



対話型AIで実現する 次世代CX戦略

August 5, 2025





中村光晴

Twilio Japan 合同会社
プリンシパルソリューションアーキテクト





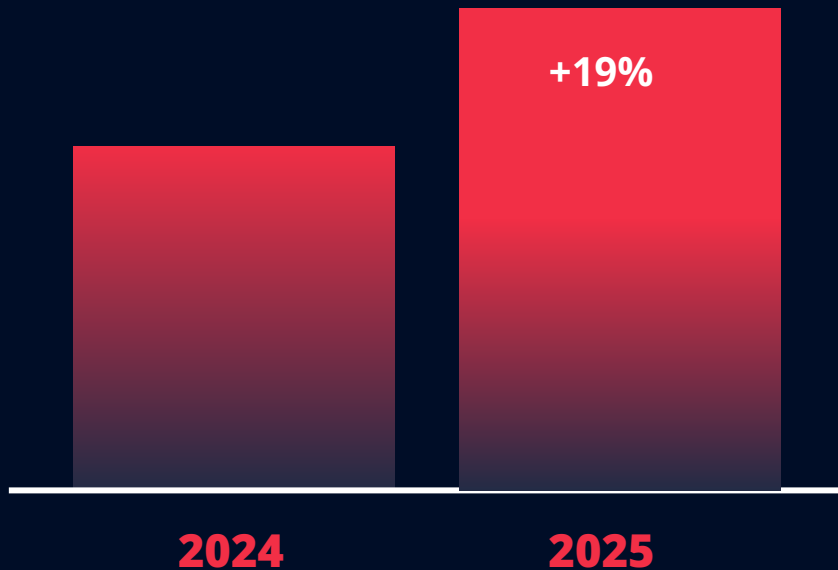
卓越したCXを提供するブランドは、**5.7**
倍の収益を生み出す



CXが悪いことによる損失額は前年比 19%増加

57.9 兆円

CXの低品質が原因で企業が失う恐れのある
年間収益総額



Source: [Qualtrics](#)



