

ニュースリリース「AOI Pro. 『JAPAN FACTORY@SXSW』の開催決定」添付資料

2017年3月7日

## 2. Featured Speaker@SXSW2017 各セッションの詳細

### 1) 大阪大学 石黒浩教授 ・ NTT シニアリサーチサイエンティスト 東中竜一郎氏



大阪大学 石黒浩教授(上段左)と NTT シニアリサーチサイエンティスト  
東中竜一郎氏(上段右)、Repliee (リプリー) S1(下段左)と  
キュートな女性型アンドロイド“U”(下段右)

2017年3月12日(日) (現地時間 3:30PM-4:30PM)

#### “Humans and Robots in a Free-for-All Discussion”

2016年の“[Androids and Future life @ Featured Session, SXSW](#)”で全米の話題をさらった大阪大学の石黒浩教授が、続く二年目となる2017年、議論対話システムの研究者である東中氏 (NTT) と2体のAIロボット “Repliee (リプリー) S1” と “アンドロイドU” を引き連れ、更に進化を遂げて登場します。

イベント詳細 (SXSW Web サイト) : <http://schedule.sxsw.com/2017/events/PP95381>

### 2) 片付けコンサルタント 近藤麻理恵氏



2017年3月11日(土) (現地時間 12:30PM-1:30PM)

#### “Organize the World: Design Your Life to Spark Joy”

世界に“Kondo-ing”と呼ばれる「こんまり式」の片付けメソッドを広め、2015年タイムマガジン誌が選ぶ「世界で最も影響力のある100人」にも選出された近藤麻理恵氏がSXSW Featured Sessionに登壇。世界的なベストセラーとなった“Spark Joy”で提示された「ときめき」を感じるかどうかによる片付け法を通じ、今日の大量消費社会にお

SXSWに初参加の近藤麻理恵氏

ける本当の心の豊かさについてスピーチします。

イベント詳細 (SXSU Web サイト) : <http://schedule.sxsw.com/2017/events/PP95885>



予防医学の研究者 石川善樹氏

この他にも、JAPAN FACTORY がプロデュースを手掛けた予防医学の研究者・石川善樹氏が SXSU のオフィシャルセッションに選出。最新の AI 技術を応用した「味覚のパーソナライゼーション」をテーマに、味覚の好みや AI が熟知する食の未来についてスピーチを行います。

2017 年 3 月 13 日 (月) (現地時間 : 9:30AM-10:30AM)

“You Are What AI Cooks: Personalization of Taste”

イベント詳細 (Web SXSU サイト) :

<http://schedule.sxsw.com/2017/events/PP96857>

### 3. JAPAN FACTORY コンテンツ 参考資料



Repliee (リプリー) S1(左) と  
キュートな女性型アンドロイド “U”(右)

#### 1) NTT & 大阪大学石黒研究所

AI ロボットが言外のコンテキストを自ら判断し世間話ができるようになる時代において、人間と AI ロボットの関係性は、また、我々「人間」のあり方とはどう変化していくのかを探ります。

大阪大学と NTT が開発する議論対話システム及び NTT の音声技術のコラボレーションにより、実現した人間と自然な対話をするリプリー S1 とアンドロイド U を通して、少し先の人間とロボットの未来を覗いてみます。

## AI ロボットと人間の対話を支える大阪大学と NTT の技術\*1)

I) 「互いの意見を交換しながら合意形成を行う」議論対話システム (A) と、大阪大学の「ロボット同士の対話に人を巻き込む形で論理的な対話を成立させる」議論対話システム (B)

### A. NTT の議論対話システム :

大規模な構造化された議論データを対話知識として用いることで、特定の話題に関する深い議論を実現します。議論対話システムは、ユーザーの発話・状態を理解し、議論の状態を適切に管理することで、タイムリーに論拠の提示や反論を行います。

本技術は、人間とシステムの合意形成や人間の意思決定支援に役立ちます。

### B. 大阪大学の議論対話システム :

従来のロボットの対話研究では、単体のロボットが人間と対面し人間の発話意図に沿うように応答する対話システムの構築に注目が集まっていた。

これに対し、大阪大学が目指しているのが「複数のロボット」が人間と対面しロボット同士の対話に人間を巻き込む「ロボット主導」の社会的対話システムの構築です。

複数のロボットが主導する対話では、人間に対するロボットの応答が人間の意図から逸れるものとなっても、別のロボットがフォローすることによって対話全体が論理的に進行していると感じさせられるため、ロボットから人間への積極的な問いかけが可能になります。

