sansan DSOC

Data Strategy & Operation Center

Presentation Material for DSOC, the Data Integration Division

Sansan, Inc.

April 14, 2021

Introduction

Satoru Joraku

Sansan, Inc.
Director, CISO (1), DPO (2), Managing Executive (DSOC (3))

Satoru cofounded Sansan Inc. in 2007 and he initially led product development of the Sansan, a B2B cloud-based contact management service. He now serves as the Director of Sansan's research hub DSOC, which conducts analysis and application of collected data, and refines the company's digitization technology. He also drives and implements internal security measures as CISO.

(1) Chief Information Security Officer
(2) Data Protection Officer
(3) Data Strategy & Operation Center

© Sansan, Inc

Table of Contents

1 DSOC Overview

2 DSOC’s Roles

3 DSOC in the Future

© Sansan, Inc

Table of Contents

1 DSOC Overview

- Mission and Roles
- History
- Organizational Structure
- Outstanding Engineers and Researchers

2 DSOC’s Role

3 DSOC in the Future

© Sansan, Inc

Mission and Roles

Create a competitive advantage in the market as a data management division

Mission

Activating Business Data

Through publicly available business data, such as business cards, documents, financial results, and stock information, we build “Encounter Database” that creates new possibilities leading to the future of business and society.

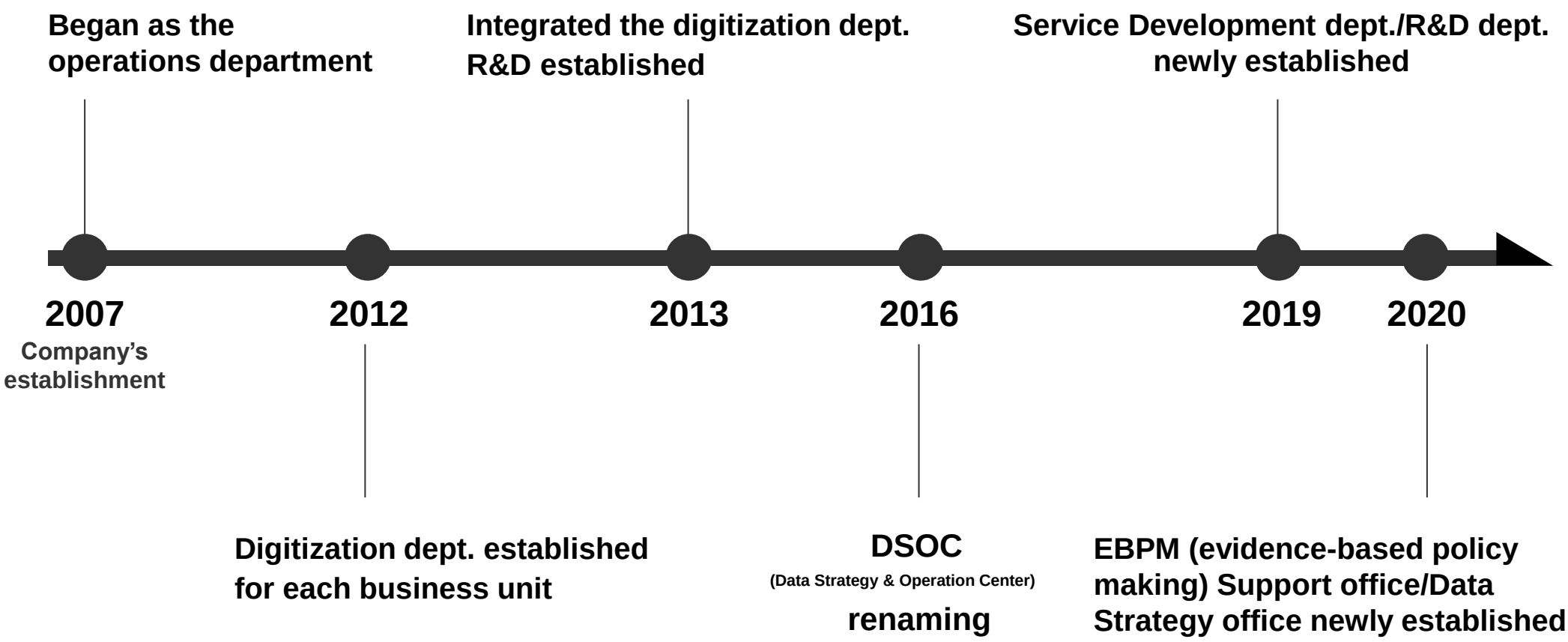
3 Roles of DSOC

Generating, Organizing, and Utilizing data

© Sansan, Inc

History

From a business card digitization department to a data management division handling all types of business data



© Sansan, Inc

Organizational Structure, Personnel