セルフサービスは期待外れ？

73%

の顧客が課題解決の過程でセルフサービスを利用

14%

の課題しかセルフサービス
で完全に解決されていない



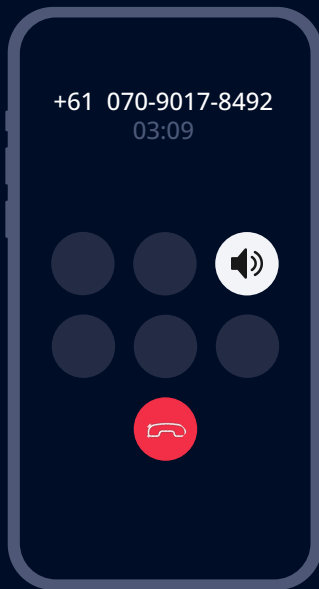
パーソナライズされない対応が、**CX**を低下させる



ご予約者様のお名前、予約確認番号、そしてご予約時に使用されたクレジットカードの下4桁をご提供ください。



メニュー内容が変更されております。ご注意の上、お聞きください。



担当者との通話をご希望かと存じますが、はじめに以下のメニューから該当する項目をお選びください



ご対応可能な担当者におつなぎいたします。



お客様のご用件は最優先で対応しております。現在の推定待ち時間は27分です。

システムが連携されておらず、データにアクセスできない状態

IVRの構成が固定されていて、変更や調整が困難

文脈や優先度に基づいたインテリジェントなルーティングが行われていない

オペレーターが顧客の状況や背景を把握できていない

「4つの恐怖」がCXチームの近代化を妨げている



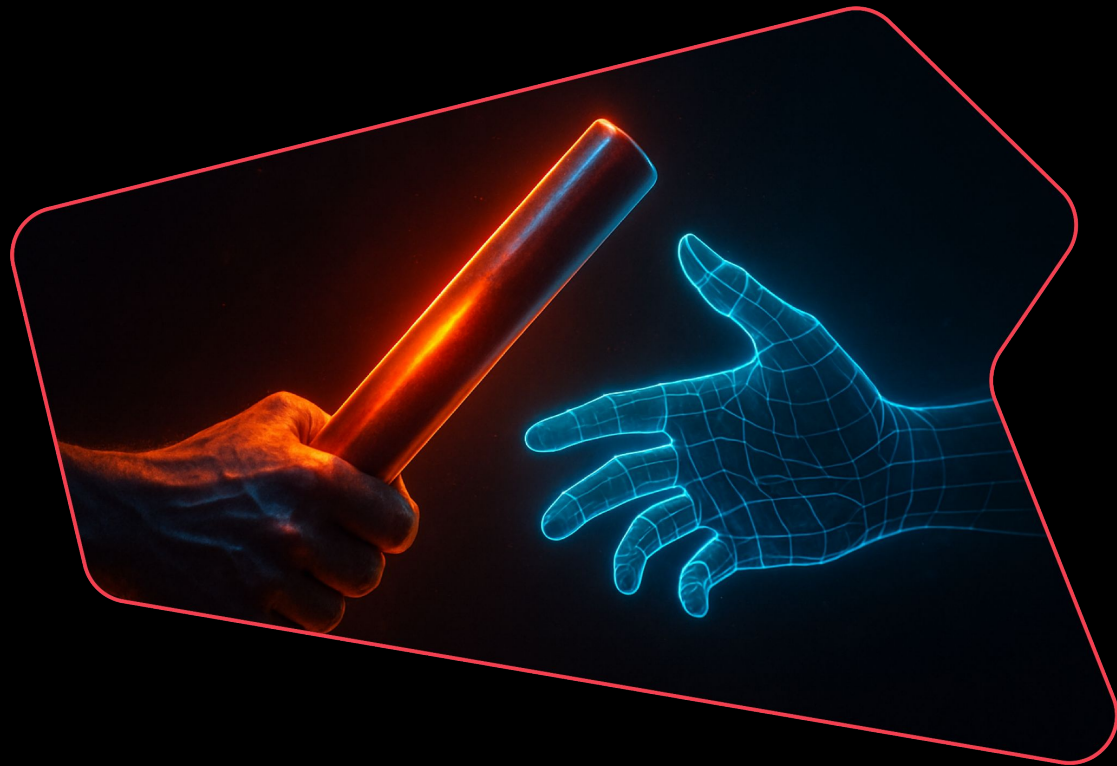
初期投資が莫大になるのではないか

既存のシステムをすべて入れ替える
必要があるのでは？

投資の効果が出るまでに時間が
かかりすぎるのでは？

AIの導入は複雑すぎて無理そう





AIで変わるCX、
そろそろ本気で考えるときです。



Twilioの描く世界

AI

適応性が高く人間のように対応する
セルフサービス



顧客体験を
リアルタイムに提供

顧客データ

DATA

一人一人の顧客を深く理解するための
文脈を備えたデータ
CDP



CPaaS

顧客とつながるための
コミュニケーションチャネル
音声、SMS、RCS、メール、ビデオ



AI-Powered 顧客エンゲージメントの柱

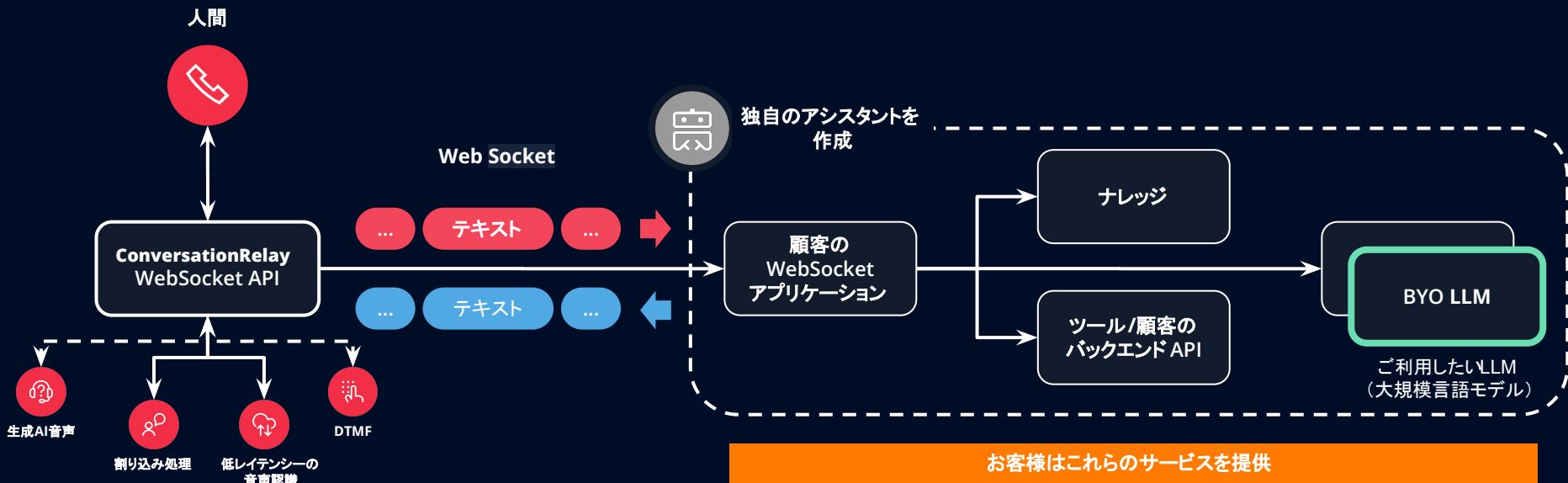


コンテキストに基づくデータとクロスチャネルオーケストレーション



ConversationRelay

好きなAIベンダーと連携





ConversationRelay

BYO-AI to Twilio Voice

- お好きなLLMを自由に接続可能 (Bring Your Own LLM)
- リアルタイムの音声認識 (STT) と音声合成 (TTS) の連携
- 超低遅延での音声対話
- 人間らしい自然な会話体験 (割り込み対応も可能)

