

# NTTデータのAI研究事例 —— チャットボットからAI記者まで

2018年3月5日

株式会社NTTデータ技術開発本部 小間 洋和

# 目次

## 1. NTTデータのR&D重点AI適用領域

## 2. 顧客接点領域

事例 —— チャットボットによる接客応対

## 3. 業務高度化領域

事例 —— AI記者によるニュース原稿自動生成

# 当社のR & D重点AI適用領域

顧客対応等の  
フロントオフィス業務

審査等のミドル・  
バックオフィス業務

大規模データ分析による  
予測や予兆検知

顧客接点領域

業務高度化領域

複合  
高度分析領域

業務領域

社会基盤領域

# 当社のR&D重点AI適用領域

顧客対応等の  
フロントオフィス業務

審査等のミドル・  
バックオフィス業務

大規模データ分析による  
予測や予兆検知

顧客接点領域

業務高度化領域

複合  
高度分析領域

業務領域

社会基盤領域

Webサイトを訪問  
してくれたお客様を  
つなぎとめたい



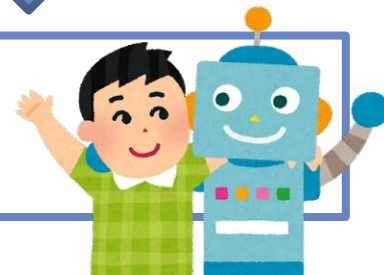
定型化できる  
簡単な問合せ  
対応を減らしたい



丁寧なクロージングで  
売上増や顧客満足度を  
向上させたい



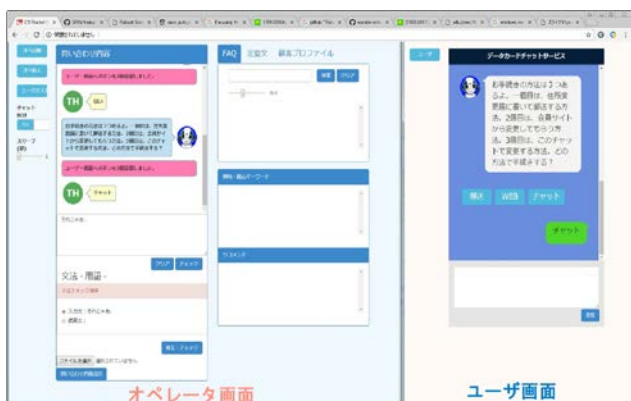
### AIチャットボット（ヒトとAIが力を合わせる）



ショッピングサイトやメーカー、金融機関等、ホームページでAIを活用したチャット対応を始めるお客様が増えています

# ヒトとボットが連携する仕組みで解決

ヒトとボットが連携することで単純な対応はボットに任せ、ヒトは売上増や顧客満足度向上に注力できます

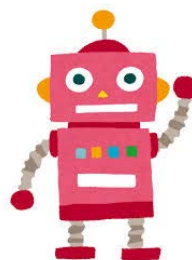


幅広いリーチには  
ボットの接客

有望顧客には  
ヒトの接客

- Webチャット
- チャットアプリ

ボット



チャット接客



Powered by corevo

※「corevo」は日本電信電話株式会社の商標です

# ボットによるチャット対応のポイント

- お客様の発言は多種多様であるため、単純なキーワードマッチ方式ではお問合せの意図を判断できません
  - ➡ • **単語ベクトルを基にした自然言語理解エンジン**
- お問い合わせ以外(雑談など)にも対応する必要があります
  - ➡ • **会話からお客様の嗜好を見つけ出す興味抽出技術**
  - **30万ペアの対話DBを基に構築した大規模ルール**

# 多様な言い回しに対する意図判定技術

NTTグループのAI「corevo™」をもちいて、お問合せの「意図」を判定して最適な文章を回答します。

## お客様の多様な言い回し

キャンセルはいつまで可能？

キャンセル期限は注文後何時間ですか？

注文を取消したいのですが、いつまで可能でしょうか？

NTT研究所技術



corevo

自然言語理解

## お問合せの意図を判定

質問の意図

キャンセル期限

商品検索

...