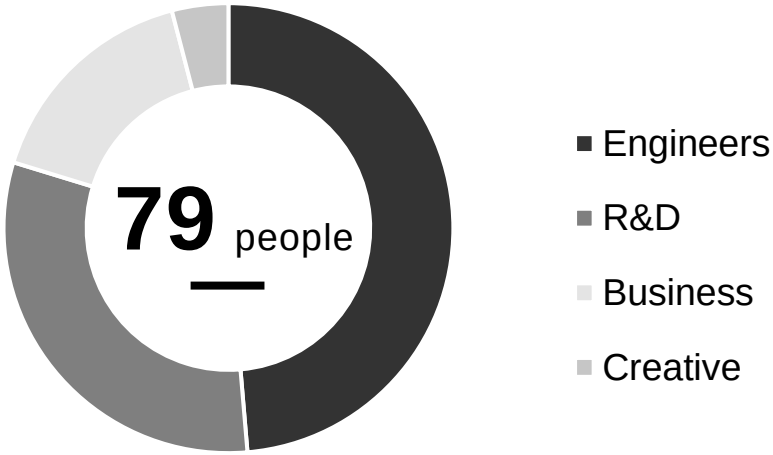
Structure centered on service operation/development and R&D related to data utilization
Most personnel are engineers and researchers

DSOC organizational structure (1)

sansan DSOC

Service Development Department	**R&D Department**
Data Strategy Office	**Creative Group**
EBPM Support	

DSOC staff (2)



(1) As of 3/1/2021
(2) As of 2/28/2021

© Sansan, Inc

Outstanding Engineers and Researchers

In addition to engineers involved in digitization technology, DSOC has specialists with diverse backgrounds, such as social science and economics researchers

- **Image Processing**
- **Natural Language Processing**
- **Deep Learning**
- **Machine Learning**
- **Complex Networks**
- **Economics**
- **Business Administration**
- **Sociology**

kaggle

Grandmaster

PhDs

Science (2), Economics (1), Engineering (2)

© Sansan, Inc

Table of Contents

1 DSOC Overview

2 DSOC's Role

- **Data Generation**
- **Data Organization**
- **Data Utilization**

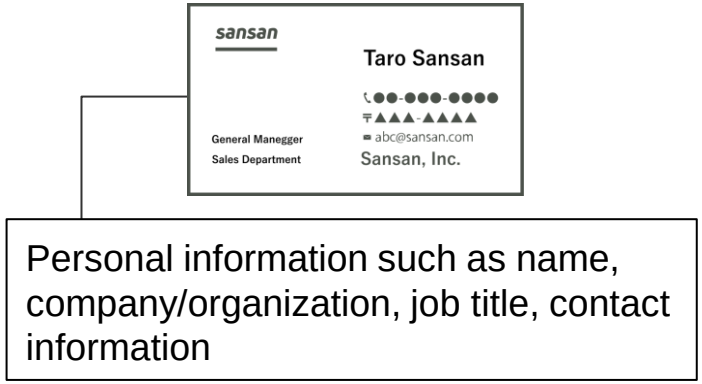
3 DSOC's in the Future

© Sansan, Inc

Data Generation: Unique Issues in Business Card Digitization

Business card information must be accurately digitized for effective data utilization
Practically difficult to correctly convert business card info into data using only OCR because of unique issues in business card digitization

Accuracy needed for business card digitization

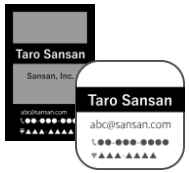


Personal information such as name, company/organization, job title, contact information

If even one character in the data is incorrect, the value as information is damaged.

