

ホワイトペーパー



Webex AI Codec

モダンコミュニケーションに対応する
AI を活用したコーデック

目次

03	はじめに
04	現代のデジタル音声通信の基礎
05	既存のテクノロジーの限界
06	Webex AI Codec
07	テストの方法と結果
09	Webex プラットフォーム全体での Webex AI Codec の影響
10	シスコの責任ある AI への取り組み
13	まとめ



概要

効果的な現代の通信システムには、明瞭な音声の提供が不可欠です。これはビデオ会議、1対1の通話、ユーザー補助とインクルージョンを強化する音声アシスタント、翻訳サービス、音声合成の展開など、さまざまなシナリオに当てはまります。

しかし、背景雑音、残響、音声キャプチャとサウンドの再生品質、ネットワーク障害などの要因によって音質が低下し、理解が困難になる場合があります。

Webex では、あらゆる人にとってインクルーシブな未来を実現するうえで、明瞭な音声がきわめて重要な役割を果たすことを理解しています。そうした理由から、お客様に信頼性と忠実度の高い優れた通信システムを提供することに取り組んでいます。背景雑音の除去、自分やあらゆる人の音声に合わせた最適化など、AI ベースの音声を明瞭化するテクノロジーに多額の投資を行った結果、Webex でのユーザー体験が大幅に向上しました。

そして今、Webex AI Codec という画期的なテクノロジーを導入しました。この AI ベースの音声コーデックは、非常に低いビットレートで優れた音質を実現し、ネットワーク障害に対する堅牢性を発揮します。その結果として、非常にレジリエンスの高い通信システムが完成し、ノイズフリー、高い忠実度、非常に明瞭な音声通信が実現されています。

音声通信の未来は新しい時代に突入しました。そして、Webex はその最前線に立ち、この変革の過程を先導しています。

現代のデジタル音声通信の基礎

現代のデジタル音声通信システムは、音声コーデックとパケット損失の軽減機能という 2 つの不可欠な構成要素に依存しています。

音声コーデック

音声コーデックは、リアルタイムの音声を伝送するのに必要なビット数の低減に役立つため、デジタル通信システムに欠かせません。これに対応するために、音声コーデックでは必要なビット数を低減します。「コーデック」という用語は、「コーダ」と「デコーダ」を組み合わせたものです。コーダ（またはエンコーダ）が音声波形を特定のビットレートに圧縮する一方で、デコーダは受信側で音声波形を再構築します。音声コーディングの目的は、マイクでキャプチャしたオーディオストリームを所定のビットレートバジェット内で圧縮し、受信側で元の音質に可能な限り近い音質で再構築できるようにすることです。

リアルタイムの通信システムでは、マイクでキャプチャした音声は一般的に、フレームへの分割、コーデックによる圧縮、パケット化を経てからネットワーク接続経由で伝送されます。こうした「音声パケット」が正常に配信されるかどうかは、ネットワーク接続の安定性と信頼性に左右されますが、常に保証されているわけではありません。

パケット損失の軽減

パケット損失とは、ネットワーク障害が原因でパケットが損失することを指します。たとえば、ネットワーク接続が部分的に中断された場合、受信側で音声は「途切れ」たり、明瞭度、ひいては生産性に影響がおよんだりする可能性があります。

音声通信システムでは、損失したパケットのオンデマンド送信、冗長なオーディオフレームの予防的伝送、パケット損失の隠蔽（PLC）技術など、そうした損失に対処するためのメカニズムを導入する必要があります。

それぞれの手法に固有のメリットとデメリットがあります。たとえば、損失したパケットの再送によって帯域幅効率は高くなりますが（メリット）、その代償として遅延が増加します（デメリット）。逆に、パケットの重複送信によって帯域幅効率は低下しますが（デメリット）、結果として遅延が少なくなります（メリット）。