支払方法



# 興味キーワード抽出技術

機械学習をもちいて、対話データからキーワードを興味スコア付きで抽出します。得られたキーワードを使って商品をレコメンドします。

## 顧客の興味対象の抽出



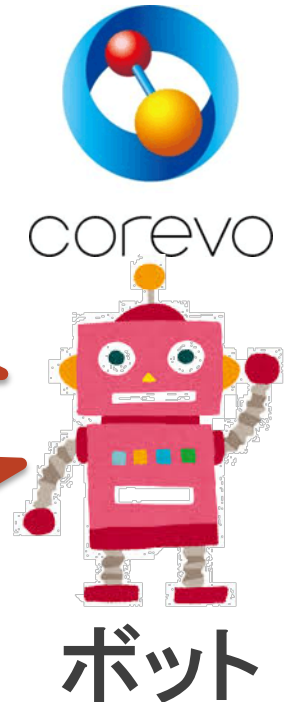
あぶり焼き  
サーモン頂戴

うん、  
大好き

それ頂戴

サーモン  
好きなの？

期間限定の  
サーモンうに軍艦  
食べた？



# 当社のR&D重点AI適用領域

顧客対応等の  
フロントオフィス業務

審査等のミドル・  
バックオフィス業務

大規模データ分析による  
予測や予兆検知

顧客接点領域

業務高度化領域

複合  
高度分析領域

業務領域

社会基盤領域

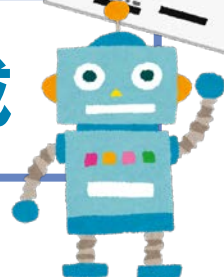
定型的な原稿を作成する作業から解放され、取材などのコア業務に集中したい

同じカテゴリの文書は書き方を統一したいがテンプレートやルールの作成が困難

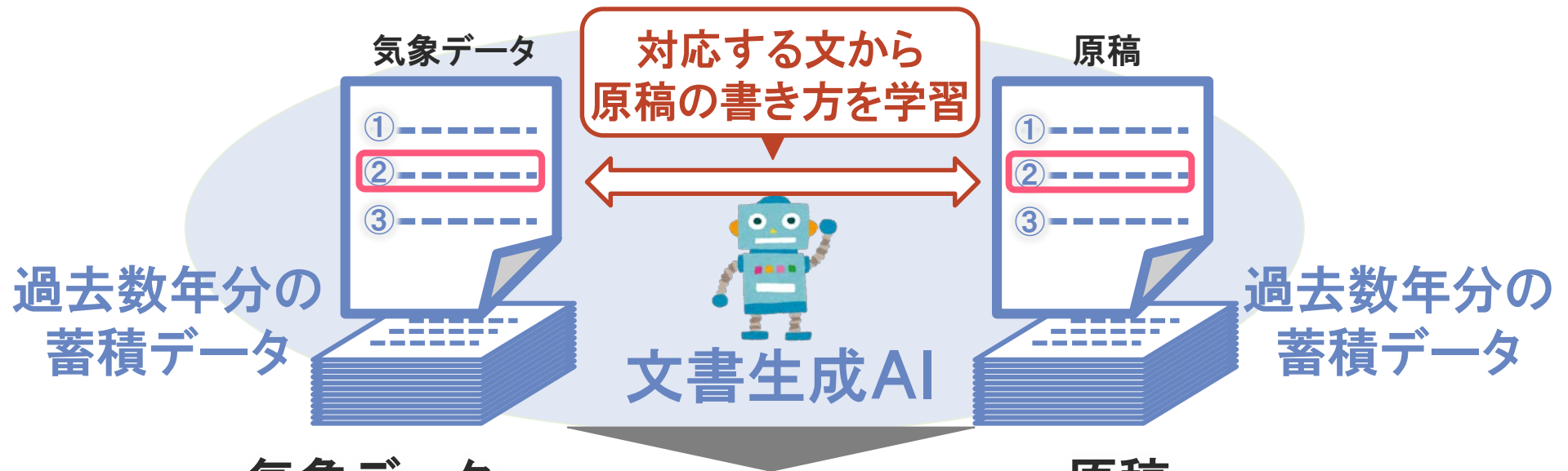


過去の大量データから文書生成ルールを自動作成

近年進化を遂げている深層学習によって文書生成や翻訳／要約の精度が高まっています



# 深層学習を使用した文書生成アプローチ



## 気象データ

- ① 前線に南から暖かい空気が流れ込み九州地方は大雨となる見込み。
- ② 15日6時までに予想される24時間雨量 九州南部 100ミリ
- ③ <防災事項>  
土砂災害

