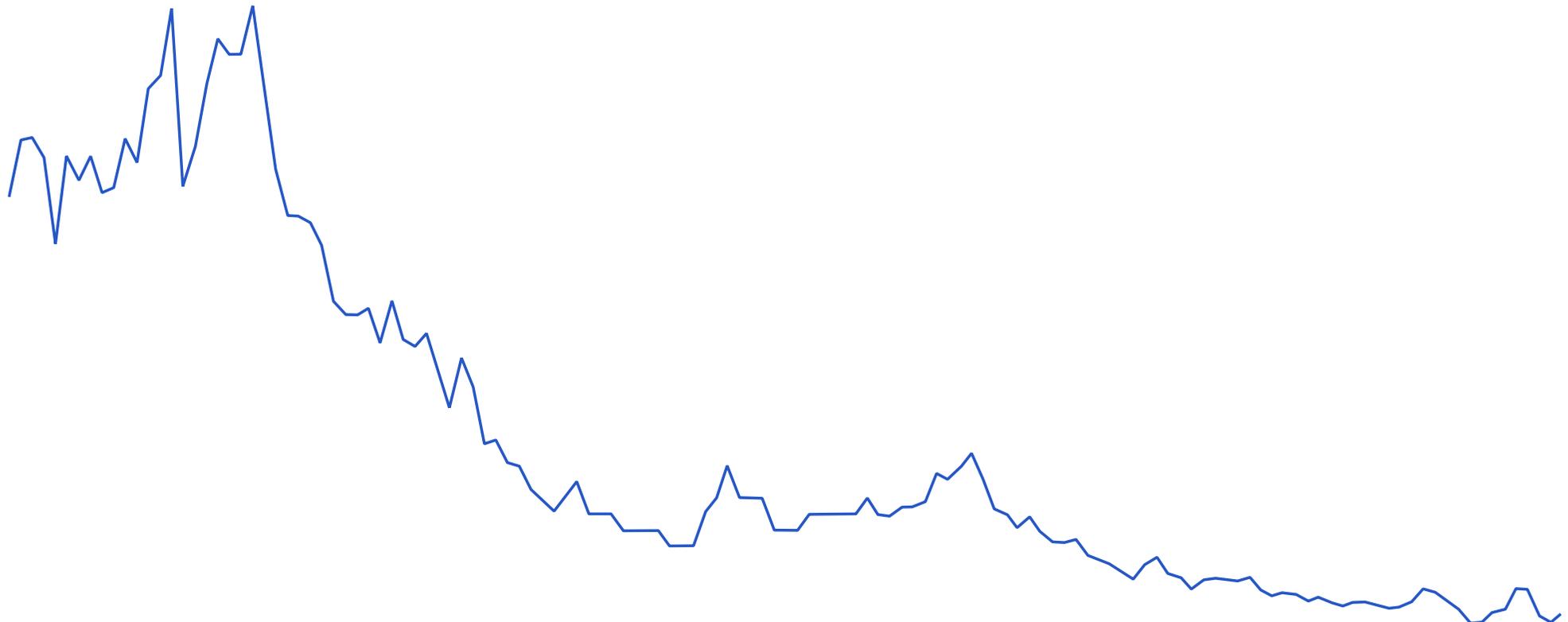
名刺に特化したOCRエンジンを独自に開発し、メールアドレスを99.7%以上の精度でデータ化  
氏名の読み取りも開始しており、名刺全項目のデータ化に向けた開発・研究を行う