一方で、PLC の目的は、新しいパケットの着信を待つ間に受信されなかった音声をマスキングすることです。その際に用いる手法の例として、最後に正常に受信したパケットから音声を再生する、損失したパケットを無音に置き換える、典型的な音声パターンに基づいて妥当な「空白部分を埋める」音声を再構築するといったことが挙げられます。

パケット損失による影響の軽減にどの手法を使用する場合でも、帯域幅の使用、遅延、計算コスト、音質の間でトレードオフが発生します。さらに、大量のパケットが損失した場合、上記のすべての手法が機能不全に陥り、会話の大部分が完全に失われるおそれがあります。

より低いビット数で優れた音質を実現することができれば（そのようにして帯域幅効率が向上すれば）、（所定のビットレートバジェット内でより多くのパケットを送信することで）パケットの重複送信や再送のデメリットを解消できるでしょう。また、パケット損失が重大な場合は、生成 AI ベースの PLC でコンテキストに基づいて失われた音声を（再）作成してもよいでしょう。結局のところ、音声コーディングとパケット損失による影響の軽減に AI を活用したテクノロジーを導入すれば、既存の手法の限界の大部分を解決できる可能性があります。これについては、次のセクションで詳しく説明します。

既存のテクノロジーの限界

音声コーデックは、数十年にわたって通信システムの開発できわめて重要な役割を果たしてきました。

音声コーディングは、ほぼすべての情報を保持しながらオーディオストリームを簡素化するというコンセプトがベースとなっていて、音声の内容を理解したり音楽を楽しんだりするのに十分な音質を維持しつつ、伝送に必要な帯域幅を低減します。「非可逆音声圧縮」と呼ばれるこのプロセスは、画像やビデオの圧縮技術に似ています。

音声コーディングの分野では、情報理論、心理音響学、神経科学、デジタル信号処理など、さまざまな分野が活用されています。デジタルテクノロジーが台頭し、（特にインターネットを介して）限られた帯域幅のチャンネルでデジタルメディアを交換する必要性が生じているため、こうした科学的進歩はますます重要になりました。その結果として、音声圧縮は不可欠になりました。これがきっかけとなり、音楽向けの MP3、AAC、Vorbis、FLAC、Opus などの音声コーディングテクノロジーおよびフォーマットが開発されました。

リアルタイムの通信システムでは、利用者は自宅、オフィス内、外出先など、場所を問わずにクリアなビデオと音声再生されることを期待しています。しかし、質の高いメディア体験の実現には、大量の帯域幅が必要であり、限られたリソースを効率的に使用する必要があります。こうした理由から、より低いビットレートで遅延を最小限に抑えながら優れた音質を実現し、帯域幅の使用を最適化できる音声コーデックとビデオコーデックが重要になります。

従来の音声コーデックは長年にわたって利用されてきましたが、優れた音質を維持しながら音声コンテンツを効果的に圧縮するには限界があります。ビットレートバジェットを下げると、音質に多大な影響がおよび、再構築された音声は不明瞭になったり、聞き取りにくくなったりします。いくつかの調査（Opus や EVS での調査など）では、この見解が支持されています [1]、[2]、[3]、[4]。

AI ベースのコーデックには、こうした限界を克服できる可能性があります。

Google [3]、Amazon [5]、Microsoft [6]、Meta [2] などの業界リーダーによる画期的なリサーチで示されたのは、AI ベースのコーディングによって、1.6 ~ 6 kbps 程度のきわめて低いビットレートでも優れた音質を実現できることです。この進歩は、リアルタイムの音声通信に変革をもたらします。

しかし、パワフルな低ビットレートのコーデックだけでは、堅牢な音声通信システムの導入には不十分です。高い音質を確保し、優れたユーザー体験を提供するには、システムでノイズ除去、残響の低減、マイクアーティファクトの補正、ネットワーク障害に対するレジリエンスなどの課題にも対処する必要があります。さらに、音声は文字起こし、翻訳、音声生体認証などのさまざまな用途で重要な役割を果たします。複数のテクノロジーをワークフローに組み込む利用者が増えるにつれて、音声に関連したタスクの処理に対して包括的なアプローチをとることが不可欠になります。