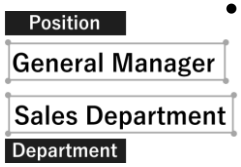
Digitization of business cards requires as much as 100% accuracy.

Issues in business card digitization



- **Atypical**
  Varying designs, shapes, fonts

- **Image Quality**
  Depends on shooting environment



- **Item Identification**
  Judgment of names, titles, etc.
  Unified notation rules

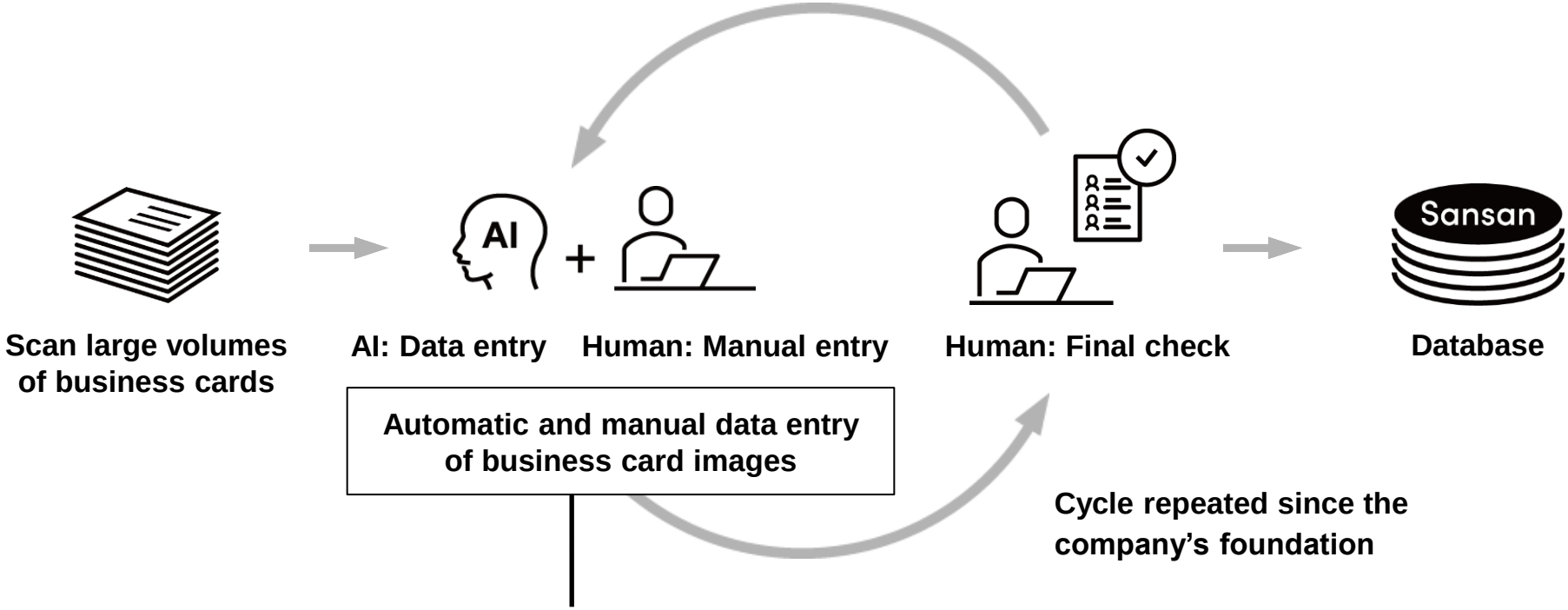
Accurate digitization is difficult if only using OCR.