## 原稿

- ① 前線の影響で九州地方で大雨となる見込みです。
- ② 九州地方では、15日朝までの24時間に多いところで150ミリの雨が降るおそれがあります。
- ③ 気象庁は土砂災害に警戒するよう呼びかけています。

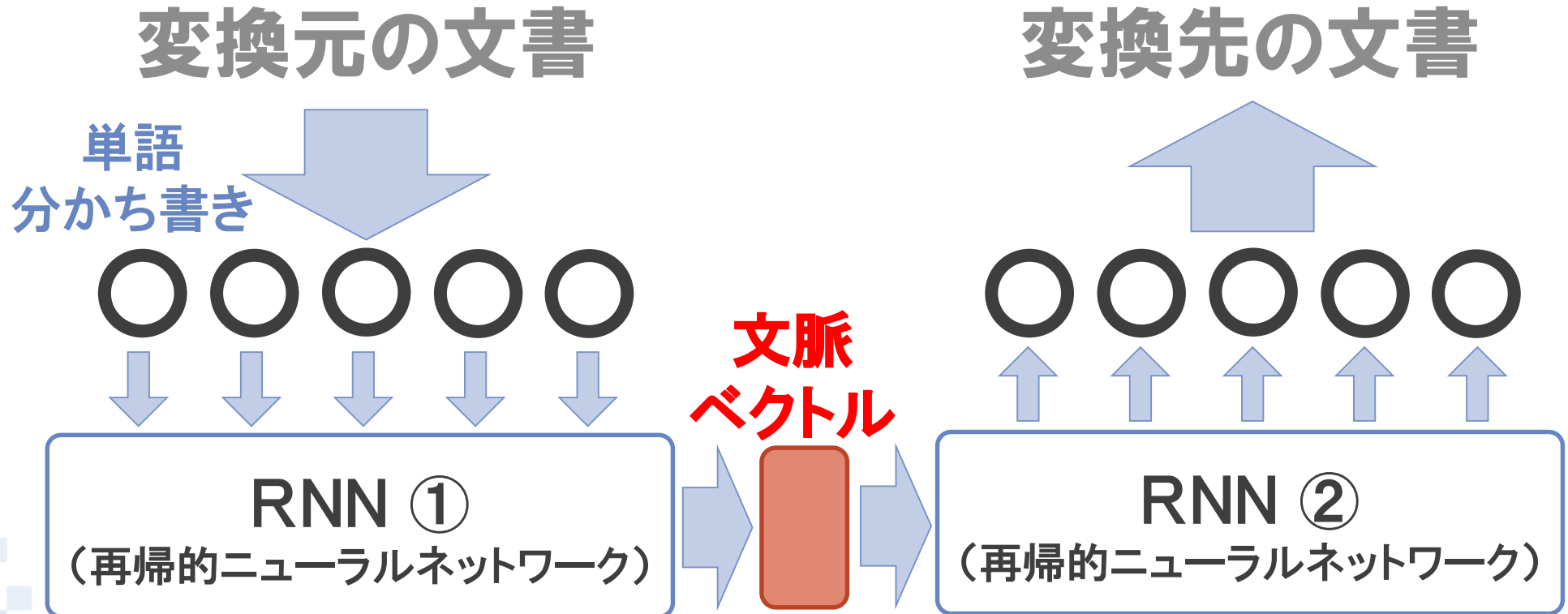
# 深層学習をもちいた文書生成技術

文書を変換するタスクにおいて、**Seq2seq**と呼ばれるニューラルネットワークが過去最高精度を達成

- 機械翻訳
- 要約
- 文書の誤り修正

# Seq2seqの概要

文書をニューラルネットワークで「**文脈ベクトル**」に変換後、更に別のニューラルネットワークで「**文脈ベクトル**」から文書を出力



# 文書生成の動向とNTTデータ手法の特長

## 文書生成における問題

重要な単語が欠損したり誤変換してしまうことが多い

→ 変換元の文書に登場するフレーズを部分的にコピーして文書を生成する方式が登場 (Copy Mechanism) [2015]