たとえば、利用者が騒がしい環境からカスタマー サービス エージェントに連絡するシナリオを想像してみましょう。この時点で、利用者は（背景雑音を最小限に抑えるために）ネットワークカバレッジが良好で静かな場所を見つける必要があるでしょう。この一般的なシナリオでは、両方の側で音質の低下が発生し、シームレスな体験が実現されません。

次のことを同時に実行できるテクノロジーがあったとしたら、素晴らしいと思いませんか？

1. 通信を行う両方の側で背景雑音を除去する。
2. どのネットワークでも完全な音声伝送を実現する。

こうしたテクノロジーがあれば、全体的な通信体験が大幅に向上し、トレードオフが解消され、利用者とエージェントの両方にシームレスなソリューションが提供されます。

これらの要件に包括的に対応するために、Webex AI Codec が開発されました。これは、音声の明瞭化とレジリエンスの高い通信を実現する単一のエンドツーエンドソリューションです。

Webex AI Codec では、最先端のテクノロジーを組み合わせて優れた音質を実現し、さまざまなシナリオで信頼性の高い通信を確保します。

Webex AI Codec : 現代の通信に対応する AI を 活用したコーデック

Webex AI Codec は、質の高い低ビットレート音声圧縮、ネットワーク障害に対するレジリエンス、音声の明瞭化の実現を通じて、リアルタイムの通信システムの課題に対処するように設計された新しい AI ベースの音声コーデックです。低ビットレートでの卓越した音声コーディングパフォーマンスを活かすことで、Webex AI Codec を使用して、ネットワーク障害に対するレジリエンスの高い通信システムを実装できます。基本的なコンセプトは、現在の音声フレームを低ビットレート（6 kbps など）でエンコードし、前のフレームのコピーをさらに低いビットレート（1 kbps など）で符号化することです。

図 1 に示すように、Webex AI Codec はエンコーダ、ベクトル量子化器（VQ）、音声デコーダで構成されていて、数千時間の音声と背景雑音を使用してエンドツーエンドでトレーニングを受けます。