人間らしく会話するために

Text to Speech Providers

- Google Voices
- Amazon Voice
- ElevenLabs Voices

Speech to Text Providers

- Google Speech API
- DeepGram



ConversationRelay

Twilio VoiceからBYO-AIへの接続

```
<?xml version="1.0" encoding="UTF-8"?>
<Response>
  <Connect>
    <ConversationRelay
      url="wss://mywebsocketserver.com/websocket"
      language="ja-JP"
      transcriptionProvider="google"
      speechModel="telephony"
      ttsProvider="google"
      voice="ja-JP-Chirp3-HD-Aoede"/>
    </Connect>
  </Response>
```

Twiml

```
{
  type: "prompt",
  callSid: "CA7f3bfc2ae470fxxxxxxxx",
  voiceId: "",
  prompt: "お名前をお伺いしてもよろしいでしょうか?",
  language: "ja-JP"
}
```

SPI

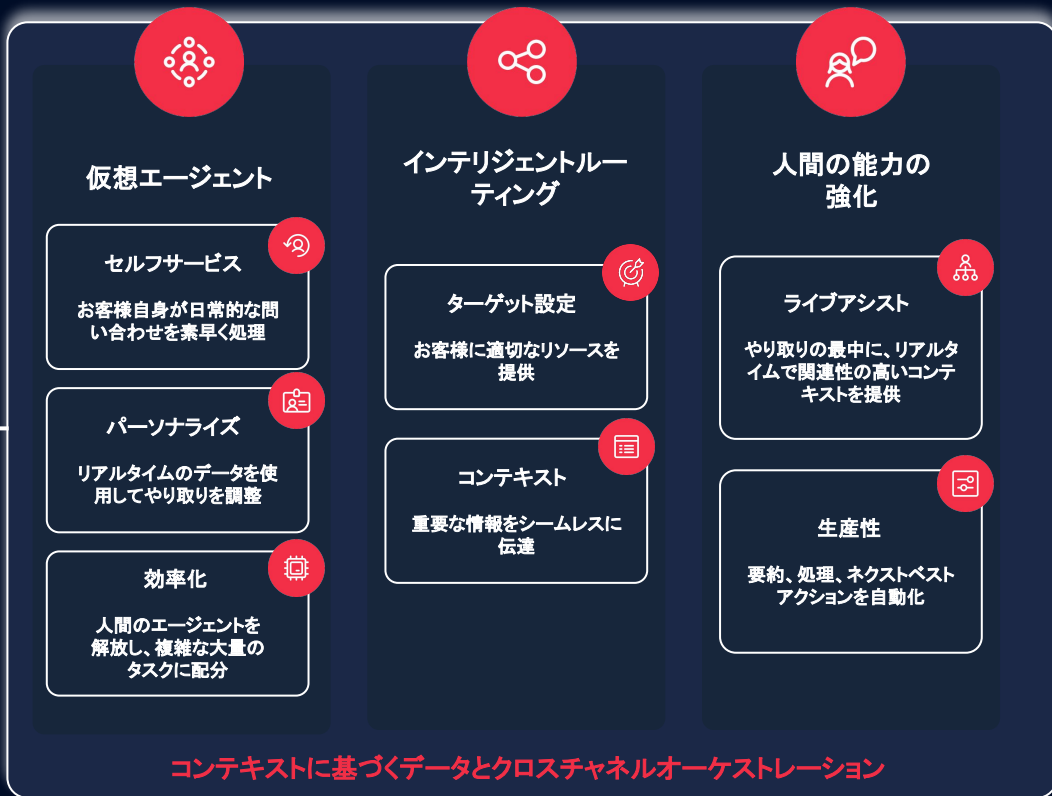


既存のシステムで動作

既存のコンタクトセンターにキャリアレベルでインテリジェンスを追加



Twilio SuperNetwork





AI 中心のオーケストレーションアーキテクチャ



このソリューションが向いているのは？



Max A. Millioncalls

Lead AI Engineer

- 開発力を持つ ISV やエンタープライズ企業に最適
- 差別化の核となる部分は自社で持ち、それ以外は効率よく外部に任せたい
- AI エージェントの構成要素をより多く自社でコントロールしたい
- 自社で選んだ LLM (大規模言語モデル) を使いたい
- すでにチャットボットを導入済みで、音声対応も加えたい
- レガシーな IVR (自動音声応答) をモダン化し、生成 AI を活用したい
- システムの細かな挙動まで自由にカスタマイズ・制御したい



