# modality.ai



## Leveraging Multimodal Dialog Technologies & AI/ML for Patient Health Diagnosis, Monitoring, and Intervention

Vikram Ramanarayanan, David Suendermann-Oeft, David Pautler, Oliver Roesler, Jackson Liscombe, Michael Neumann, Doug Habberstad & Andrew Cornish

Using conversational speech and video technologies in conjunction with signal processing and machine learning algorithms can allow patients to interact with an engaging virtual agent in the comfort of their homes, producing various automatically-computed, clinically-validated measures of disease onset and progression that can help those patients, as well as their supervising physicians and caregivers in diagnosis, monitoring and understanding.

### ality.ai Health Monitoring Session

Health Monitoring by Artificial Intelligence

is conducting alpha-testing on a product prototype that will be used to monitor health conditions. You will be asked to use a computer t make a video-call with an automated system in which your facial movements and speech will be recorded. There is no clinician involved in the study. gally-authorized representative (LAR): Unfortunately we cannot accept participants who need a legally-authorized representative (LAR). If you are a LAI

please exit by closing this page.

Our study: In this study, you interact with a conversational agent, named Nina, using this website. Both audio and video of the conversation will be recorded. Once you have completed your responses, Nina will say goodbye, then she'll hang-up and the video and audio recordings will stop. One call will take between 5-10 minutes. You will not receive feedback about your responses. There is no monetary compensation.

Your privacy: Your responses, including audio and video recordings, will be used to inform future research and development. Personally identifying information collected will include a video recording and an audio recording. Additionally, participants might be asked to specify other demographics such as their gender age, and the city where they have lived for the majority of their life. Identifying information, such as names or phone numbers, will not be collected or shared, although we may provide an option for you to provide your email address for us to follow-up with you. Individual responses will only be shared with Modality employees and other Modality partners who are compliant under the Health Insurance Portability and Accountability Act (HIPAA).

Your rights: Your participation is voluntary, and you are free to withdraw at any time by hanging-up and closing this page.

Study confidentiality: By participating in this study, you must agree to maintain the full confidentiality of all study materials. Specifically, you must agree not to reproduce and/or disclose any study materials in whole or in part to any person or entity.

Contact us! If you have questions or concerns, please contact us by clicking the email address below.

When your call is complete, we will press the Hangup button for you and go to the survey.

This activity collects your speech and video data, and hence requires you to have a working webcam, microphone, and speaker installed on your computer.

By calling into our system, you affirm that you are 18 years of age or older and agree to the conditions printed and referenced above. You understand that if you reproduce the study materials in any manner, you may be subject to legal action(s) Modality.AI may take in order to protect its intellectual property rights.

### Get Started

After your call, there will be a survey where you can tell us if there were any technical problems. If you can't get that far, you can also tell us about them by emailing us at help@modality.ai and giving us this session ID: 389d1689-16c7-4dd0-8a65-669673339fe6

**Health Monitoring Session** Health Monitoring Session Video call guestionnaire Did you have any technical problems during your call? Video call No Yes, what happened was If an underlined word isn't familiar, move your mouse over it to reveal a definition. Health Monitoring Session The Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) Normal speech processes Detectable speech disturbance Intelligible with repeating Speech combined with nonvocal communication Loss of useful speech

If an underlined word isn't familiar, move your mouse over it to reveal a definition

## Diagnosis & monitoring of neurological health in patients remain a critical need and challenge

- lack of access to neurologists or psychiatrists
- lack of awareness of a given condition and the need to see a
- lack of an effective standardized diagnostic or endpoint for many of
  - these health conditions
- substantial transportation and cost involved in conventional or
  - traditional solutions