## NTTデータ手法の特長

研究所技術「リッチインデクサ」により「地名」等の固有名詞を抽出し、変換元文書から変換先文書へコピーすることが可能

前線に南から暖かい空気が流れ込み九州地方は大雨となる見込み。



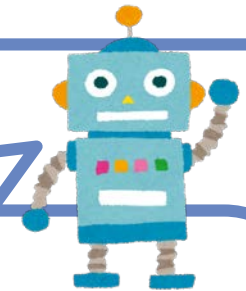
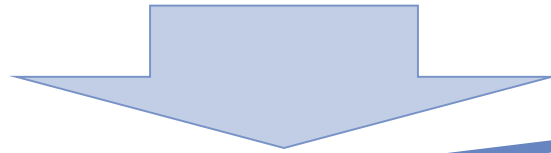
前線の影響で九州地方で大雨となる見込みです。

# 実際の文書変換例

前線に南から暖かい空気が流れ込み九州地方は大雨となる見込み。

20日18時までに予想される24時間雨量

<九州北部> 250ミリ



気象台の発表によりますと、前線に湿った空気が流れ込む影響で九州地方で20日にかけて大気の状態が不安定になり見込みです。

20日18時までに降る24時間雨量は九州北部で250ミリになると予想されています。



# 効果測定と課題

## 生成原稿の正確性を「文法」と「内容」の観点で4段階評価

### 正確性評価

生成した文に含まれる誤りの個数でAIの性能をスコアリングしました

#### 評価指標

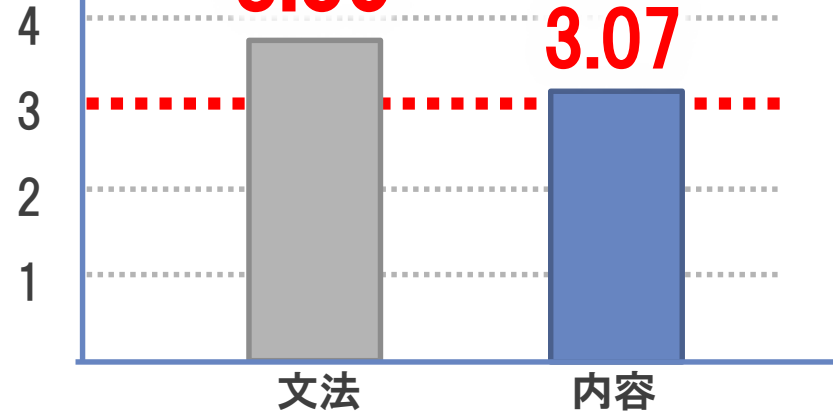
修正すべき箇所がない

1～2単語の修正が必要

3～5単語の修正が必要

6単語以上の修正が必要

正しさのスコア **3.86**



### 課題

- 正確性を高める補正ルールとの併用による高精度化
- 誤り箇所の可視化による人力修正の効率化

# まとめ

効果検証を通じて以下の課題が得られました。

- ◆ チャットボット
  - 人間オペレータへスムーズに繋ぐための対話破綻検知
  - 煩雑なルールベース管理の省力化
- ◆ AI記者による原稿自動生成
  - 正確性を高める補正ルールとの併用による高精度化
  - 誤り箇所の可視化による人力修正の効率化
- ◆ 共通課題
  - 学習データ(大量のラベル付きデータ)収集の効率化

以上の課題に対して引き続き検討を行い、実ビジネスでの適用を目指します。



# NTT DATA

Global IT Innovator