Sansan DSOC 独自OCRエンジン

DSOC  
OCR

## データの生成：名刺1枚当たりのデータ化費用の推移

名刺データ化オペレーションの進化とともに、1枚当たりデータ化費用は創業時の20分の1以下に



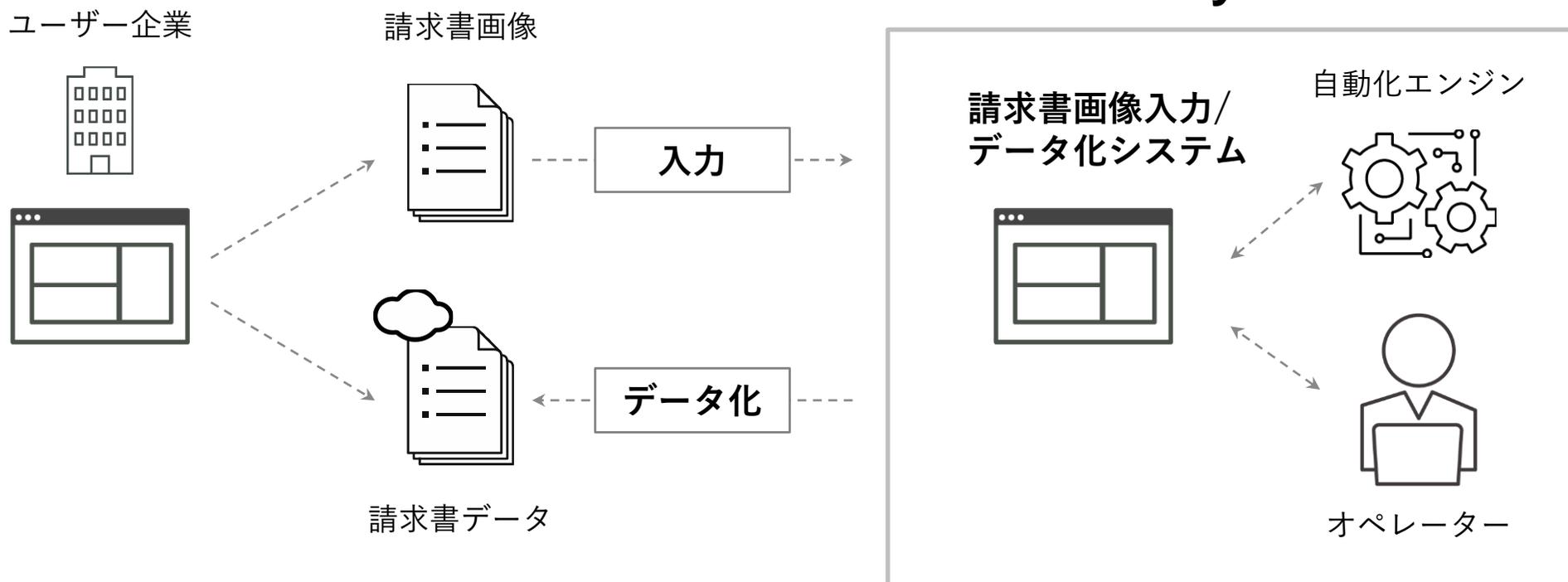
現在

# データの生成：名刺データ化オペレーションの他分野への活用

正確な名刺データ化オペレーションを請求書のデータ化等の他分野に横展開



## Bill One Entry



## データの整理：データの有効活用における課題

蓄積されたデータの一部は、社名が古い、入力漏れがある等のデータの欠落があるため、営業やマーケティングに最大限活用ができない状態

取引先名	三三株式会社	社名が古い
HP		入力漏れ
電話番号	03-6758-0033	
所在地	東京都渋谷区神宮前5-52-2	
担当社名	山田 太郎	
部署	デジタルマーケティング室	
役職	室長	役職が古い
E-mail	Yamada.taro@33i.co.jp	

## データの整理：データ基盤の構築

名刺データを正確かつ最新化した上でさまざまなビジネス情報を加えてリッチ化し、  
ビジネスで有効活用できるデータ基盤を整備



## 正規化・最新化

取引先名	Sansan株式会社	社名を正規化
HP	<a href="https://jp.corp-sansan.com/">https://jp.corp-sansan.com/</a>	
所在地	東京都渋谷区神宮前5-52-2	入力漏れを補足
電話番号	03-6758-0033	
担当者名	山田 太郎	
部署	デジタルマーケティング室	
役職	室長	役職を最新化
E-mail	yamada.taro@sansan.com	

## さまざまなビジネス情報を加えてリッチ化

登記社名	Sansan株式会社	設立	2007年6月
法人番号	4010001120965	決算期	5月
TDBコード	989671019	代表者役職	代表取締役
主業	情報処理サービス業	代表者氏名	寺田 親弘
従業	パッケージソフト業	株式公開	公開
資本金(円)	10億~100億		
従業員	500人~1,000人		
売上高	100億~300億		

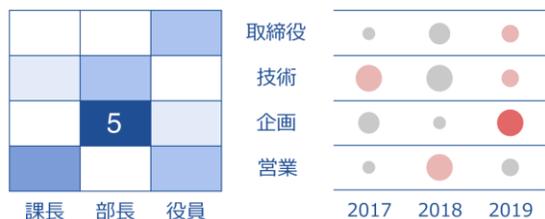
業績情報/関係会社/ニュース配信/  
拠点情報 Etc.