アメリカの医療費支払い体験の改善に特化したスタートアップ企業

患者体験とカスタマーサポートの満足度を向上

「TwilioのAIは、文脈を理解して会話を効率化することで、人間のオペレーターが最適なタイミングで介入できるようにします。その結果、十分な情報をもとに、より意味のあるパーソナライズされた対応が可能になります。」



Duncan Winkie
Head of Commercial Strategy

AIが日常的なやり取りを代行しつつ、必要なときは人に繋ぐという“ハイブリッド対応”で、患者にもオペレーターにもやさしいCXを実現

Automated

2025年末までに、受電の30%をAIで自動対応予定

Improved

AIが定型的な対応を担うことで、CSATスコア(満足度)や待ち時間が改善

Personalized

請求通知や予約リマインダーを、個別最適化して提供

ConversationRelay

AIと人がリアルな会話を自然に交わせる体験を、簡単に・高品質に・迅速に実現

皆さんの課題

- 音声認識・音声合成の選定と連携が複雑
- 割り込み処理や話速調整など、リアルな会話体験の実装が困難
- STT/TTSベンダーやインフラ構築に多大なコストと時間
- 開発者が本来集中すべきAIロジックに手が回らない

ConversationRelay

ConversationRelayは、人とAIが、まるで自然な会話をしているようなリアルタイム音声体験を簡単に実現するためのソリューションです。これを実現するために必要な技術をwilio上でパッケージングし、低レイテンシーかつ柔軟性の高い電話音声にAIに届けます。

価格

定価: **\$0.07 / 分** (約9.98円※2025/07時点)

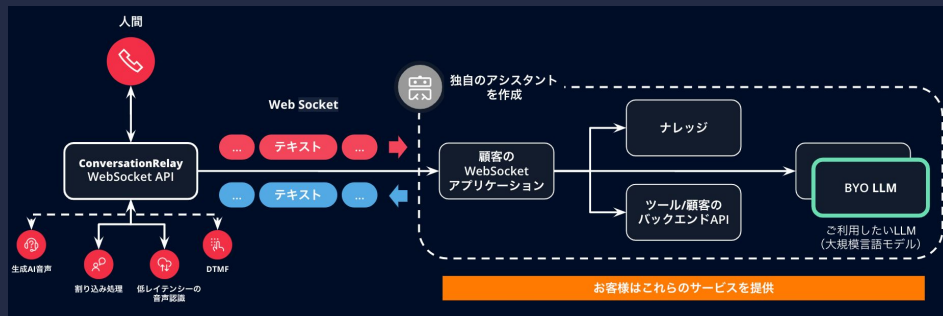
ConversationRelayと同じ機能を自社で契約し構築するよりも安価にTTS/STT/WebSocketの機能が利用できます。

詳細情報

- [ConversationRelayのHP](#)
- [ConversationRelayの技術仕様](#)

Twilioの解決策

- ConversationRelayなら、**音声制御・会話オーケストレーション・STT/TTS連携**までをTwilioがすべて提供。
- 開発者はWebSocketひとつで接続し、AIの開発・LLMの調整にだけ集中できます。



想定

費用対効果

オペレーターの費用

約26.67円 / 分

日本の関東圏におけるオペレーターの平均時給は 1600円です。最大時給は 2400円となります。
平均時給から1分あたりの分給を計算すると約 26.67円となります。

AI Agentの費用

約10.02円 / 分

Twilio ConversationRelayの費用
\$0.07 / min

例: Gemini 2.5 Flashの料金
\$0.00003 / 100 token 入力
\$0.00025 / 100 token 出力

1分の日本語の会話の平均 Token数は200と見積もった場合に、1分あたりのコストは、
入力100トークン、出力 100トークンで、\$0.00028です。

合計でTwilio+AIの合計は\$0.07028/minなので日本円にすると約 10.02円となります。
また為替レートは 2025年7月のTwilio内のレートで計算しています。
USD-JPY : 142.589円 / \$

約62.4%の人的費用削減



開発をもっと自由自在に

Twilioで、最も大胆で革新的なアイデアを実現

Thank you.

letsBuildSomethingAmazing();