specialist

Patients/clinicians

secure web link to

Users guided to set

up their webcam and

interacting with agent

Users interact with

out clinical surveys

an agent and fill

microphone before

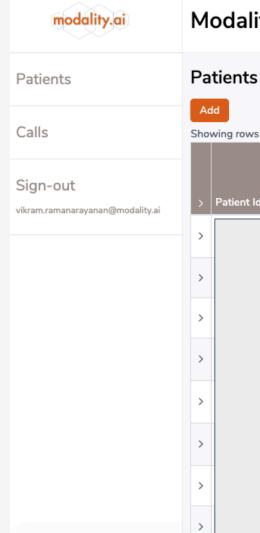
the assessment

• severe shortage of medical specialists in these fields to begin with

- - over the phone, smartphone app, or web browser

### Analytics modules extract speech and video metrics automatically are provided with a

- Speech: Video: • opening, width, displacement, velocity, • speaking rate • intelligibility acceleration and jerk of the upper and lower lips • mean pausing duration • mouth surface • mean F0 • mean symmetry ratio between right and left mouth duration surfaces • vertical positions of eyebrows Verification of speech • eye opening and video metrics • head tilt translation and rotation vectors Verification corpus (11 interactions) collected using AMT Automatic extraction of speech metrics through voice activity Modality:Demos modality.a and pitch detection algorithms Patients 🖓 Patients • Automatic video brightness and blurriness checks to instruct Add Calls Showing rows users to fix low video quality • Automatic face and facial landmark detection Sign-out Calculation of facial metrics using 14 of the 68 detected facial landmarks
- Ground truth labels obtained through manual annotations of:
  - number of words and speech duration
  - facial markers required for facial metrics calculation
- Preliminary results show the automatically-computed metrics are very similar, i.e. above 80%, to the ground truth



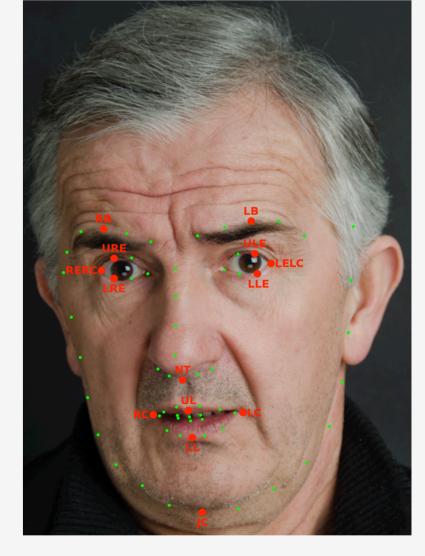
## Optimize drug trials through multimodal AI



ALCHEMIST ACCEL FRATOR

## NEMSI bridges this gap!

• NEurological and Mental health Screening Instrument cloud-based multimodal dialog system that conducts automated screening interviews • elicits evidence required for detection or progress monitoring • NEMSI makes novel contributions along three significant directions: • uses end point devices available to everyone everywhere • deployed in an automatically scalable cloud environment allowing it to serve an arbitrary number of end users at a very small cost per interaction natively equipped with real-time speech and video analytics modules that extract a variety of features of direct relevance to clinicians



### Dashboard provides easy visualization!

to	19	of	19	

1↑↓	Access code ↑↓	Patient type ↑↓	Provide access ↑↓	Session Date (UTC) ↑↓	ALSFRS-R score ↑↓	Speaking rate (words/minute including pauses) ↑↓	Articulation rate (words/minute excluding pauses) ↑↓
	j462w0	Crowdsourced testers (xx2x)	Re-invite 🖸	11/27/2019 18:44 <b>&gt;</b>	44/44 🗸	211.99 🗸	231.68 🗸
	w9egv3	Modality tester (xx0x)	Invite 🖸	02/08/2020 17:52			
	4f03au	ePHI (xx3x)	Re-invite 🖸	01/06/2020 20:49	No survey responses (1)	No speech metrics (1)	
	t3cx06	Modality tester (xx0x)	Re-invite 🖸	12/16/2019 22:46 (Þ)	48/48	222.02	297.17
	sr1cyz	Internal tester (xx1x)	Re-invite 🖸	01/24/2020 23:50 <b>&gt;</b>	48/48	214.39	232.08
	cc0db8	Modality tester (xx0x)	Re-invite 🖸	11/21/2019 23:33 (Þ)	No survey responses (4)	150.56	170.32
	xt034d	Modality tester (xx0x)	Re-invite 🖸	02/03/2020 23:20	No survey responses (4)	No speech metrics (2)	
	o6gjxq	Internal tester (xx1x)	Re-invite 🖸	01/21/2020 23:10 <b>&gt;</b>	48/48	145.63	197.33