# データの活用：未来のビジネスシーンを変える「Sansan Labs」

最先端のデータ活用によってビジネス課題を解決する実験的な機能を「Sansan」で提供

名刺をスキャンするだけで、営業活動の効率化や社内人材のナレッジ・人脈活用をサポートするDSOCの実験的な分析機能を体験できる

## Sales Tech -営業活動を支援-

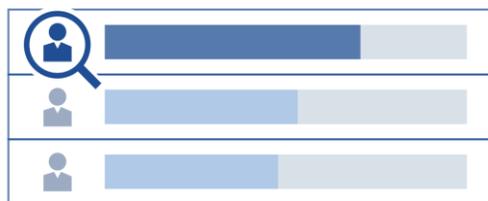


## HR Tech -タレントマネジメント-



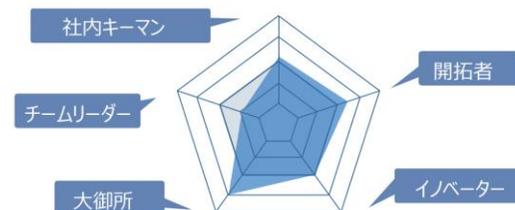
## 顧客との接点をヒートマップで可視化

ABMダッシュボード3 (β)



## ナレッジから同僚を探す

同僚ナレッジサーチ (β)



## 営業先のキーパーソンを探す

ホットリード・レコメンド (β)

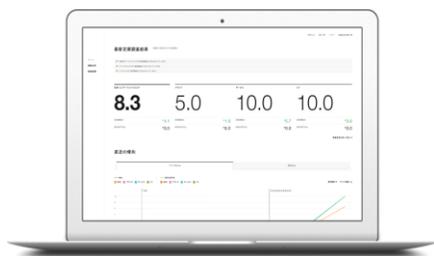
## 所有名刺でタイプを分析

ビジネスパーソンタイプ分析 (β)

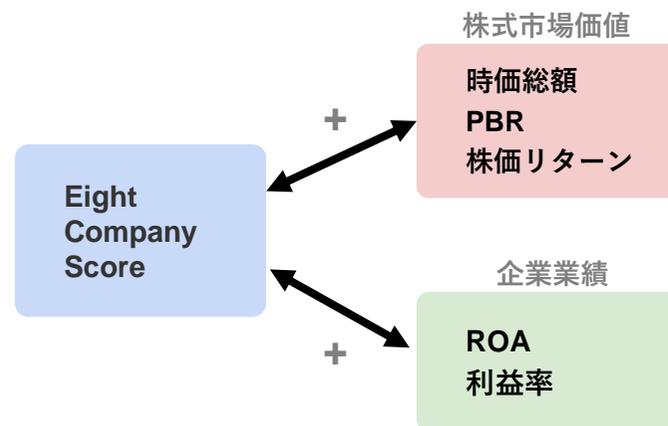
## データの活用：EBPM支援室の取り組み①「Eight Company Score」

名刺アプリ「Eight」のビジネスネットワークを活用し、企業評価を定量化する独自指標 ESGのSocialとの関わりが深いスコアであり、投信インデックス開発に向けた研究を開始