© Sansan, Inc

Data Generation: Business Card Digitization Operations

Mechanism & tech to achieve 99%+ accuracy in digitization – source of competitive advantage

Prompt and accurate digitization of a large volume of business card information



Combination of Technologies

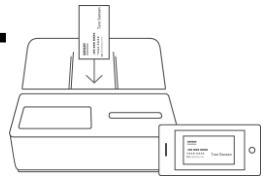
- Automatic detection and microtasking of business card information
- Automatic detection of business card language
- Automatic sorting to operators
- Diverse network of data entry operators
- Image processing technology

© Sansan, Inc

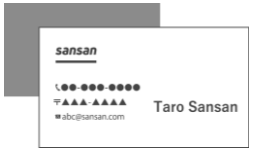
Data Generation: Ways of Supplementing OCR -1- GEES Digitization System

A unique operation system that accurately and efficiently digitizes high volumes of business cards

1



Business card input



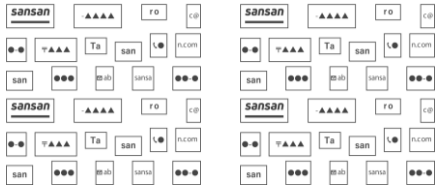
Separates background

Image correction

2

Item separation

3



Security item subdivision, item input

Builds a secure environment

4

Merging

5

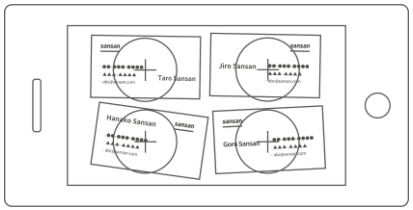
Company	Sansan.inc
Name	Taro Sansan
TEL	●●-●●●-●●●●
Mobile	▲▲▲-▲▲▲▲
Email	abc@sansan.com

Check and correct

© Sansan, Inc

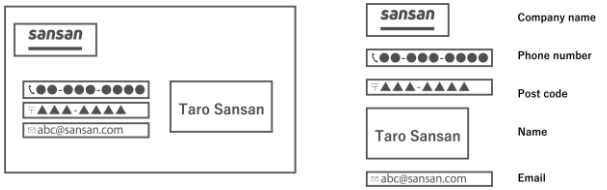
Data Generation: Ways of Supplementing OCR -2- AI/Image Recognition Tech

High-speed and high-precision digitization of business cards using unique image recognition technology with AI