Webex AI Codec

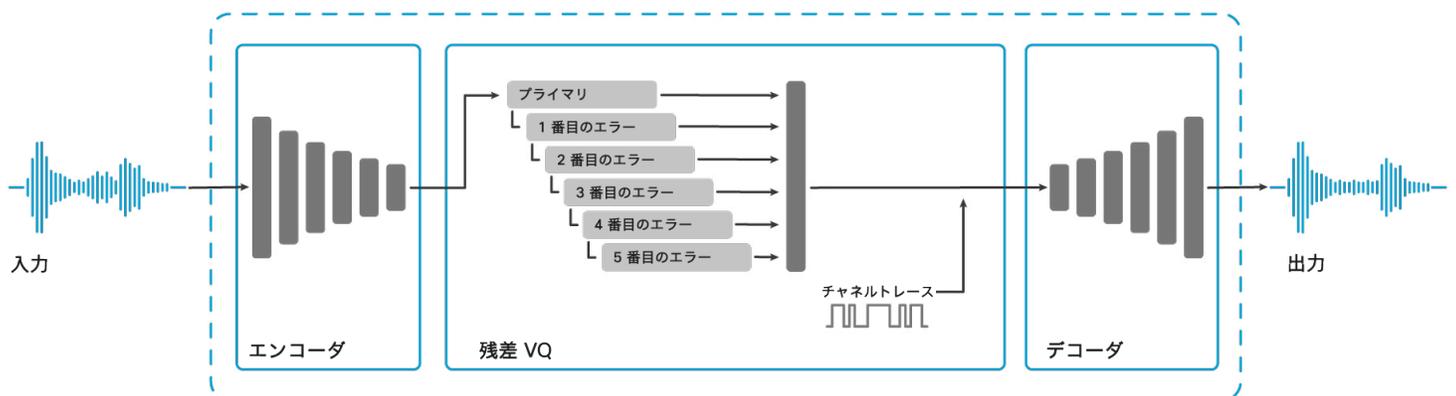


図 1 : Webex AI Codec のコアテクノロジー

エンコーダ

搭載されているエンコーダは、アルゴリズムの遅延がゼロのニューラルネットワークです。音声エンコーダで未加工の音声入力を受信し、音声のすべてのフレームにおける特徴ベクトルの埋め込みを抽出します。また、エンコーダで入力ストリームからノイズ、残響、その他のアーティファクトを除去することもできます。この「クリーニングされた」音声によって、さらに密度の高いエンコーディングが可能になります。

ベクトル量子化器

ベクトル量子化（VQ）はよく知られた量子化手法であり、音声/ビデオコーディング、クラスタリング、類似性に基づくベクトル検索など、さまざまな分野で幅広く使用されてきました。残差 VQ [6] は多段 VQ [7] とも呼ばれ、複数の VQ レイヤを使用します。それぞれの VQ レイヤでは、前のレイヤから残差信号を取得して、図 1 に示すように量子化を順次進めます。Webex AI Codec システムでは、伝送前に残差 VQ を使用して特徴ベクトルの埋め込みをさらに圧縮し、最小限のビット使用量でコンテンツを送信することを狙っています [8]。

デコーダ

デコーダのアーキテクチャは「ミラーリング」方式ですが、エンコーダ内と同様の設計に従っています。デコーダでは、損失した音声フレームの埋め合わせを含め、受信した音声ベクトルから音声を再構築します。

ニューラルコーデックのトレーニング

ニューラル コーデック システムをトレーニングするために、背景雑音、残響、帯域制限、パケット損失など、さまざまなアーティファクトをクリーンな音声信号に導入します。このコーデックのトレーニングは、10,000 時間超のクリーンな音声とノイズのサンプルで構成された、ノイズを含む数百万時間単位の独自の音声を使用して行われました。大規模なトレーニングデータによって、音声とアーティファクトの全体でモデルを幅広く一般化できるようになります。このトレーニングの成果として、音声エンコーダで音声入力を受信し、ディープ ニューラル ネットワークを活用して、複雑な音声と背景雑音の特性をまとめて、または個別にカプセル化する包括的な一連の特徴を抽出することが可能になります。抽出された音声特徴は音量、ピッチの変調、アクセントのニュアンスなどの属性を表します。

これが意味するのは、手動で設計された特徴に依存する従来の音声コーデックからの脱却です。ニューラルエンコーダはその広範で多様なデータセットから特徴抽出プロセスを学習して改善するため、より汎用性が高く、一般化された表現が得られます。

パフォーマンス検証

シスコでは、オンライン アンケート プラットフォームの Qualtrics と Amazon MTurk を使用して、クラウドソーシングによる MUSHRA (Multi Stimuli with Hidden Reference and Anchor) リスニング テスト フレームワークを開発しました [9]。このツールを使用すると、必要な数の参加者で MUSHRA テストをより頻繁に実施できます。ただし、このアプローチでは結果の一貫性が保証されない場合があります。こうした不一致に対処するために、リアルタイムとテスト後の両方のスクリーニングプロセスを導入しました。

一貫性のある再現可能な結果を得られるようにするには、この徹底的なスクリーニングプロセスが不可欠であり、テストを複数回にわたって反復する場合は特に欠かせません。

これを実現するために、リスナーとそのスコアの適格性を判断し、当該のスコアを検証するための 2 段階のスクリーニングプロセスを構築しました。

Qualtrics でリスニングテストを設計する際には、特定のルールが課せられます。そのルールによってリスナーが不適格となり、リスナーのスコアがスクリーニング後の段階から除外される場合があります。