# Eight Company Score



BtoB企業の  
「関係性」の強さを数値化、  
定量的に把握できる



株式価値と企業業績  
との関連性を示す研究結果

## データの活用：EBPM支援室の取り組み②「ビジネス関係人口」

地域とビジネスとの関わりの度合いを示す独自指標

地方創生や災害復興の政策での活用を想定し、行政の取り組みを支援



出典：「『関係人口』ポータルサイト」(総務省) (<http://www.soumu.go.jp/kankeijinkou/>) を加工して作成

# 目次

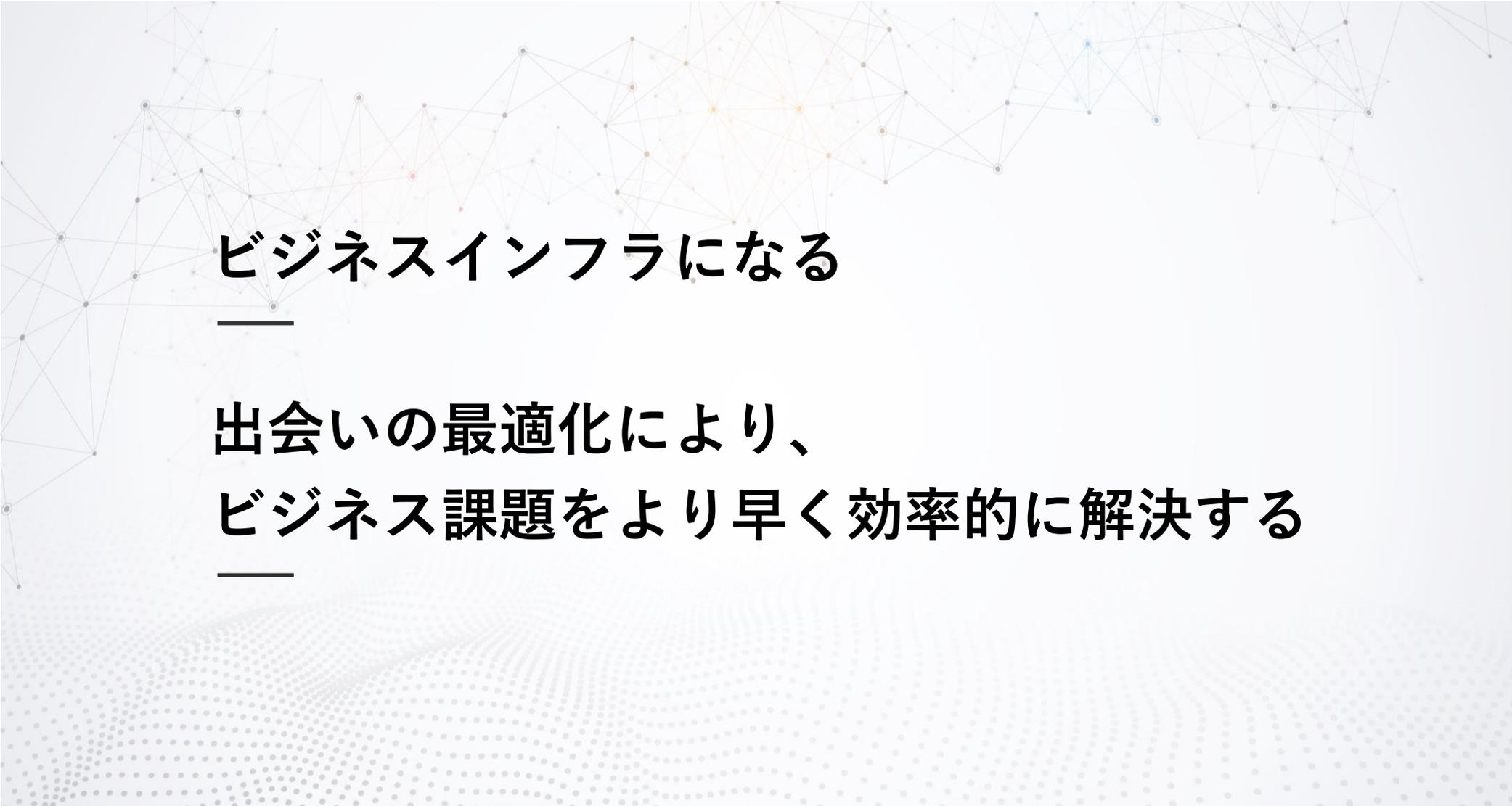
---

1 DSOCの概要

2 DSOCの役割

**3 DSOCの将来像**

– DSOCの目指す未来



**ビジネスインフラになる**

**出会いの最適化により、  
ビジネス課題をより早く効率的に解決する**

***sansan***