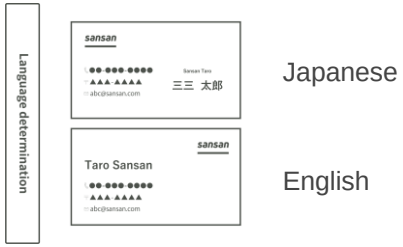
Smart Capture

Technology that enables users to receive results



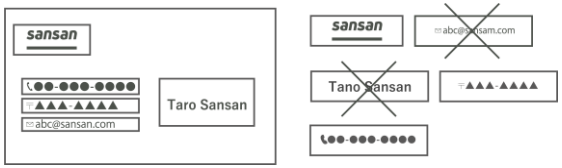
Item Segmentation

Distinguishes items from the business card design without reading letters



Language determination

Judges language without reading letters



Mistake Detector

Learns errors trends and predicts potential mistakes

© Sansan, Inc

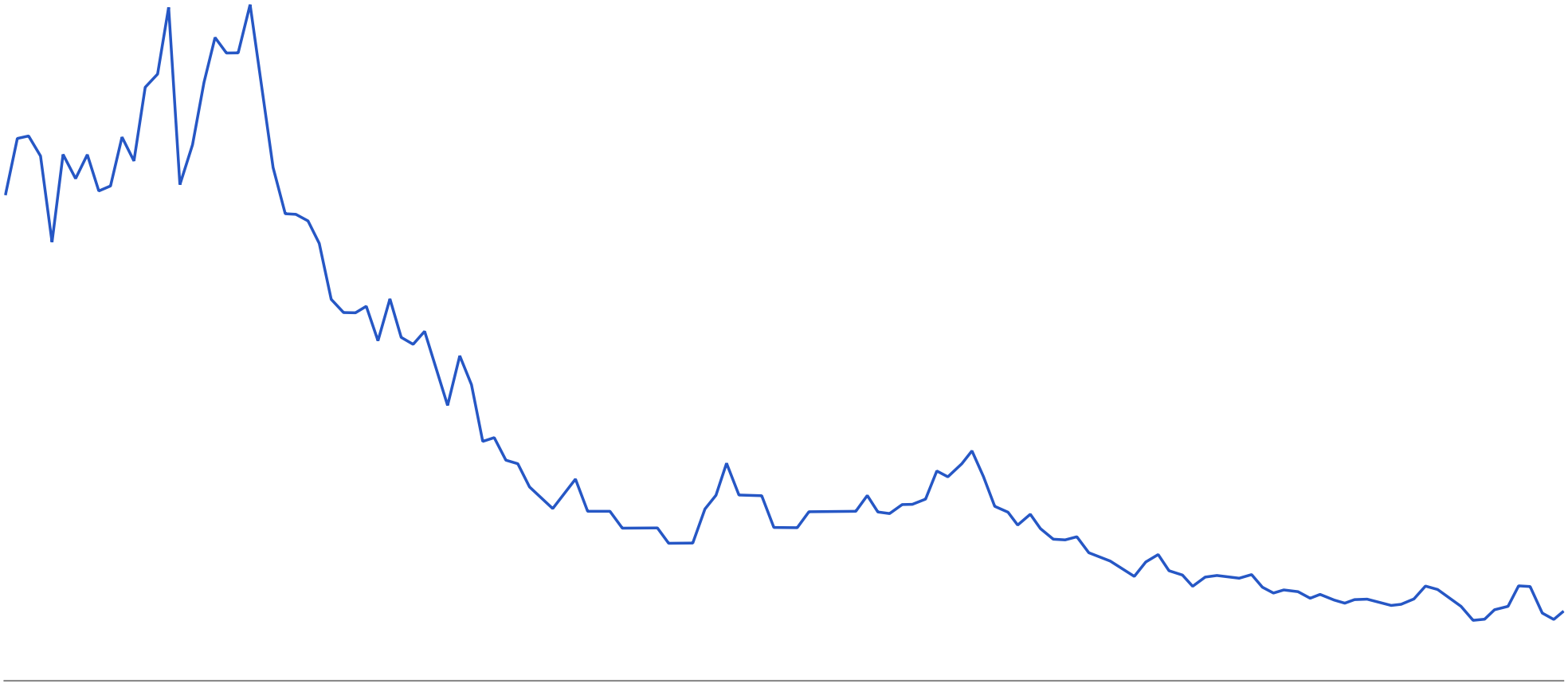
Data Generation: Ways of Supplementing OCR -3- DSOC OCR

We independently developed an OCR engine specializing in business cards, and that can read email addresses with 99.7%+ accuracy. We will continue researching and developing digitization for all business card items.

© Sansan, Inc

Data Generation: Changes in Digitization Cost per Business Card

With the evolution of business card digitization, the cost per card will be less than 1/20 what it originally was.