テストの方法

最初の段階では、リスナーに対して聴力テストを実施します。このテストに合格すると、「トレーニングテスト」に進みます。このテストでは、実際のリスニングテストに類似したサンプルを確認できます。リスニングテストでは、リスナーのスコアが綿密にモニタリングされ、回答に一貫性がない場合は不適格となる場合があります。

テスト結果

図 2 は、1 kbps および 6 kbps で動作する Webex AI Codec と 16 kbps で動作する Opus を比較したリスニングテストの結果を示しています。最も注目すべきは、6 kbps で動作する Webex AI Codec が 16 kbps で動作する Opus を上回っていることです。また、1 kbps で動作する Webex AI Codec は優れた音質を実現します。

なお、結果は、すべてのスコアを共通の「アンカー」スコアと「隠れ基準」スコアに正規化することで統合されます。

わずかな帯域幅で最適な音質を実現

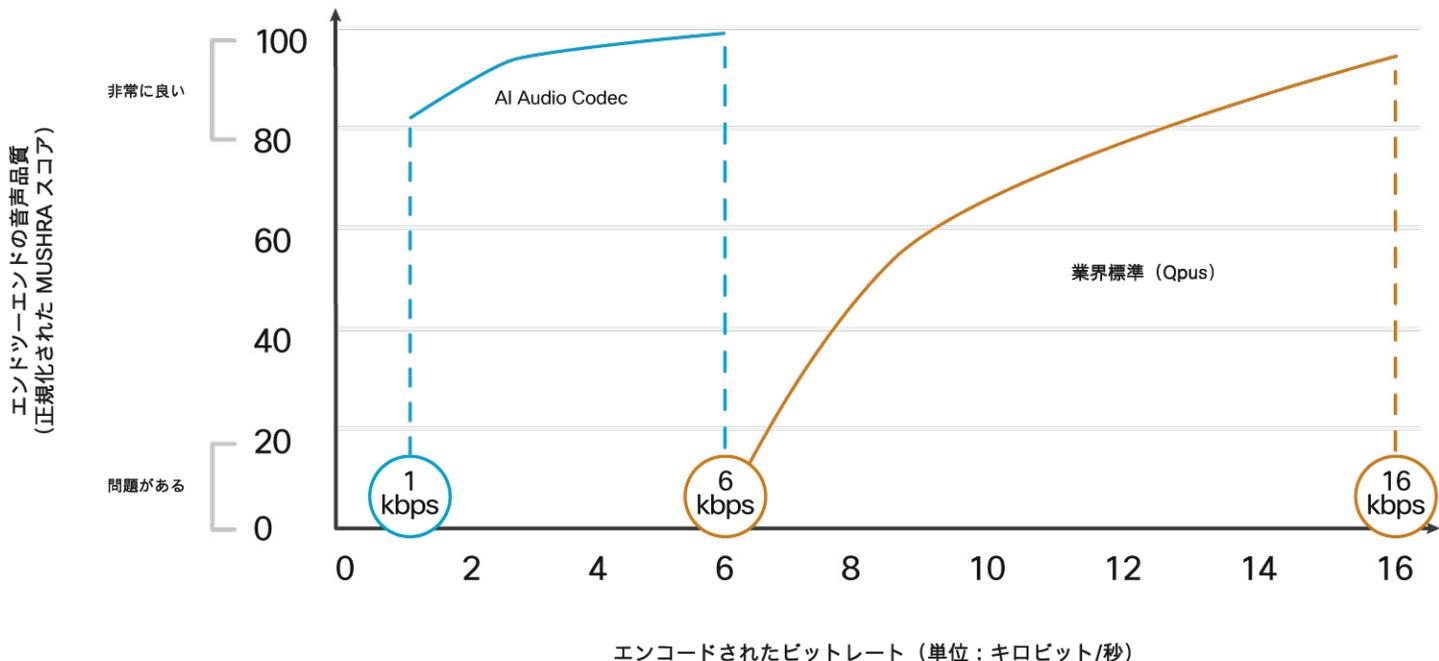


図 2 : Webex AI Codec のパフォーマンス

94%

通話品質が同等の業界標準のコーデックと比べて低減された帯域幅の使用量の割合

80 ポイント超

1、3、6 kbps で動作する Webex AI Codec の MUSHRA スコア

Webex プラットフォーム全体での Webex AI Codec の影響や用途

Webex AI Codec は汎用性が高いため、Webex Suite（会議、通話、ウェビナー）や Webex Contact Center など、あらゆる Webex のユースケースに適しています。また、Webex アプリ、Room システム、IP Phone、Desk Series デバイスなど、さまざまな Webex 対応エンドポイントとシームレスに統合できます。

さらに、Webex AI Codec の利点は Webex だけにとどまりません。音声合成（TTS）、自動音声認識（ASR）、言語翻訳、音声生体認証などの他の用途でも、その機能を活用できます。Compact Speech Representation は Webex AI Codec エンコーダで生成される音声フォーマットであり、他のシステムと簡単に統合できます。これにより、音声情報に依存するサービス間で低遅延かつ高精度のインタラクションが可能になり、異なるプラットフォーム間で効率的かつシームレスな通信が促進されます。

シスコの責任ある AI への取り組み

