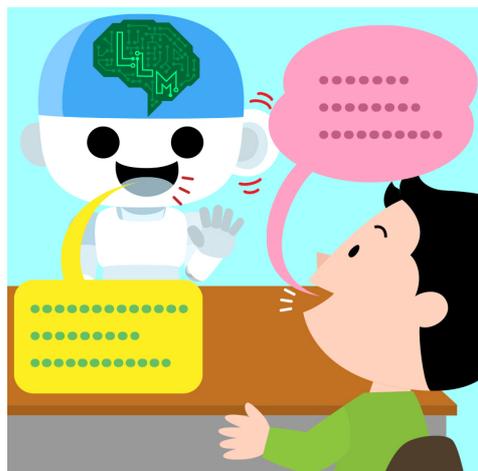


聴覚・音声を持つ「tsuzumi」

IOWN Pick Up NTT版大規模言語モデル



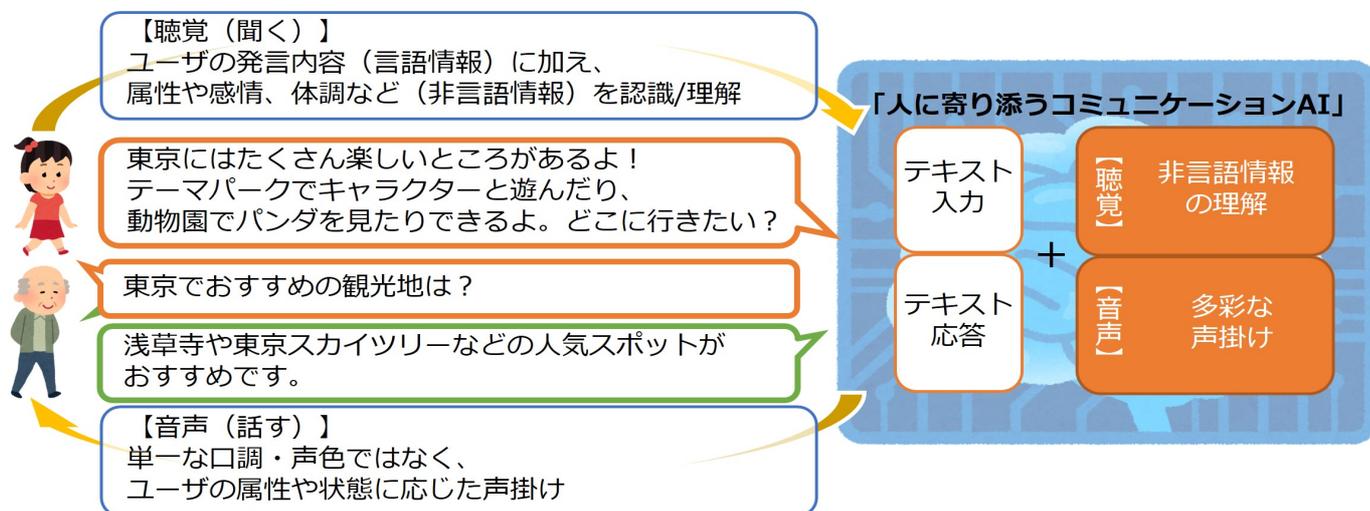
背景

人間同士のコミュニケーションでは、相手の感情／体調をくみ取って声の掛け方や話しかける内容を変えたり、年齢によって言葉遣いや声色を変えたりしています。一方、現状のLLMは主にテキスト情報を扱うため、相手に寄り添うコミュニケーションはできていません。

成果の概要

人に寄り添うコミュニケーションAIの実現に向け、下記を自在に扱えるLLMに取り組んでいます。

- ・聴覚（聞く）：発言内容（言語情報）に加え、属性や感情、体調など（非言語情報）の認識／理解
- ・音声（話す）：単一口調・声色ではなく、ユーザの属性や状態に応じた声掛け



技術のポイント

- 人と円滑にコミュニケーション可能なAIの実現をめざした、非言語情報を含む音声を自在に操れるLLM
- 利用者の背景やコンテキスト情報の活用に向けた、スマートデバイスでも常時リアルタイム動作可能な音声認識技術
- 発話内容や感情ラベルに沿って多種多様な口調・声色の話し方を推定／生成可能な音声合成技術

この研究がもたらす未来

AIによる豊かな社会の実現に向け ①カウンセリングなどで「人の内面」を理解・共感し寄り添えるLLM
②使用者の状態の常時把握により、暮らしやすくするLLM の実現をめざします。

出展企業

日本電信電話株式会社

問い合わせ先

rdforum-exhibition@ml.ntt.com