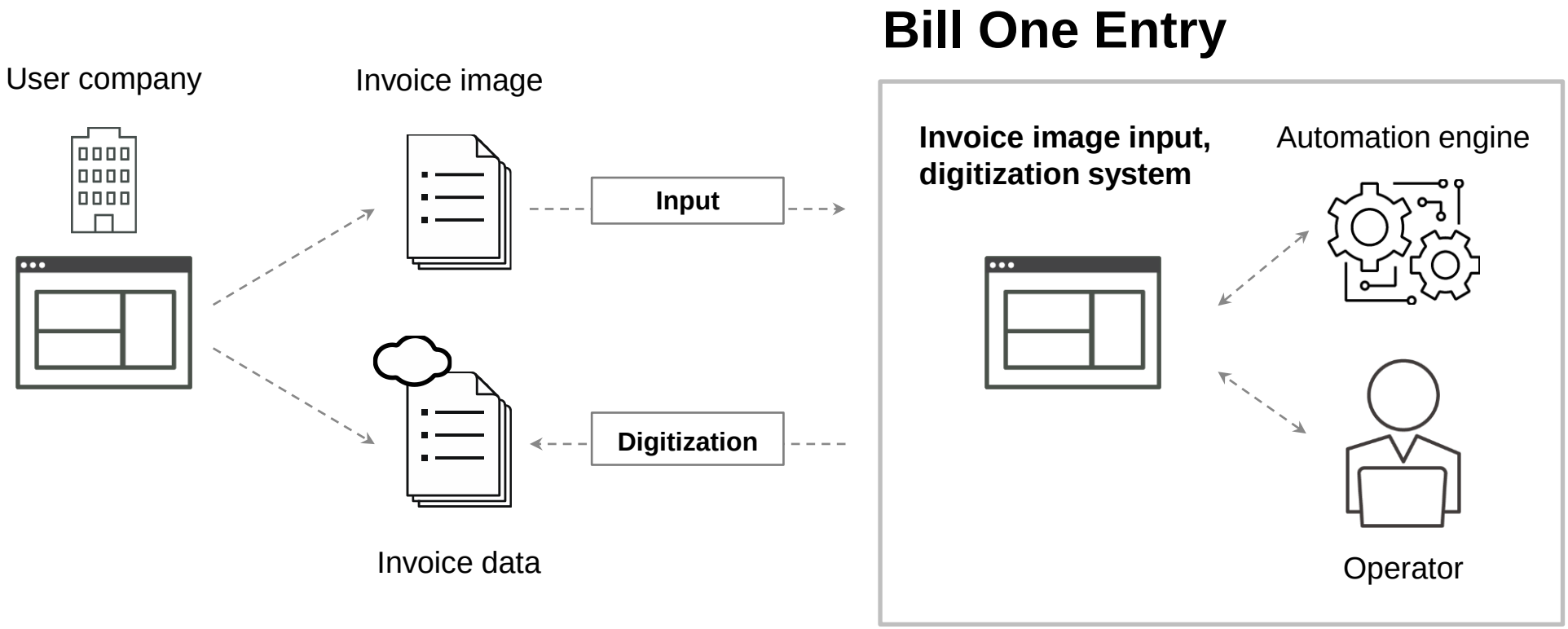


© Sansan, Inc

Data Generation: Using Digitization in Operations for Other Fields

Expanding accurate business card digitization operations to other fields, such as invoice digitization

BillOne
powered by Sansan



© Sansan, Inc

Data Organization: Challenges in Effective Data Use

Some accumulated data cannot be fully utilized for sales and marketing because of data problems such as outdated information and empty data items.

Company name	SAN SAN, Inc.
HP	
Tel.	+81-3-6758-0033
Address	5-52-2 Jingumae, Shibuya-ku, Tokyo, 150-0001, Japan
Contact name	Taro Yamada
Department	Digital Marketing
Job title	Director
E-mail	Yamada.taro@33i.co.jp

Old business name

Empty data field

Old job title

© Sansan, Inc

Data Organization: Building Data Infrastructure

Use accurate and up-to-date business card data, enrich with other items, establish data infrastructure for effective business use

Normalizing & updating

Company name	Sansan, Inc.
URL	https://www.corp-sansan.com/
Tel.	+81-3-6758-0033
Address	5-52-2 Jingumae, Shibuya-ku, Tokyo, 150-0001, Japan
Contact name	Taro Yamada
Department	Digital Marketing
Job title	Director
Email	Yamada.taro@33i.co.jp

Correct name

Fill in missing info

Updated job title

+

Enrichment with added info

Reg. name	Sansan, Inc.	Est.	June 2007
Corp #	4010001120965	Fiscal end	May
TDB Code	989671019	Position	Representative Director & CEO
Main business	Information processing	Name	Chika Terada
Industry	Package software	IPO	Yes
Capital (yen)	1–10 billion		
Employees	500–1,000		
Revenue (yen)	10–30 billion		

+

Performance information, affiliated companies, company news, etc.