Webex AI Codec は AI ベースのイノベーションとして、[透明性](#)、公平性、説明責任、プライバシー、セキュリティ、信頼性を含む[シスコの責任ある AI の原則](#)に準拠しています。これらの原則は、要件の収集とモデルの評価から展開と文書化に至るまで、Webex AI Codec 製品ライフサイクルの全段階の指針となります。

表 1. シスコの責任ある AI の原則

シスコの責任ある AI の原則

透明性

AI は、大規模なデータセットと高度なアルゴリズムに依存しています。多くの場合、AI が意思決定にいつどのように関与するかは、ユーザーにとって不明瞭です。透明性はシスコの信頼の原則に含まれていて、このフレームワークの中核を成しています。そのためシスコでは、重大かつ必然的にお客様に影響をおよぼす意思決定で AI を使用している場合は、お客様に通知します。お客様とユーザーは通知を受けた後に、懸念事項や決定に同意しない旨をシスコに通知することができます。シスコは、コミュニケーションチャンネルをオープンに保つことで、AI サービスに対するお客様、ユーザー、従業員、その他の関係者からの信頼を構築、維持、強化することを目指しています。

シスコの目標は、シスコのテクノロジーに AI を採用する際のユーザーへの通知において、明確性と一貫性を確保することです。具体的には、AI の意図、モデルクラス、データのデモグラフィック、アクセス性と透明性、わかりやすさを備えた方法でモデルに適用されるセキュリティ、プライバシー、人権の監督について通知します。また、AI の使用に関する詳細情報を取得する方法も共有します。

公平性

AI を使用すると、人間の有害な偏見が技術的なシステムによって根付いたり、増強されたりする可能性が生じます。同時に、意思決定における有害な偏見や差別的な結果をより深く理解して軽減し、インクルージョンを促進するテクノロジーを開発する機会も生じます。より適切に意思決定を行うには、AI が適用される多様なあらゆる領域の全体で、トレーニングデータに個人または集団のデモグラフィックが確実に反映されるようにする必要があります。

シスコは、重大な判断（個人や集団に法的な影響または人権に関する影響をおよぼす可能性のある判断）に直接関与するアルゴリズム、トレーニングデータ、アプリケーション内の有害な偏見を特定して修復するよう努めています。また、シスコの責任ある AI フレームワークに必須の要素として、お客様がフィードバックを提供し、懸念を提起して、インシデント対応チームがレビューとアクションを実施するためのメカニズムも開発しました。シスコは、AI における技術進歩を含めた最新の技術進歩を反映するために、これらのプラクティスを定期的に更新しています。