複数のロボットが主導する対話をベースとすることで、可能なときだけ人の意図に沿い、そうでないときはロボットどうして話題展開をするという形で、人間とロボットのより自然な対話が実現されていきます。

II) ロボットが正しく「聞き取り」「認識をし」「発話する」ための音声技術 (NTT) - NTT の「耐騒音集音技術」「音声認識技術」「音声合成技術」

アンドロイドが人間と自然な対話をするためには、人間の発話内容を正しく聞き取り、発話の意味合いを文脈に沿って認識した上で人間らしい声で発話をする必要があります。

本展示における2つの議論対話システムを支える NTT の音声技術を紹介します。

### A. 耐騒音集音技術

NTT 研究所の開発した「インテリジェントマイク技術」を使用しています。これは複数の汎用小型マイクと音響信号処理技術の組み合わせで実現されています。音響信号処理技術では、音の到来方向、周波数特性、時間的な変動特性を利用することで、目的音の劣化を抑えつつ、周囲雑音のパワーを低減させ、目的音のクリアな集音を可能とします。この技術を利用することで、騒音環境下でも高精度な音声認識を可能にしました。

### B. 音声認識技術

最新の Deep Learning (深層学習) を活用した音響モデリングと、WFST (重み付き有限状態トランスデューサ) による最適な探索ネットワークにより、大語彙でも高速・高精度な音声認識を可能にしました。

### C. 音声合成技術

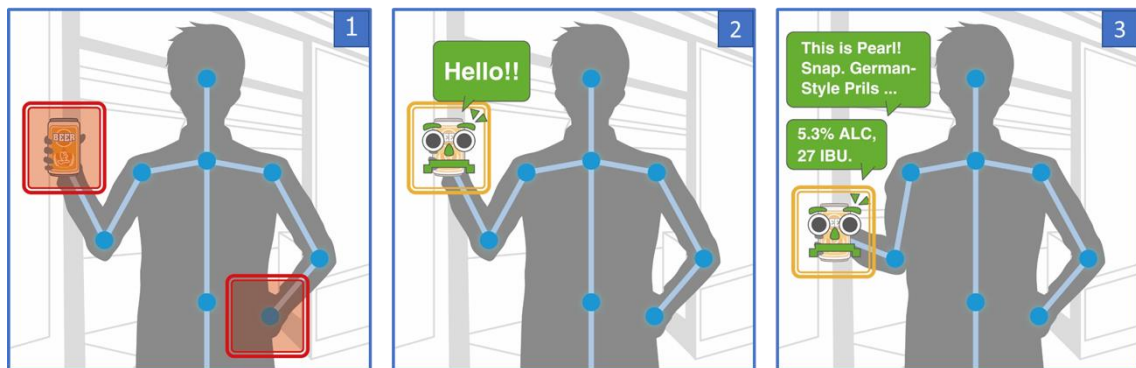
人の声を特徴付ける要素である「話者」と「口調」を自由に組み合わせることで、バリエーション豊かな音声作成を実現する技術です。

本展示では、これらの音声技術と2つの議論対話システムを組み合わせることで、アンドロイドと人間の自然な対話を実現します。

### 2) NTT DATA 参考資料

本デジタルサイネージの前に立つと、モーションキャプチャ技術が手の部分を認識してビールの銘柄を特定します。本サービスを活用することで、ユーザーに意識させることなくその人に合った情報の提供ができるため、記憶に残る新たな広告体験を生み出します。本技術を支えるのが、NTTグループのAI技術である corevo<sup>®</sup>\*2) を構成する技術のひとつである「アングルフリー物体検索技術」です。

ビールの銘柄が特定されるとその銘柄に対応したキャラクターがAR技術を活用してビール画像上に登場し、当該ビールの情報を紹介します。また、本サービスは複数ビールの同時認識も可能で、各ビールに対応したキャラクター同士が挨拶する等、異なる商品が互いに協調しながらビールの商品情報を楽しく提供します。ユーザーが身につけている情報とNTTの物体認識・AR技術を結びつけることで、新たな広告体験を創り出すことができます。



「Linked Virtual Ad. ~Talking Beer~」の実現イメージ

### 3) JAPAN FACTORY Original Installation

タンジブルなインタラクションを介在させること、またデジタルとアナログの境界をあいまいにしシンクロさせることで、「世の中のあらゆる事象は対峙するふたつの概念を同時に孕んでいる」ということ。そして、「それらはいつも人々の身体的、感情的な動きによって変化し続ける」ということをシンボライズしました。4つのスクリーンで同時多発的に起きるアイデンティティの生成・解体・再構成が、人が実際に触れることでさらに加速し、拡散していきます。