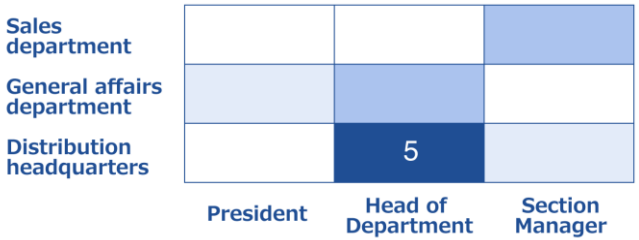
© Sansan, Inc

Data Utilization: "Sansan Labs" changing the future of business

"Sansan" provides experimental functions to solve business problems via use of cutting-edge data

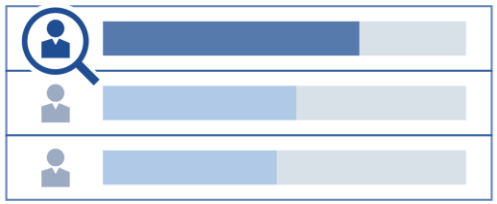
Simply scanning a business card lets the user experience DSOC's experimental analysis functions that support efficient sales activities and in-house staff's knowledge and network utilization.

Sales Tech: supporting sales activities



An Overview of Points of Contact for Each Customer

Account-based marketing (ABM) dashboard (β)



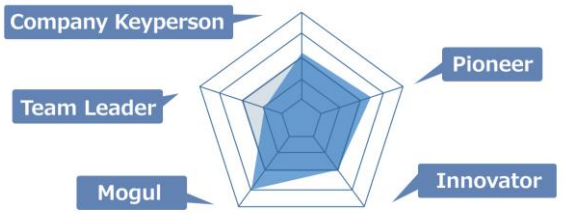
Find a Key Person in a Business Area

Hot lead recommendations (β)

HR Tech: talent management

Search Your Colleagues Based on Their Expertise

Search internal knowledge (β)



Understanding Strengths through Trends in Connections

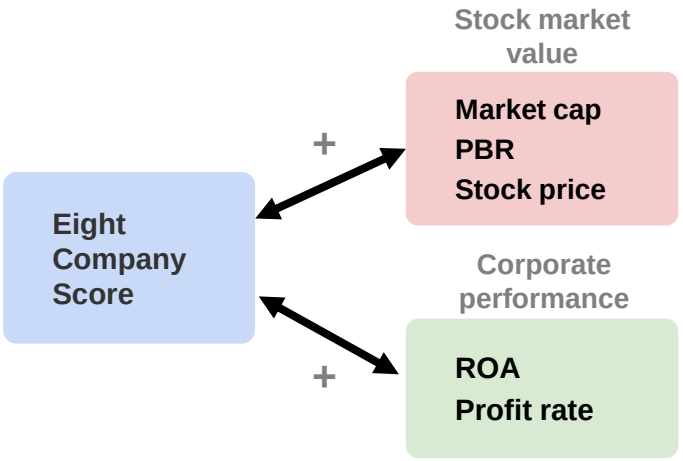
Businessperson type analysis (β)

© Sansan, Inc

Data Utilization: Work by the EBPM Support -1- "Eight Company Score"

This unique index quantifies corporate assessments using the business network of the business card management app "Eight." The score is closely related to the social aspect of Environment, Social, Governance (ESG), and we have started research on developing an investment trust index.