説明責任

AI ソリューションとその開発チームに関する説明責任は、AI ライフサイクルの全体で責任ある開発と運用を行うために不可欠です。多くの場合、AI ツールには、開発時に予測できなかった想定外のユースケースや用途も含め、複数の用途があります。AI ソリューションを開発、展開、使用する企業に求められるのは、AI ソリューションが意図したとおりに動作することを保証し、不適切な使用を防止できるように、適切なガバナンスと管理機能の導入を通じて、こうした領域での業務に責任を負うことです。

Cisco Public 3 シスコは、そのグローバル人権ポリシーに明記されているように、あらゆる人の人権を守り、尊重することに取り組んでいます。シスコの責任ある AI フレームワークでは、開発の最初期から AI ライフサイクルの終わりまでわたって、プライバシー、セキュリティ、人権への影響に対して説明責任を負うことをチームに義務付けています。説明責任を負うための手段としては、AI のユースケースの文書化を義務付けること、影響の評価を実施すること、部門横断的なリーダーのグループが監督することなどがあります。

プライバシー

AI の利用時には個人データが使用されることが多く、適切に管理されなければ個人のプライバシーや市民の自由に影響がおよぶ可能性があります。AI で個人データを使用する場合、または個人を対象とした意思決定を行う場合、個人データの使用が許可され、目的に合致し、釣り合いが取れていて、公平であることを保証するための補助的なテクノロジーに、プライバシー管理機能を組み込む必要があります。こうした管理機能は、データとソリューションのライフサイクルの全体を通じて維持する必要があります。

シスコは、Cisco Secure Development Lifecycle (CSDL) にプライバシー エンジニアリングのプラクティスを組み込んでいます。これらのプラクティスは、シスコでプライバシーを強化する機能やプロセスを設計、構築、運用し、シスコの製品やサービスに組み込んでいることを保証するうえで役立ちます。シスコは、個人情報処理の際に、個人情報保護とプライバシーに関するグローバルポリシーに明記された原則に従うことに取り組んでいます。同ポリシーは、適用される国際的なプライバシー法および基準に準拠しています。

セキュリティ

AI システムに求められるのは、レジリエンスを備えていることと、標準的なソフトウェア開発と同様のセキュアな開発ライフサイクル管理機能を使用して、悪意のある攻撃者から保護されていることです。セキュリティ脅威からの保護には、従来型および敵対的な機械学習攻撃に対する AI システムのレジリエンステスト、脆弱性とサイバー攻撃に関する情報の共有、個人データのプライバシー、完全性、機密性の保護が含まれます。

シスコは、最先端のセキュリティプラクティスを使用して AI テクノロジーを構築し、セキュアな開発ライフサイクルを活用してレジリエンスと信頼性を最大限に高めています。また、AI に固有の特性に対応するために、独特なセキュリティ管理機能を AI に追加しました。これにより、攻撃へのレジリエンス、データ保護、プライバシー、脅威モデリング、モニタリング、サードパーティ コンプライアンスが向上します。

信頼性

AI ソリューションの有効性は、AI のさまざまな用途の全体にわたって、トレーニングに使用したデータセットと継続的な学習に使用したデータに基づき、当該のソリューションがどの程度確実に目的の出力を生成するかによって測定されます。AI ソリューションの主な特長の 1 つは正確性の向上であり、これはソリューションが体系的にテストされ、再現可能な結果が得られるように設計されている場合にのみ実現します。

シスコはイノベーションを優先し、AI システムとそのコンポーネントの設計とテストを行って信頼性を確保しています。また、シスコの責任ある AI アセスメントの一環として、さまざまな条件やユースケースで運用する際に目的と意図の一貫性を維持するために、AI ベースのソリューションのレビューを実施し、そのライフサイクルに管理機能が埋め込まれていることを確認しています。AI ソリューションが安全性に影響をおよぼす可能性があることを特定した場合は、さらに完全性の管理機能を埋め込みます。

透明性と説明責任を確保するために、Webex AI Codec がさまざまな Webex 製品にわたって導入される際に、Webex AI Codec の責任ある AI アセスメントに関する詳細を公開する予定です。これにより、テクノロジーとその責任ある AI プラクティスの遵守状況に関する理解が深まり、ユーザー間で信頼と確信が育まれます。