### 4) CYBER TELEPORTATION TOKYO at SXSW powered by Kirari!<sup>®</sup>\*2)

本展示では、NTTのイマーシブテレプレゼンス技術 Kirari!<sup>®</sup>\*2)による任意背景リアルタイ

ム被写体抽出技術および超高臨場感メディア同期技術が活用されており、東京側でパフォーマンスを行うアーティストを、グリーンスクリーン等のスタジオ設備を用いない任意背景下においてリアルタイムに被写体抽出することや、また、抽出された複数アーティスト映像や複数音声といった多数のメディアをオースティンへ同期を保ったままで低遅延に伝送することが可能となります。

また、東京よりオースティンへ伝送されたアーティストの映像は、会場内に設置された多数の透過型スクリーンを使用し、裏表の表現、奥行き感、ダイナミックな動きを伴いながらプロジェクションされます。このような空間映像演出により、東京のアーティストと、オースティンの starRo とが同期し、一体となったパフォーマンスを行うという、これまでにない遠隔ライブ体験を提供いたします。



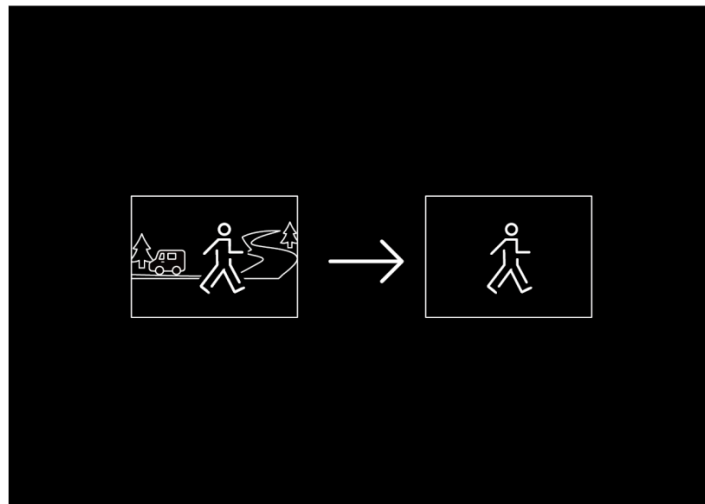
オースティンのライブ会場イメージ

なお、本ショーケースに用いる東京-オースティン間的高速ネットワークは、NTT が保有している研究開発用ネットワーク“GEMNet2”に加えて、米国の大学コンソーシアム Internet2 が運用している全米研究教育ネットワーク、テキサス州の地域研究教育ネットワーク LEARN (Lonestar Education and Research Network)、テキサス大学システム (University of Texas System)、及びオースティン地区広域ネットワーク (Greater Austin Area Telecommunication Network) の協力のもとで実現されています。

## CYBER TELEPORTATION TOKYO at SXSW を支える NTT の技術

### I) 任意背景リアルタイム被写体抽出技術

グリーンスクリーン等のスタジオ設備を用いずに、撮影した映像中からリアルタイムに被写体を抽出する画像処理技術です。今回のアーティスト映像の被写体抽出では、スタジオ壁面等の背景に特別な加工を施さず、ライブパフォーマンスの演出に従って照明等が変化するような映像を入力として、被写体部分のみリアルタイムかつ精緻に抽出します。この抽出した映像を、被写体が演じている場所とは異なる場所に等身大で投影することで、伝送先でアーティストをリアルに再現します。



## II) 超高臨場感メディア同期技術 (Advanced MMT)

複数の映像や音声をネットワーク経由で伝送する際に、それらの同期を厳密に保ったまま伝送先で再生する技術です。従来に比べて送出処理時間を低減し、再生機器処理時間まで含めた遅延時間の大幅な削減を可能としました。

今回は、東京でパフォーマンスするアーティストを複数アングルから撮影した映像を、音声とともに本技術を用いてオースティンに国際伝送し、両面を含む多数の透過型スクリーンにフレーム単位で同期させながらプロジェクションすることにより、アーティストのパフォーマンスを自然な形で再現します。またオースティン会場の模様を低遅延で東京に伝送することでグローバル・インタラクティブ遠隔ライブパフォーマンスを実現します。



※1) NTTグループのAI技術「corevo®」を活用しています。

※2) 「corevo®」、「Kirari!®」は日本電信電話株式会社の登録商標です。