Eight Company Score



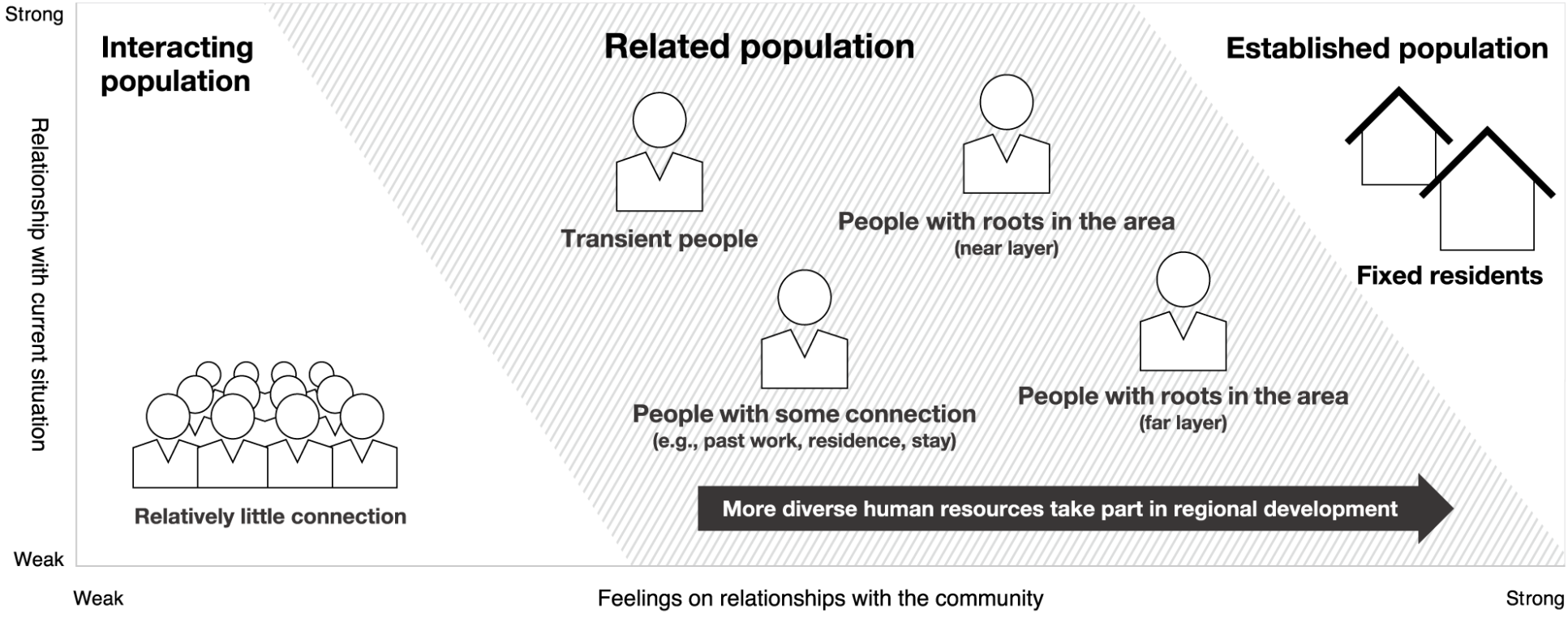
Can quantify and quantitatively grasp the strength of relationships between B2B companies

Research results showing relationship between stock market value and corporate performance

© Sansan, Inc

Data Utilization: Work by the EBPM Support -2- "Business-related population"

This unique index shows the degree of the relationship between the region and business. It supports administrative efforts, looking at use in regional revitalization and disaster-recovery policies.



Source: Created by processing the "Related population" portal site (Ministry of Internal Affairs and Communications)

© Sansan, Inc

Table of Contents

1 DSOC Overview

2 DSOC's Role

3 DSOC in the Future

– **DSOC's Aims as We Move Forward**

© Sansan, Inc

DSOC’s Aims as We Move Forward

Become business infrastructure

Optimize business encounters to solve business problems faster and more efficiently

© Sansan, Inc

sansan