まとめ

通信システムの主な目的は、ノイズ、残響、マイクやヘッドセットのアーティファクトといった問題に対処し、ネットワーク障害に対する堅牢性を発揮して、クリアな音声を確保することです。

こうした課題に対処するために、シスコは Webex AI Codec を開発しました。これは、ノイズ除去機能やその他の音声の明瞭化機能を統合した AI ベースの音声コーデックです。Webex AI Codec は、1 kbps 程度の低ビットレートで、ほぼクリアな音質を実現します。これにより、レジリエンスの高い通信システムの導入が可能になります。

Webex AI Codec には独自のアプローチが取り入れられていて、エンコードされたパケットには、現在のフレームからの音声と、さらに低いビットレートで符号化された前のフレームのコピーが含まれています。この設計により、重大なパケット損失が発生した場合でも音声を再構築できるため、会議、通話、その他のインタラクションにおける情報の損失を最小限に抑えられます。

現在、Webex AI Codec では、複数の参加者を擁する Webex Meetings と特定のタイプの Webex Calling 通話においてトライアルを実施しています。

シスコはこの画期的なソリューションによって、AI の力を活用して音声通信における基本的な課題に対処し、信頼性と質が高く、コラボレーションを円滑化するテクノロジーを提供します。さらに、Webex AI Codec を含むどの AI イノベーションでも責任ある AI の原則が確実に遵守されるようにしています。

Webex AI Codec は、あらゆる人にとってインクルーシブな未来の実現を追求するうえでの最新の進歩そのものであり、多様な環境における質の高い通信に不可欠な構成要素となっています。

参考資料：

1. A. Rämö and H. Toukoma, "Subjective quality evaluation of the 3GPP EVS codec," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 2015, pp. 5157-5161, doi: 10.1109/ICASSP.2015.7178954.
2. J.-M. Valin, U. Isik, P. Smaragdis, A. Krishnaswamy, "Neural speech synthesis on a shoestring: Improving the efficiency of LPCNET," Preprint, February 2022
3. J. Dani and S. Srinivasan, "Satin: Microsoft's latest AI-powered audio codec for real-time communications," Microsoft Teams Blog webpage, February 2021, URL: <https://techcommunity.microsoft.com/t5/microsoft-teams-blog/satin-microsoft-s-latest-ai-powered-audio-codec-for-real-time/ba-p/2141382>
4. N. Zeghidour, A. Luebs, A. Omran, J. Skoglund, M. Tagliasacchi, "SoundStream: An End-to-End Neural Audio Codec," Preprint, 2021, <https://doi.org/10.48550/arXiv.2107.03312>
5. A. Défossez, J. Copet, G. Synnaeve, Y. Adi, "High Fidelity Neural Audio Compression," Preprint, October 2022, <https://doi.org/10.48550/arXiv.2210.13438>
6. A. Vasuki and P. Vanathi, "A review of vector quantization techniques," IEEE Potentials, July 2006, doi: 10.1109/MP.2006.1664069
7. R. Gray, "Vector quantization," in IEEE ASSP Magazine, vol. 1, no. 2, pp. 4-29, April 1984, doi: 10.1109/MASSP.1984.1162229
8. M. Kolundžija, M. Kavalekalam, I. Balić, M. Mao and R. Casas, "Low Bitrate Loss Resilience Scheme for a Speech Enhancing Neural Codec," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, 2024, pp. 1031-1035, doi: 10.1109/ICASSP48485.2024.10445746.
9. L. Lechler and K. Wojcicki, "Crowdsourced Multilingual Speech Intelligibility Testing," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, Republic of, 2024, pp. 1441-1445, doi: 10.1109/ICASSP48485.2024.10447869.

2024年9月



詳細情報

以下をご覧ください：webex.com/products/collaboration-ai.html